



---

# **Week 8 Lecture - Bivariate Correlation**

Undergraduate Research Methods in Psychology

---

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

Department of Psychology

## Table of Contents

<b>1</b>	<b>Last Week Review &amp; Announcements</b>	<b>2</b>
1.1	Announcements and Due Dates . . . . .	2
1.2	Last Week Content . . . . .	2
<b>2</b>	<b>Exam Review</b>	<b>2</b>
2.1	Areas for Review . . . . .	2
<b>3</b>	<b>Quiz 7</b>	<b>2</b>
3.1	Quiz Content . . . . .	2
3.2	Quiz Rules . . . . .	3
<b>4</b>	<b>Learning Objectives</b>	<b>3</b>
4.1	Textbook Objectives . . . . .	3
4.2	Professor's Objectives . . . . .	3
<b>5</b>	<b>Chapter Overview</b>	<b>4</b>
5.1	Chapter Overview . . . . .	4
<b>6</b>	<b>Bivariate Correlation</b>	<b>4</b>
6.1	Overview . . . . .	4
6.2	Association Between Two Continuous Variables . . . . .	4
6.3	Association Between Categorical and Continuous Variables . . . . .	5
6.4	Fully Correlation Design . . . . .	7
<b>7</b>	<b>Critical Association Claims</b>	<b>7</b>
7.1	Construct Validity . . . . .	7
7.2	Statistical Validity . . . . .	8
7.2.1	Effect Size . . . . .	8
7.2.2	Precision . . . . .	8
7.2.3	Replication . . . . .	8
7.2.4	Outliers . . . . .	9
7.2.5	Restriction of Range . . . . .	9
7.2.6	Curvilinear Relationships . . . . .	10
7.3	Internal Validity . . . . .	10
7.4	External Validity . . . . .	12

# 1 Last Week Review & Announcements

## 1.1 Announcements and Due Dates

- Keep turning in reading evidence if you are reading the textbook and taking good notes (which you should be doing either way!)
- Remember that you can give anonymous instructor feedback to me at anytime via the Qualtrics survey link under the "Getting Started & Course Overview - BEGIN HERE" header in Blackboard.
- No class on Tuesday, October 22nd - Fall Break! Make sure to spend part of it relaxing and recharging, everyone.
- I will have office hours again at 3:00pm - 6:00pm EST in AuSable 1307 on Friday 10/18/2024. Please consider coming if you have recently been struggling.

## 1.2 Last Week Content

- We took the exam 1 / midterm (bleh!) - we will discuss momentarily.
- We discussed various sampling methods and techniques, and their relative strengths and weaknesses in extracting a representative sample from a population of interest.
- We talked about how to understand external validity as a function of the sampling technique used, and when this validity is most important.

# 2 Exam Review

## 2.1 Areas for Review

- Will go over in class!

# 3 Quiz 7

## 3.1 Quiz Content

- Covers all content from 10/8 class meeting, including but not limited to:
    - Chapter 7 of Morling Textbook
    - Lecture on Chapter 7
  - *Any last minute questions?*
-

## 3.2 Quiz Rules

- *From the Syllabus:*
  - Each quiz is 10 multiple-choice questions, 1 point for each question
  - Quizzes will be taken at the start of the class period on the Blackboard LMS
  - Quizzes will be on content covered in the previous lecture and the associated reading for that lecture
  - Quizzes are timed, 23 minutes only (previously was 15 minutes). If you finish before time is up, please remain in class and find another activity to work on quietly
  - Quizzes are open-note and open-book, that is, you are allowed to use those resources during the quizzes. Thus, they reward good structure in thoughtfulness in your notes and preparation
  - You may not collaborate with others during the quizzes, or discuss questions with other students after the quiz. You cannot use AI tools or the internet to help you during the quiz
  - Quizzes and exam will be ended early if all students are clearly finished and content with their answers
  - Quizzes will be graded promptly and reviewed the following week

## 4 Learning Objectives

### 4.1 Textbook Objectives

- Explain that measured variables - not any particular statistic - make a study correlational.
- Interrogate the construct validity and statistical validity (and, of lower priority, external validity) of an association claim.
- Explain why a correlational study can support an association claim, but not a causal claim.

