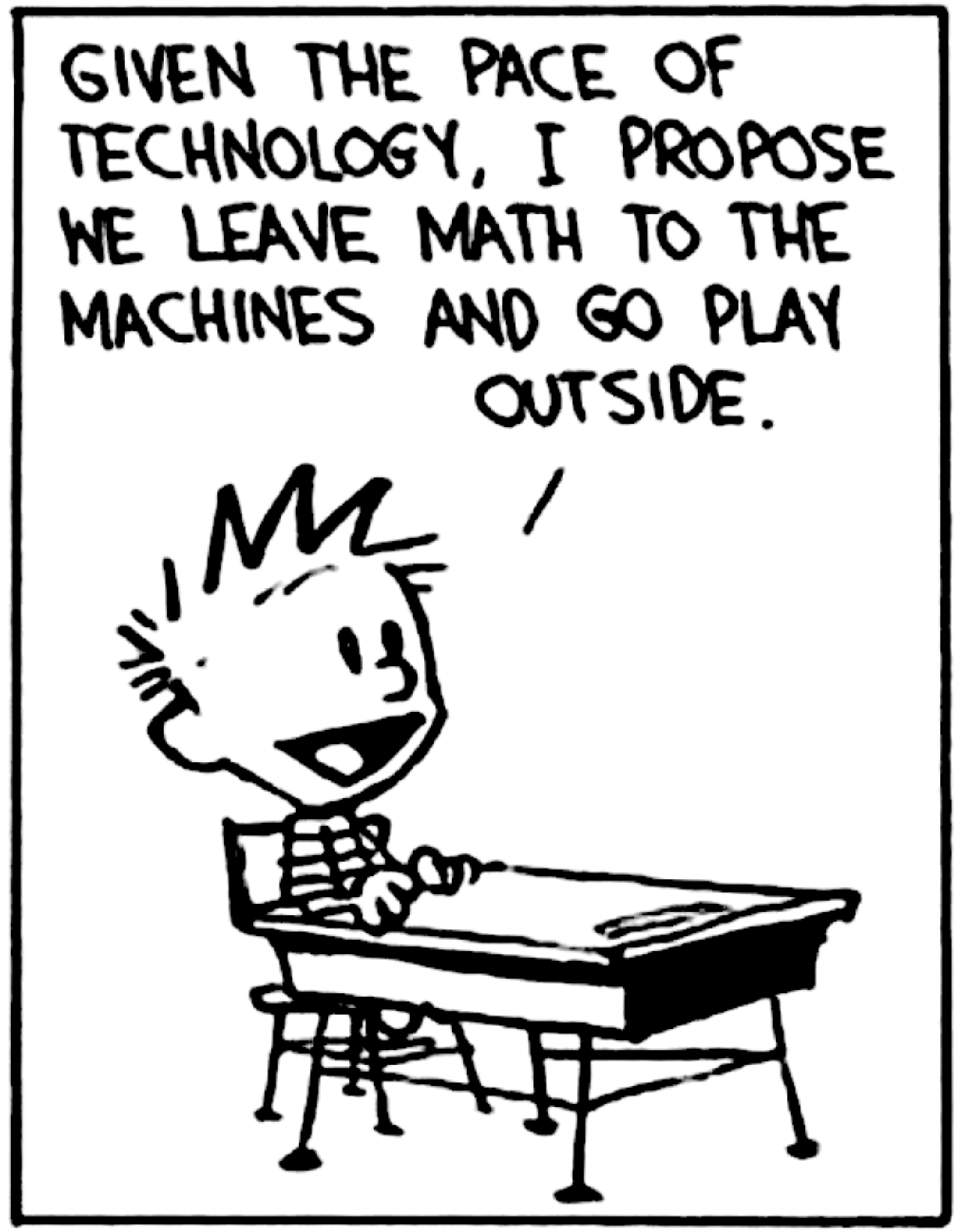
Mediation and Moderation

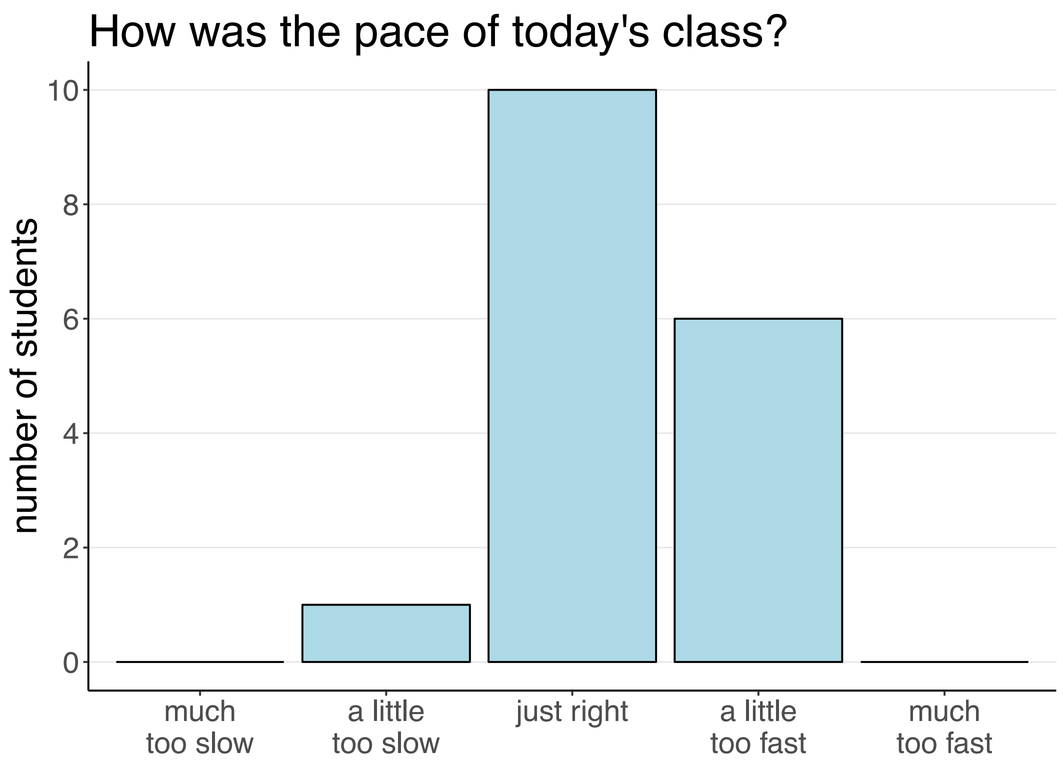


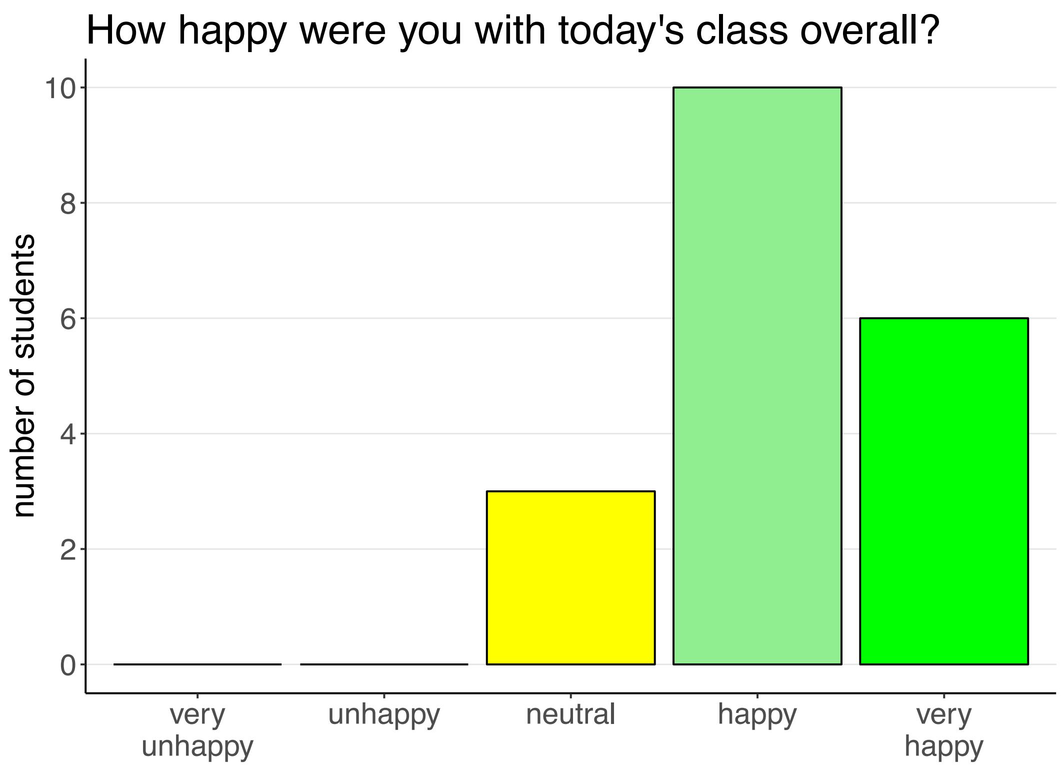
03/11/2019

Logistics

Your feedback

Your feedback





Your feedback

Definitely a drinking from the fire hose lecture but well-structured. Will go through r codes to get a better digestion

