



---

# **Week 9 Lecture - Multivariate Correlation**

Undergraduate Research Methods in Psychology

---

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

Department of Psychology

## Table of Contents

<b>1</b>	<b>Last Unit Review &amp; Announcements</b>	<b>3</b>
1.1	Announcements and Due Dates . . . . .	3
1.2	Last Unit Content . . . . .	3
<b>2</b>	<b>Quiz 7 Review</b>	<b>3</b>
2.1	Areas for Review . . . . .	3
<b>3</b>	<b>Quiz 8</b>	<b>4</b>
3.1	Quiz Content . . . . .	4
3.2	Quiz Rules . . . . .	4
<b>4</b>	<b>Learning Objectives</b>	<b>4</b>
4.1	Textbook Objectives . . . . .	4
4.2	Professor's Objectives . . . . .	5
<b>5</b>	<b>Chapter Overview</b>	<b>5</b>
5.1	Chapter Overview . . . . .	5
<b>6</b>	<b>Review of the Casual Criteria</b>	<b>5</b>
6.1	Overview . . . . .	5
<b>7</b>	<b>Longitudinal Designs and Temporal Precedence</b>	<b>5</b>
7.1	Overview . . . . .	5
7.2	Cross-sectional Correlation . . . . .	6
7.3	Autocorrelations . . . . .	6
7.4	Cross-lag Correlation . . . . .	6
7.5	Longitudinal and Causation Criteria . . . . .	7
7.6	Why Not Experiment . . . . .	7
<b>8</b>	<b>Multiple Regression and Third Variables</b>	<b>7</b>
8.1	Overview . . . . .	7
8.2	Criterion and Predictor Variables . . . . .	7
8.3	More Than Two Variables . . . . .	8
8.4	Statistics for Third Variables . . . . .	9
8.5	Beta coefficients . . . . .	9
8.6	More Predictors, The Merrier? . . . . .	9
8.7	In Popular Media . . . . .	9
8.8	Still Not Causal . . . . .	9
<b>9</b>	<b>Pattern and Parsimony</b>	<b>10</b>
9.1	Overview . . . . .	10
9.2	Meaning . . . . .	10

9.3	In Popular Media . . . . .	10
<b>10</b>	<b>Mediation</b>	<b>10</b>
10.1	Overview . . . . .	10
10.2	Compared to “Third Variable” . . . . .	11
10.3	Compared to “Moderators” . . . . .	11
<b>11</b>	<b>Analysis with the Four Validities</b>	<b>12</b>
11.1	Overview . . . . .	12

# 1 Last Unit Review & Announcements

## 1.1 Announcements and Due Dates

- Don't forget about the **reading evidence**!
- Make sure to be making consistent progress on your project work (**article critique and research proposal**) outside of class!
- I will have office hours again at 3:00pm - 6:00pm EST in AuSable 1307 on Friday 11/01/2024. Please consider coming if you have recently been struggling or would like to discuss the projects. I am available via email as well.

## 1.2 Last Unit Content

- We learned about bivariate, correlational design appropriate for association claims
- We discussed some of the basics on analyzing and graphing data in this design, as well as using the 4 claim validities to interrogate these claims
- We also introduced the basics of the two major take-home projects for the course

# 2 Quiz 7 Review

## 2.1 Areas for Review

- In my opinion, stated in class, we should *always* try to maximize external validity regardless of claim type.
  - Cluster Sampling Example
    - I randomly choose hospitals within the Midwest, and then sample all of the patients within the selected hospitals
    - Hospitals = cluster
    - Patient = participant
    - Because the hospitals were sampled randomly, this is a probabilistic sampling method that will result in an unbiased sample
-

## 3 Quiz 8

### 3.1 Quiz Content

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

- State why simple bivariate correlations are not sufficient for establishing causation.
  - Explain how longitudinal correlational designs can establish temporal precedence.
  - Explain how multiple-regression analyses can rule out some (but not all) third variables.
  - Describe the value of pattern and parsimony, in which a variety of research results support a single, parsimonious causal theory.
-

- Explain the function of a mediating variable.

## 4.2 Professor's Objectives

- Understand the relative strengths and weaknesses of a multivariate design, compared to bivariate designs
- Be able to describe interpreting longitudinal and multiple regression designs, and interpret statistical values from these designs

# 5 Chapter Overview

## 5.1 Chapter Overview

- \_\_\_\_\_ → multiple (more than two) variables
- This correlational research is *still* comprised purely of \_\_\_\_\_ variables
- These techniques will get us **closer** to a causal claim, but we are still just short of reaching the necessary criteria, and still in range of \_\_\_\_\_ claim

# 6 Review of the Casual Criteria

## 6.1 Overview

- **Covariance**: are two (or more) variables varying in some consistent pattern?
- **Temporal Precedence**: does one variable come before another in time?
- \_\_\_\_\_: is the relationship not better explained by a third variable?