### 4.2 Professor's Objectives

- Be able to understand the defining characteristics of a correlational, bivariate research design
  - Be able to use appropriate vocabulary and logic used in describing and assessing bivariate designs
-

## 5 Chapter Overview

### 5.1 Chapter Overview

- Moving forward, much of the content will be about applying chapters 1 - 7 to different research designs. Make sure you review vocab from the first half of class.
- Bivariate → \_\_\_\_\_ variables
- Two, **measured** variables → likely \_\_\_\_\_ claim
- *Review*: Remember the vocabulary we should expect with associative claims between
  - E.g., “linked”, “associated with”, “correlated with”, etc.
  - *Discuss*: Any other examples of vocab normally used with associative claims?

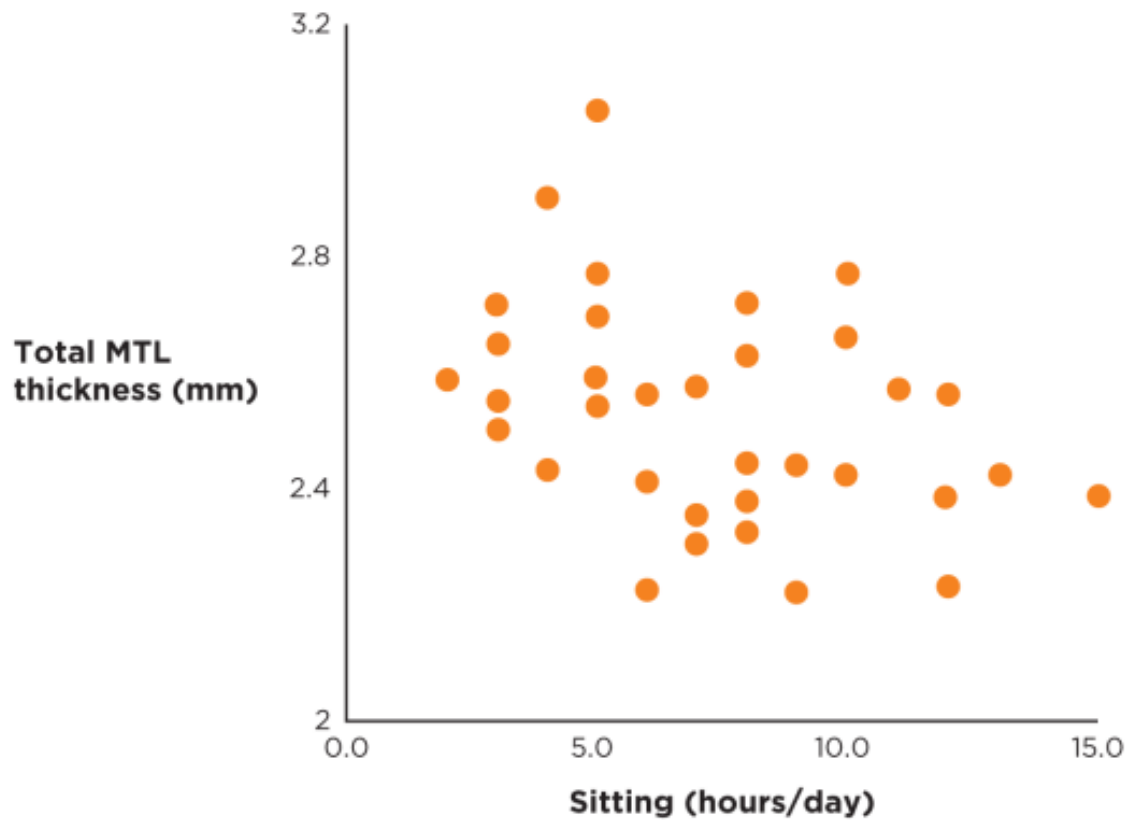
## 6 Bivariate Correlation

### 6.1 Overview

- *Review*: Recall that there are three possible associations / relationship descriptions between two variables:
  - Positive
  - Negative
  - \_\_\_\_\_

