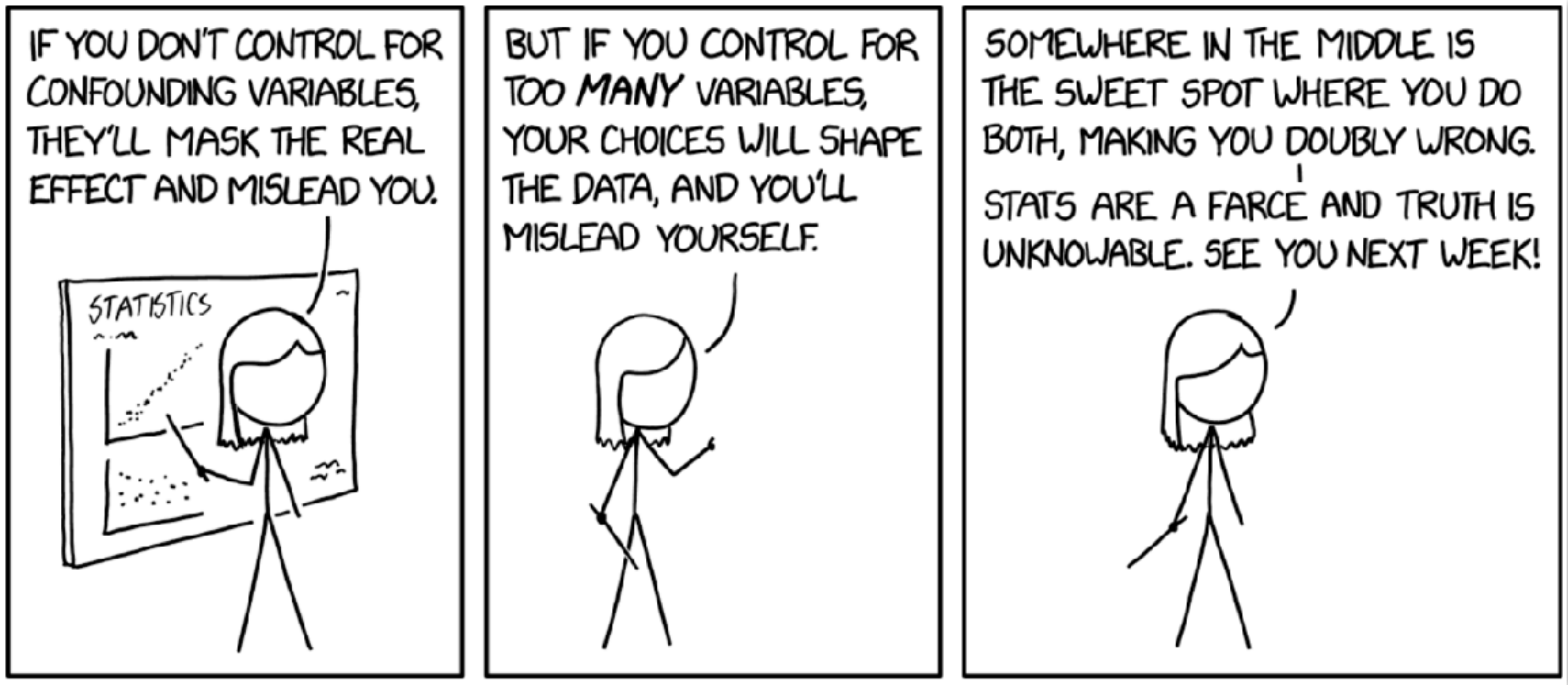
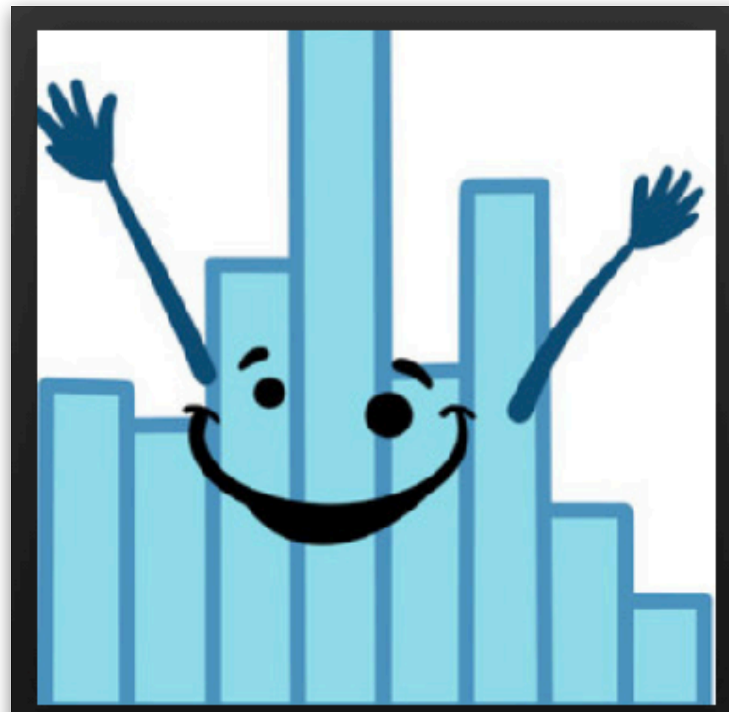
Causation





COLLABORATIVE PLAYLIST

psych252

https://tinyurl.com/psych252spotify25

PLAY

02/28/2025

Feedback

Responses cannot be edited

Psych 252 anonymous feedback form

Please use the text box below if you'd like to provide anonymous feedback for Psych 252.

* Indicates required question

*

I wish more posts were visible on Ed Discussion. I have heard from instructors and students that questions are being asked and answered, but few of them are made public.

Logistics

Homework 6

Linear mixed effects models

My name and the names of the people and AI I have worked with go here

2025-02-26 11:15:59.062313

Instructions

This homework is due by **Thursday, March 6th, 8:00pm**.

Note:

- When asked to report results, please do so like you would in a scientific article (see examples from lectures, as well as in 'Reporting Results.pdf' on Canvas under Files > papers).
- Some code chunks contain some skeleton code. The code chunk option for these chunks is set to `eval=F` so that knitting the RMarkdown document doesn't throw any errors. Make sure to set these chunks to `eval=T` when you knit your homework, so that your calculations are shown in the pdf.
- Make sure to show the results of your calculations in the knitted pdf, for example, by using the `print()` function at the end of a code chunk.
- Some questions ask for a short written response as indicated by this prompt: **Your answer:**

Part 1 (4 points)

Load and visualize data

Here we have (simulated) data on measurements of extroversion from students nested within classes. Extroversion is one of the big-5 personality traits. We would like to know whether a student's openness and agreeableness predicts their social behavior (extroversion). We know the data are non-independent because the students were sampled from 4 different classes, and we might expect that students from the same class are more similar to each other than students from different classes. We need to take this dependence into account when modeling our data.

First, let's Load the data set.

```
df.class = read_csv("data/extra_class.csv")
```

1.1 (0.5 points)

Below is a plot that shows the relationship between openness (`open`) on the x-axis and extroversion (`extra`) on the y-axis for you, separated by class ("a"-"d"). Briefly describe the plot.

```
ggplot(data = df.class,
       mapping = aes(x = open,
                     y = extra)) +
  geom_point(alpha = .5) +
  geom_smooth(method = "lm") +
  facet_wrap(~ class)
```

1.4 (1 point)

Add a fixed effect of agreeableness (`agree`) to the linear mixed effects model (with random intercepts for `class`). Is a model that takes both aggreeableness and openness into account when trying to predict extroversion significantly better than a mixed effects model that only considers openness?

> Tip: You can use a likelihood ratio test (via the `anova()` function), or use the `joint_tests()` function from the "emmeans" package.

```
### YOUR CODE HERE ###

######################
```

Your answer:

1.5 (1 point)

In addition to the random intercept, add random slopes for the agreeableness-extraversion relation to the linear mixed effects model that has both agreeableness and openness as fixed effects. Fit the model and compare it to the model without random slopes. Is adding the random slopes worth it?

```
### YOUR CODE HERE ###

######################
```

Your answer:

Part 2 (4 points)

A recent paper in Psychological Science reported that children as young as 4 years of age use probability to infer how good an outcome is. You can check out the original paper here. They share the data on OSF so we can try to reproduce their analysis using the tools we have learned in class.

In Experiment 4, the `Mostly Yummy` vs. `Mostly Yucky` condition was manipulated within participants (and the order was counterbalanced). At the end of each experiment, children gave a rating from -3 to +3 to either a **happiness** question ("How does the girl feel about the gumballs that she got?") or a **quality** question ("How good was that?"). The type of the question was manipulated between participants. You can read the experiment procedures in more details in the figure below.

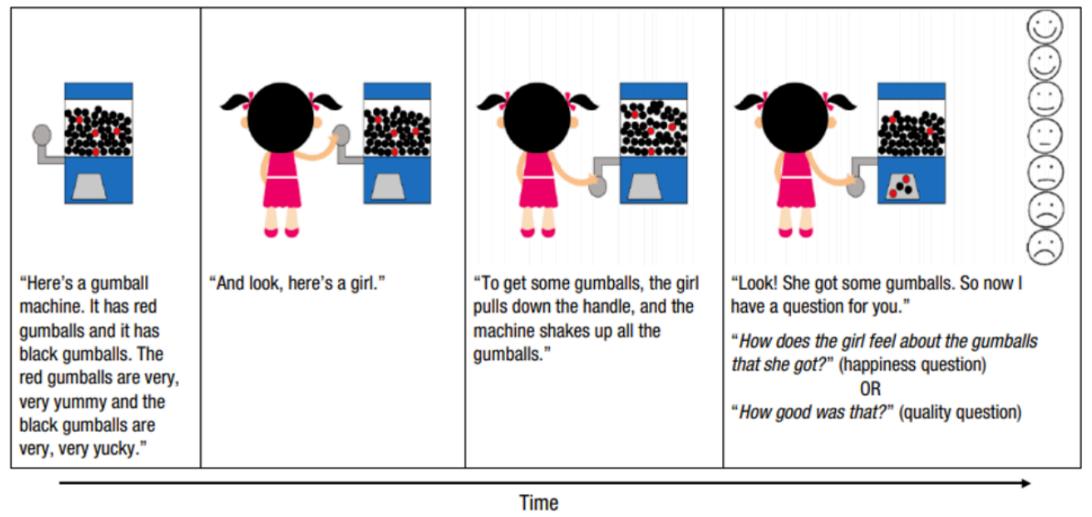


Fig. 6. Sample slides and script for the mostly yucky condition in Experiment 4. The mostly yummy condition was identical, but the distribution of yummy and yucky gumballs was reversed.

Plan for today

- Quick recap
- Linear Mixed Model
  - Different random effect structures
  - lmer() standard operating procedures
- Causation
  - Causation vs. correlation
  - Controlling for variables
  - Mediation
  - Moderation

Quick recap

Quick recap: Simulation

Let's simulate an `lmer()`

```r
# make example reproducible
set.seed(1)

# parameters
sample_size = 100
b0 = 1
b1 = 2
sd_residual = 1
sd_participant_intercept = 0.5

# generate the data
df.mixed = tibble(participant = rep(1:sample_size, 2),
                  condition = rep(0:1, each = sample_size)) %>%
  group_by(participant) %>%
  mutate(participant_intercepts = rnorm(n = 1, sd = sd_participant_intercept)) %>%
  ungroup() %>%
  mutate(value = b0 + b1 * condition + participant_intercept + rnorm(n(), sd = sd_residual)) %>%
  arrange(participant, condition)
```

participant	condition	participant_intercept	value
1	0	-0.31	0.07
1	1	-0.31	3.10
2	0	0.09	1.13
2	1	0.09	4.78
3	0	-0.42	-0.33
3	1	-0.42	4.17
4	0	0.80	1.96
4	1	0.80	3.47
5	0	0.16	0.51
5	1	0.16	0.88

$$\text{value}_{ij} = b_0 + b_1 \cdot \text{condition}_{ij} + U_i + e_{ij}$$

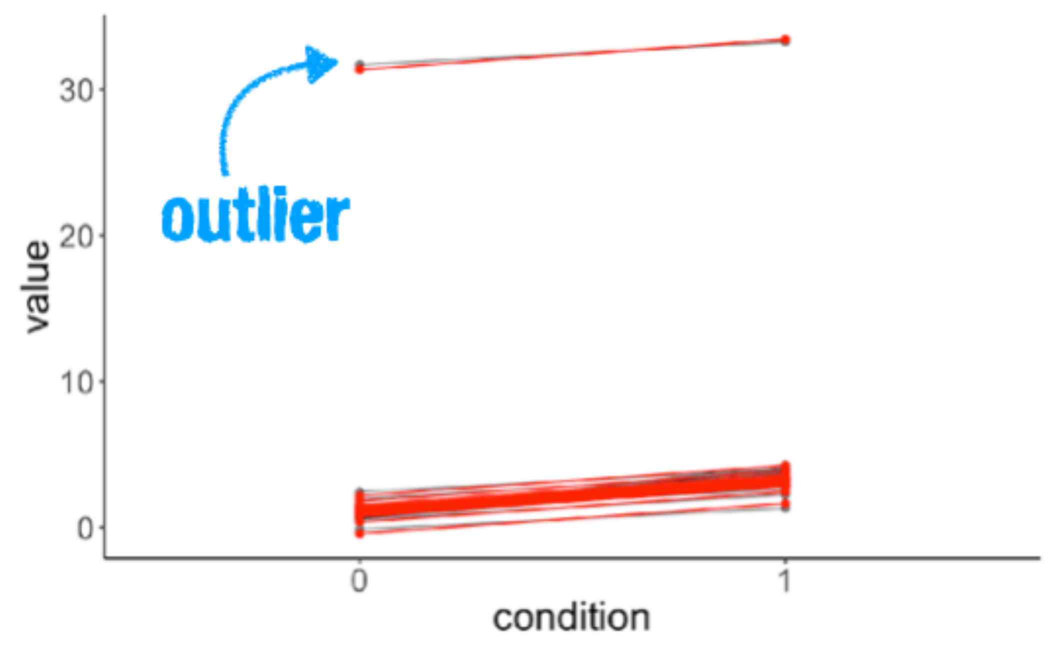
$$e_{ij} \sim \mathcal{N}(\text{mean} = 0, \text{sd} = s_{\text{error}})$$

$$U_i \sim \mathcal{N}(\text{mean} = 0, \text{sd} = s_{\text{U}})$$

simulating data from a model and trying to recover the parameters is a great way to check one's understanding of what the model does

15

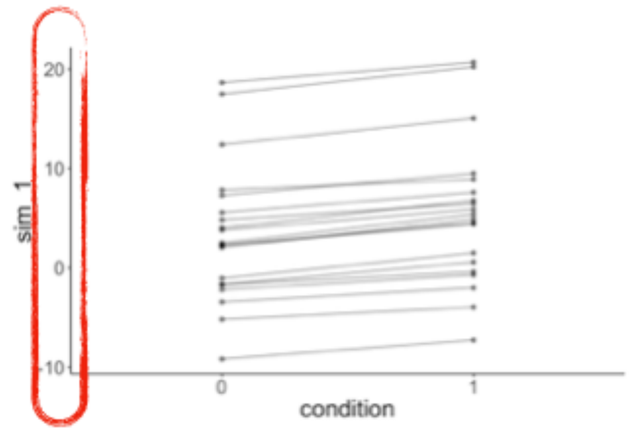
With outlier




```
Linear mixed model fit by REML ['lmerMod']
Formula: value ~ 1 + condition + (1 | participant)
   Data: df.test

REML criterion at convergence: 171.7

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.4038 -0.4678 -0.0094  0.5800  1.3930

Random effects:
 Groups      Name        Variance Std.Dev.
 participant (Intercept) 46.198   6.7969
 Residual                 0.227   0.4764
Number of obs: 40, groups:  participant, 20

Fixed effects:
            Estimate Std. Error t value
(Intercept)   2.5920     1.5236   1.701
condition1    2.0726     0.1507  13.758

Correlation of Fixed Effects:
           (Intr)
condition1 -0.049
```