# 7 Longitudinal Designs and Temporal Precedence

## 7.1 Overview

- **Longitudinal Research**:
    - Research conducted over multiple time points (usually over an extended period of time)
-

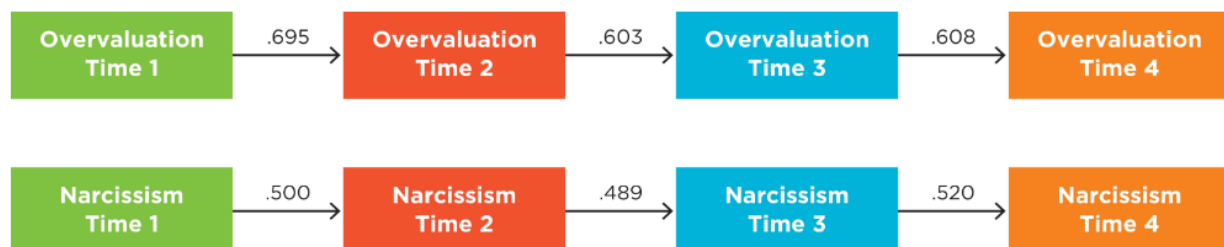
- Normally follows the same group of people taking *same* multiple measures at the many time points
- With this design, comes multiples types of \_\_\_\_\_
- The multivariate nature of this comes from the 2 or more measurements happening at \_\_\_\_\_ time points

## 7.2 Cross-sectional Correlation



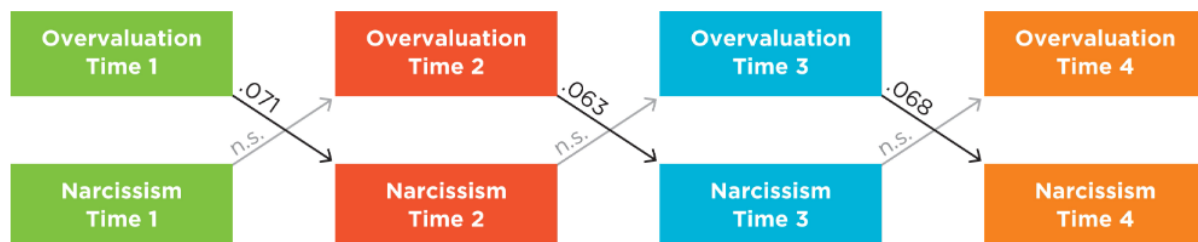
- Are the variables at the same time point correlating with one another?

## 7.3 Autocorrelations



- Are the variables correlated with \_\_\_\_\_ at the other time points?

## 7.4 Cross-lag Correlation



- Are the variables correlated with each other across time points?
- This is sufficient for establishing \_\_\_\_\_

- This is also the key outcome we are looking at for longitudinal designs

## 7.5 Longitudinal and Causation Criteria

- Generally, a good longitudinal study can establish covariance and temporal precedence well.
- However, there needs to be an \_\_\_\_\_ effort in order to rule out third variables
  - This can be done by including additional variables to track which may have an impact on one or both of the variables

## 7.6 Why Not Experiment

- You can't always easily manipulate variables that may be causing something
  - *Discuss*: What are examples of variables that are difficult to manipulate for practical reasons?
- You can't always ethically manipulate a variable
  - *Discuss*: What are examples of variables that can't be ethically changed?
- Sometimes we can do a \_\_\_\_\_ study that would be unethical to over a longer period of time. Then we can combine those experimental designs with longitudinal designs.

# 8 Multiple Regression and Third Variables

## 8.1 Overview

- Multiple \_\_\_\_\_ regression is when we predict one continuous variable by way of multiple \_\_\_\_\_ variables
  - Stats sidebar: your book temporarily refers to this as “multivariate regression” - don't call it that, because that is a separate, more complicated technique
- Technically, we can predict with categorical or continuous variables, but our **single** outcome must be continuous

## 8.2 Criterion and Predictor Variables

- When we use multiple regression, we have:
    - \_\_\_\_\_: the “dependent” variable we are trying to predict
    - Predictor variables: Those that we are using to predict
-



Multiple-Regression Results from a Study Predicting Pregnancy from Sexual Content on TV and Age