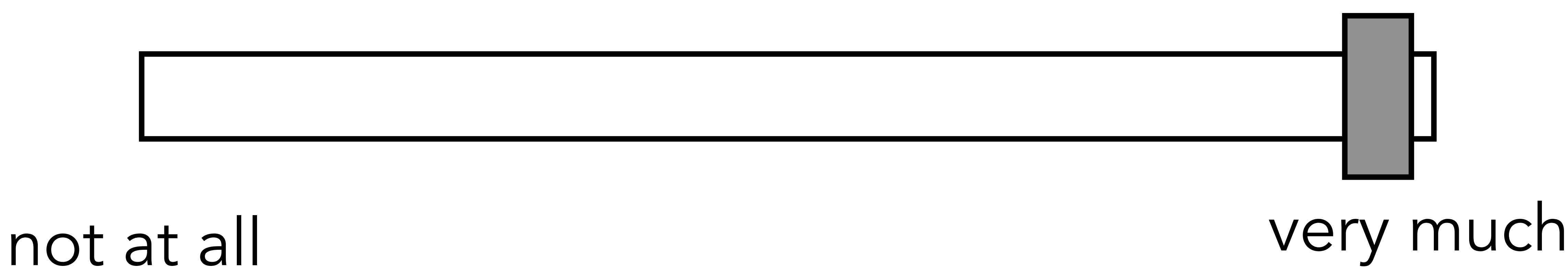
I'll make the R code available shortly

Things that came up

Using sliders

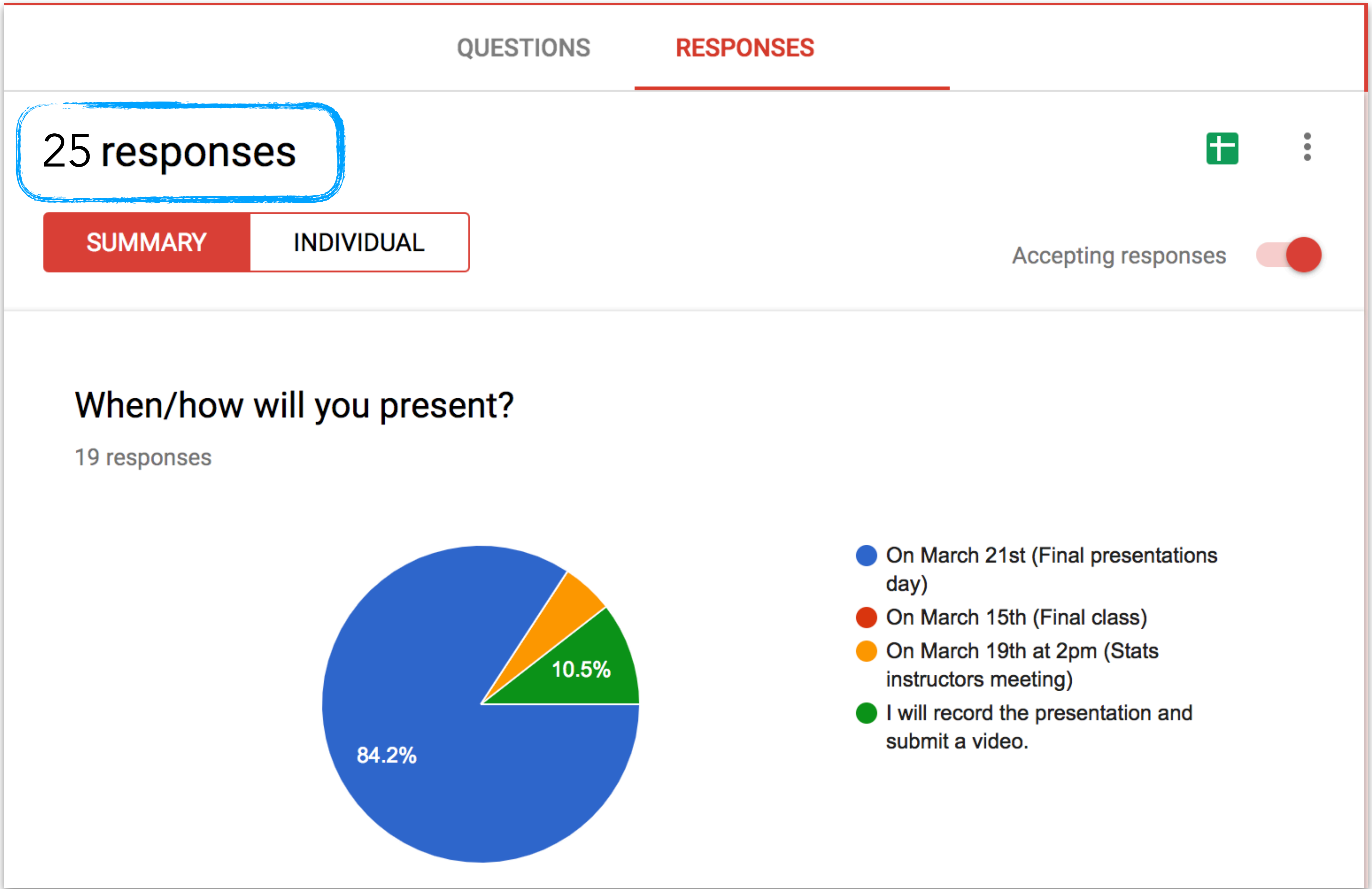
I noticed you said you use sliders. Jon Krosnick teaches a course here on survey methodology, and he recommends that people always avoid sliders because of anchoring effects. Just thought you might like to know he has that piece of advice which he tells students.

How much do you like Psych 252?



Final presentation

Final presentations



https://tinyurl.com/psych252presentation

Final presentations

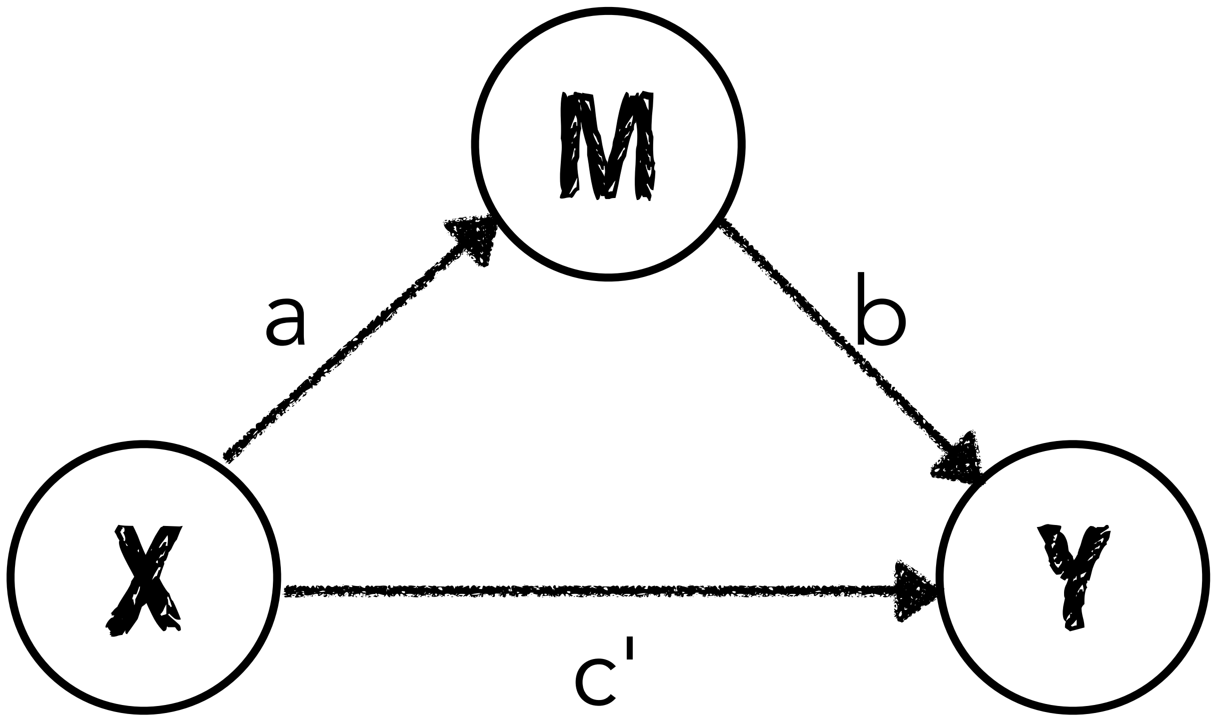
7 min presentation + 3 min questions

Plan for today

- Mediation
- Moderation
- Reporting results
- What we've learned
- Where next?

Mediation

Definition

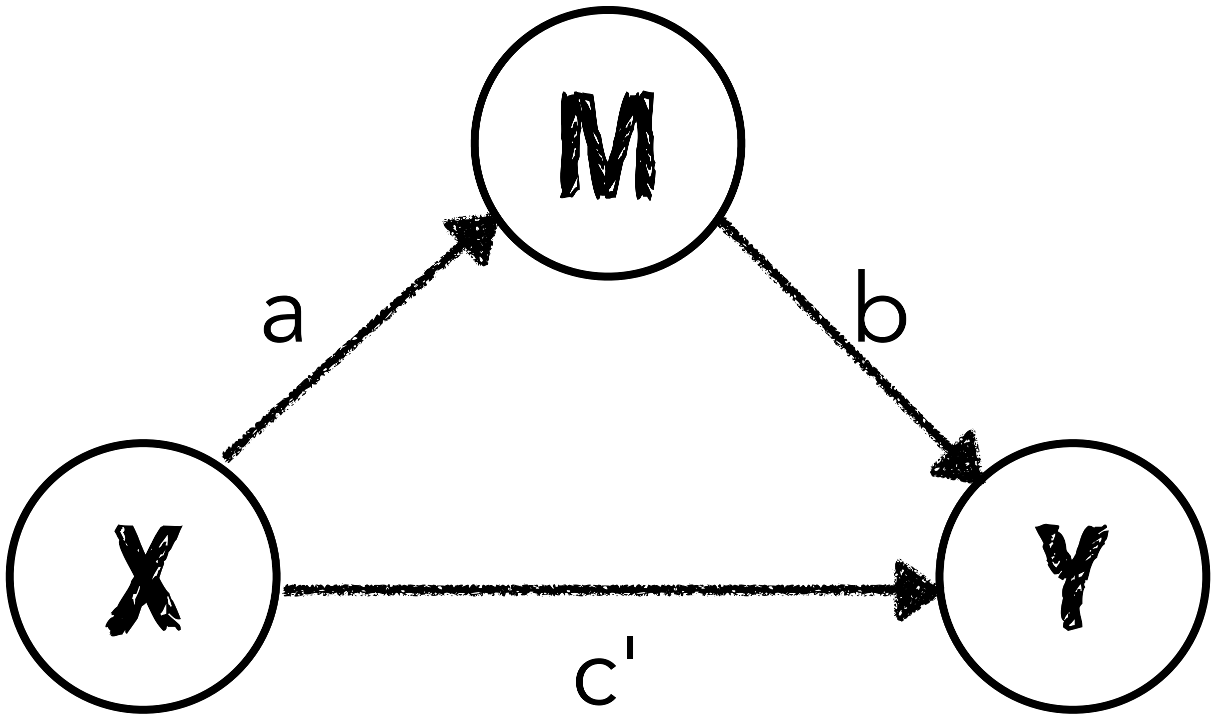


Rather than a direct causal relationship between **X** and **Y**, a mediation model proposes that **X** influences the mediator variable **M**, which in turn influences **Y**. Thus, the mediator variable serves to clarify the nature of the relationship between **X** and **Y**.

Adapted from Wikipedia

https://en.wikipedia.org/wiki/Mediation_(statistics)

Example



X = grades in Psych 252

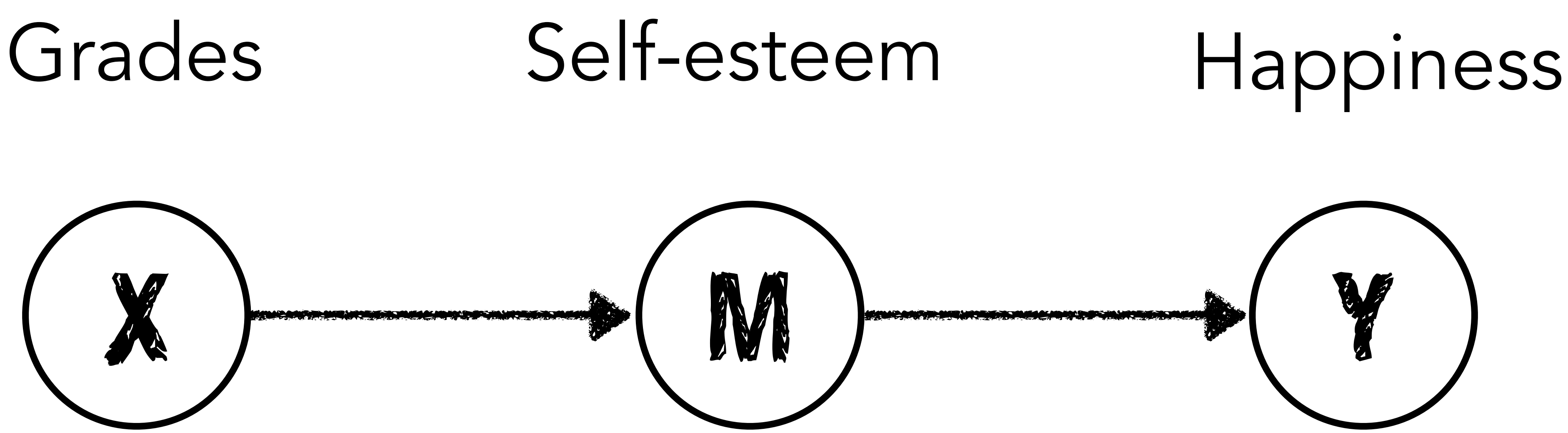
M = feelings of self-esteem

Y = happiness

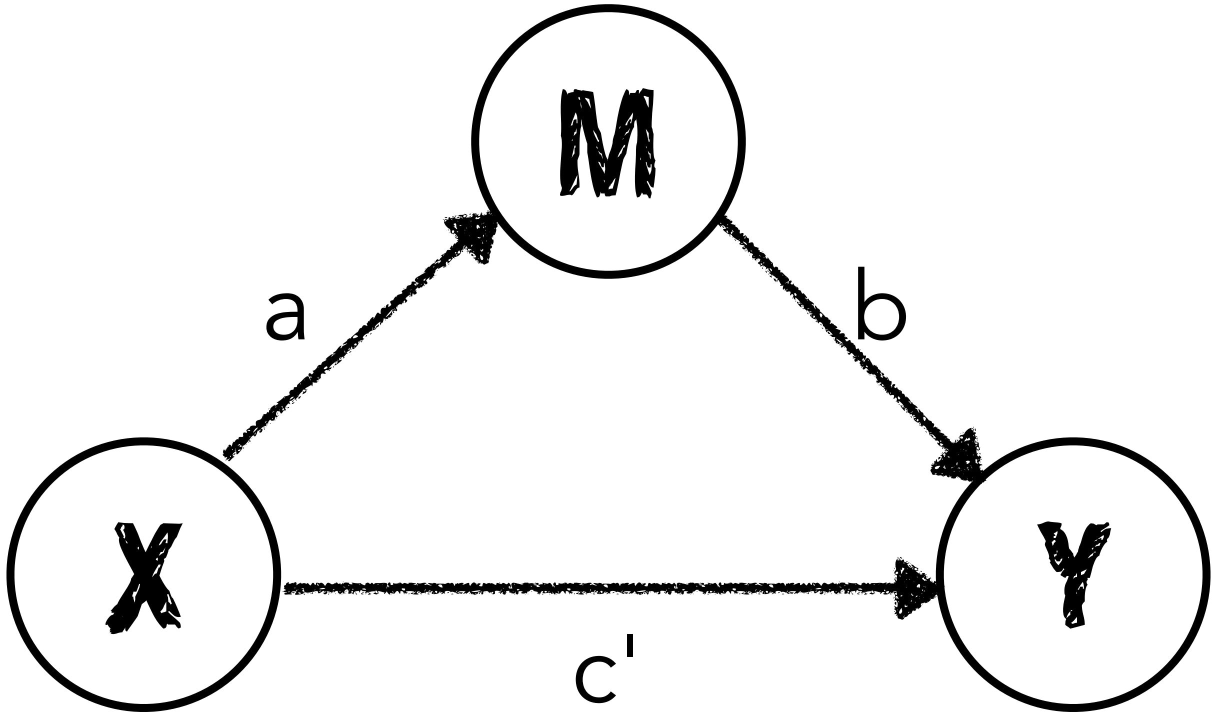
Is the relationship between grades in Psych 252 and happiness mediated by feelings of self-esteem?