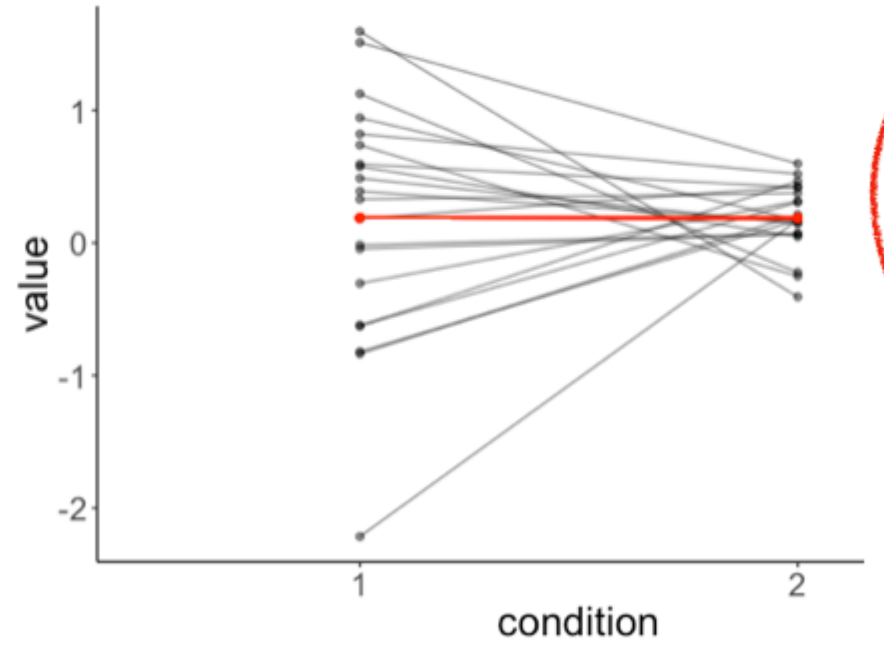
```r
# fit model
fit.test = lmer(formula = value ~ 1 + condition + (1 | participant),
                data = df.test)

# simulate data
fit.test %>%
  simulate()
```

simulated data



18

Non-equal variance




```
singular fit
Linear mixed model fit by REML ['lmerMod']
Formula: value ~ 1 + condition + (1 | participant)
   Data: df.test

REML criterion at convergence: 83.6

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.5808 -0.3184  0.0130  0.4551  2.0913

Random effects:
 Groups      Name        Variance Std.Dev.
 participant (Intercept) 0.0000   0.0000
 Residual                0.4512   0.6717
Number of obs: 40, groups:  participant, 20

Fixed effects:
             Estimate Std. Error t value
(Intercept)  0.190524   0.150197   1.268
condition2  -0.001941   0.212411  -0.009

Correlation of Fixed Effects:
           (Intr)
condition2 -0.707
convergence code: 0
singular fit
```

clearly there are interindividual differences though?!

19

Quick recap: Worked examples

Sleep data

fit the model

```r
fit = lmer(formula = reaction ~ 1 + days + (1 | subject),
           data = df.sleep %>%
             mutate(days = as.factor(days)))

contrast = list(first_vs_second = c(-1, 1, rep(0, 8)),
                early_vs_late = c(rep(-1, 5)/5, rep(1, 5)/5))

fit %>%
  emmeans(specs = "days",
          contr = contrast) %>%
  pluck("contrasts")
```

define the contrasts

test the contrasts

contrast	estimate	SE	df	t.ratio	p.value
first_vs_second	7.82	10.10	156	0.775	0.4398
early_vs_late	53.66	4.65	155	11.534	<.0001

Degrees-of-freedom method: kenward-roger

days	reaction
0	257.54
1	265.73

index	reaction
early	271.67
late	325.39

https://aosmith.rbind.io/2019/03/25/getting-started-with-emmeans/

Weight loss data

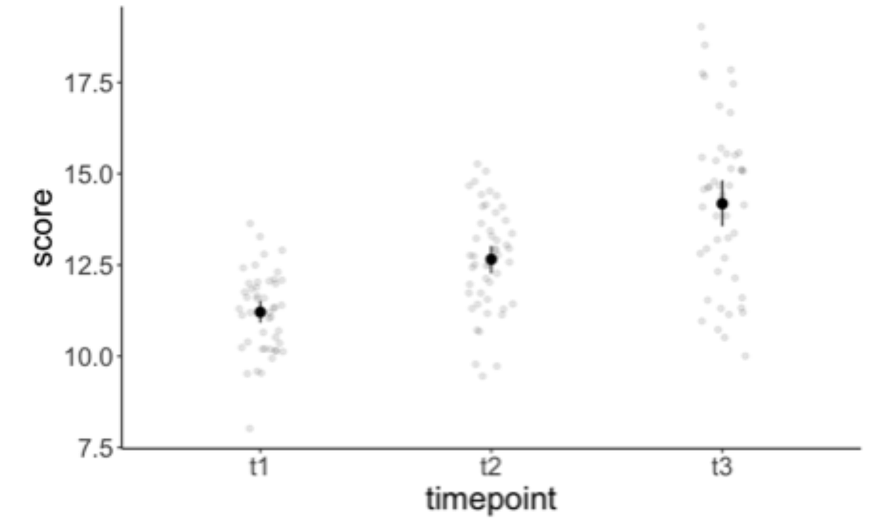
1. Is the score at the third time point different from the other two time points?
2. Is there a linear increase across time points?

df.weightloss

id	diet	exercises	timepoint	score
1	no	no	t1	10.43
1	no	no	t2	13.21
1	no	no	t3	11.59
1	yes	no	t1	10.20
1	yes	no	t2	12.51
1	yes	no	t3	14.60
2	no	no	t1	11.59

```r
fit = aov_ez(id = "id",
       dv = "score",
       between = "exercises",
       within = c("diet", "timepoint"),
       data = df.weightloss)

contrasts = list(first_two_vs_last = c(-0.5, -0.5, 1),
                 linear_increase = c(-1, 0, 1))

fit %>%
  emmeans(spec = "timepoint",
          contr = contrasts)
```

contrast	estimate	SE	df	t.ratio	p.value
first_two_vs_last	2.24	0.200	44	11.194	<.0001
linear_increase	2.97	0.231	44	12.820	<.0001



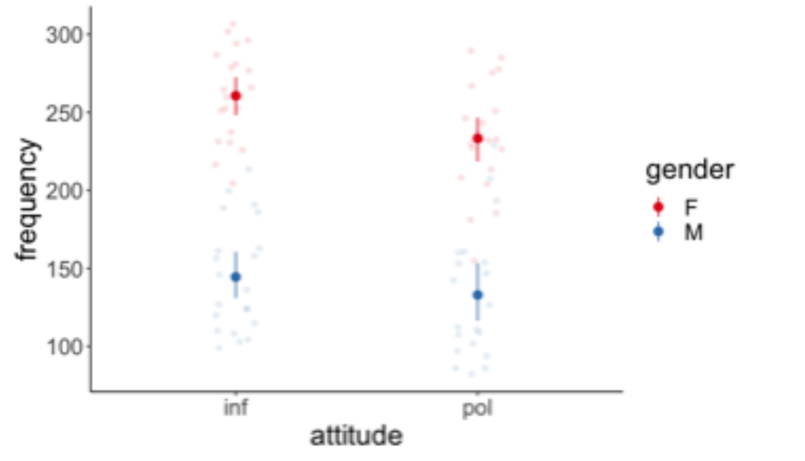
31

Politeness

Was there an effect of gender and attitude on pitch?

```r
lmer(formula = frequency ~ 1 + attitude * gender + (1 + attitude | subject) + (1 | scenario),
     data = df.politeness) %>%
  joint_tests()
```

model term	df1	df2	F.ratio	p.value
attitude	1	3.99	12.411	0.0244
gender	1	4.00	26.570	0.0067
attitude:gender	1	3.99	1.959	0.2342



main effect of attitude, main effect of gender, no significant interaction effect

Linear mixed effects model

Different random effect structures

R <- PSYCHOLOGIST

About Posts Visualizations

Contents

Using R and lme/lmer to fit different two- and three-level longitudinal models

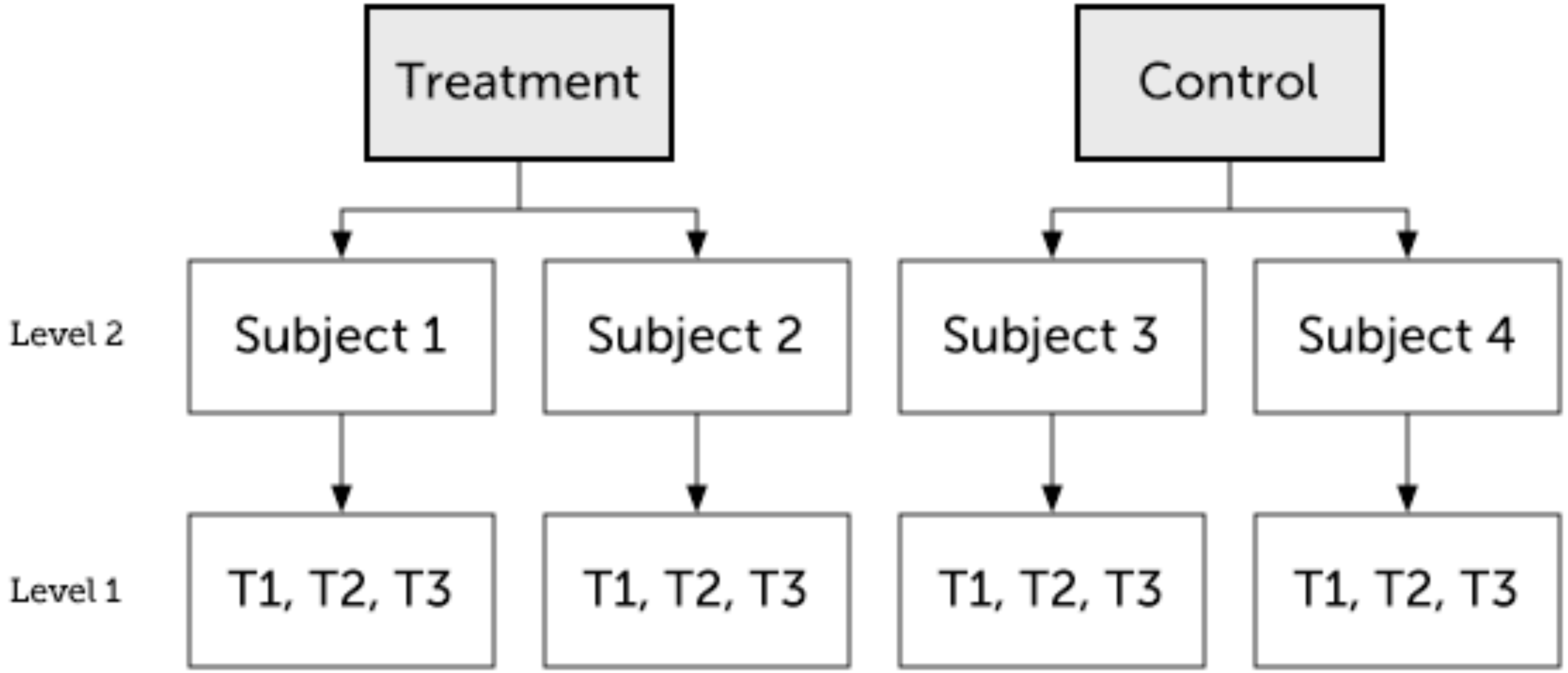
April 21, 2015

I often get asked how to fit different multilevel models (or individual growth models, hierarchical linear models or linear mixed-models, etc.) in R. In this guide I have compiled some of the more common and/or useful models (at least common in clinical psychology), and how to fit them using `nlme::lme()` and `lme4::lmer()`. I will cover the common two-level random intercept-slope model, and three-level models when subjects are clustered due to some higher level grouping (such as therapists), partially nested models were there are clustering in one group but not the other, and different level 1 residual covariances (such as AR(1)). The point of this post is to show how to fit these longitudinal models in R, not to cover the statistical theory behind them, or how to interpret them.

https://rpsychologist.com/r-guide-longitudinal-lme-lmer#power-analysis-and-simulating-these-models

Two-level model

Graphical representation

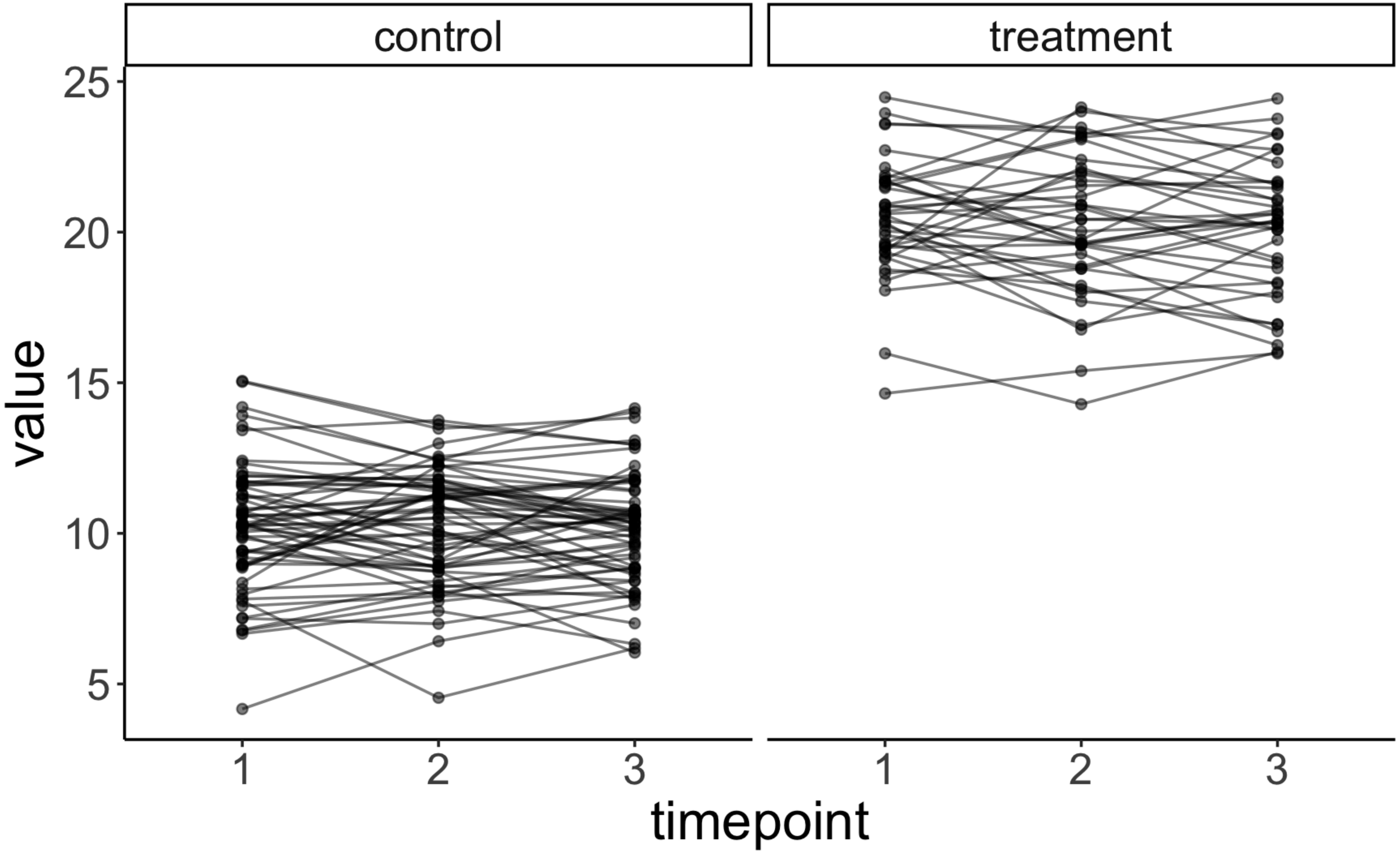


Simulate data

	participant	timepoint	intercept_participant	condition	value
1	1	1	-1.25290762	0	9.197279
2	1	2	-1.25290762	0	8.728533
3	1	3	-1.25290762	0	8.429024
4	2	1	0.36728665	0	9.437925
5	2	2	0.36728665	0	8.879826
6	2	3	0.36728665	0	9.292094
7	3	1	-1.67125722	1	19.328772
8	3	2	-1.67125722	1	17.707476
9	3	3	-1.67125722	1	16.944316
10	4	1	3.19056160	0	15.059852
11	4	2	3.19056160	0	13.615662
12	4	3	3.19056160	0	12.951915

```r
set.seed(1)

n_participants = 100
n_timepoints = 3
n_conditions = 2
p_condition = 0.5
b0 = 10
b1 = 10
sd_participant = 2
sd_residual = 1

df.data = tibble(participant = rep(1:n_participants, each = n_timepoints),
                 timepoint = rep(1:n_timepoints, times = n_participants),
                 intercept_participant = rep(rnorm(n_participants, sd = sd_participant),
                                             each = n_timepoints)) %>%
  group_by(participant) %>%
  mutate(condition = rbinom(n = 1, size = 1, prob = p_condition)) %>%
  ungroup() %>%
  mutate(value = b0 + b1 * condition + intercept_participant +
           rnorm(n_participants * n_timepoints, sd = sd_residual))
```