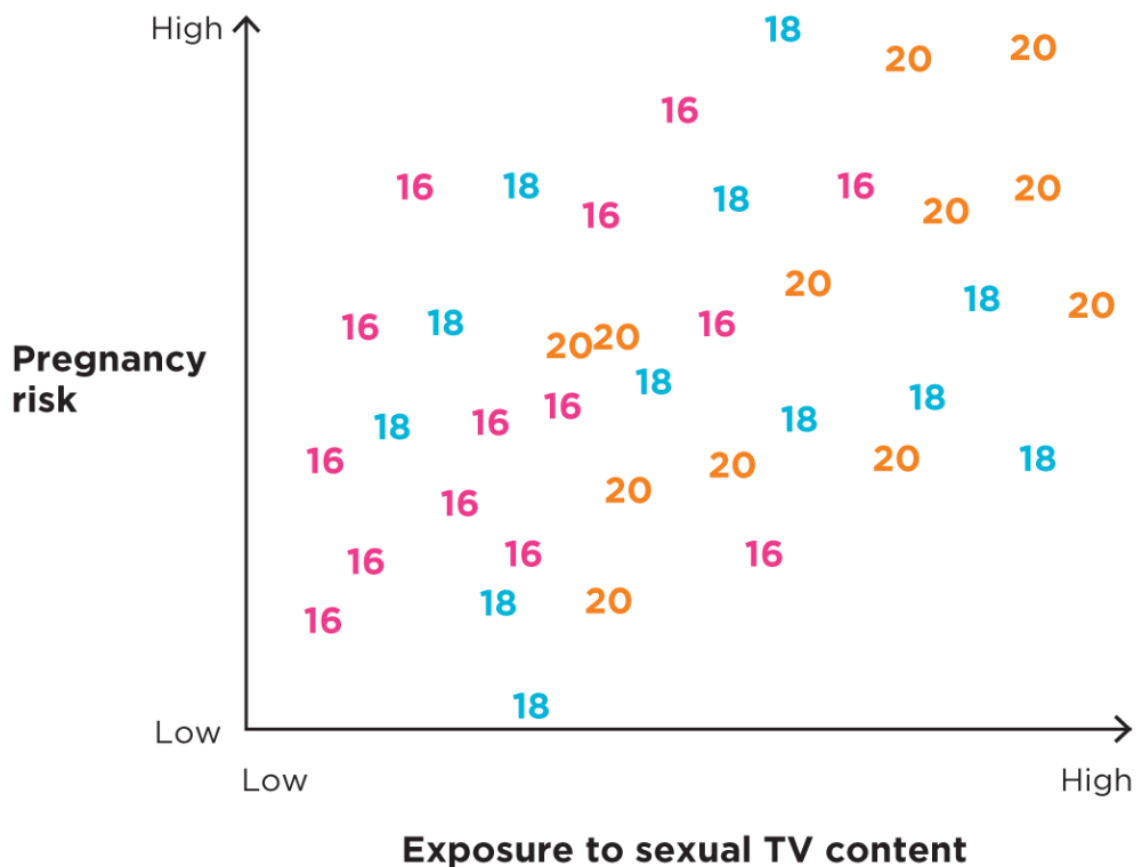
CRITERION (DEPENDENT) VARIABLE: PREGNANCY RISK	BETA	95% CI FOR BETA	Statistical Significance
<b>Predictor (independent) variables:</b>			
Exposure to sex on TV	0.25	[.14, .36]	*
Age	0.33	[.20, .46]	*

Note: Data are fabricated, based on imagined results if the researchers had used only two predictor variables.

\* $p < .05$ , meaning the result is statistically significant and the 95% CI does not include zero.

### 8.3 More Than Two Variables

- The benefit to this is we can *statistically control* for the effects of other variables and isolate the effects of the variables we care about the most



## 8.4 Statistics for Third Variables

- When we talk about *controlling* for some variable in a model, we are talking about “holding it still” or eliminating the effect it has on the \_\_\_\_\_
- That way, we can see the individual impacts of each predictor variable on the criterion

## 8.5 Beta coefficients

- Beta, represented as  $\beta$ , is a linear effect that one unit on the predictor variable changes on the outcome variable
- For example, if our age  $\beta$  is +1.25 on pregnancy risk - that means that for every 1 year of age, predicted pregnancy risks raises by 1.25.
- A beta further from 0 signifies a \_\_\_\_\_ effect, and we can look for statistical significance in  $\beta$  just like in other stats.

## 8.6 More Predictors, The Merrier?

- There is no natural limit to however many predictors to put into a model, with some caveats:
  - A good rule of thumb is that sample should be at least \_\_\_\_\_ times the number of predictor variables
  - Too many predictors can complicate effects if the predictors correlate with one another!

## 8.7 In Popular Media

- “Controlling for”, “Adjusting for”, “Considering” are all phrases that often indicate use of a regression model. Be on the lookout for these.