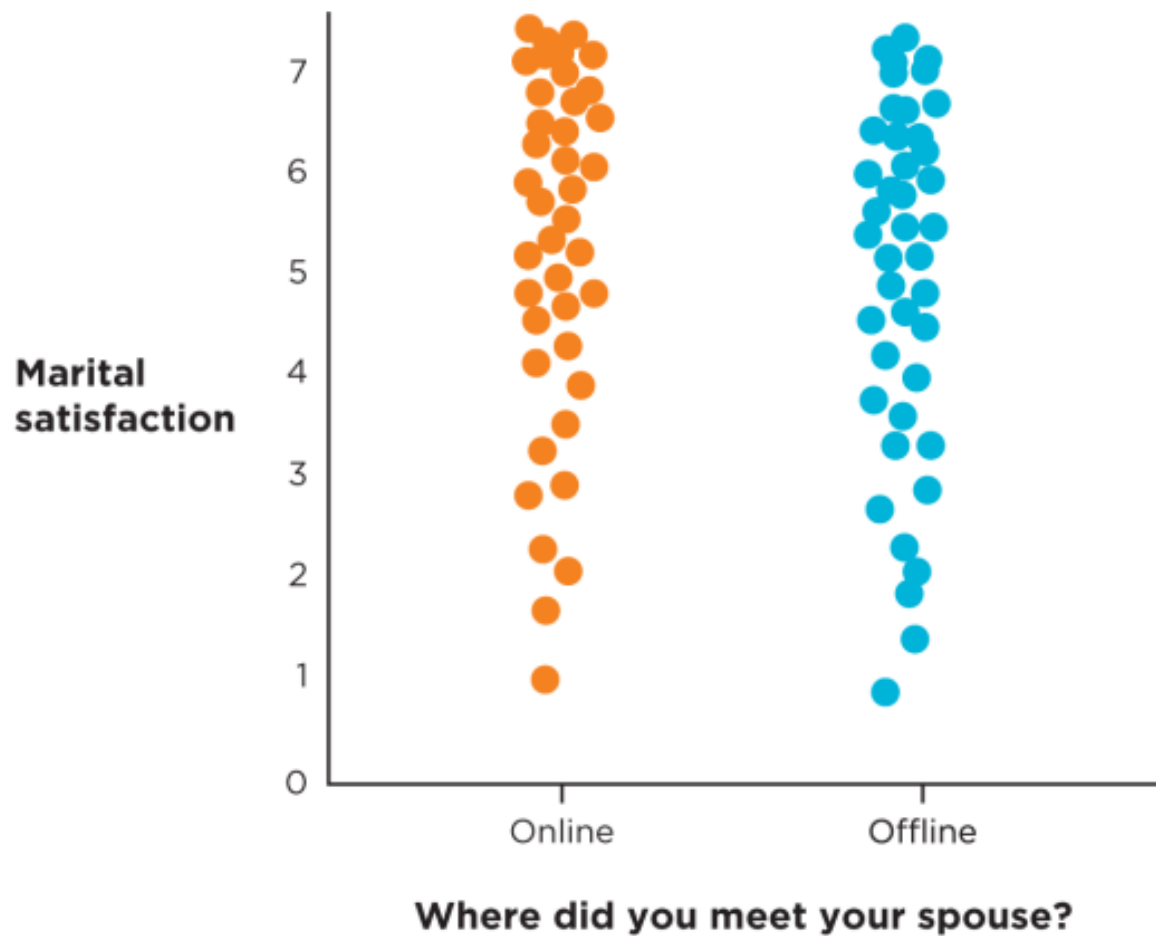
### 6.2 Association Between Two Continuous Variables

- Example:
    - Two self-report measurements are strongly, positively correlated with one another
    - *Review*: what type of validity is this possibly associated with? \_\_\_\_\_
    - *Discuss*: what statistic do we use for association between two continuous variables, and what value would indicate a strong, positive relationship?
  - *Review*: When using  $r$  correlation coefficient, \_\_\_\_\_ is the most appropriate graphing method
-