Plot data



Fit the model

```
fit = lmer(formula = value ~ 1 + condition + (1 | participant),
           data = df.data)
```

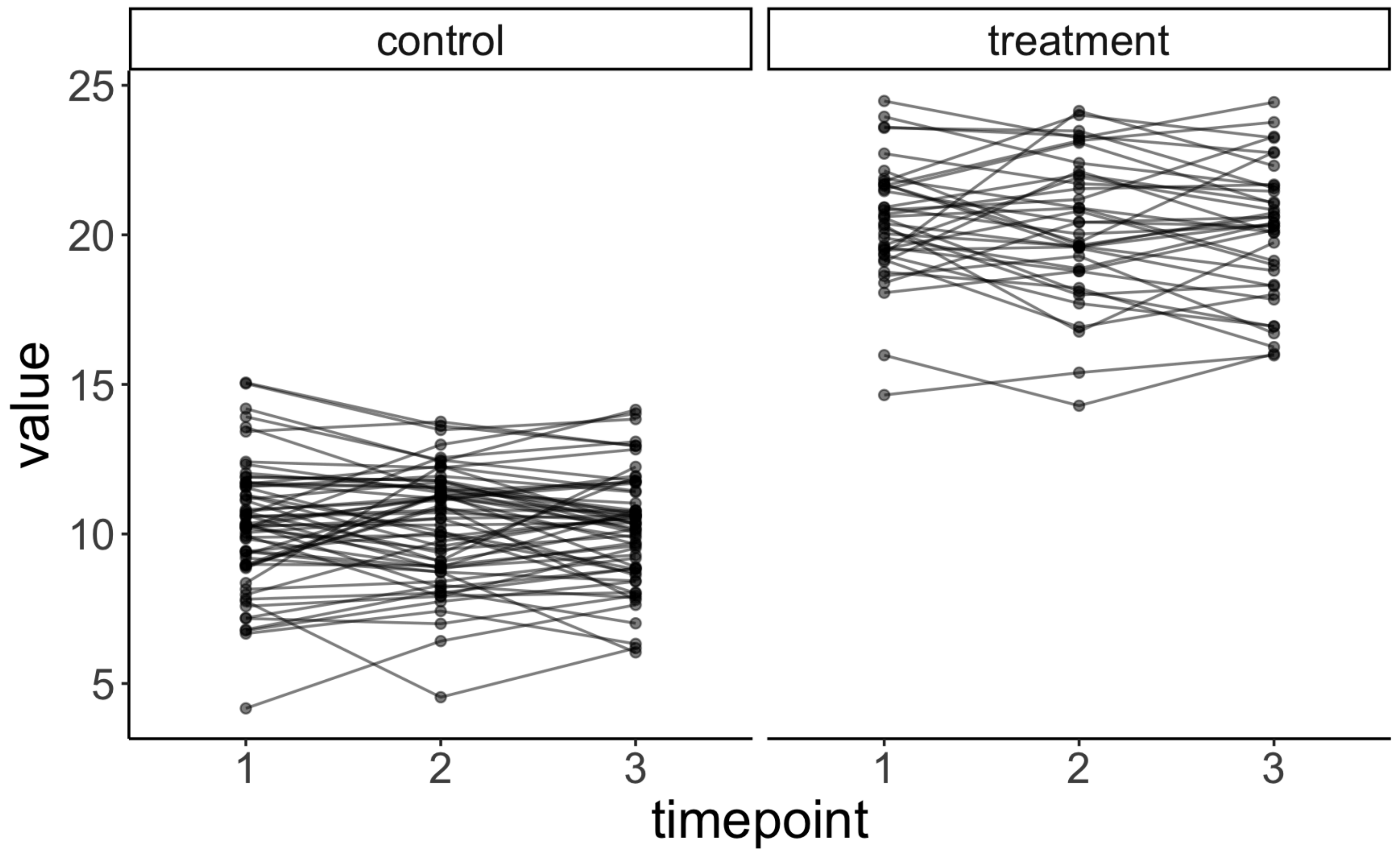
```
b0 = 10
b1 = 10
sd_participant = 2
sd_residual = 1
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: value ~ 1 + condition + (1 | participant)
   Data: df.data

REML criterion at convergence: 1102

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.30522 -0.57146  0.03152  0.56826  2.28135

Random effects:
 Groups      Name        Variance Std.Dev.
 participant (Intercept) 3.106    1.762
 Residual                1.087    1.043
Number of obs: 300, groups:  participant, 100

Fixed effects:
            Estimate Std. Error t value
(Intercept)  10.2199     0.2365   43.21
condition    10.0461     0.3837   26.18

Correlation of Fixed Effects:
          (Intr)
condition -0.616
```

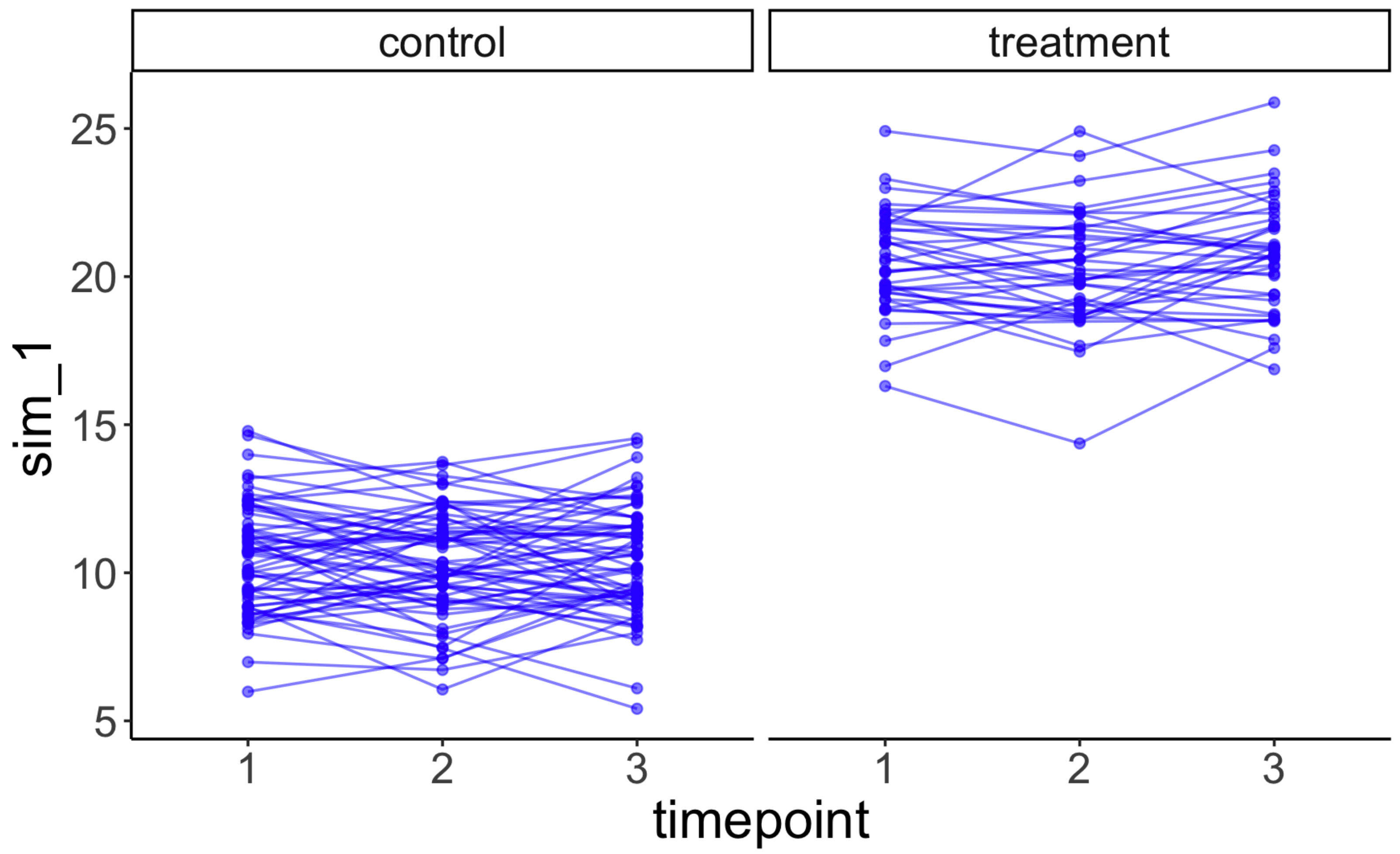
Simulate new data from the model

```r
set.seed(1)

fit %>%
  simulate() %>%
  bind_cols(df.data) %>%
  mutate(condition = factor(condition,
                            levels = c(0, 1),
                            labels = c("control", "treatment")),
         timepoint = as.factor(timepoint)) %>%
  ggplot(data = .,
         mapping = aes(x = timepoint,
                       y = sim_1,
                       group = participant)) +
  geom_point(alpha = 0.5,
             color = "blue") +
  geom_line(alpha = 0.5,
            color = "blue") +
  facet_grid(~ condition) +
  labs(x = "timepoint")
```