Simulate a mediation analysis

```r
# number of participants
n = 100

# generate data
df.mediation = tibble(
  x = rnorm(n, 75, 7),                  # grades
  m = 0.7 * x + rnorm(n, 0, 5), # self-esteem
  y = 0.4 * m + rnorm(n, 0, 5)   # happiness
)
```



Baron and Kenny's (1986) steps for mediation



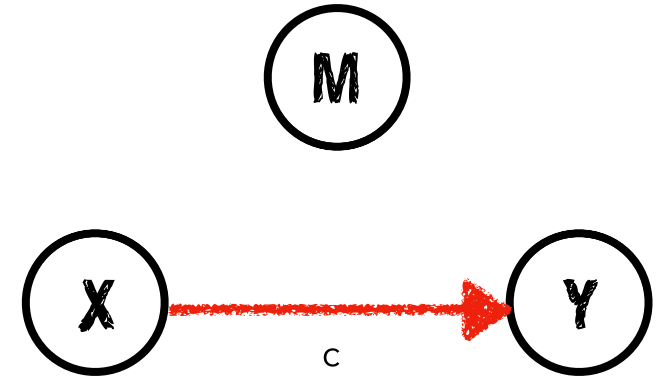
Sequence of regression models

1. Is there a relationship between **X** and **Y**?
2. Is there a relationship between **X** and **M**?
3. Does the relationship between **X** and **Y** change, once we control for **M**?

Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.

Is there a relationship between **X** and **Y**?

$$\hat{y} = b_0 + b_1 \cdot x$$



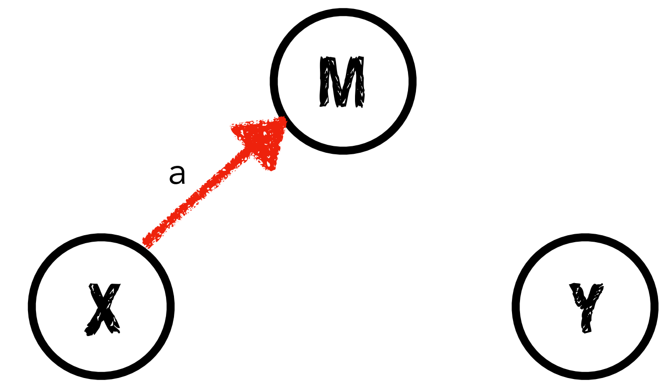

```r
# fit the model
fit.y_x = lm(formula = y ~ 1 + x,
             data = df.mediation)

# summarize the results
fit.y_x %>% summary()
```

```
Call:
lm(formula = y ~ 1 + x, data = df.mediation)

Residuals:
    Min      1Q  Median      3Q     Max
-10.917  -3.738  -0.259   2.910  12.540

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  8.78300    6.16002   1.426   0.1571
x            0.16899    0.08116   2.082   0.0399 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.16 on 98 degrees of freedom
Multiple R-squared:  0.04237,   Adjusted R-squared:  0.0326
F-statistic: 4.336 on 1 and 98 DF,  p-value: 0.03993
```

significant relationship

Is there a relationship between **X** and **M**?

$$\hat{m} = b_0 + b_1 \cdot x$$



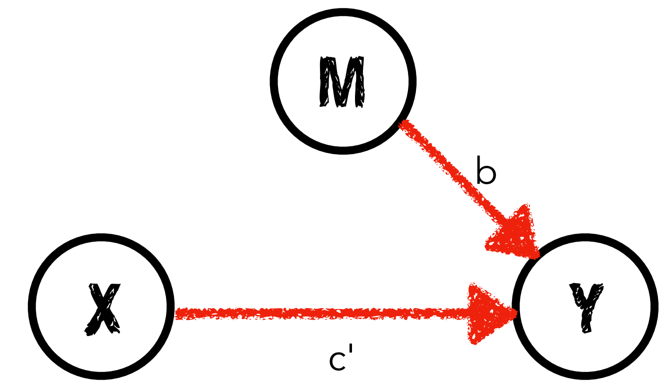

```r
# fit the model
fit.m_x = lm(formula = m ~ 1 + x,
             data = df.mediation)

# summarize the results
fit.m_x %>% summary()
```

```
Call:
lm(formula = m ~ 1 + x, data = df.mediation)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.5367 -3.4175 -0.4375  2.9032 16.4520 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  6.04494   13.41692   0.451    0.653    
x            0.66252    0.07634   8.678 8.87e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.854 on 98 degrees of freedom
Multiple R-squared:  0.4346,	Adjusted R-squared:  0.4288 
F-statistic: 75.31 on 1 and 98 DF,  p-value: 8.872e-14
```

significant relationship

Is there a relationship between **X** and **Y,** controlling for **M**?

$$\hat{y} = b_0 + b_1 \cdot m + b_2 \cdot x$$

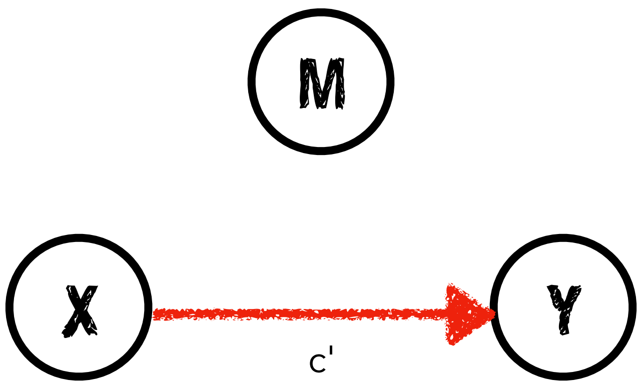



```
# fit the model
fit.y_mx = lm(formula = m ~ 1 + m + x,
              data = df.mediation)

# summarize the results
fit.y_mx %>% summary()
```

```
Call:
lm(formula = y ~ 1 + m + x, data = df.mediation)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.3651 -3.3037 -0.6222  3.1068 10.3991 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  7.80952    5.68297   1.374    0.173    
m            0.42381    0.09899   4.281 4.37e-05 ***
x           -0.11179    0.09949  -1.124    0.264    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.756 on 97 degrees of freedom
Multiple R-squared:  0.1946,	Adjusted R-squared:  0.1779 
F-statistic: 11.72 on 2 and 97 DF,  p-value: 2.771e-05
```

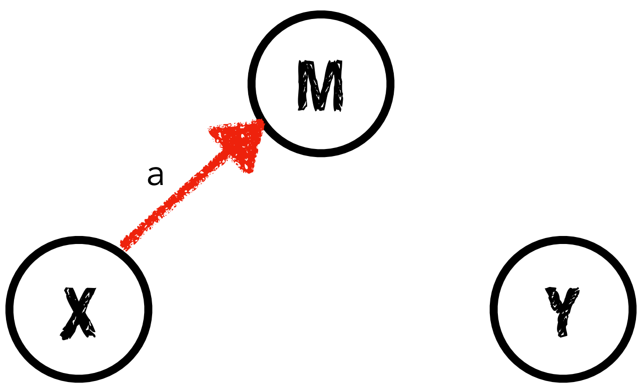
not significant

3 Step procedure



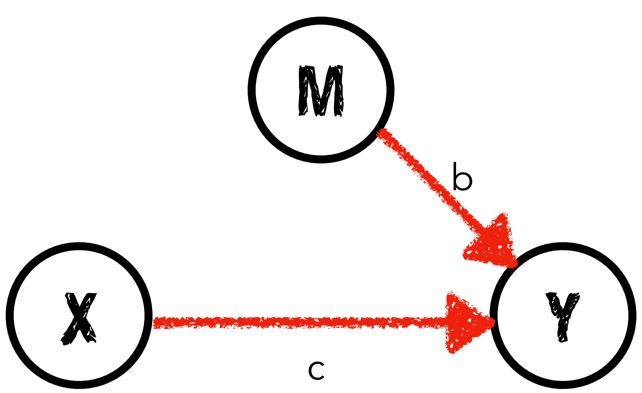
Relationship between **X** and **Y**?

$$\hat{y} = b_0 + \boxed{b_1 \cdot x}$$



Relationship between **X** and **M**?

$$\hat{m} = b_0 + b_1 \cdot x$$



Relationship between **X** and **Y**, controlling for **M**?

$$\hat{y} = b_0 + b_1 \cdot m + \boxed{b_2 \cdot x}$$

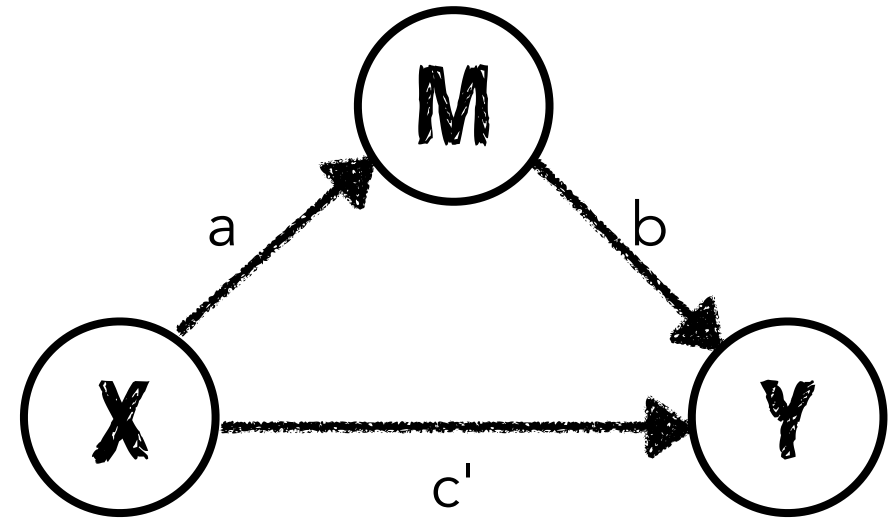
significant change?

just because it changes from significant to not significant, does not mean the change was significant!

Is the mediation significant?

1. **Sobel test**
   - assumes normally distributed data
   - has low power
2. **Bootstrapping**
   - no assumption about how the data is distributed
   - has higher power
3. **Bayesian mediation analysis**

1. Sobel test




```r
library("multilevel")

# run the sobel test
fit.sobel = sobel(pred = df.mediation$x,
                  med = df.mediation$m,
                  out = df.mediation$y)

# calculate the p-value
(1 - pnorm(fit.sobel$z.value))*2
```

product of the coefficients

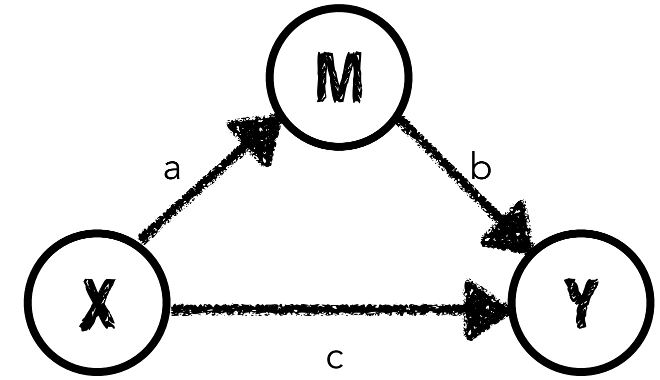
$$Z = \frac{ab}{a^2\sigma_b^2 + b^2\sigma_b^2 + \sigma_a^2\sigma_b^2}$$

standard error of a x b

"It is becoming increasingly more difficult to publish tests of mediation based purely on the Baron and Kenny method or tests that make distributional assumptions such as the Sobel test."