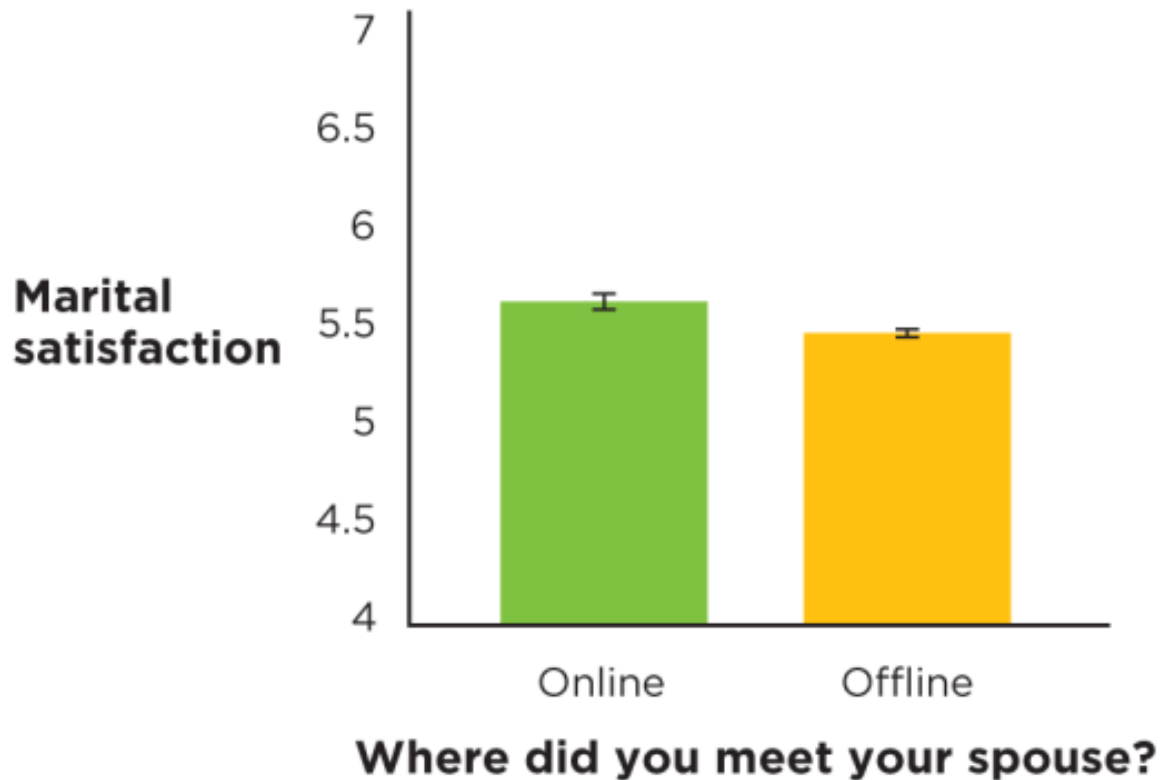


### 6.3 Association Between Categorical and Continuous Variables

- What do we do when we have one categorical/discrete and one continuous variable?
- One option: A sort of “split” scatterplot by the categorical variable



- *Discuss:* What is the direction and strength of relationship here?
- In most cases, we will probably prefer other methods to graphical display information when we have categorical and continuous variables.
- Another option: A \_\_\_\_\_



- *Discuss:* What do you all think those small, black bars mean? What sorts of helpful information could they tell us about differences between the groups?

## 6.4 Fully Correlation Design

- A \_\_\_\_\_ design is one that results from all relevant variables being measured, *not* based on what scale of measurement was used, statistic, or graph.
- The designs that we discuss are largely defined by the whether variables are measured or manipulated *and* the relationship we hypothesize between the variables (i.e., what claim we are attempting to make)

# 7 Critical Association Claims

## 7.1 Construct Validity

- *Review:* construct validity is all about the measurement reliability and validity of the tools we use to measure a phenomenon.
  - 3 reliabilities: test-retest, \_\_\_\_\_, and internal
  - 5 validities: face, content, criterion, convergent, divergent/discriminant



- Use the terms and suggestions from chapters 5 and 6 to assess this!
- *Discuss*: One of the items on one of my measures may not be measuring the same constructs as the others due to questionable use of the word “not” in the item.
  - How would we identify an item not well correlated with other items of the same construct? \_\_\_\_\_
  - What vocabulary do we use to describe this problem? \_\_\_\_\_

## 7.2 Statistical Validity

- Statistical validity has multiple components to be looking for:

### 7.2.1 Effect Size

- Assessing  $r$  correlation coefficient effect
  - General guideline: 0.10 (weak), 0.20 (moderate), 0.30 (powerful), >0.40 (very powerful) - Table 8.4 on pg 474
- We may also use \_\_\_\_\_
  - This refers to the amount of variance that one variable accounts for in another
- Low effect → possible negligible or non-meaningful effect
- High effect → possible meaningful/impactful effect

### 7.2.2 Precision

- *Review*: What do we call the intervals in which 95% of estimates will fall? \_\_\_\_\_
- Common notation: [Low-end Estimate, High-end Estimate]
- Larger sample size → \_\_\_\_\_ confidence intervals
- More precision = “better”, more accurate estimate