original data

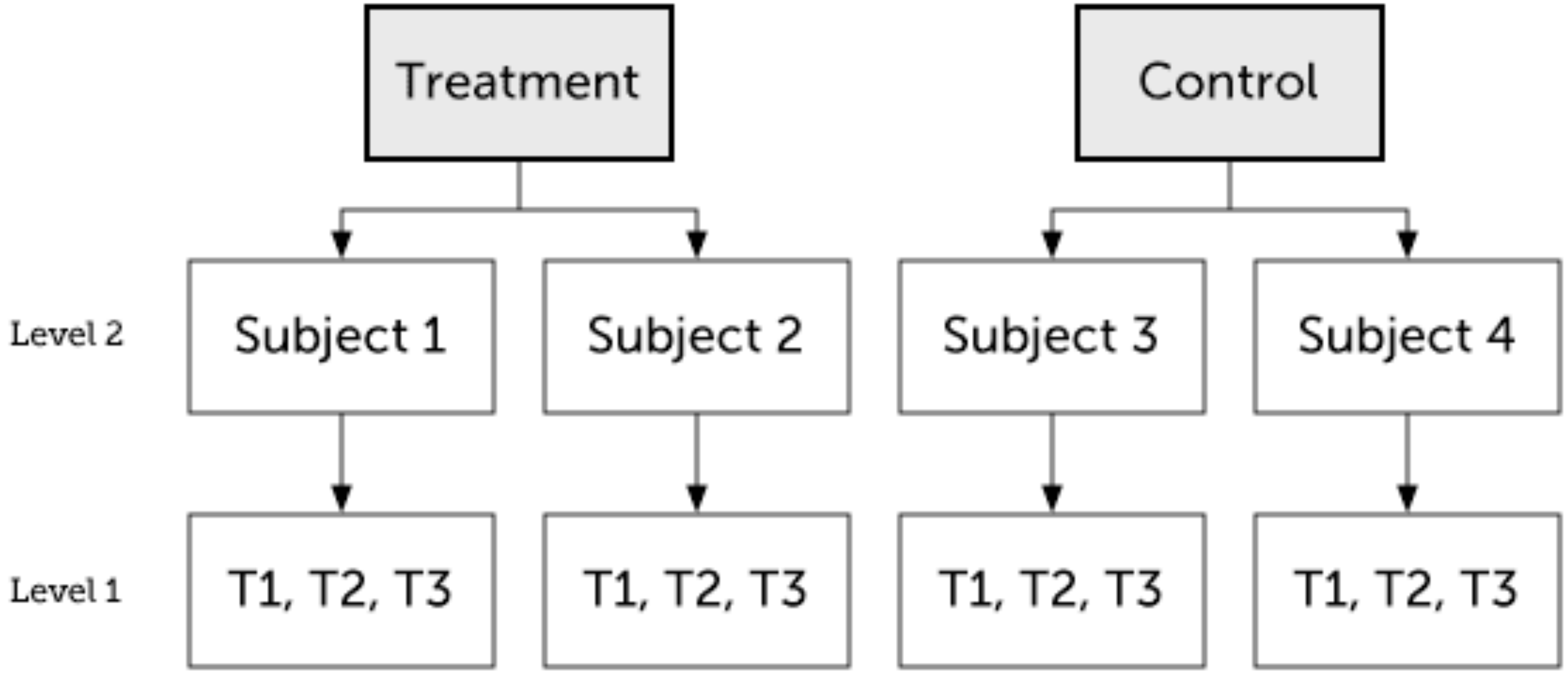


simulated data



Two-level growth model

Graphical representation



assume a linear effect of time

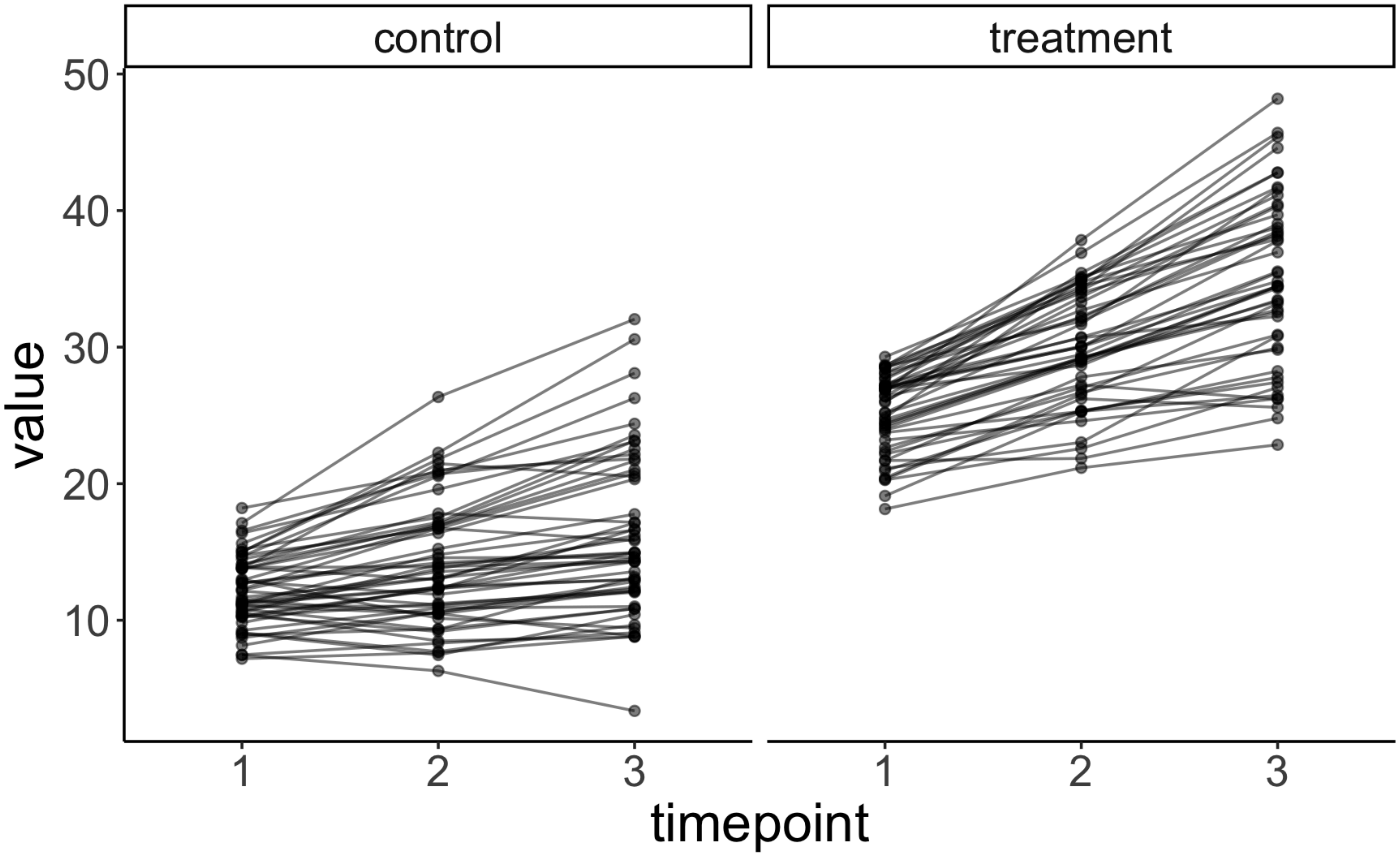
Simulate data

	participant	timepoint	intercept_participant	time_participant	condition	value
1	1	1	-1.252907621	-0.310183339	1	20.573131
2	1	2	-1.252907621	-0.310183339	1	22.844077
3	1	3	-1.252907621	-0.310183339	1	24.367254
4	2	1	0.367286648	0.021057937	0	12.119622
5	2	2	0.367286648	0.021057937	0	15.062837
6	2	3	0.367286648	0.021057937	0	17.513515
7	3	1	-1.671257225	-0.455460824	1	17.470186
8	3	2	-1.671257225	-0.455460824	1	22.446022
9	3	3	-1.671257225	-0.455460824	1	24.248006
10	4	1	3.190561604	0.079014386	1	24.844308
11	4	2	3.190561604	0.079014386	1	28.220589
12	4	3	3.190561604	0.079014386	1	28.880339
13	5	1	0.659015544	-0.327292322	1	22.047393
14	5	2	0.659015544	-0.327292322	1	25.189133
15	5	3	0.659015544	-0.327292322	1	28.051351

```r
set.seed(1)

n_participants = 100
n_timepoints = 3
n_conditions = 2
p_condition = 0.5
b0 = 10 # intercept
b1 = 10 # condition
b2 = 2 # time
b3 = 3 # interaction
sd_intercept_participant = 2
sd_time_participant = 2
sd_residual = 1

df.data = tibble(participant = rep(1:n_participants, each = n_timepoints),
                 timepoint = rep(1:n_timepoints, times = n_participants),
                 intercept_participant = rep(rnorm(n_participants, sd = sd_intercept_participant),
                                             each = n_timepoints),
                 time_participant = rep(rnorm(n_participants, sd = sd_time_participant),
                                        each = n_timepoints)) %>%
  group_by(participant) %>%
  mutate(condition = rbinom(n = 1, size = 1, prob = p_condition)) %>%
  ungroup() %>%
  mutate(value = b0 + intercept_participant +
           b1 * condition +
           (b2  + time_participant) * timepoint +
           b3 * condition * timepoint +
           rnorm(n_participants * n_timepoints, sd = sd_residual))
```

random slopes

Plot data

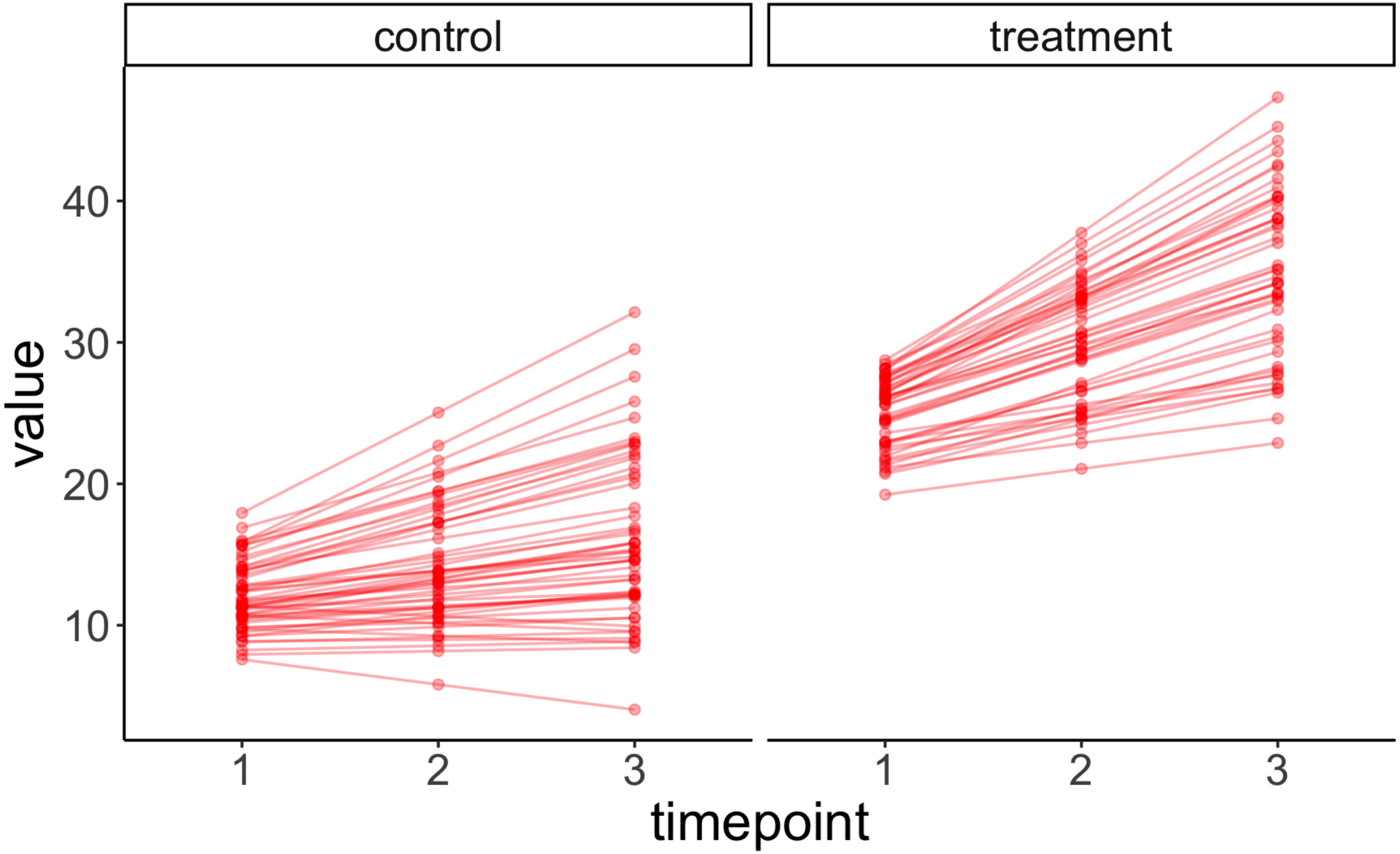


Fit the model

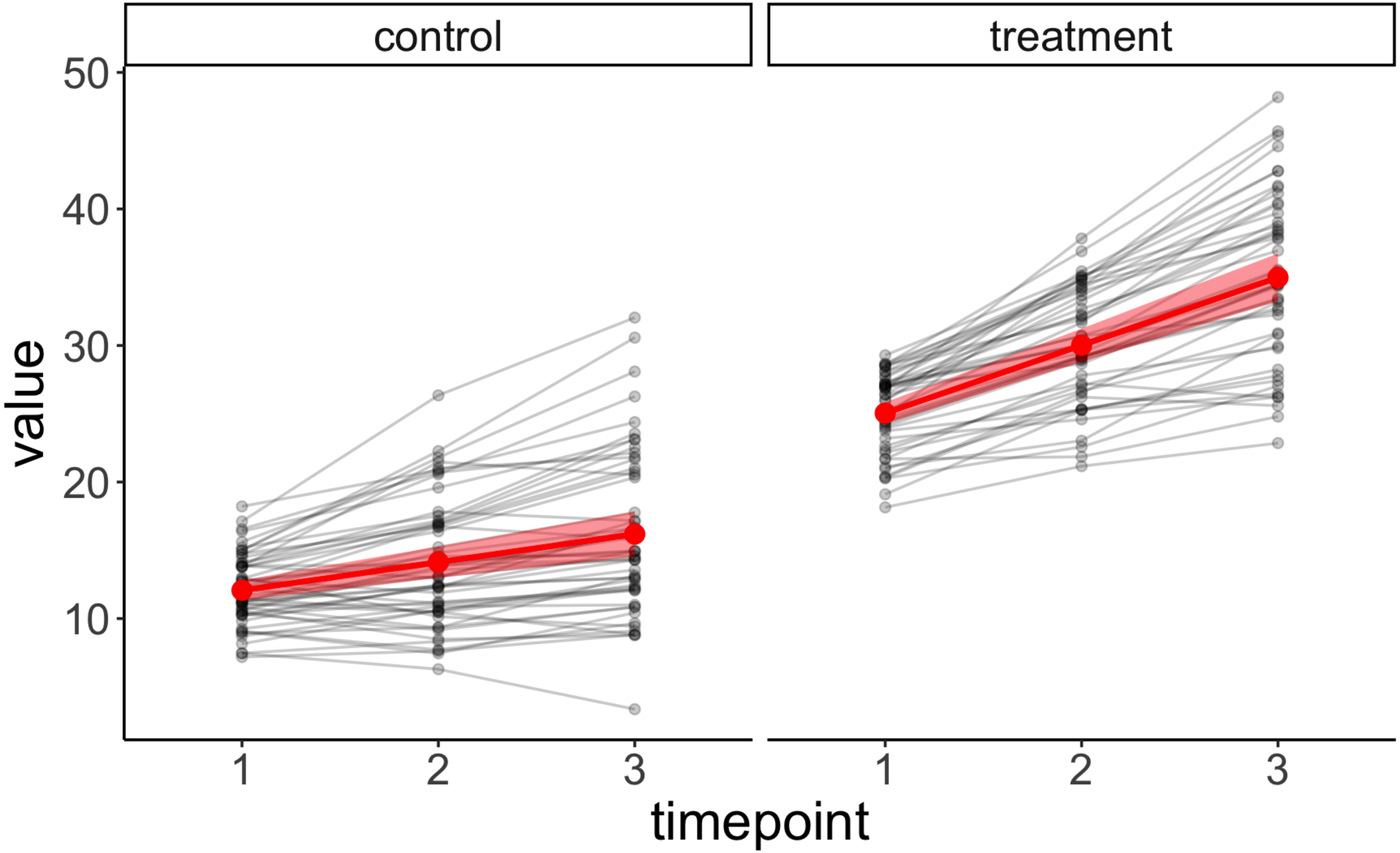
```r
fit = lmer(formula = value ~ 1 + condition * timepoint + (1 + timepoint | participant),
           data = df.data)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: value ~ 1 + condition * timepoint + (1 + timepoint | participant)
   Data: df.data

REML criterion at convergence: 1360.3

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.14633 -0.46360  0.03902  0.42302  2.82945

Random effects:
 Groups      Name        Variance Std.Dev. Corr
 participant (Intercept) 3.190    1.786
             timepoint   3.831    1.957    -0.06
 Residual                1.149    1.072
Number of obs: 300, groups:  participant, 100

Fixed effects:
                    Estimate Std. Error t value
(Intercept)          10.0101     0.3328  30.079
condition            10.0684     0.4854  20.741
timepoint             2.0595     0.2883   7.143
condition:timepoint   2.9090     0.4205   6.917

Correlation of Fixed Effects:
            (Intr) condtn timpnt
condition   -0.686
timepoint   -0.266  0.182
cndtn:tmpnt  0.182 -0.266 -0.686
```

Visualize model predictions (individual)



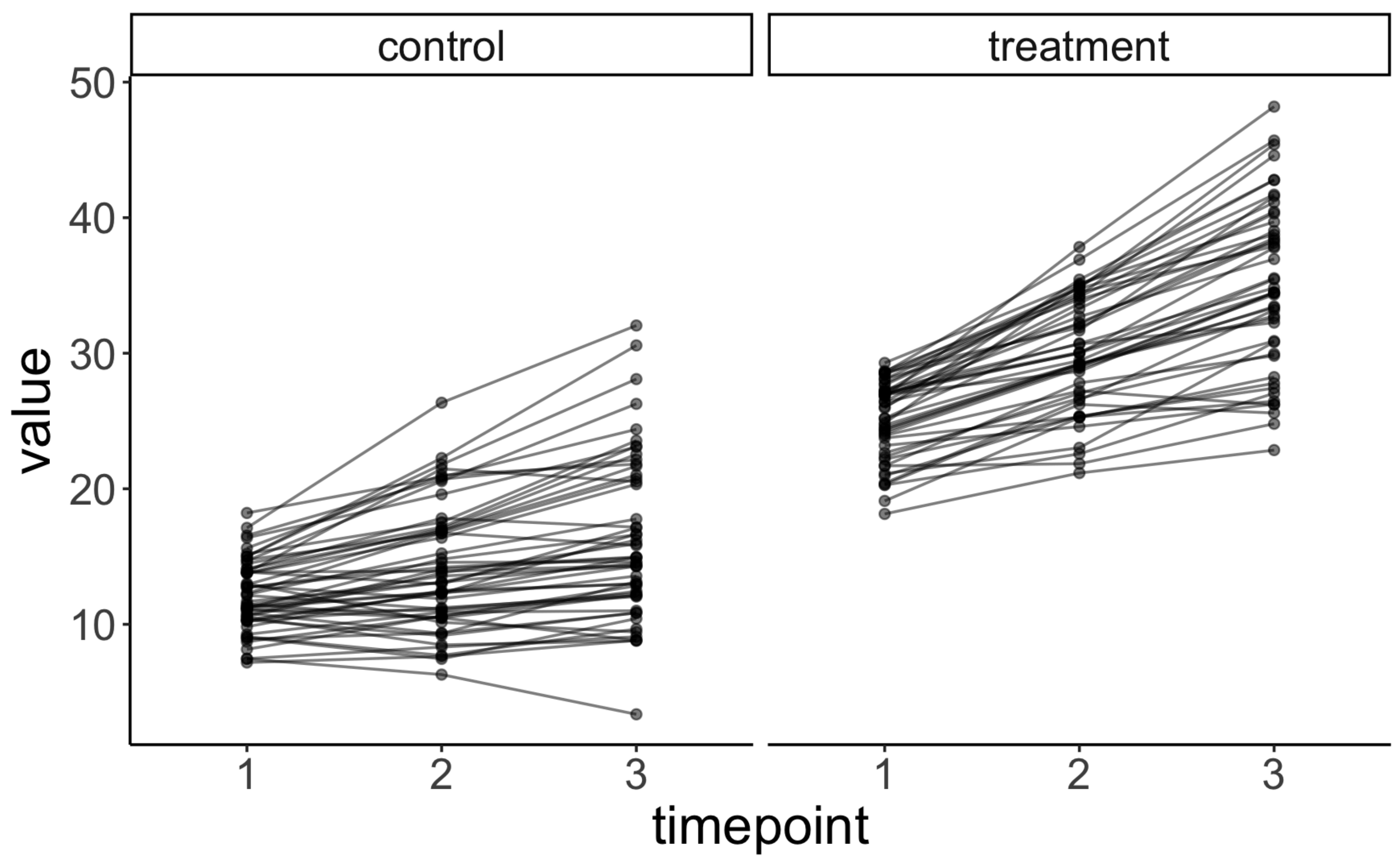
Visualize model predictions (overall)