A Wikipedia person

2. Bootstrapping



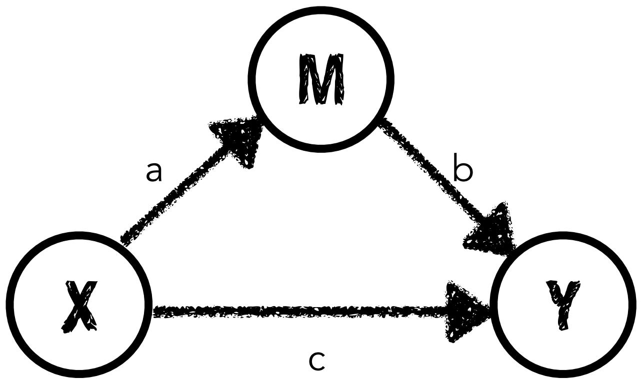

```r
library("mediation")

# bootstrapped mediation
fit.mediation = mediate(model.m = fit.m_x,
                        model.y = fit.y_mx,
                        treat = "x",
                        mediator = "m",
                        boot = T)

# summarize results
fit.mediation %>% summary()
```

$$\hat{m} = b_0 + b_1 \cdot x$$

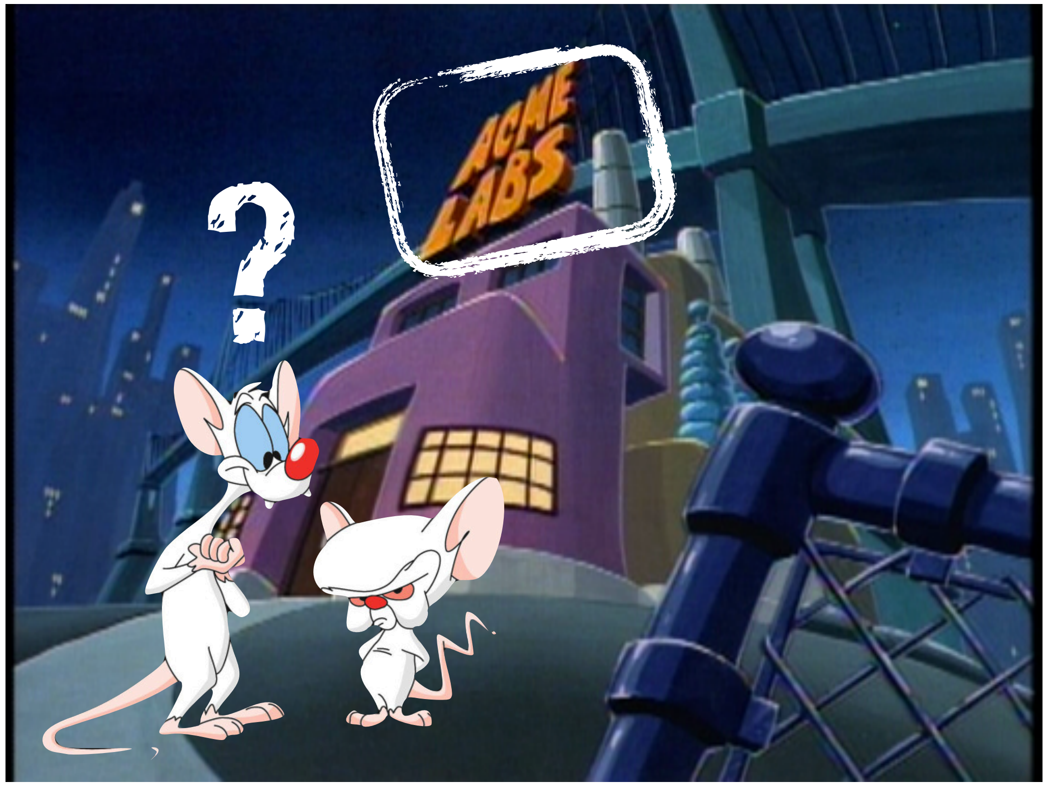
$$\hat{y} = b_0 + b_1 \cdot m + b_2 \cdot x$$

```
Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

               Estimate 95% CI Lower 95% CI Upper p-value    
ACME            0.28078      0.14059         0.42  <2e-16 ***
ADE            -0.11179     -0.29276         0.10   0.272    
Total Effect    0.16899     -0.00415         0.34   0.064 .  
Prop. Mediated  1.66151     -3.22476        11.46   0.064 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 100


Simulations: 1000
```

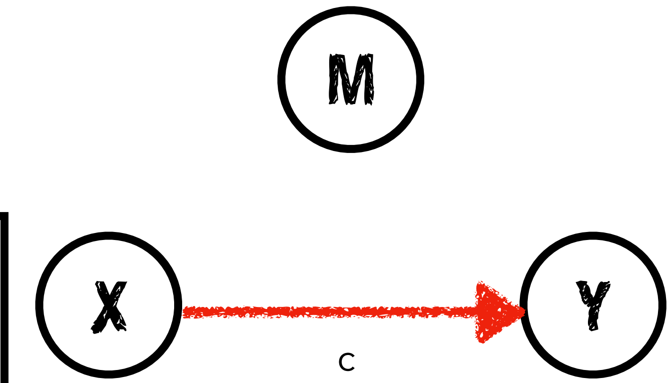
2. Bootstrapping




```
Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

                 Estimate 95% CI Lower 95% CI Upper p-value    
ACME              0.28078      0.14059         0.42  <2e-16 ***
ADE              -0.11179     -0.29276         0.10   0.272    
Total Effect      0.16899     -0.00415         0.34   0.064 .  
Prop. Mediated    1.66151     -3.22476        11.46   0.064 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 100


Simulations: 1000
```



2. Bootstrapping

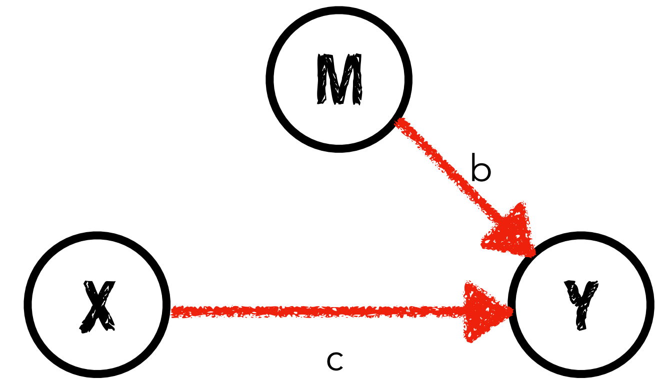



```
Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

               Estimate 95% CI Lower 95% CI Upper p-value    
ACME            0.28078      0.14059         0.42  <2e-16 ***
ADE            -0.11179     -0.29276         0.10   0.272    
Total Effect    0.16899     -0.00415         0.34   0.064 .  
Prop. Mediated  1.66151     -3.22476        11.46   0.064 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 100 


Simulations: 1000
```

$$\hat{y} = b_0 + b_1 \cdot x$$

```
Call:
lm(formula = y ~ 1 + x, data = df.mediation)

Residuals:
    Min      1Q  Median      3Q     Max 
-10.917  -3.738  -0.259   2.910  12.540 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)  8.78300    6.16002   1.426   0.1571  
x            0.16899    0.08116   2.082   0.0399 *
```

2. Bootstrapping




```
Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

                 Estimate 95% CI Lower 95% CI Upper p-value    
ACME              0.28078      0.14059         0.42  <2e-16 ***
ADE              -0.11179     -0.29276         0.10   0.272    
Total Effect      0.16899     -0.00415         0.34   0.064 .  
Prop. Mediated  1.66151     -3.22476        11.46   0.064 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Sample Size Used: 100 


Simulations: 1000
```

$$\hat{y} = b_0 + b_1 \cdot m + b_2 \cdot x$$

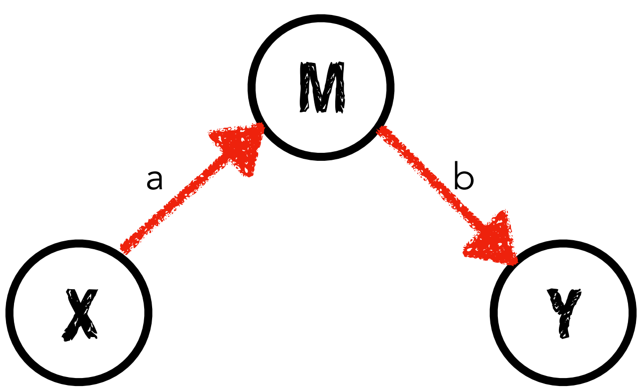
ADE: Average direct effect

```
Call:
lm(formula = y ~ 1 + m + x, data = df.mediation)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.3651 -3.3037 -0.6222  3.1068 10.3991 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  7.80952    5.68297   1.374    0.173    
m            0.42381    0.09899   4.281 4.37e-05 ***
x           -0.11179    0.09949  -1.124    0.264    
```

2. Bootstrapping

```
Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

               Estimate 95% CI Lower 95% CI Upper p-value
ACME            0.28078      0.14059         0.42  <2e-16 ***
ADE            -0.11179     -0.29276         0.10   0.272
Total Effect    0.16899     -0.00415         0.34   0.064 .
Prop. Mediated  1.66151     -3.22476        11.46   0.064 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 100


Simulations: 1000
```

ACME



indirect effect: a * b

ADE: Average direct effect

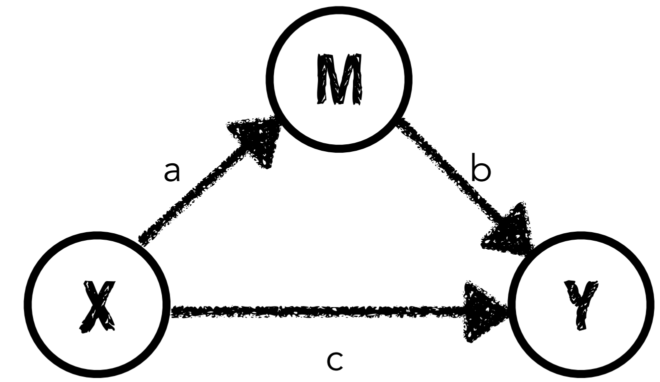
ACME: Average causal mediation effect

ACME = Total effect - ADE

3. Bayesian approach

- Instead of doing bootstrapping (where we resample the data), we fit a multi-variate Bayesian regression model and use the posterior to express our uncertainty

3. Bayesian approach



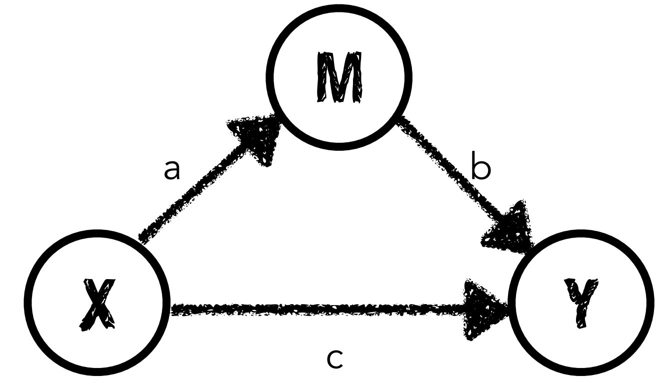

```
# model specification
y_mx = bf(y ~ 1 + m + x)
m_x = bf(m ~ 1 + x)

# fit the model
fit.brm_mediation = brm(
  formula = y_mx + m_x + set_rescor(FALSE),
  data = df.mediation,
  file = "brm_mediation",
  seed = 1
)

# summarize the result
fit.brm_mediation %>% summary()
```