### 7.2.3 Replication

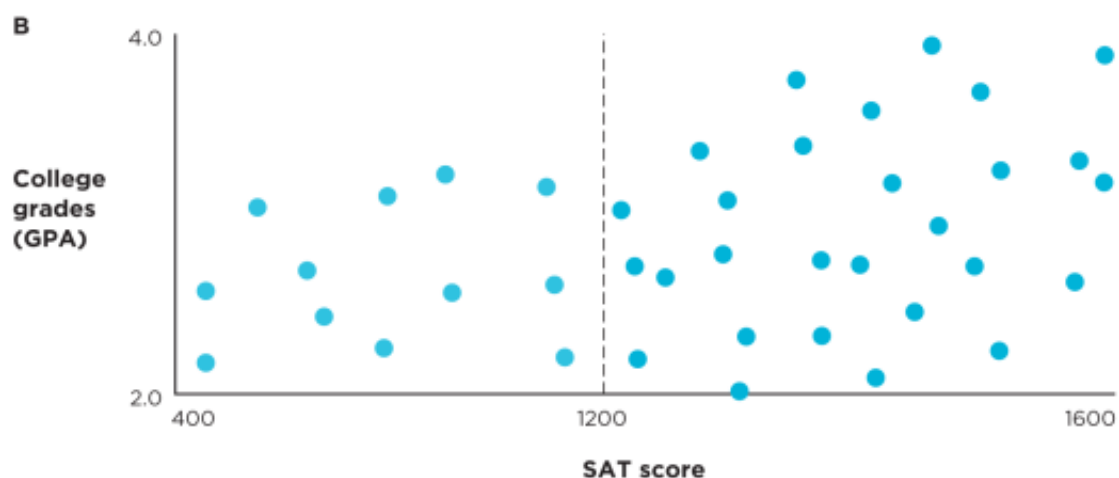
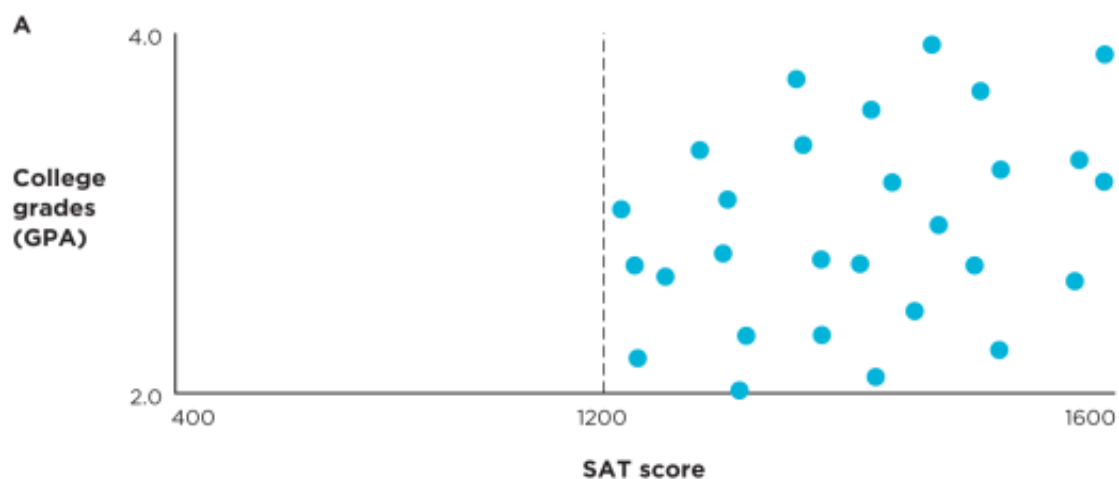
- Has this association been shown before in the literature?
  - We may \_\_\_\_\_ the study and measurements again on a new group and see if we get a similar value
  - *Review*: This is a good use case for literature review and meta-analysis articles, as they can compile different studies investigating the same association
-

### 7.2.4 Outliers

- *Discuss:* Based on what you all know from previous classes, what is the definition of an outlier?
- An outlier, especially one \_\_\_\_\_ on both variables, can greatly sway the relationship of two variables.
- However, we should not simply throw out outliers because they are inconvenient findings.
  - Instead, we may consider **robust/non-parametric** statistics that can work around them as is