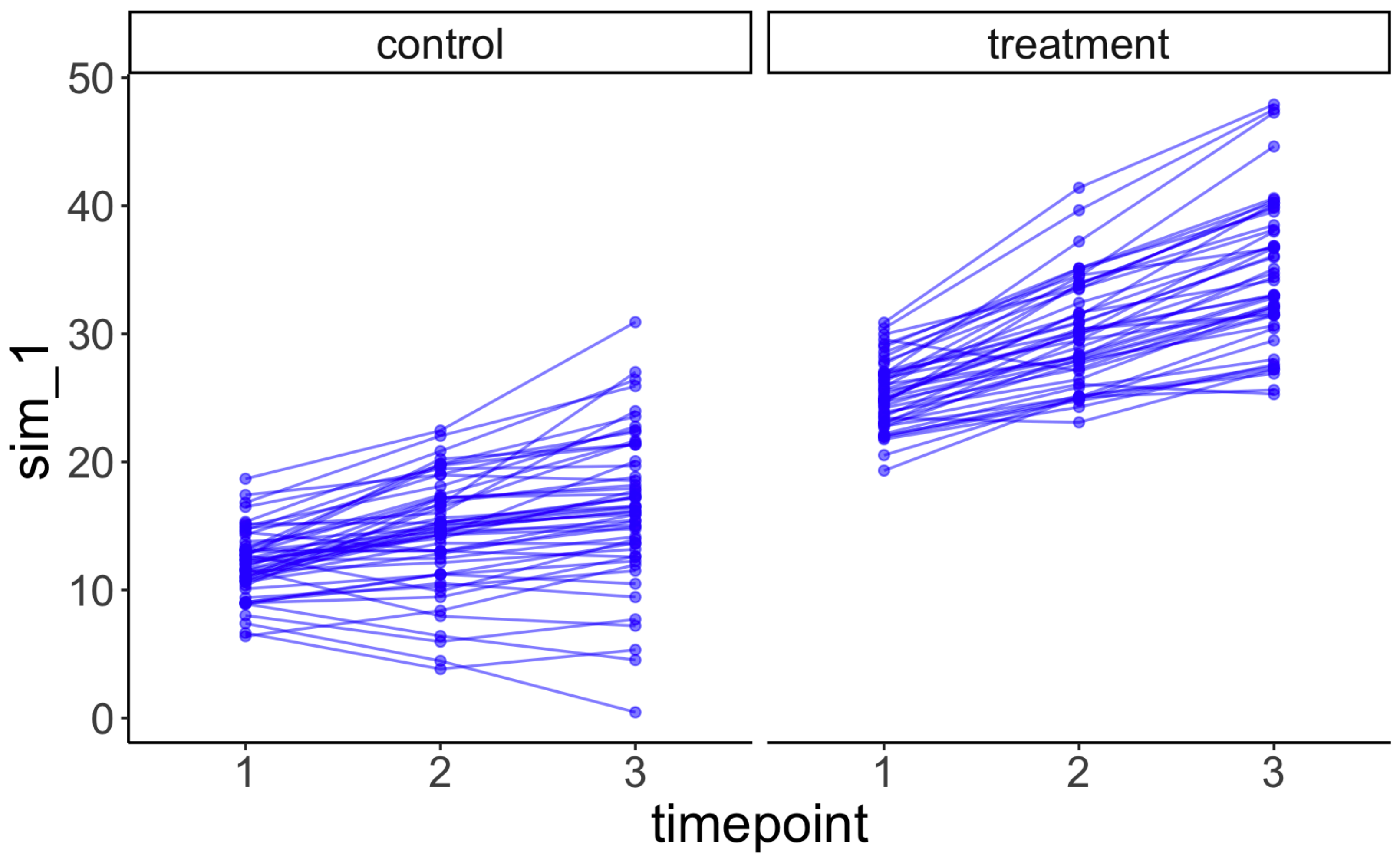
Simulate new data from the model

```r
set.seed(1)

fit %>%
  simulate() %>%
  bind_cols(df.data) %>%
  mutate(condition = factor(condition,
                            levels = c(0, 1),
                            labels = c("control", "treatment")),
         timepoint = as.factor(timepoint)) %>%
  ggplot(data = .,
         mapping = aes(x = timepoint,
                       y = sim_1,
                       group = participant)) +
  geom_point(alpha = 0.5,
             color = "blue") +
  geom_line(alpha = 0.5,
            color = "blue") +
  facet_grid(~ condition) +
  labs(x = "timepoint")
```

original data

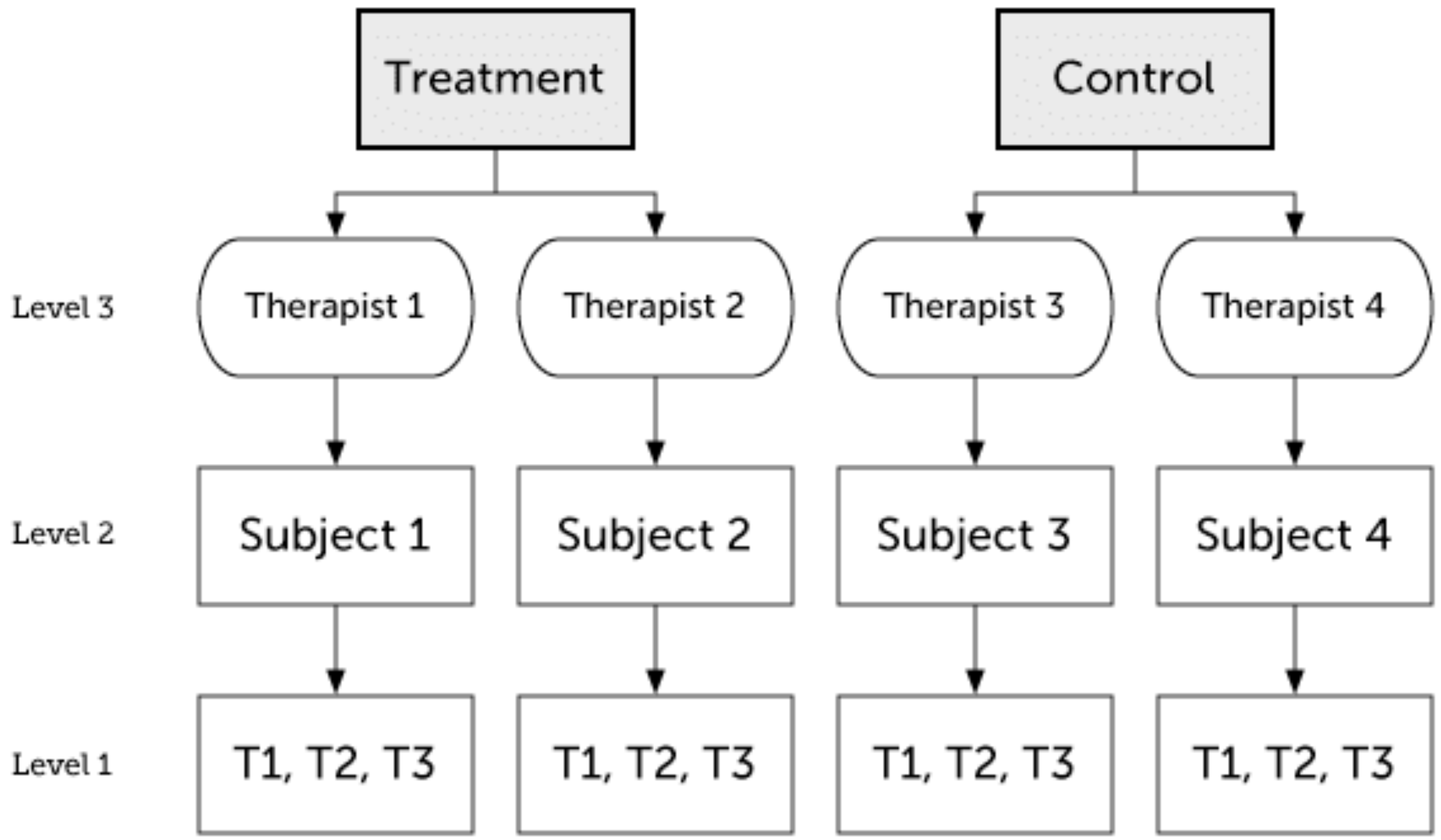


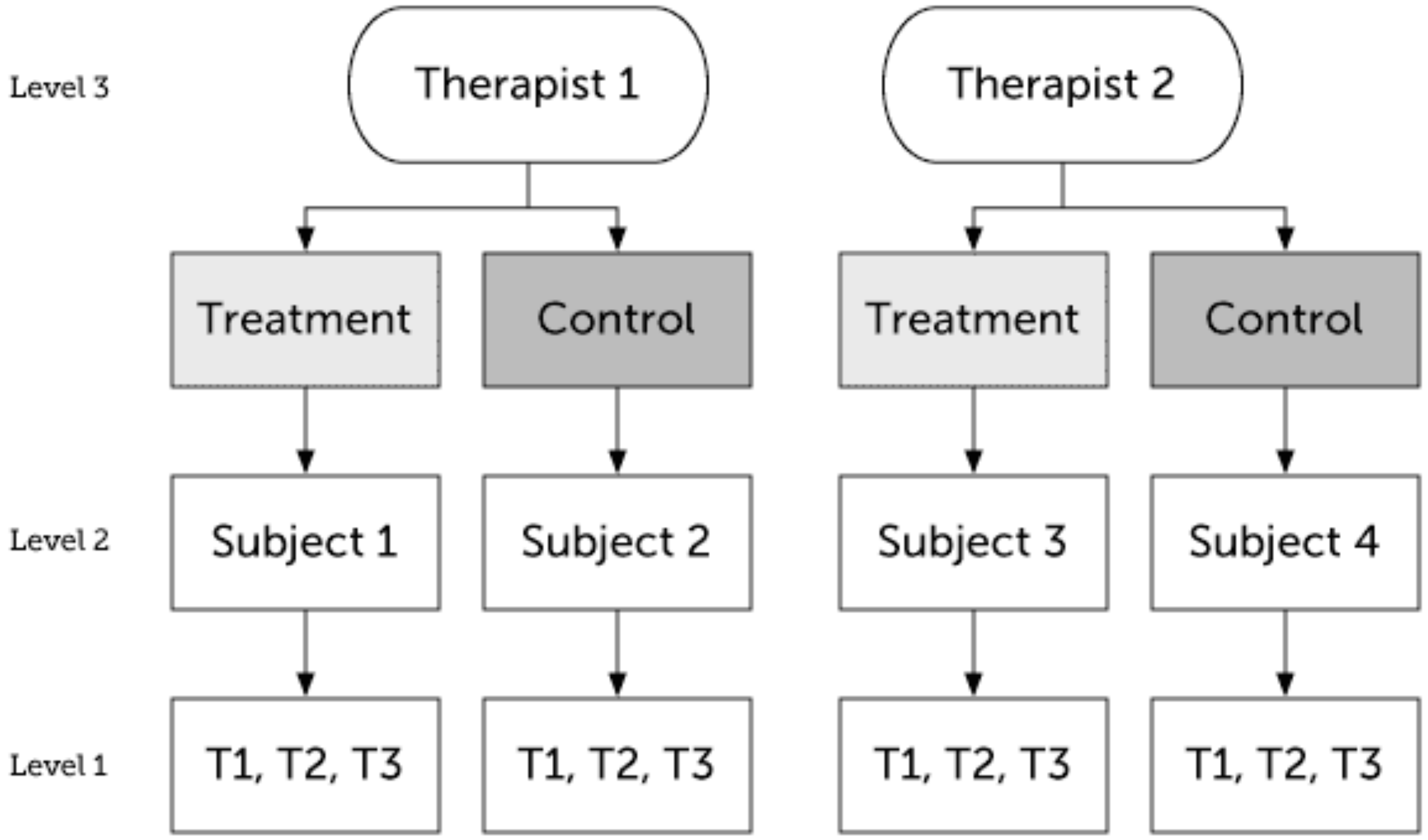
simulated data



Three-level model

Graphical representation





https://rpsychologist.com/r-guide-longitudinal-lme-lmer#power-analysis-and-simulating-these-models

`lmer()` standard operating procedures

Standard Operating Procedures For Using Mixed-Effects Models

A Principled Workflow from the Decision, Development, and Psychopathology (D2P2) Lab
document version 1.0.0 -- 28 June 2020
[This document will be continuously updated and expanded; it may contain typos and other errors--both unintentional errors and errors based on incorrect or outdated knowledge--we will try to improve these things in future versions. Feel free to let us know if you spotted such things, how to further improve this document!]

Authors (in alphabetical order except that the youngsters were so kind to put the oldest guy in the lab first; BF)

Bernd Figner, Johannes Algermissen, Floor Burghoorn, Leslie Held, Afreen Khalid, Felix Klaassen, Farnaz Mosannenzadeh, Julian Quandt

Content/Analysis Steps

http://decision-lab.org/wp-content/uploads/2020/07/SOP_Mixed_Models_D2P2_v1_0_0.pdf

What shall I include as random effects?

- mixed opinions on the topic
- go maximal!

Journal of Memory and Language 68 (2013) 255–278

Contents lists available at SciVerse ScienceDirect

Journal of Memory and Language

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

Random effects structure for confirmatory hypothesis testing: Keep it maximal

Dale J. Barr [a,*], Roger Levy [b], Christoph Scheepers [a], Harry J. Tily [c]

[a] Institute of Neuroscience and Psychology, University of Glasgow, 58 Hillhead St., Glasgow G12 8QB, United Kingdom
[b] Department of Linguistics, University of California at San Diego, La Jolla, CA 92093-0108, USA
[c] Department of Brain and Cognitive Sciences, Massachussetts Institute of Technology, Cambridge, MA 02139, USA

"Through theoretical arguments and Monte Carlo simulation, we show that LMEMs generalize best when they include the maximal random effects structure justified by the design. ...

Maximal LMEMs should be the 'gold standard' for confirmatory hypothesis testing in psycholinguistics and beyond."

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255--278.

What shall I include as random effects?

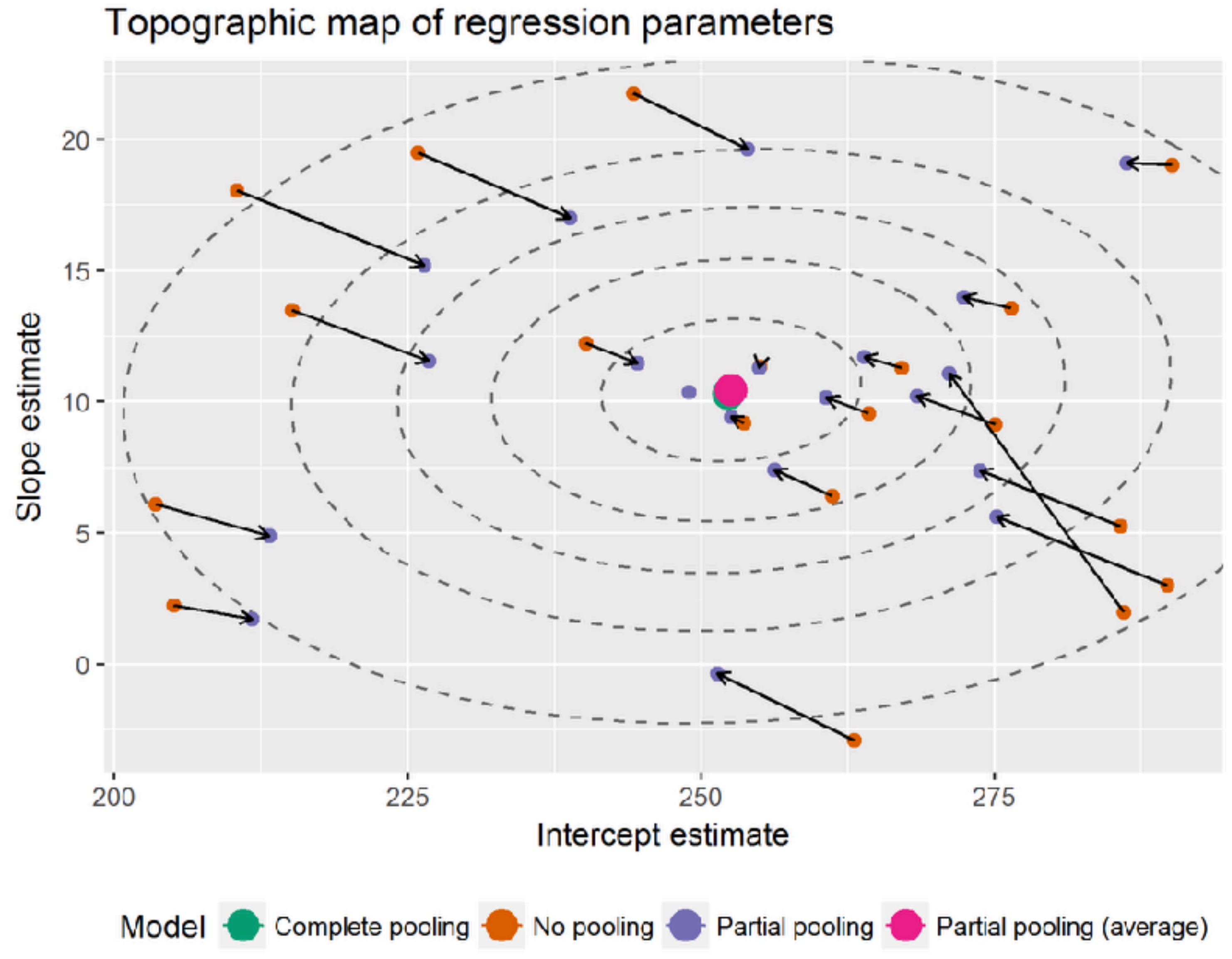
- general advice:
  - start maximal (as supported by the design)
    - random intercepts and slopes for participants
    - random intercepts for items
  - reduce complexity of the random effects structure step by step
    - remove the correlations between random effects first

