



TEXAS TECH UNIVERSITY™

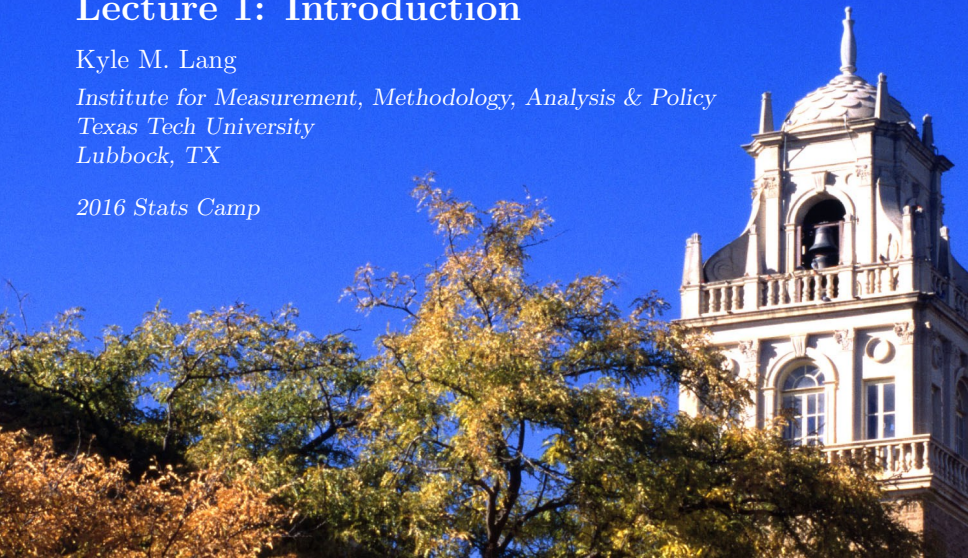


Lecture 1: Introduction

Kyle M. Lang

Institute for Measurement, Methodology, Analysis & Policy
Texas Tech University
Lubbock, TX

2016 Stats Camp



Please give the following information:

1. Name
2. Home-base
3. Prior experience with linear regression
4. Prior experience with structural equation modeling
5. Prior experience with R
6. Interest in this course
7. Favorite animal with a name beginning with the letter B

- Framing mediation, moderation, and conditional process analysis
- Rough outline of what we'll cover this week
- Get everyone up and running with R

What do we mean by *mediation* and *moderation*?

What do we mean by *mediation* and *moderation*?

Mediation and moderation are types of hypotheses, not statistical methods or models.

- Mediation tells us *how* one variable influences another.
- Moderation tells us *when* one variable influences another.

Say we wish to explore the process underlying exercise habits.

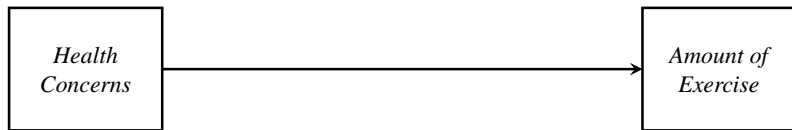
Our first task is to operationalize “exercise habits”

- DV: Hours per week spent in vigorous exercise (*exerciseAmount*).

We may initial ask: what predicts devoting more time to exercise?

- IV: Concerns about negative health outcomes (*healthConcerns*).

The *healthConcerns* \rightarrow *exerciseAmount* relation is our *focal effect*

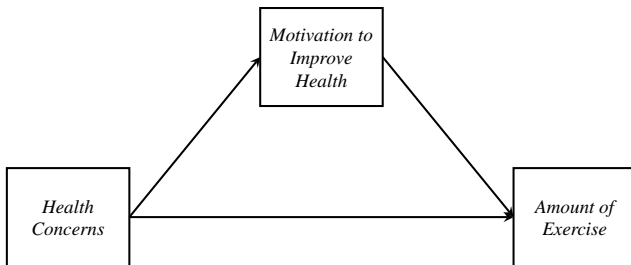


- Mediation, moderation, and conditional process analysis all attempt to describe the focal effect in more detail.
- We always begin by hypothesizing a focal effect.

The Mediation Hypothesis

A mediation analysis will attempt to describe how health concerns affect amount of exercise.

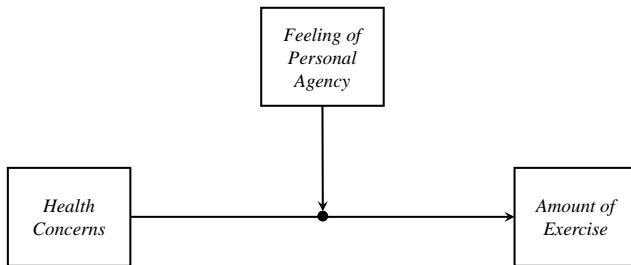
- The *how* is operationalized in terms of intermediary variables.
- Mediator: Motivation to improve health (*motivation*).



Moderation Hypothesis

A moderation hypothesis will attempt to describe when health concerns affect amount of exercise.

- The *when* is operationalized in terms of interactions between the focal predictor and contextualizing variables
- Moderator: Sense of personal agency relating to physical health (*agency*).



Conditional process analysis combines the mediation and moderation hypotheses into models of moderated mediation.

- Given a mediation model describing *how* health concerns affect exercise amount, what other variables may modulate the indirect effect.

SOFTWARE SETUP TIME!

What is R?



R is a holistic (open-source) software system for data analysis and statistical programming.

- Introduced by Ihaka and Gentleman (1996) and currently maintained by a core group of statistical programmers known as the *R Core Team*.
- Support by thousands of world-wide contributors.
 - Anyone can contribute an R package to the *Comprehensive R Archive Network* (CRAN) if it conforms to the licensing and formatting/packaging requirements of the R-Project.
- Considered a dialect/implementation of the S language developed by John Chambers (Becker & Chambers, 1984; Becker, Chambers, & Wilks, 1988; Chambers, 1998; Chambers & Hastie, 1992).

What is R?



I prefer to think about R as a *statistical programming language*, rather than as a data analysis program.

- R **IS NOT** its GUI (no matter which GUI you use).
- You can write R code in whatever program you like (e.g., RStudio, EMACS, Notepad, directly in the console/shell/command line).
- R can be used for basic (or advanced) data analysis, but its real strength is its flexible programming framework.
 - Tedious tasks can be automated.
 - Computationally demanding jobs can be run in parallel.
 - R-based research *wants* to be reproducible.
 - Analyses are automatically documented via their scripts.

R can be downloaded, for free, from the following webpage:

- <https://www.r-project.org/>

You will also need a proper text editor. For those who are just learning R, I recommend **RStudio**:

- <https://www.rstudio.com/>

- Becker, R. A., & Chambers, J. M. (1984). *S: an interactive environment for data analysis and graphics*. Monterey, CA: Wadsworth and Brooks/Cole.
- Becker, R. A., Chambers, J. M., & Wilks, A. R. (1988). *The new S language*. London: Chapman & Hall.
- Chambers, J. M. (1998). *Programming with data: A guide to the S language*. New York: Springer Science & Business Media.
- Chambers, J. M., & Hastie, T. J. (1992). *Statistical models in s*. London: Chapman & Hall.
- Ihaka, R., & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299–314.