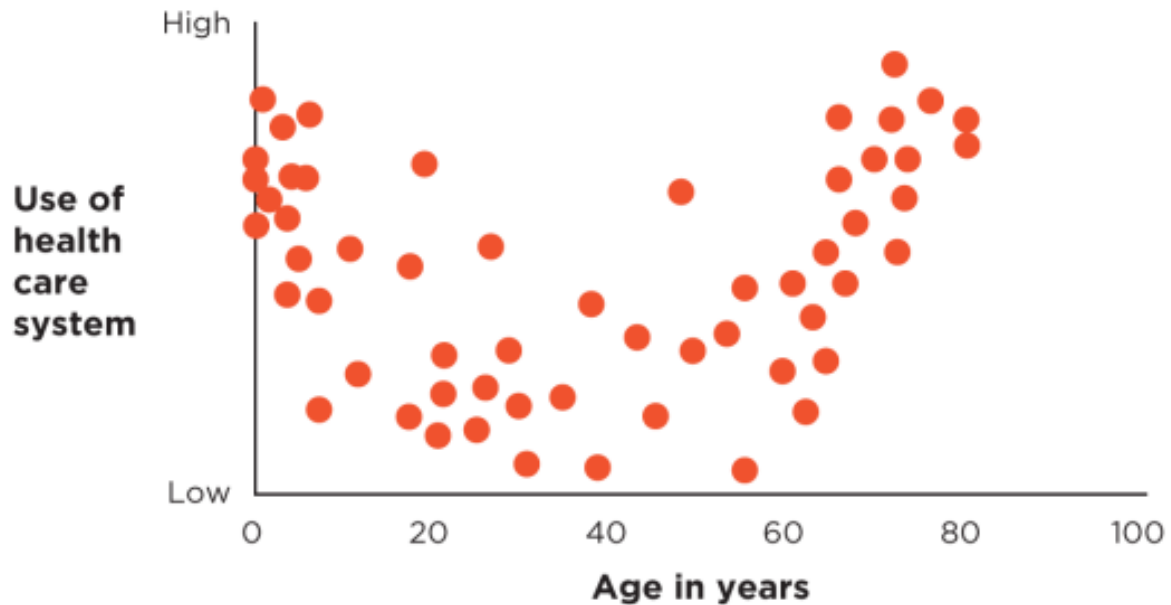
### 7.2.5 Restriction of Range

- This is when we do not have points across the full number of possible levels on one or both of our variables



### 7.2.6 Curvilinear Relationships

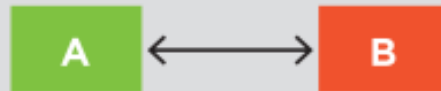
- $r$  assumes a \_\_\_\_\_ relationship between the two variables, but this is not true of all relationships in reality



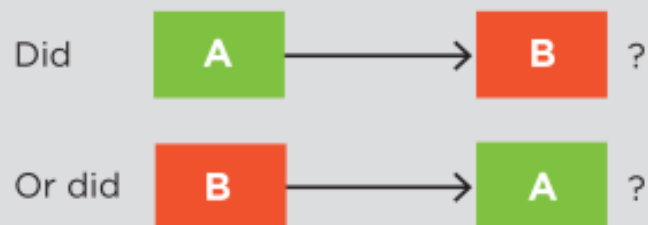
### 7.3 Internal Validity

- For association claims, we do not need to be concerned with meeting standards of internal validity, because a correlative design cannot explore causal relationships.
  - Review: What design can investigate causation? \_\_\_\_\_
- Review: What is the common phrase we use to describe the relationship between correlation and causation?
  - \_\_\_\_\_
  - Put simply, association alone does not tell us one “caused” the other!
- Discuss: What were the 3 causation criteria to establish a causal claim? Describe each.
  - Covariance
  - \_\_\_\_\_ - **directionality problem**
  - Internal validity - **third-variable problem**

1. *Covariance*: Do the results show that the variables are correlated?



2. *Temporal precedence* (directionality problem): Does the method establish which variable came first in time?



(If we cannot tell which came first, we cannot infer causation.)

3. *Internal validity* (third-variable problem): Is there a C variable that is associated with both A and B, independently?



(If there is a plausible third variable, we cannot infer causation.)

- Example: I am measuring the relationship between exercise and feeling of self-worth, both measured as continuous
  - High, positive correlation between the two - can I say exercise causes higher self-worth? \_\_\_\_\_
  - Say energy level is correlated with both exercise and feeling of self-worth, so what really is the cause?

- What is the so-called *third-variable* here? \_\_\_\_\_

## 7.4 External Validity

- What matters more for “good”, representative sampling - size of sample or methodology? \_\_\_\_\_
  - Bad external validity does not necessarily invalidate a claim, but just limits its generalizability
  - *Discuss:* I use quota sampling to gather my participants, and find an  $r = 0.60$  relationship between the variables
    - What direction and strength of relationship is this?
    - Is this a biased or unbiased sample?
  - \_\_\_\_\_ are some third variable that plays a partial role in the relationship between two variable
    - These may result in a broader, more complicated theory around certain constructs
    - But most psychological constructs are complex!
-