Remove the correlation component from your model

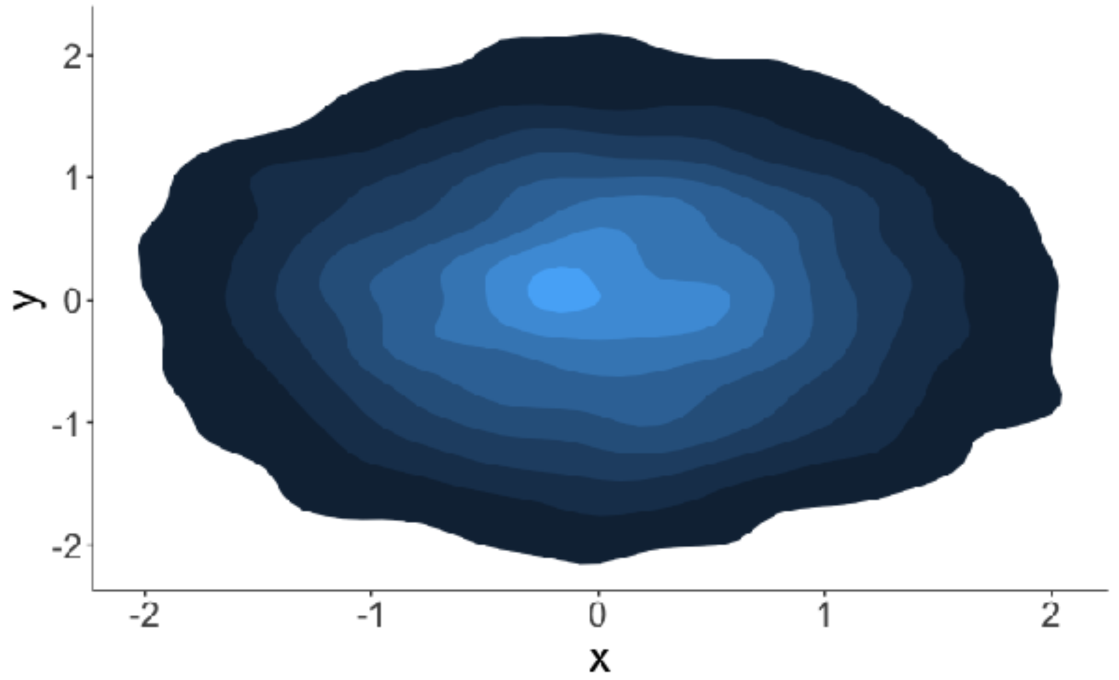
```r
# fit the model
fit.lmer = lmer(formula = reaction ~ 1 + days + (1 + days | subject),
                data = df.sleep)
# model summary
fit.lmer %>%
  summary()
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: reaction ~ 1 + days + (1 + days | subject)
   Data: df.sleep

REML criterion at convergence: 1771.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.9707 -0.4703  0.0276  0.4594  5.2009

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 subject  (Intercept) 582.73   24.140
          days         35.03    5.919   0.07
 Residual             649.36   25.483
Number of obs: 183, groups:  subject, 20

Fixed effects:
            Estimate Std. Error t value
(Intercept)  252.543      6.433  39.256
days          10.452      1.542   6.778

Correlation of Fixed Effects:
     (Intr)
days -0.137
```

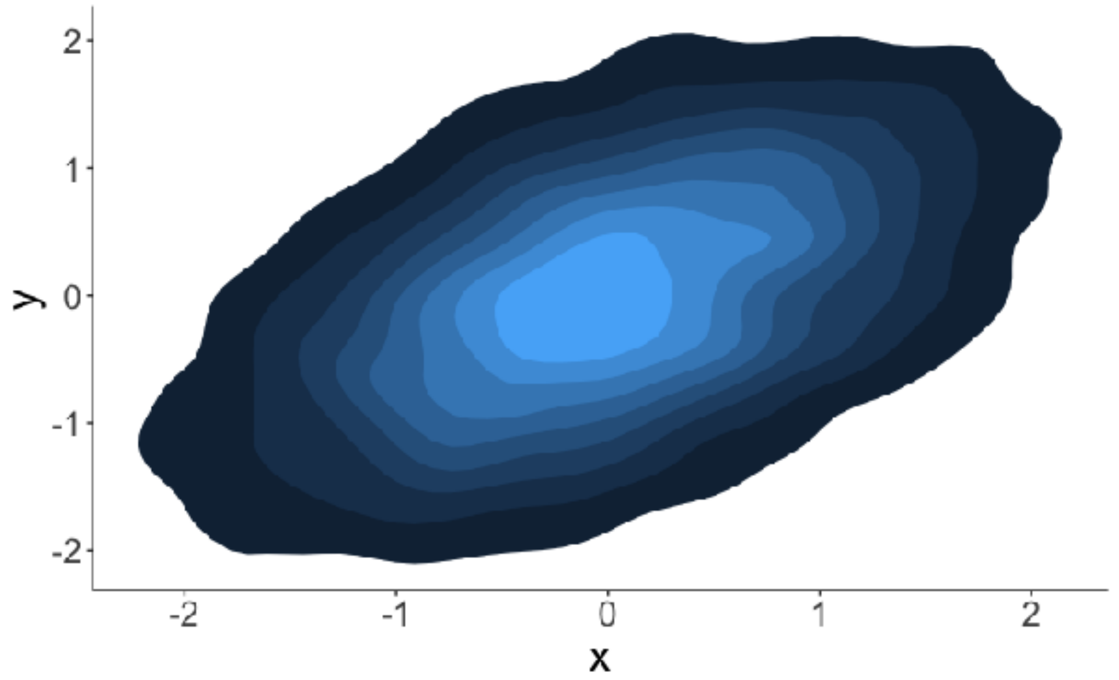
multivariate Gaussian



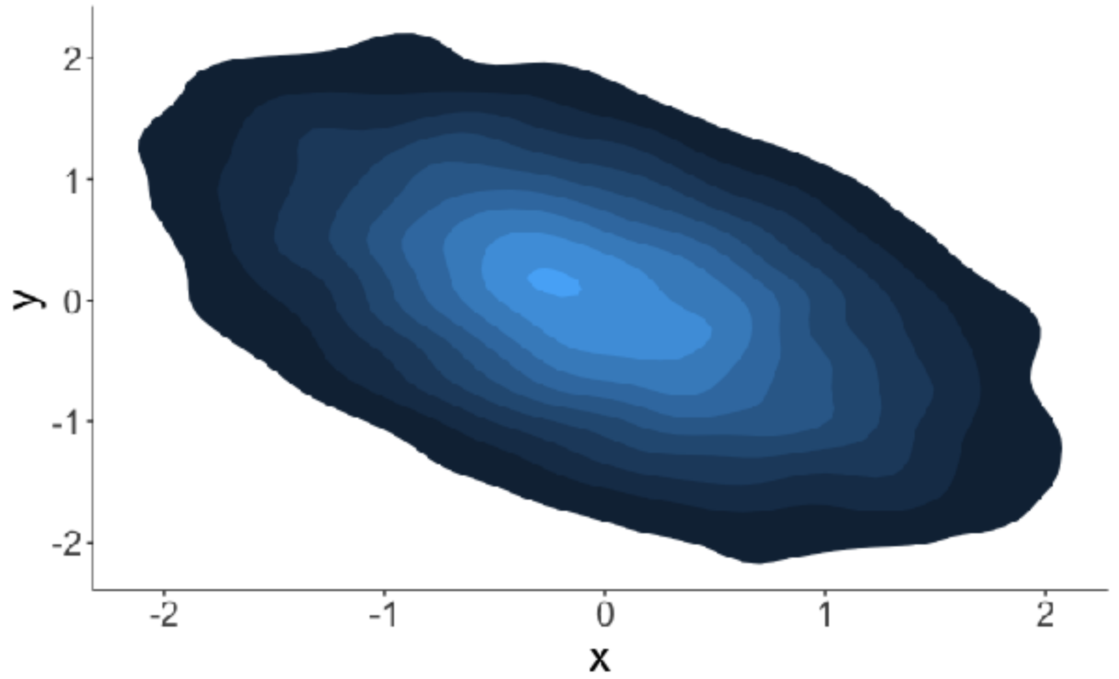
Remove the correlation component from your model



uncorrelated



positively correlated



negatively correlated

Remove the correlation component from your model

```r
# fit the model
fit.lmer = lmer(formula = reaction ~ 1 + days + (0 + days | subject) + (1 | subject),
                data = df.sleep)
# model summary
fit.lmer %>%
  summary()
```

random slopes

random intercepts

```
Linear mixed model fit by REML ['lmerMod']
Formula: reaction ~ 1 + days + (0 + days | subject) + (1 | subject)
   Data: df.sleep

REML criterion at convergence: 1771.5

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.9805 -0.4673  0.0250  0.4589  5.2083

Random effects:
 Groups    Name        Variance Std.Dev.
 subject   days         35.88    5.99
 subject.1 (Intercept) 598.11   24.46
 Residual              647.90   25.45
Number of obs: 183, groups:  subject, 20

Fixed effects:
            Estimate Std. Error t value
(Intercept)  252.550      6.491  38.907
days          10.439      1.556   6.708

Correlation of Fixed Effects:
     (Intr)
days -0.184
```

independent Gaussians

Remove the correlation component from your model

```r
# fit the model
fit.lmer = lmer(formula = reaction ~ 1 + days + (1 + days || subject),
                data = df.sleep)
# model summary
fit.lmer %>%
  summary()
```

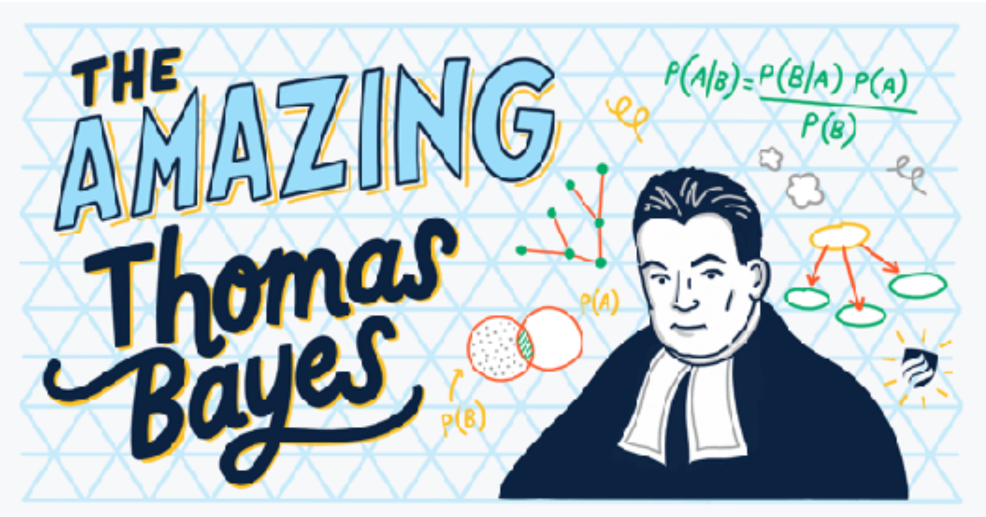
alternative syntax (doesn't model correlation between random effects)

```
Linear mixed model fit by REML ['lmerMod']
Formula: reaction ~ 1 + days + (0 + days | subject) + (1 | subject)
   Data: df.sleep

REML criterion at convergence: 1771.5

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.9805 -0.4673  0.0250  0.4589  5.2083

Random effects:
 Groups    Name        Variance Std.Dev.
 subject   days         35.88    5.99
 subject.1 (Intercept) 598.11   24.46
 Residual              647.90   25.45
Number of obs: 183, groups:  subject, 20

Fixed effects:
            Estimate Std. Error t value
(Intercept)  252.550      6.491  38.907
days          10.439      1.556   6.708

Correlation of Fixed Effects:
     (Intr)
days -0.184
```

independent Gaussians

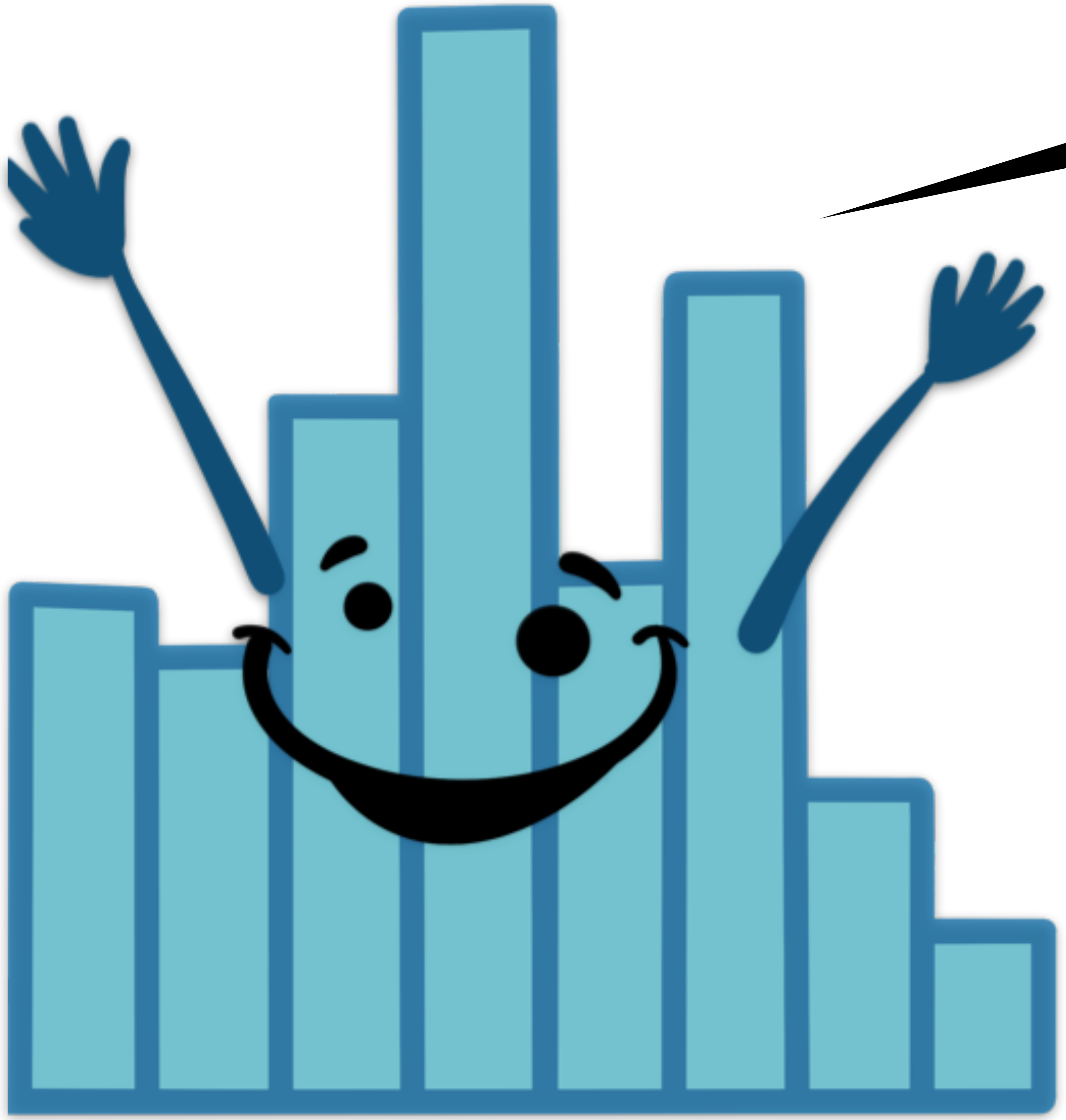
What if `lmer()` fails to converge?