$$\hat{y} = b_0 + b_1 \cdot m + b_2 \cdot x$$

$$\hat{m} = b_0 + b_1 \cdot x$$

makes it such that correlation between residuals is not modeled

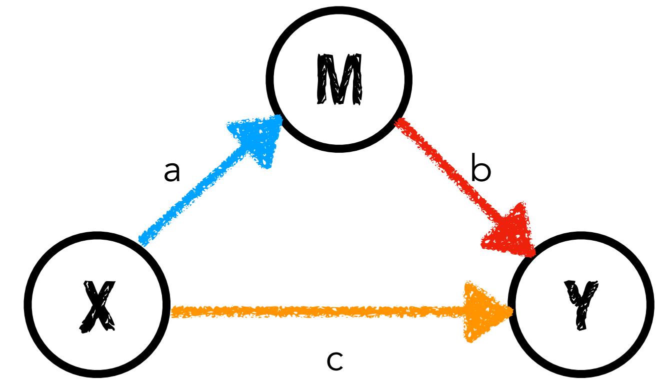
3. Bayesian approach



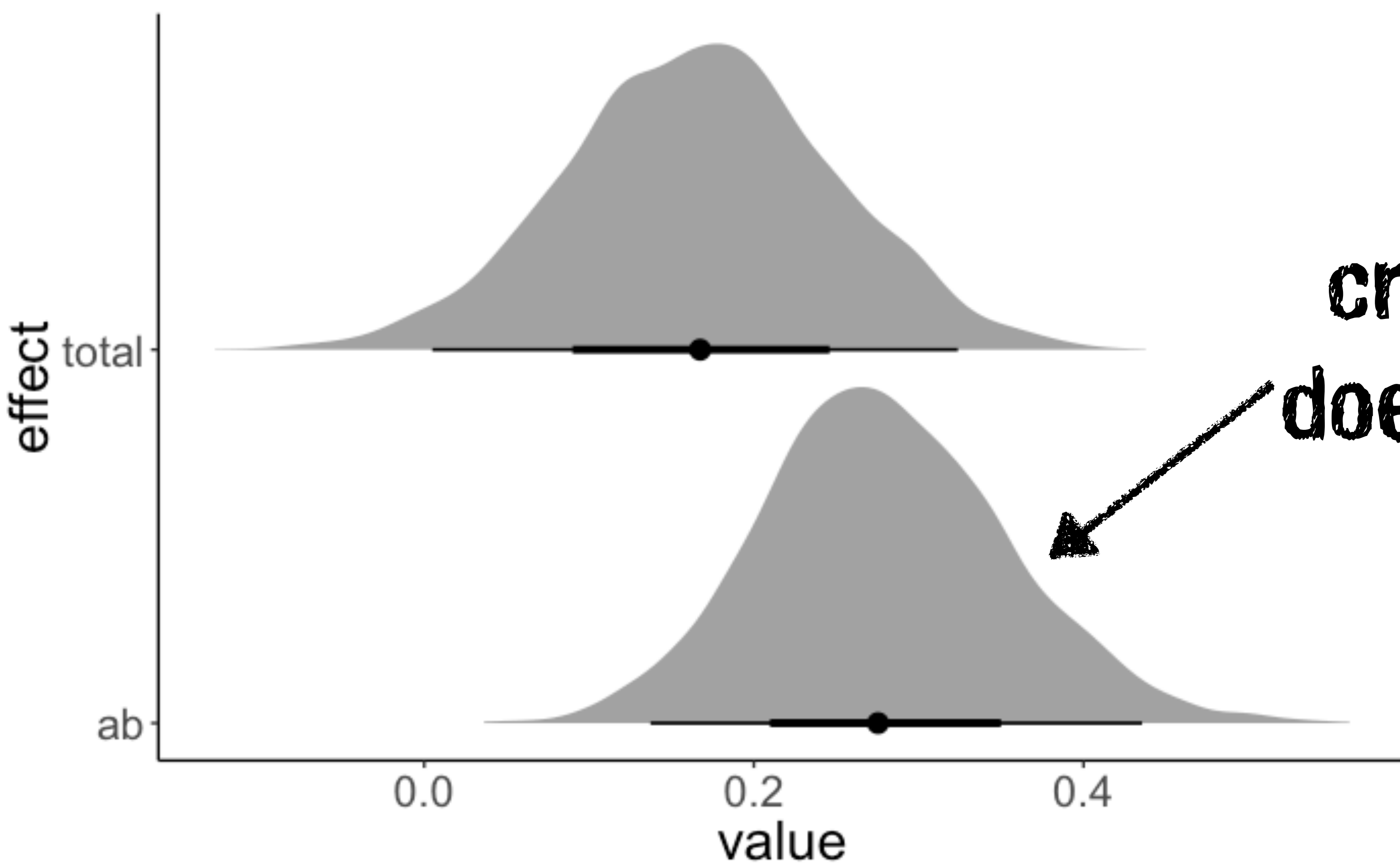

```
 Family: MV(gaussian, gaussian)
  Links: mu = identity; sigma = identity
         mu = identity; sigma = identity
Formula: y ~ 1 + m + x
         m ~ 1 + x
   Data: df.mediation (Number of observations: 100)
Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Population-Level Effects:
            Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
y_Intercept    17.60     13.28    -8.30    44.38       5643 1.00
m_Intercept     6.23     13.54   -20.83    32.41       5249 1.00
y_m             0.42      0.10     0.22     0.62       3970 1.00
y_x            -0.11      0.10    -0.31     0.09       3808 1.00
m_x             0.66      0.08     0.51     0.82       5260 1.00

Family Specific Parameters:
        Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sigma_y     4.81      0.35     4.17     5.53       5065 1.00
sigma_m     4.90      0.36     4.25     5.69       7683 1.00

Samples were drawn using sampling(NUTS). For each parameter, Eff.Sample
is a crude measure of effective sample size, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

How do we get the indirect effect (aka ACME)?

3. Bayesian approach