## 8.8 Still Not Causal

- Despite the fact that longitudinal and multiple regression studies are excellent at statistically control of third variable and establishing temporal precedence...
  - They don't quite reach the methodological gold standard of experiments
-

## 9 Pattern and Parsimony

### 9.1 Overview

- Sometimes the weight of evidence from substantial correlation studies appears to be sufficient for establishing causality... is it?

### 9.2 Meaning

- **Parsimony:** is the principle that says we should seek the simplest possible (and still accurate) description of a phenomenon or relationship
- The “pattern” we are speaking of here is the pattern of results coming from numerous studies - often summarised as a literature review or meta-analysis

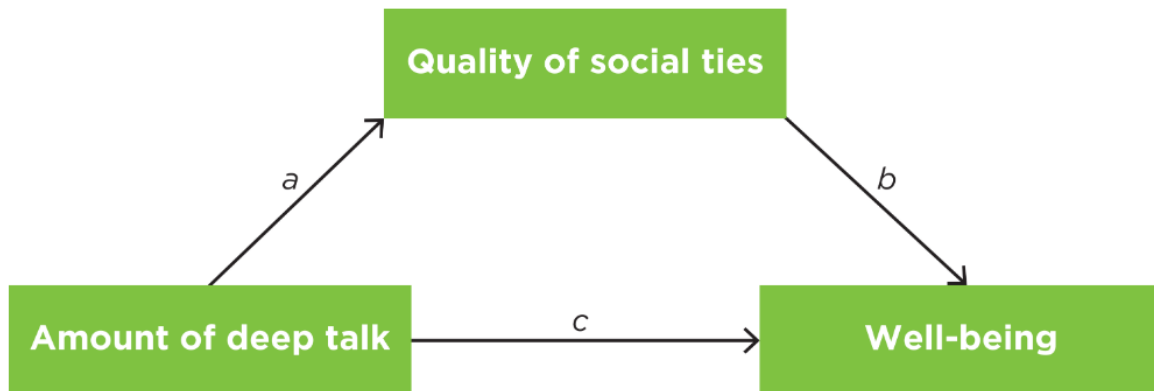
### 9.3 In Popular Media

- Popular media tends to like big consensus papers that seem to point out a prominent pattern in research
- However, remember the issues associated with putting too much trust in scientific journalism

## 10 Mediation

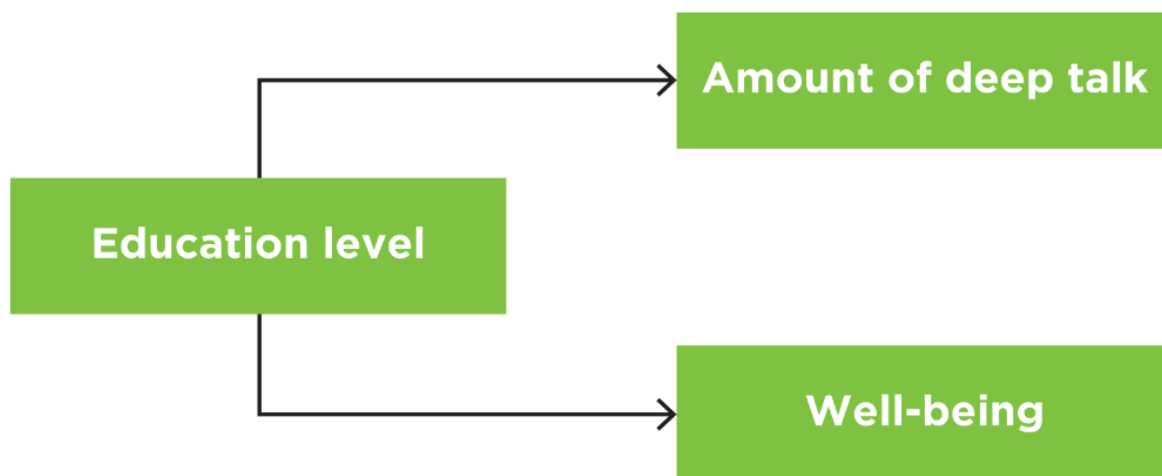
### 10.1 Overview

- **Mediation:** a claim that one variables relationship with another is explained by another variable, i.e. some mechanism through which an effect occurs a.k.a the “why”
  - Technically this is a causal claim, but is often hinted at via correlational designs.
-



## 10.2 Compared to “Third Variable”

- A “third variable” moreso refers to a variable unrelated to a mechanism that just happens to be related to both variables of interest



► More Information

## 10.3 Compared to “Moderators”

- A **Moderator** is best described as a variable that's state changes the relationship between two others
- E.g. The link between depression and anxiety are stronger when trauma is high
- Excellent table on pg 265 to show differences between the last three types

## **11 Analysis with the Four Validities**

### **11.1 Overview**

- How do we investigate these designs? Largely the same as the bivariate designs!