1. We <u>drop random effects</u> in the following order: random correlations, random slopes of covariates (where significance is of no interest), random intercepts ("0+" instead "1+") (following [Barr et al., 2013]()). We never remove the random slopes of the variables of interest (i.e., the ones for which we want to conduct significance tests).
   Please note that removing random correlation terms can be tricky if random slopes are estimated for factors with 3 or more levels. In that case, it is probably easiest to use `afex::mixed()` with `expand_re = TRUE` (an alternative option is to create manually the relevant contrasts yourself and add them as predictors to your model, which allows you to suppress the random corrections using the double pipe symbol ||).
2. We try to run <u>separate analyses</u>: For example, one model to only test the fixed and random effect of A (with fixed effect of B present); then one model to only test the effect of B.
   If we really have to drop random slopes, we follow the next step:
3. We follow the <u>PCA approach</u> suggested by **rePsychLing** (see [Bates et al., 2015]()) that is performing a PCA on the random effects and following the guidelines described in the paper.
   a. We use a likelihood ratio test to test whether the model fit becomes significantly worse. As we prefer a more conservative approach here (i.e., rather err on the side of keeping too many random effects; we prioritize avoiding inflated Type 2 errors for this kind of decision), we use larger alpha-level of .2 ([Matuschek et al., 2017]()).
   b. Alternatively, we suggest an <u>Information criterion approach</u> to avoid using a *p* value for our inclusion/exclusion decision, but choose the best model based on *BIC* or *AIC*.



3.2.2. Or we choose a Bayesian approach

As an alternative to targeting convergence issues within **lme4**, we suggest fitting the same model with **brms** and <u>comparing</u> it to the **lme4** fit. We assume that both provide similar results when

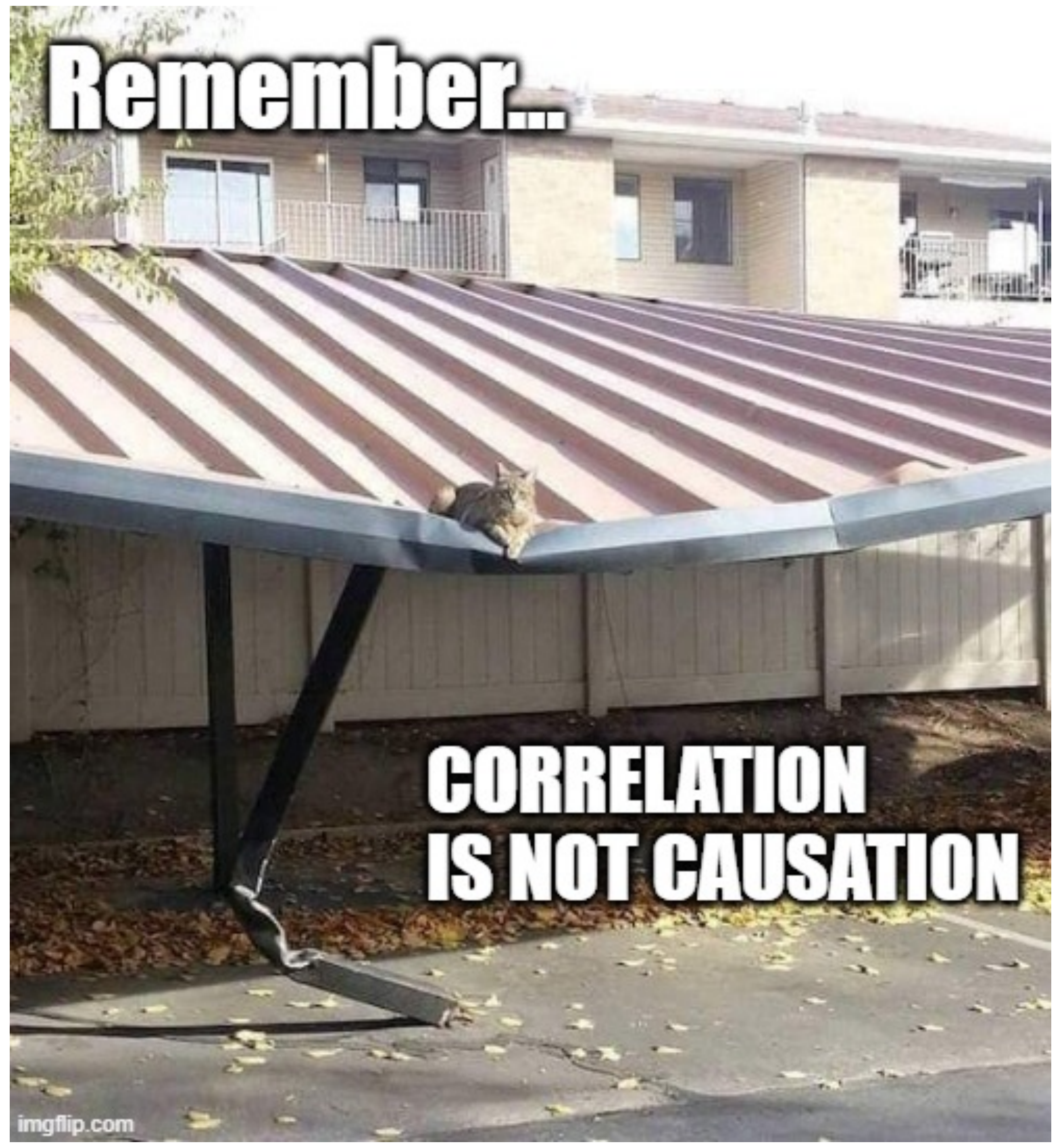


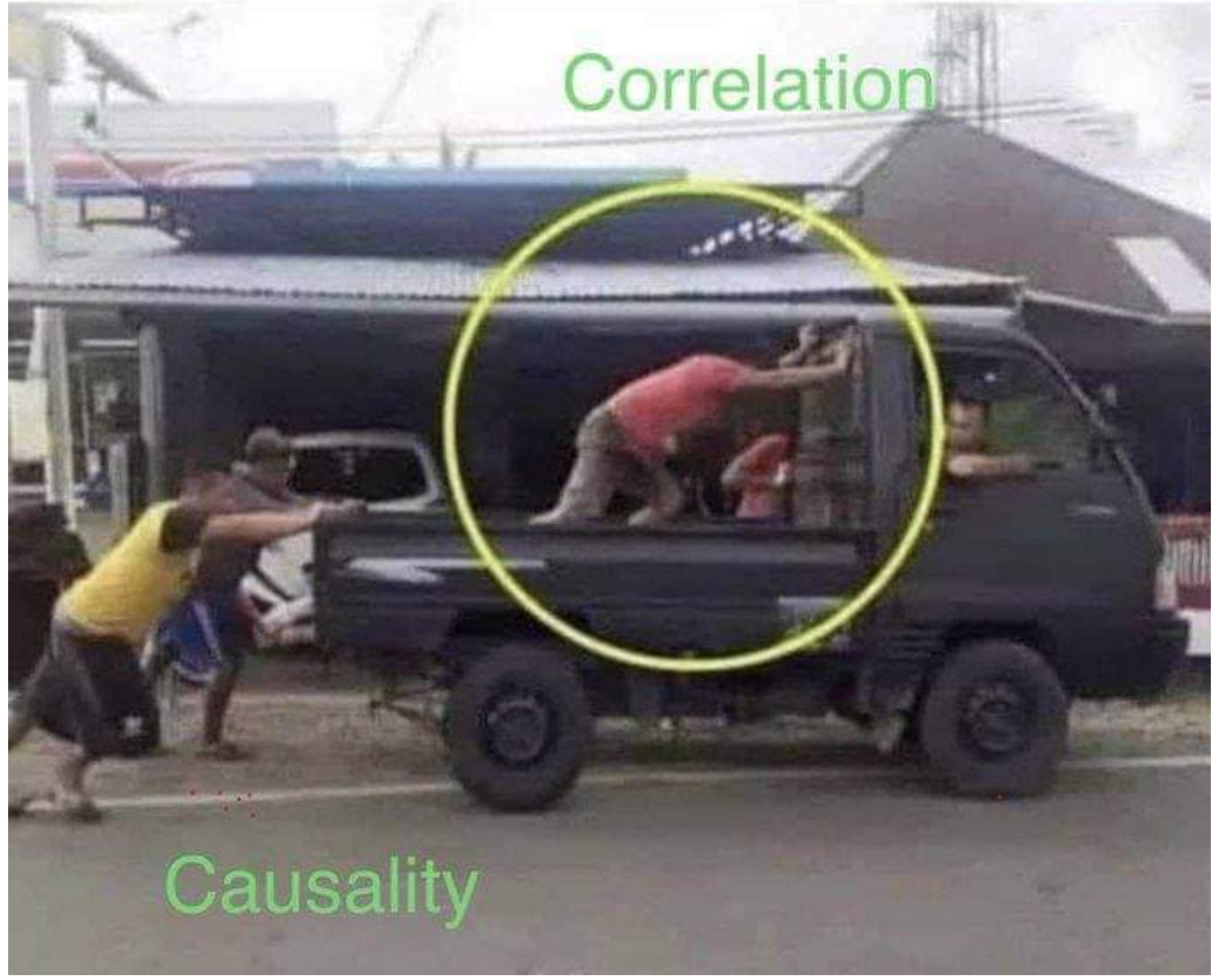
02:00

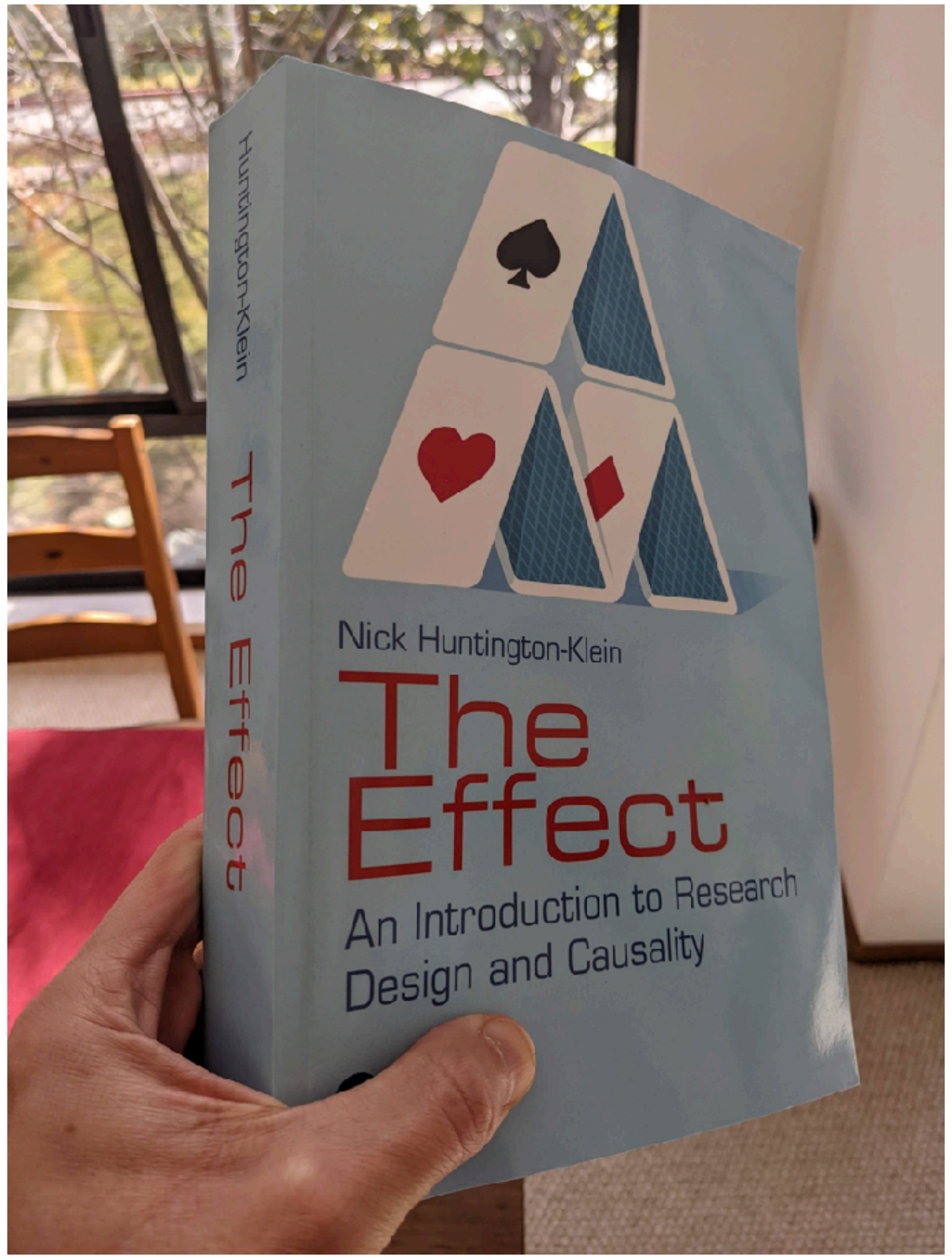
stretch break!