```
df.samples = fit.brm_mediation %>%
  posterior_samples() %>%
  mutate(ab = b_m_x * b_y_m,
         total = ab + b_y_x)
```

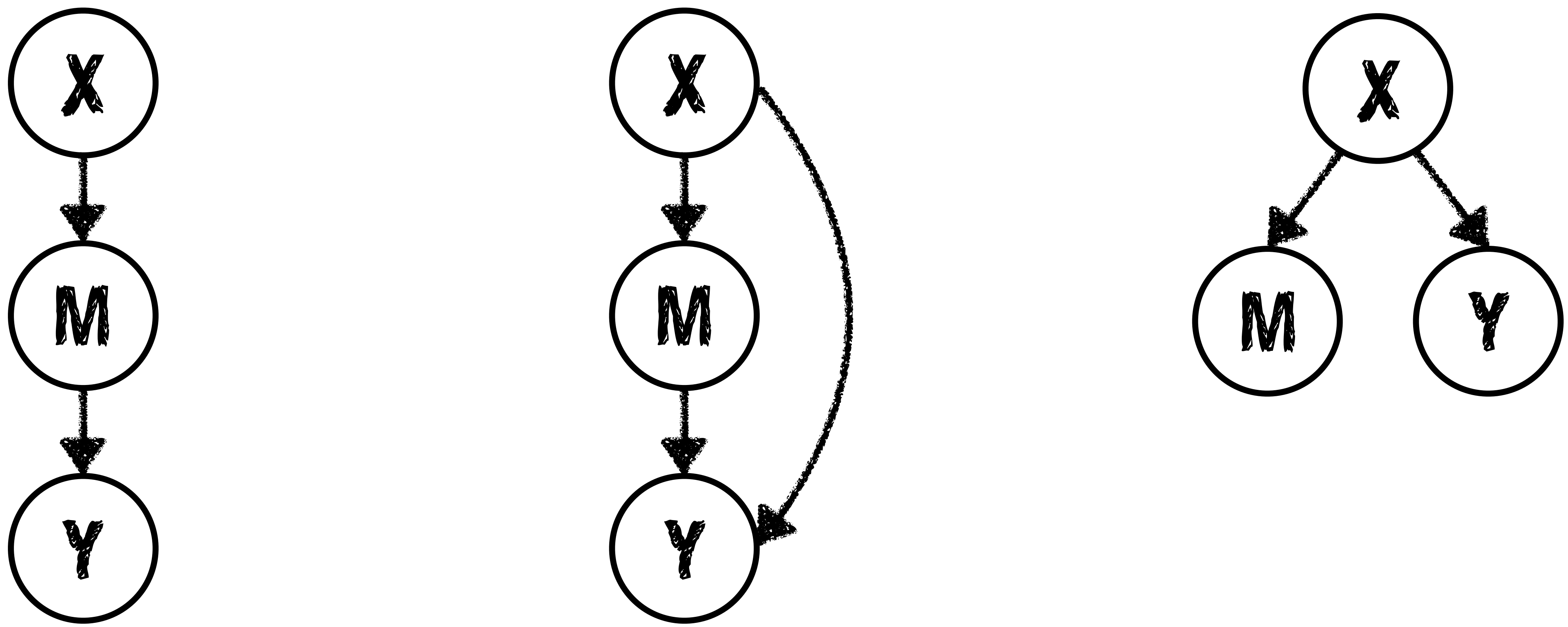


credible interval does not include 0

As usual: Arithmetic on our posterior samples!

Underlying causal model

Full mediation | **Partial mediation** | **No mediation**



Full mediation: When the effect of **X** on **Y** completely disappears, **M** fully mediates between **X** and **Y**.

Partial mediation: When the effect of **X** on **Y** still exists, but in a smaller magnitude, **M** partially mediates between **X** and **Y**.

Limitations

- correlational analysis
    - we need theories / experiments to tease apart causes and effects to properly map our variables onto the diagram

ELSEVIER

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Reports

What mediation analysis can (not) do☆

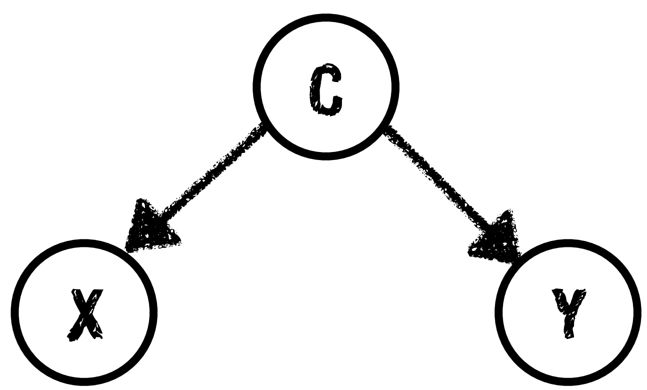
Klaus Fiedler [a,*], Malte Schott [a], Thorsten Meiser [b]

[a] University of Heidelberg, Germany
[b] University of Mannheim, Germany

Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, 47(6), 1231-1236.

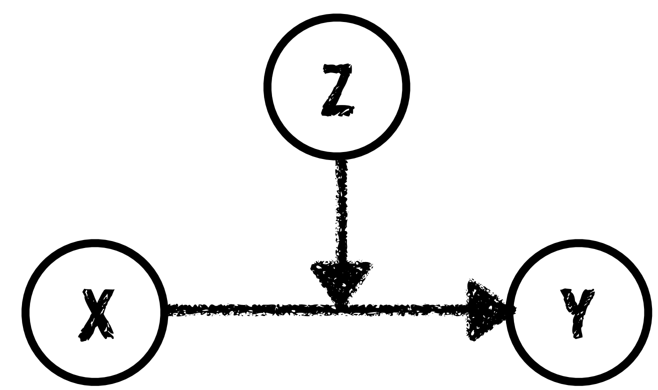
Other third variables

Confounding



C influences both **X** and **Y**

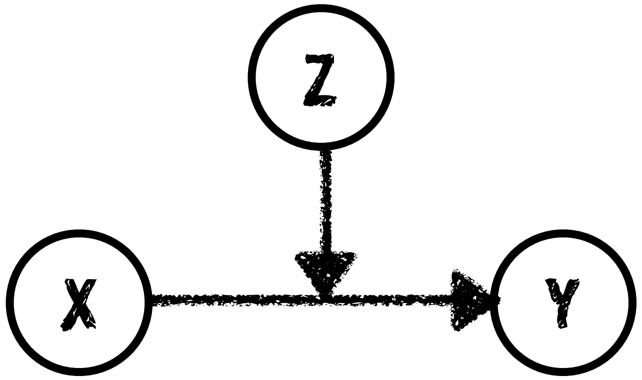
Moderators



Z affects the relationship between **X** and **Y**

Moderation

Definition

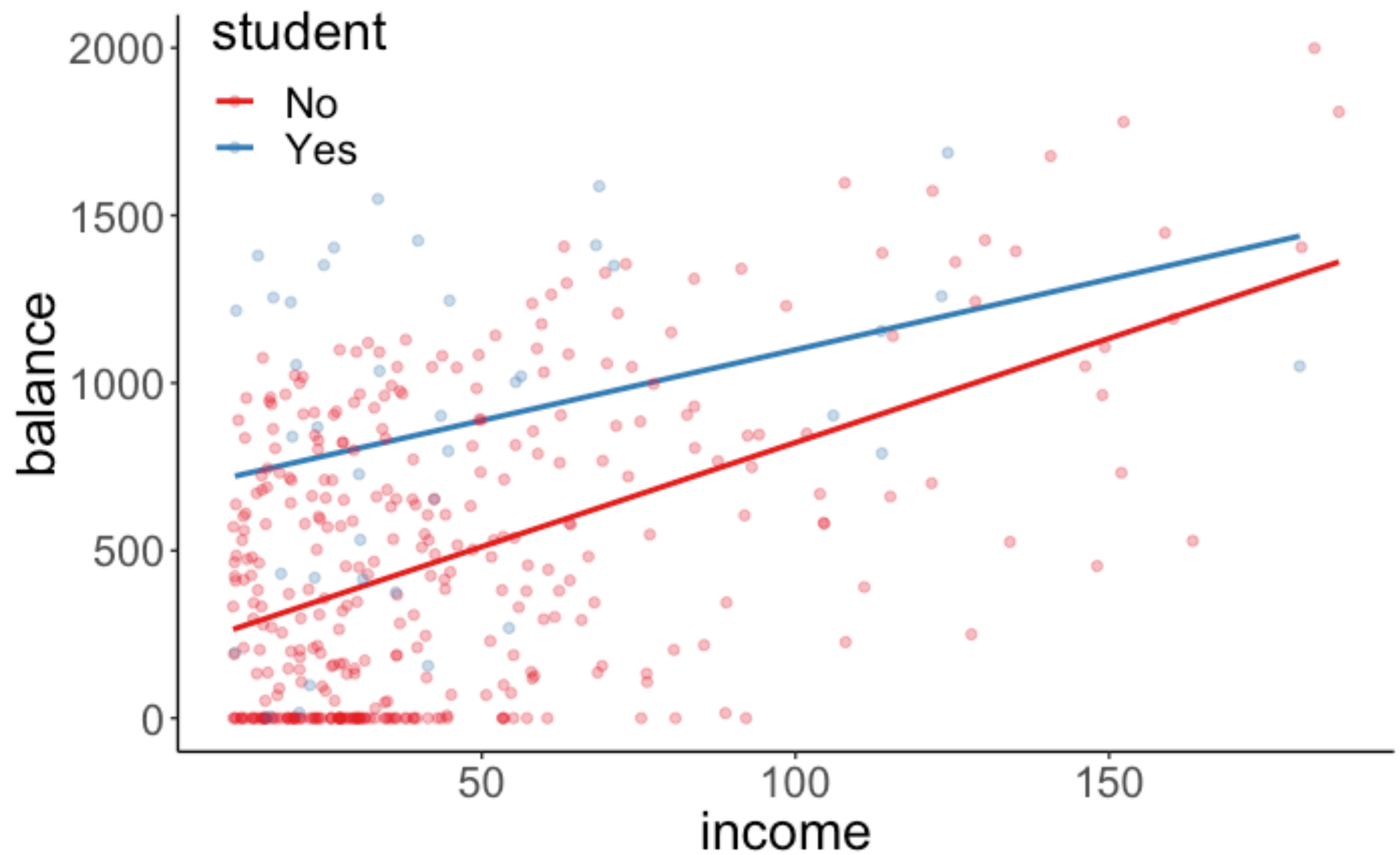


Moderation means that the effect of a predictor depends on the value of another.

Here, the nature of the relationship between **X** and **Y** depends on **Z**.

Have we come across moderation already?

Relationship between credit card balance, income, and whether the person is a student.



$$\widehat{\text{balance}}_i = 200.62 + 6.22 \cdot \text{income}_i + 476.68 \cdot \text{student}_i - 2.00 \cdot (\text{income}_i \times \text{student}_i)$$

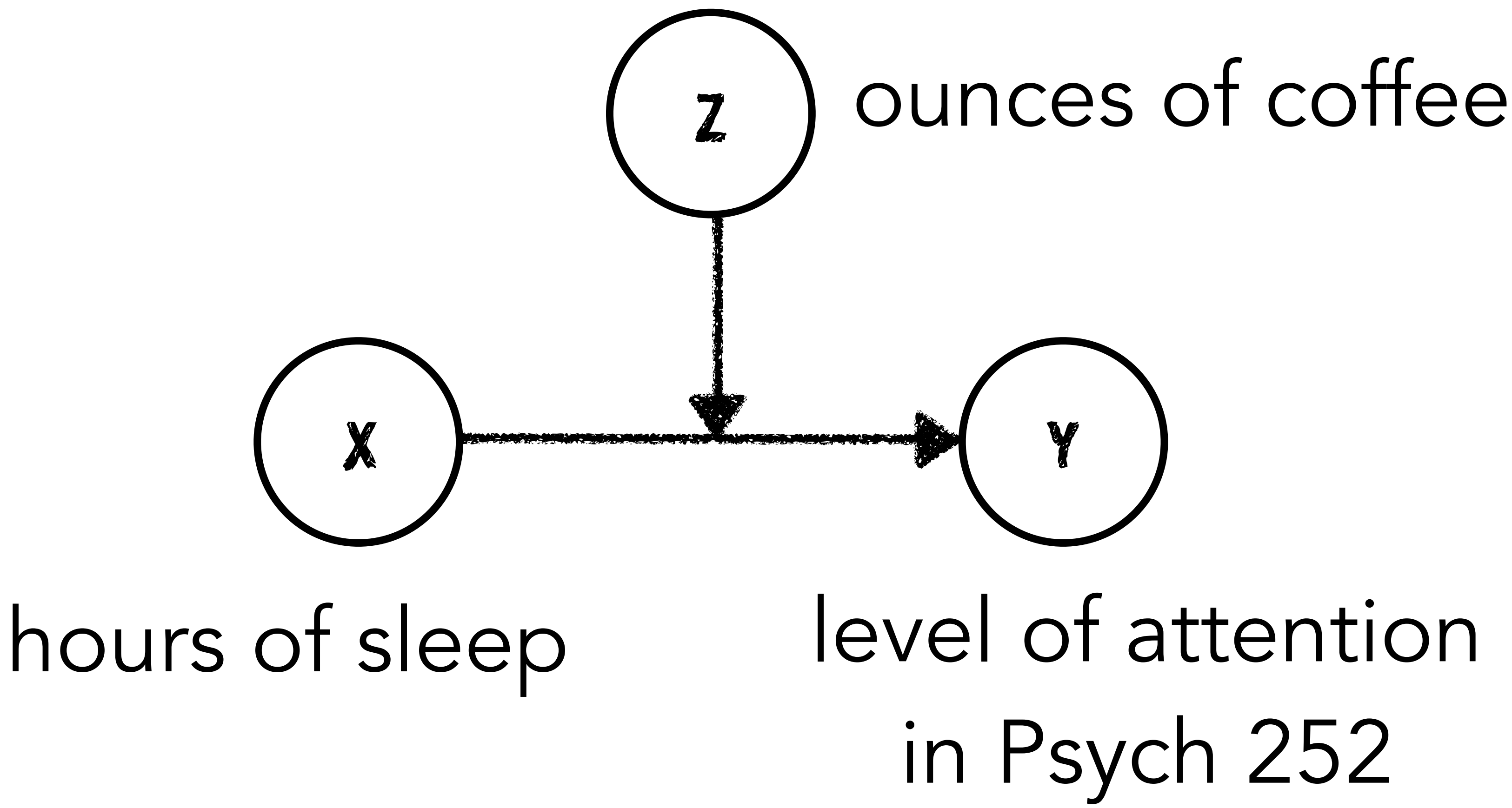
if student = "No" $\widehat{\text{balance}}_i = 200.62 + 6.22 \cdot \text{income}_i$

if student = "Yes"

$$\begin{aligned}\widehat{\text{balance}}_i &= 200.62 + 6.22 \cdot \text{income}_i + 476.68 \cdot 1 - 2.00 \cdot (\text{income}_i \times 1) \\ &= 677.3 + 6.22 \cdot \text{income}_i - 2.00 \cdot \text{income}_i \\ &= 677.3 + 4.22 \cdot \text{income}_i\end{aligned}$$

Simulating a moderation

```r
# number of participants
n  = 100

df.moderation = tibble(
  x  = abs(rnorm(n, 6, 4)), # hours of sleep
  x1 = abs(rnorm(n, 60, 30)), # adding some systematic variance to our DV
  z  = rnorm(n, 30, 8), # ounces of coffee consumed
  y  = abs((-0.8 * x) * (0.2 * z) - 0.5 * x - 0.4 * x1 + 10 + rnorm(n, 0, 3)) # attention Paid
)
```



Simulating a moderation

```r
# scale the predictors
df.moderation = df.moderation %>%
  mutate_at(vars(x, z), "scale")

# run regression model with interaction
fit.moderation = lm(formula = y ~ 1 + x * z,
                    data = df.moderation)

# summarize result
fit.moderation %>%
  summary()
```

```
Call:
lm(formula = y ~ 1 + x * z, data = df.moderation)

Residuals:
    Min      1Q  Median      3Q     Max 
-21.466  -8.972  -0.233   6.180  38.051 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   48.544      1.173  41.390  < 2e-16 ***
x             17.863      1.196  14.936  < 2e-16 ***
z              8.393      1.181   7.108 2.08e-10 ***
x:z            6.094      1.077   5.656 1.59e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.65 on 96 degrees of freedom
Multiple R-squared:  0.7661,	Adjusted R-squared:  0.7587 
F-statistic: 104.8 on 3 and 96 DF,  p-value: < 2.2e-16
```

significant interaction

Simulating a moderation

```
# generate data grid with three levels of the moderator
df.newdata = df.moderation %>%
  expand(x = c(min(x),
               max(x)),
         z = c(mean(z) - sd(z),
               mean(z),
               mean(z) + sd(z))) %>%
  mutate(moderator = rep(c("low", "average", "high"), nrow(.)/3))

# predictions for the three levels of the moderator
df.prediction = fit.moderation %>%
  augment(newdata = df.newdata) %>%
  mutate(moderator = factor(moderator, levels = c("high", "average", "low")))
```

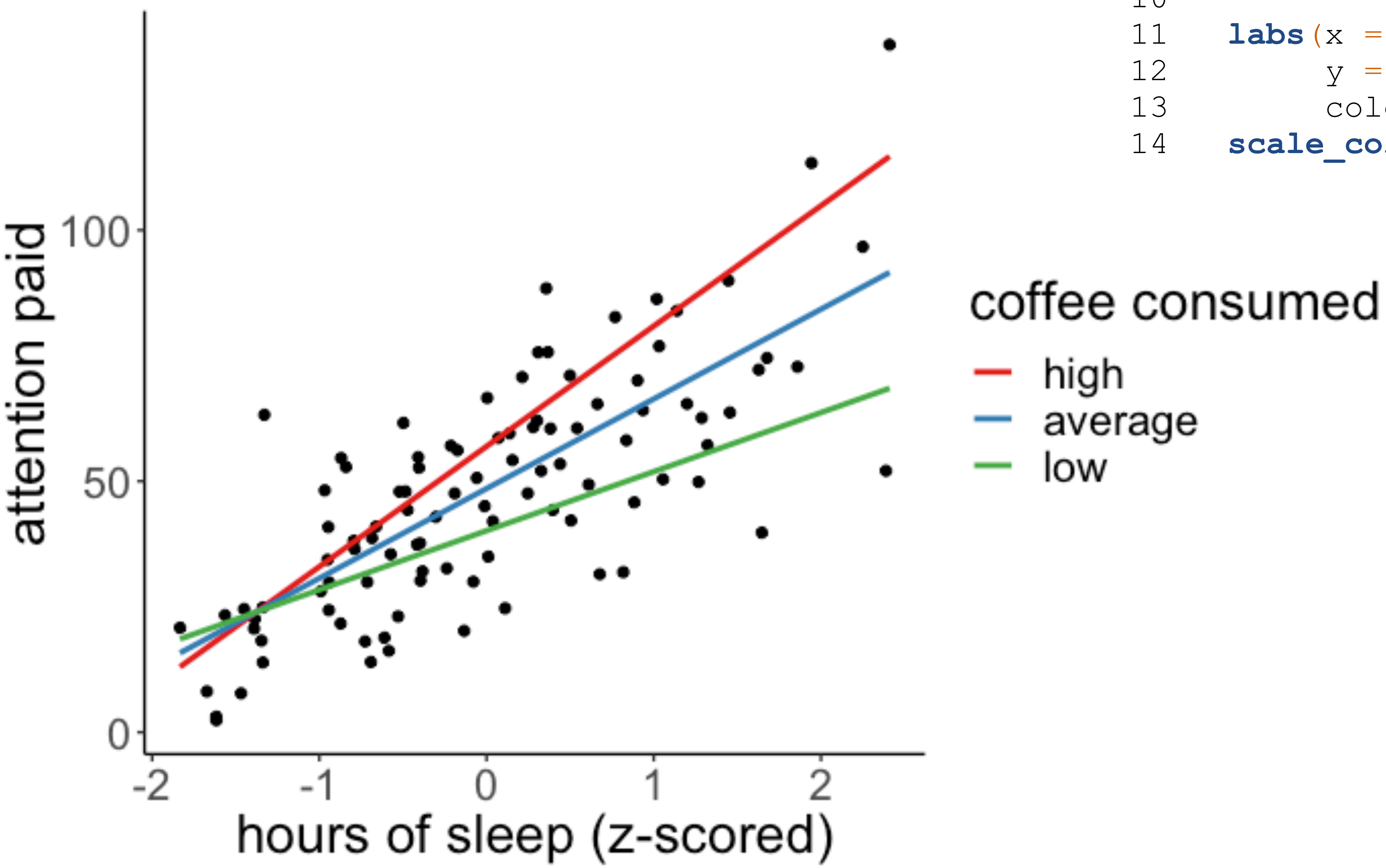
minimum and maximum

average ± 1SD

x	z	moderator	.fitted	.se.fit
-1.83	-1	low	18.58	3.75
-1.83	0	average	15.80	2.51
-1.83	1	high	13.02	2.99
2.41	-1	low	68.52	4.32
2.41	0	average	91.60	3.09
2.41	1	high	114.68	4.12

Simulating a moderation

```
# visualize the result
df.moderation %>%
  ggplot(aes(x = x,
             y = y)) +
  geom_point() +
  geom_line(aes(y = .fitted,
                group = moderator,
                color = moderator),
            data = df.prediction,
            size = 1) +
  labs(x = "hours of sleep (z-scored)",
       y = "attention paid",
       color = "coffee consumed") +
  scale_color_brewer(palette = "Set1")
```



Reminder

```
Call:
lm(formula = y ~ 1 + x * z, data = df.moderation)

Residuals:
    Min      1Q  Median      3Q     Max 
-21.466  -8.972  -0.233   6.180  38.051 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   48.544      1.173  41.390  < 2e-16 ***
x             17.863      1.196  14.936  < 2e-16 ***
z              8.393      1.181   7.108 2.08e-10 ***
x:z            6.094      1.077   5.656 1.59e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.65 on 96 degrees of freedom
Multiple R-squared:  0.7661,	Adjusted R-squared:  0.7587 
F-statistic: 104.8 on 3 and 96 DF,  p-value: < 2.2e-16
```

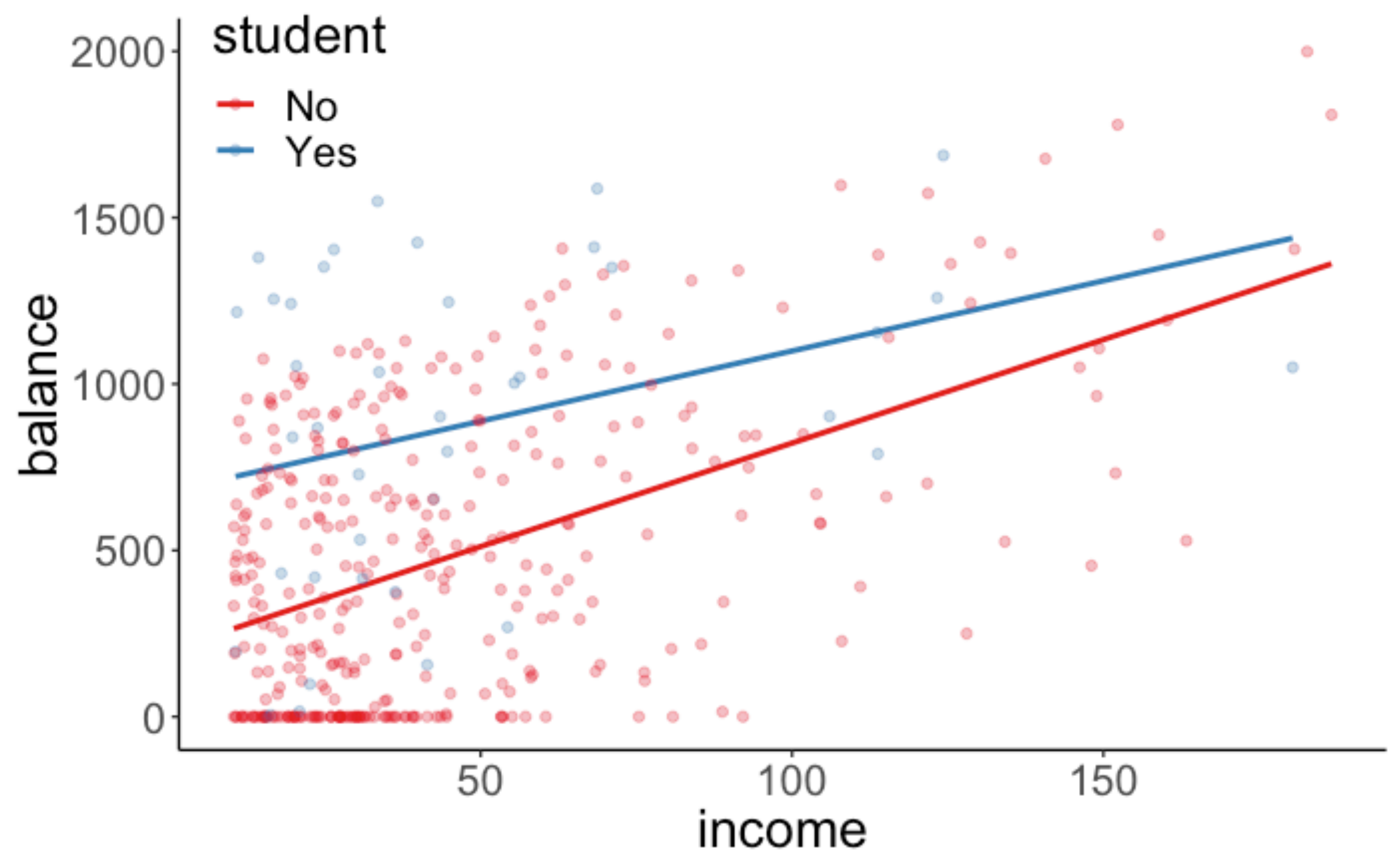
Don't interpret simple effects directly when your model features an interaction!

How to report results of interaction

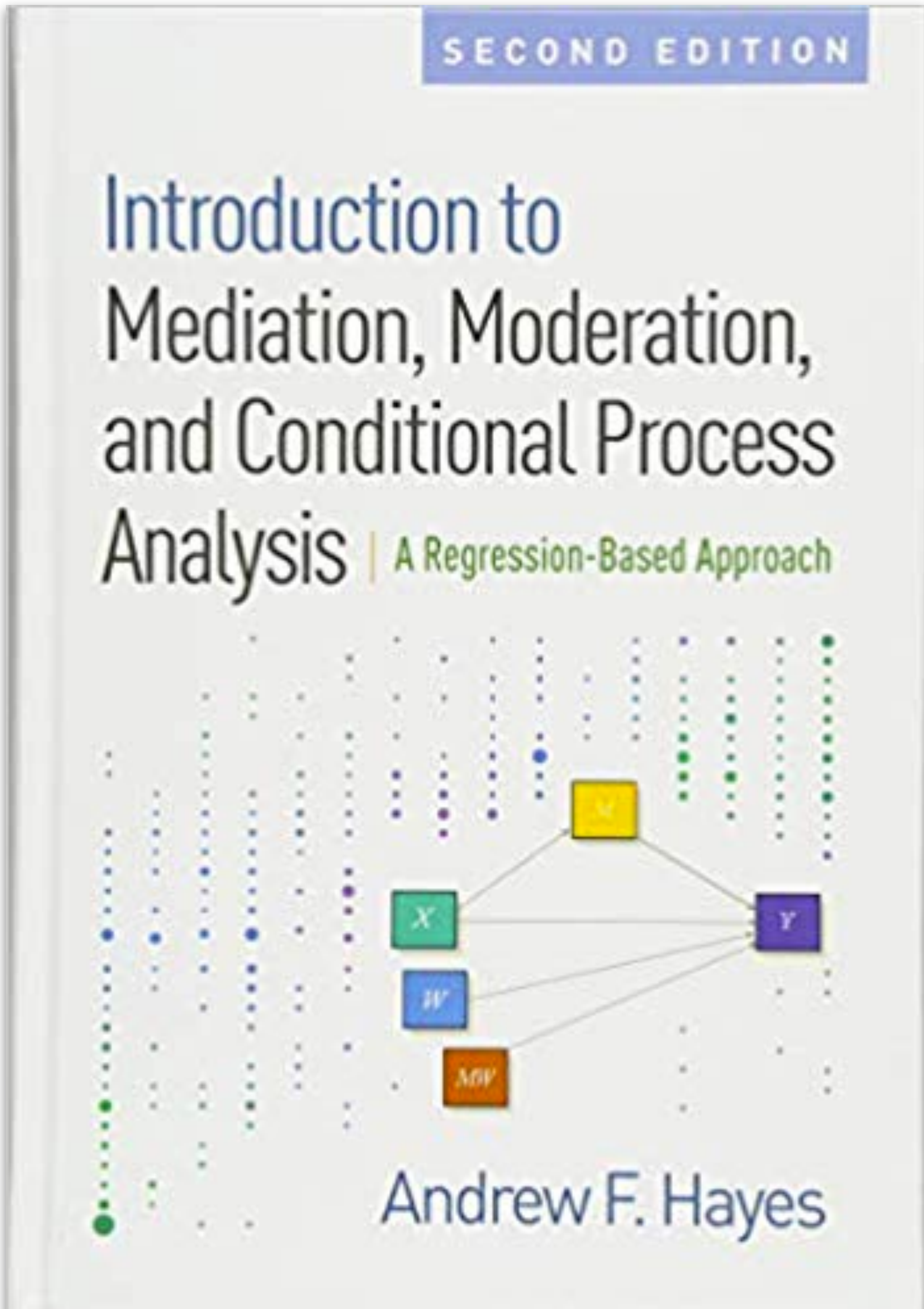
There is no significant difference in the relationship between income and balance for students versus non-students, $F(1, 396) = 1.33$, $p = 0.25$.

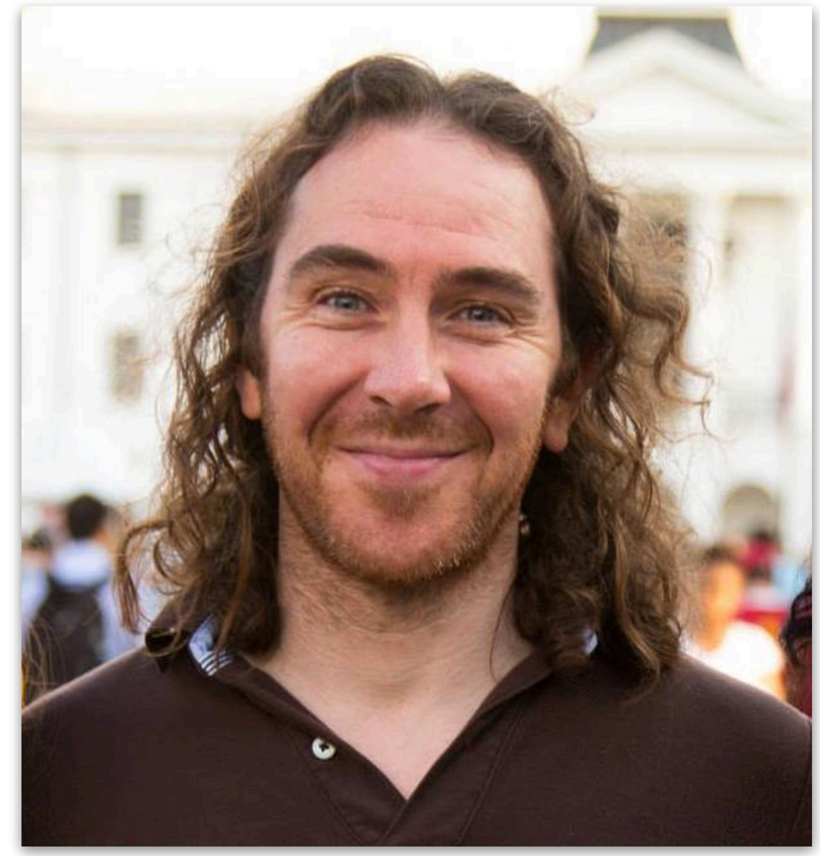
For *students*, an increase in \$1000 income is associated with an increase in \$4.21 of average credit card balance.

For *non-students*, an increase in \$1000 income is associated with an increase in \$6.22 of average credit card balance.



Learn more about mediation and moderation

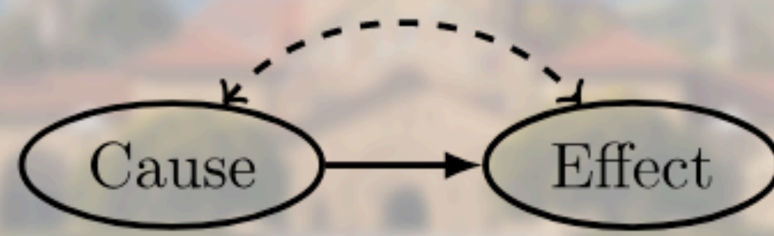




Recoded with `brms` by Solomon Kurz here: https://bookdown.org/connect/#/apps/1523/access

Beyond Curve Fitting: Causation, Counterfactuals, and Imagination-based AI

AAAI Spring Symposium, March 25-27, 2019, Stanford, CA



Motivation

In recent years, Artificial Intelligence and Machine Learning have received enormous attention from the general public, primarily because of the successful application of deep neural networks in computer vision, natural language processing, and game playing (more notably through reinforcement learning). We see AI recognizing faces with high accuracy, Alexa answering English spoken questions efficiently, and Alpha-Zero beating Go grandmasters. These are impressive achievements, almost unimaginable a few years ago. Despite the progress, there is a growing segment of the scientific community that questions whether these successes can be extrapolated to create general AI without a major retooling. Prominent scholars voice concerns that some critical pieces of the AI-puzzle are still pretty much missing. For example, Judea Pearl, who championed probabilistic reasoning in AI and causal inference, recently said in an interview: "To build truly intelligent machines, teach them cause and effect" (link). In a recent OpEd in the New York Times, Cognitive Scientist Gary Marcus noted: "Causal relationships are where contemporary machine learning techniques start to stumble" (link).

Reporting results

Reporting results

- Visualization
  - show the raw data (or as close to raw as you can)
- Report statistical procedures
  - test statistic, degrees of freedom
- Report appropriate effect sizes
- For more complicated models (e.g. `lmer()`), report the function call itself
- Interpret the model parameters via concrete examples (particularly when you have interactions in the model; and for logistic regression results)
- Bayesian data analysis:
  - report summaries of the posterior distribution
  - report hypothetical outcome plots in the appendix
- **Most importantly**: Use RMarkdown and make your analysis file available online on github (and link to OSF)
- I will make an RMarkdown file available with examples ...

What we've learned

Learning goals

What you will learn

You will learn how to **use R** to …

- read, wrangle, and analyze data
- make publication-ready plots

Understand the philosophy behind null **hypothesis significance testing (NHST)** and **Bayesian statistics** through …

- running computer simulations and visualizing the results

Formulate **research questions as statistical models** and …

- determine which models work for different situations
- check that the model's assumptions are met, how much it matters, and what to do if assumptions aren't met

Communicate what you have learned about your data …

- in short presentations in class, showcasing your visualization and analysis
- in written reports

Contribute to open and **reproducible science** through …

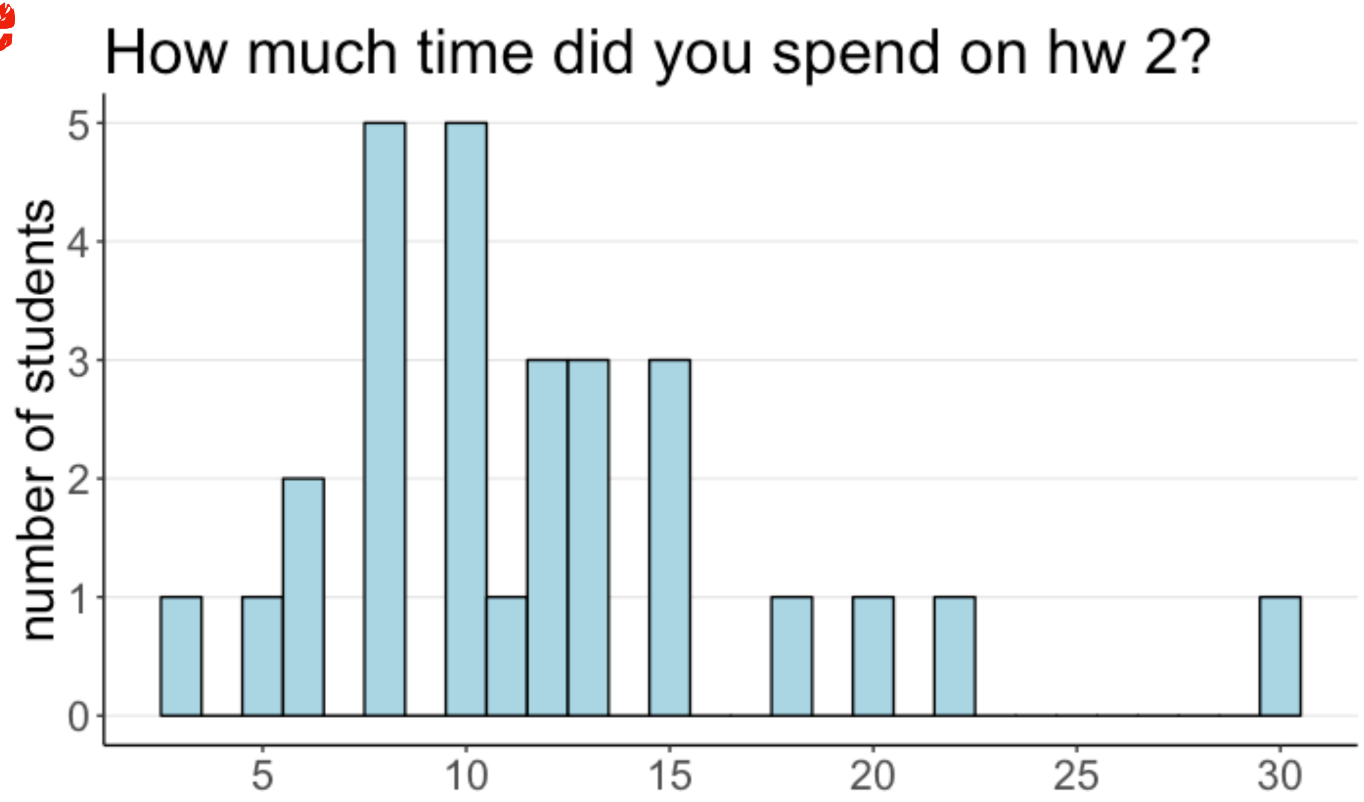
- adopting good coding practices
- sharing your data and research reports online

R Markdown

- if we use a word document ... :
  - figures change
  - results change
  - copy and paste is error prone
- in R Markdown ... :
  - figures and statistics are updated
  - no need for copy and paste
  - everything in one place
  - even better with version control (e.g. via github)

~~Fear~~ of statistics

no more



Thanks for adopting a growth mindset!

fixed mindset:
students believe their basic abilities, their intelligence, their talents, are just fixed traits.

growth mindset:
students *understand* that their talents and abilities can be developed through effort, good teaching and persistence

Vision for this class

In "A Vision for Stanford", university president Marc Tessier-Lavigne states that Stanford wants to be

"an inspired, inclusive and collaborative community of diverse scholars, students and staff, where all are supported and empowered to thrive."

Thanks for making it happen!

I'm looking forward to your presentations!

What shall I do now ?

Psych 253

Basics / Course Schedule & Syllabus / Software Tools

High-Dimensional Methods for Behavioral and Neural Data
Spring 2018, Stanford University

Introduction to high-dimensional data analysis and machine learning methods for use in the behavioral and neurosciences, including: supervised methods such as SVMs, linear and nonlinear regression and classifiers, and regularization techniques; statistical methods such as bootstrapping, signal detection, factor analysis, and reliability theory; metrics for model/data comparison such as representational similarity analysis; and unsupervised methods such as clustering. Students will learn how to both use existing statistical data analysis packages (such as scikit-learn) as well to build, optimize, and estimate their own custom models using the Tensorflow optimization framework.

Time: Tue./Thu. 12:00p - 1:20p

- focus on high-dimensional data
- classification
- representational similarity analysis
- clustering ...
- learn Python!

http://web.stanford.edu/class/cs109/schedule.html

CS109: Probability for Computer Scientists

The class starts by providing a fundamental grounding in combinatorics, and then quickly moves into the basics of probability theory. We will then cover many essential concepts in probability theory, including particular probability distributions, properties of probabilities, and mathematical tools for analyzing probabilities. Finally, the last third of the class will focus on data analysis and Machine Learning as a means for seeing direct applications of probability in this exciting and quickly growing subfield of computer science.

- learn more about probability theory through programming
- gain a deeper understanding of the fundamental underlying concepts

http://web.stanford.edu/class/cs109/schedule.html

Advanced regression analysis

EDUC 326: Advanced Regression Analysis

Social science researchers often deal with complex data and research questions that traditional statistics models like linear regression cannot adequately address. This course offers the opportunity to understand and apply two widely used types of advanced regression analysis that allow the examination of 1) multilevel data structures (multilevel models) and 2) multivariate research questions (structural equation models).

Terms: Spr | Units: 3-4 | Grading: Letter or Credit/No Credit

Instructors: Smith, S. (PI)

Schedule for EDUC 326

- multilevel models
- structural equation models

Data Challenge Lab

Where students develop their data skills by solving a progression of increasingly difficult challenges

ENGR 150: Data Challenge Lab
Terms: Win, Spr | Units: 5
Instructors: Bill Behrman, Hadley Wickham

Prof. Tidyverse

https://datalab.stanford.edu/challenge-lab

Thanks to you!

Psych 252 Team

Pam

Andrew

Mona

All of you!

Almeida Santos, Luiza
Auelua-Toomey, Sai Laisene
Beller, Aaron
Benson, Ian
Boles, Danielle Zoe
Borchers, Lauren Rebekah
Cachia, Julie Youko Anne
Chase, Elyse Deanna Zaino
Chen, Xinjie
Dauer, Tysen Drew
deMayo, Benjamin Edward
Demszky, Dora
Dietz, Griffin
E, Jane Little
Finzi, Dawn
Gonzalez, Eric Jordan
Gurrapu, Shravya
Harrison, Marc Brookes
Horii, Rina Imari
Hosseini, Zainab
Hsu, Shaw
Hsu, Tiffany
Hussein, Mohamed
Kahhale, Isabella
Kreiss, Elisa
Li, Jingyi
Louis, Kengthsagn
Metaxa, D
Miller, Mark Roman
Miri, Pardis
Park, So Yeon
Portelance, Eva
Rajagopal, Tara
Reinhart, Ellen C.
Rosier, Soren Rousseau
Sami, Sohrab
Samuels, Natalie
Siu, Alexa Fay
Srinivasan, Preeti
Strasnick, Evan Nicholas
Tay, Isabelle Q.
Tsui, Angeline Sin Mei
Wang, Karen Dan
Wang, Michelle Momo
Weitz, Elizabeth Alexandra
Wilcox, John Eric
Zhang, Jinxiao
Zhang, Marianna Yunjia
Zhang, Vivian

Thank you!

Pooja

Mr Fluffington

Feedback

Your feedback

Dear Tobias Gerstenberg,

Axess is now open for students to provide course feedback until 08:00 AM on Mon, Mar 25, 2019 PDT. Please direct your students to complete their feedback at Axess > Student > Course and Section Evaluations. Students who complete all of their feedback will see their grades as soon as they have been submitted by the faculty. Students who do not complete all of their feedback will not be able to see their grades in Axess until the day after the grade release deadline.

Here are some tips to encourage students to complete their feedback:

- Take a few minutes of class time to allow students to complete feedback;
- Explain that you, the instructor, will only see aggregated, anonymous responses; and,
- Tell students that their feedback is very important and helps you to make improvements to the course for future students.

Axess > Student > Course > Section Evaluations

If you provide feedback, you'll see your grades
as soon as they have been submitted.
Otherwise, you'll have to wait a little longer.

I will only see aggregated, anonymous responses

Your feedback will help us make improvements next year!

Thank you!