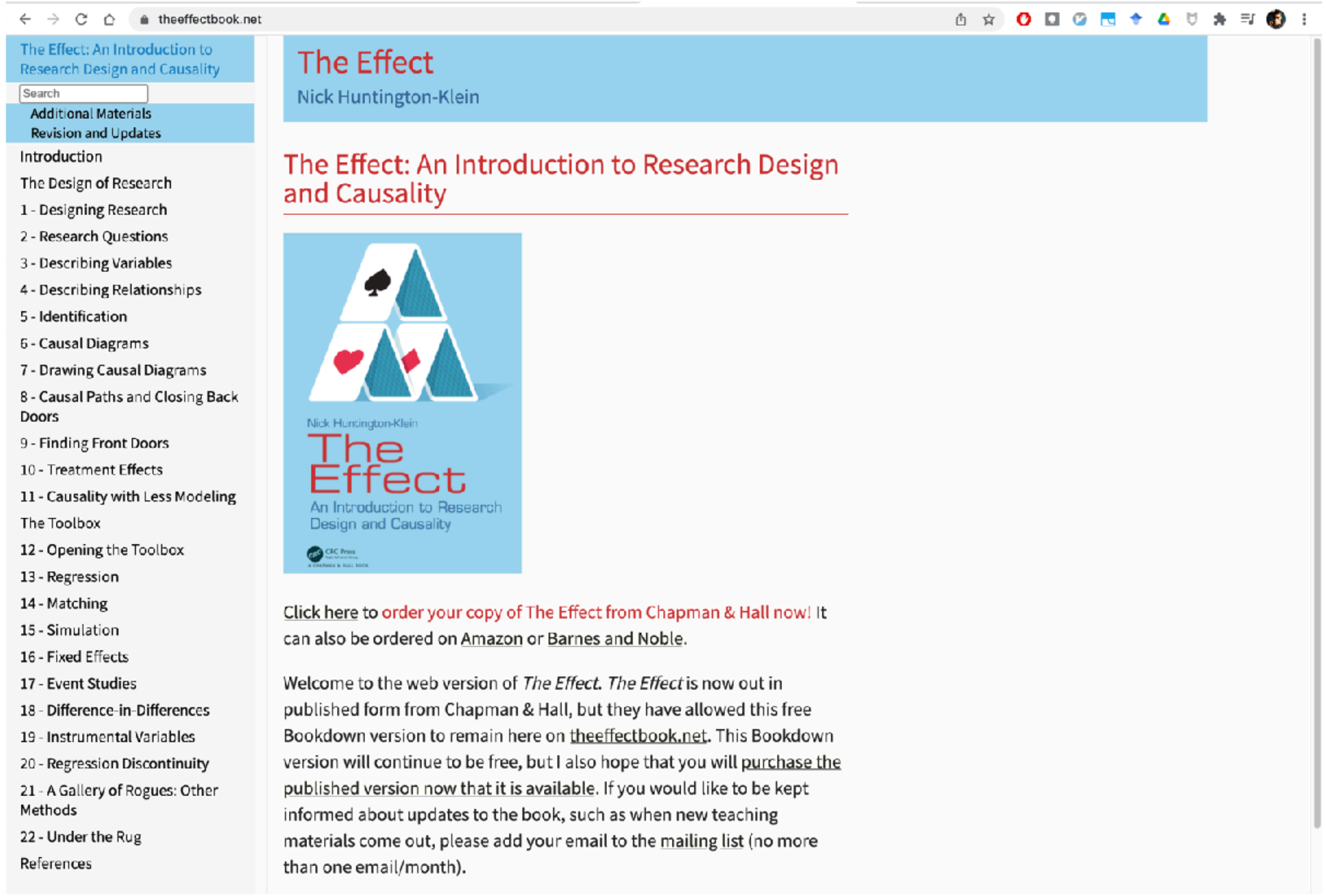
Causation vs. correlation

correlation does not imply causation



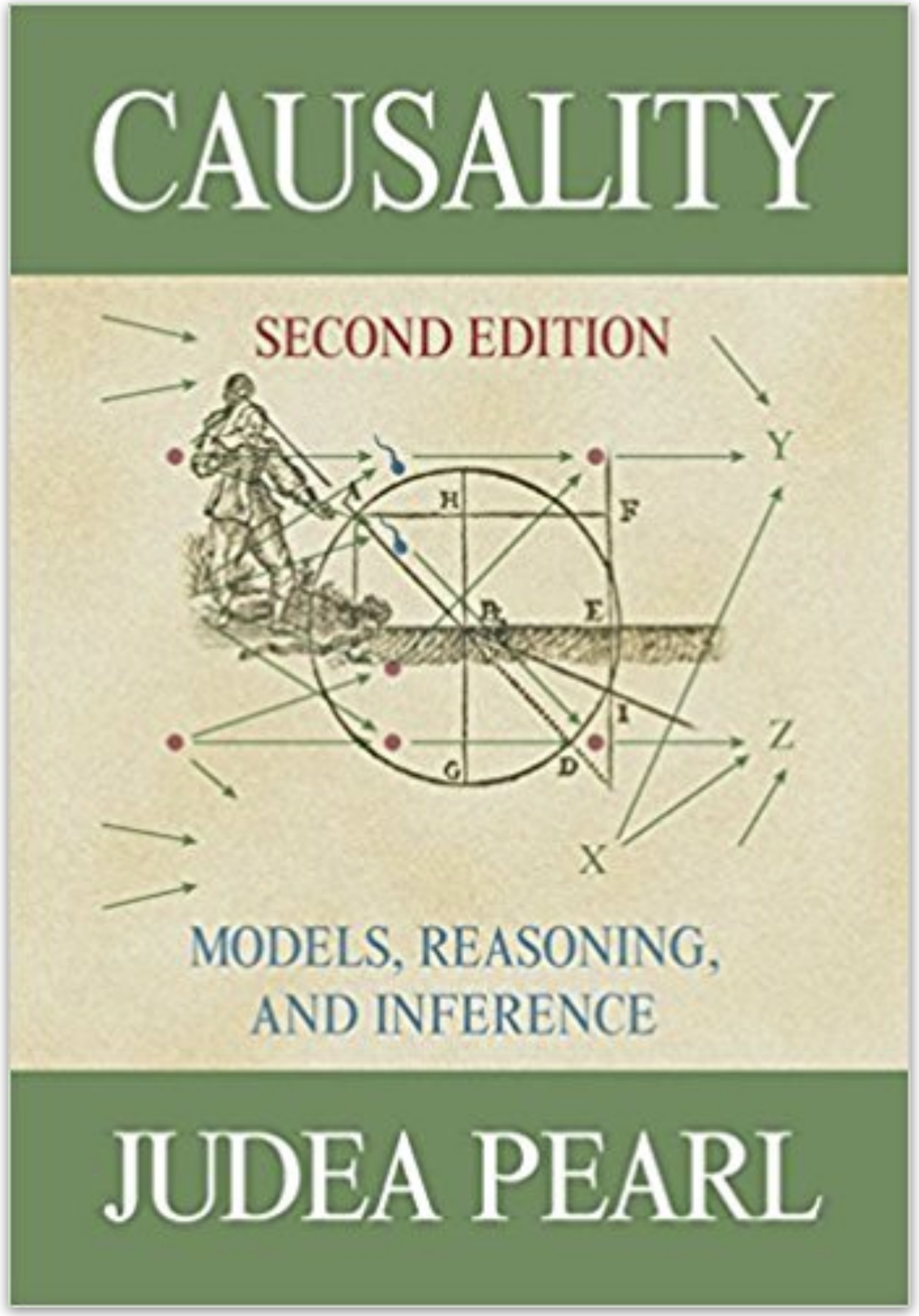


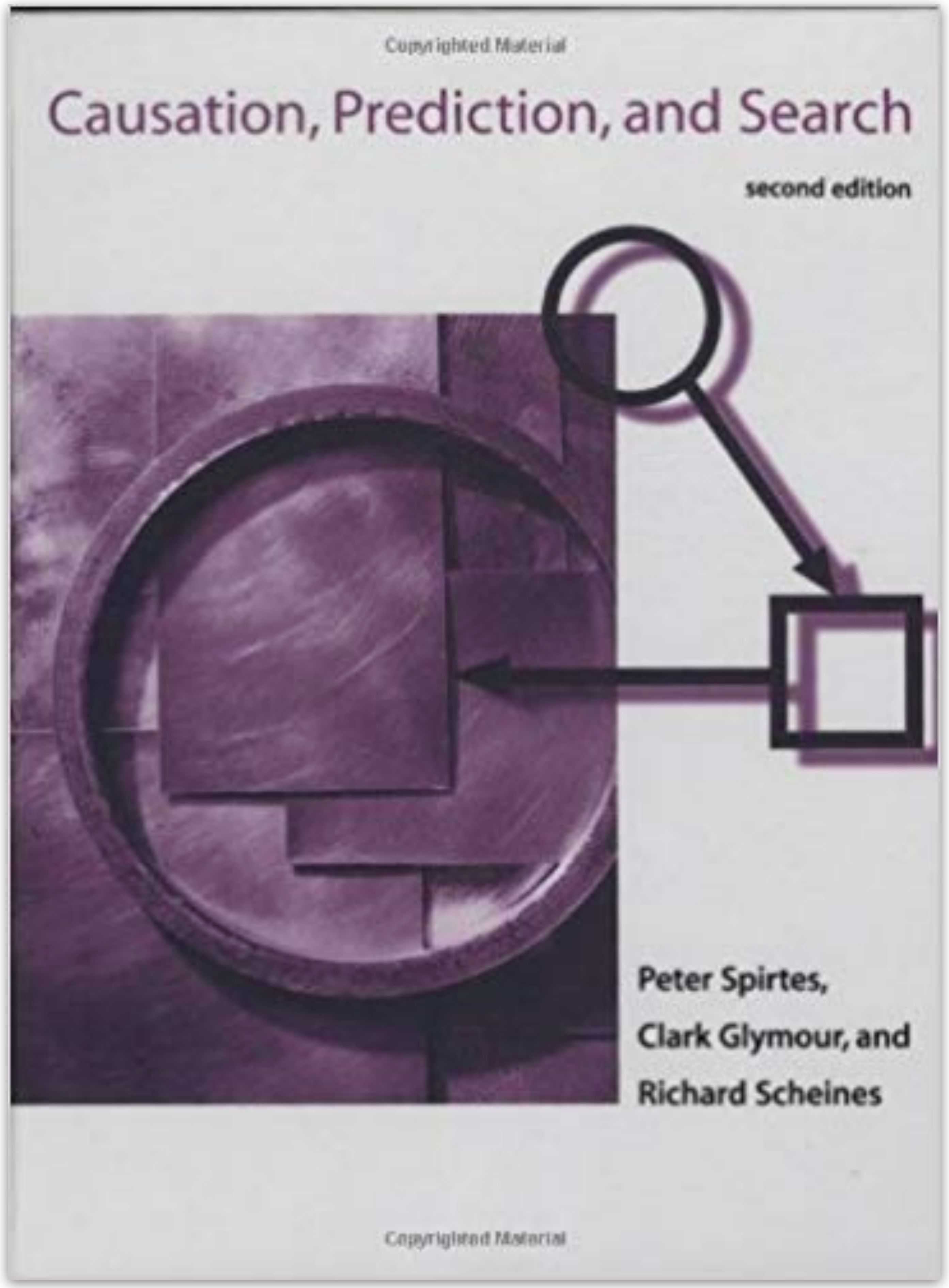




made with RMarkdown

https://theeffectbook.net/





Pearl, J. (2000). Causality: Models, reasoning and inference. Cambridge, England: Cambridge University Press.

Spirtes, P., Glymour, C. N., & Scheines, R. (2000). Causation, prediction, and search. The MIT Press.

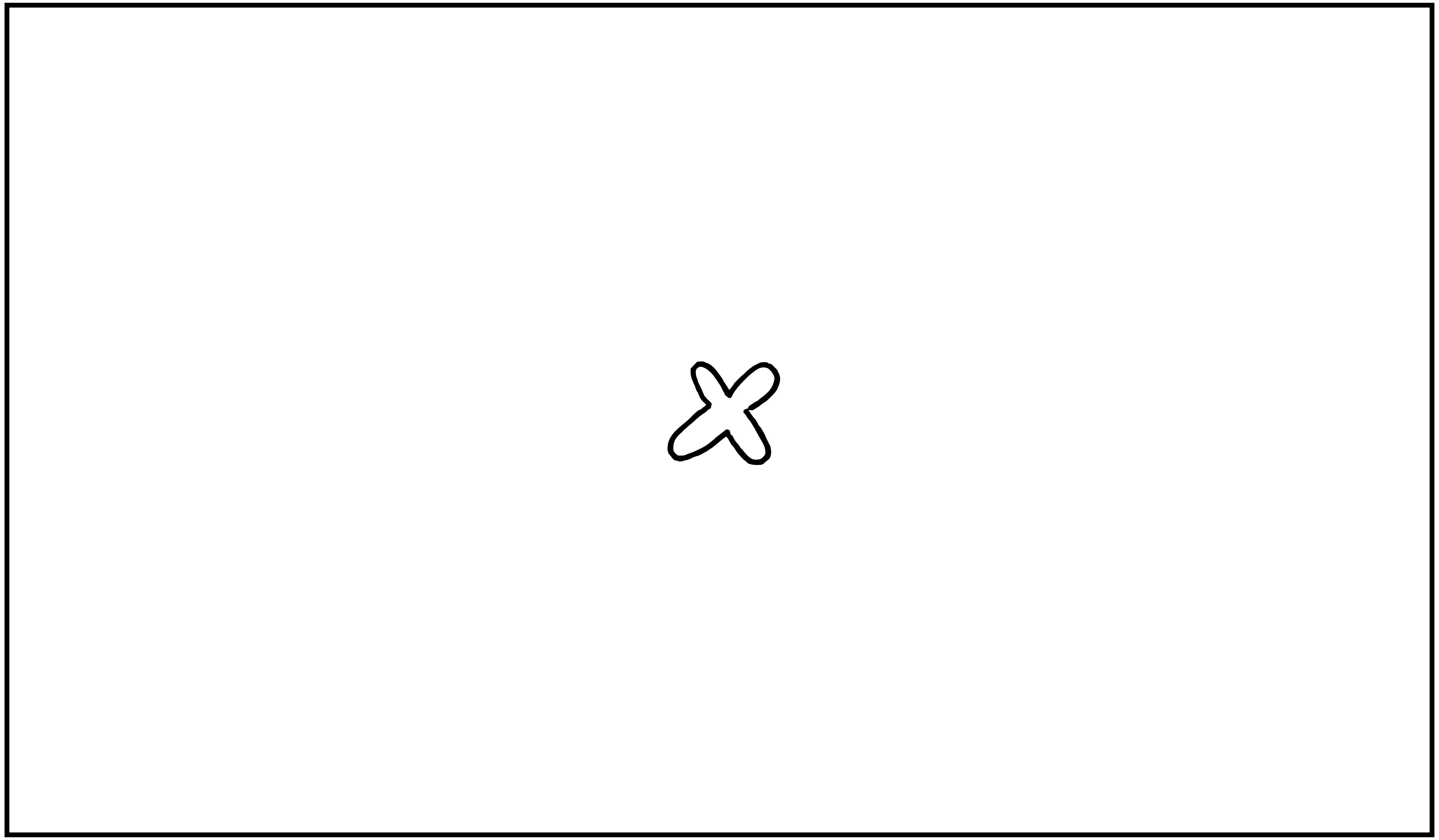
What do you think?

Suppose there is a robust, statistically significant, and long-term correlation between the color of cars and the annual rate at which they are involved in accidents.

To be concrete, assume that red cars, in particular, are involved in accidents year after year at a higher rate than cars of any other color. When you go to buy a new car, should you avoid the color red in your quest to remain safe on the road?

https://bostonreview.net/articles/tim-maudlin-you-are-smarter-your-data/

Possible causal mechanisms



running experiments

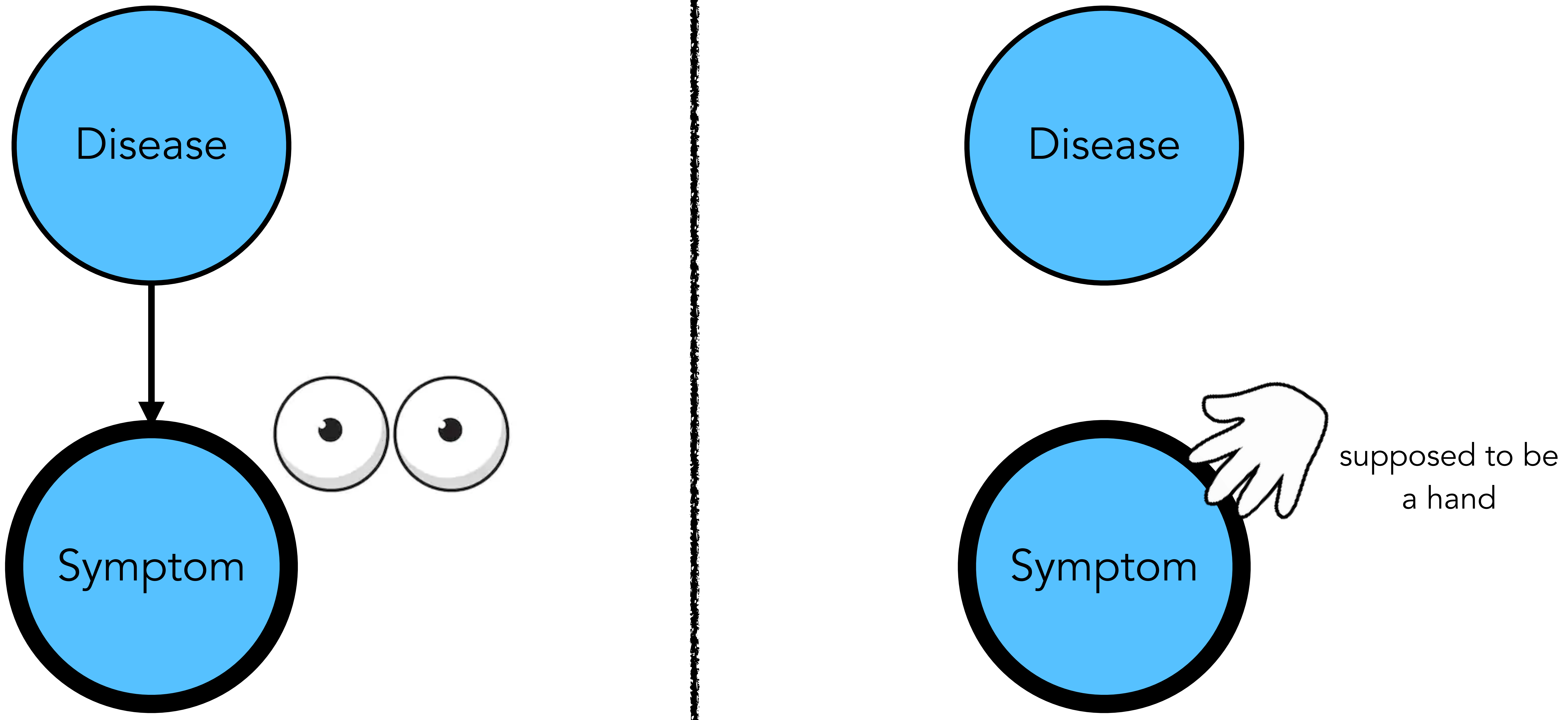
controlling for variables

can help/hinder

Seeing vs. doing

Observation vs. Intervention

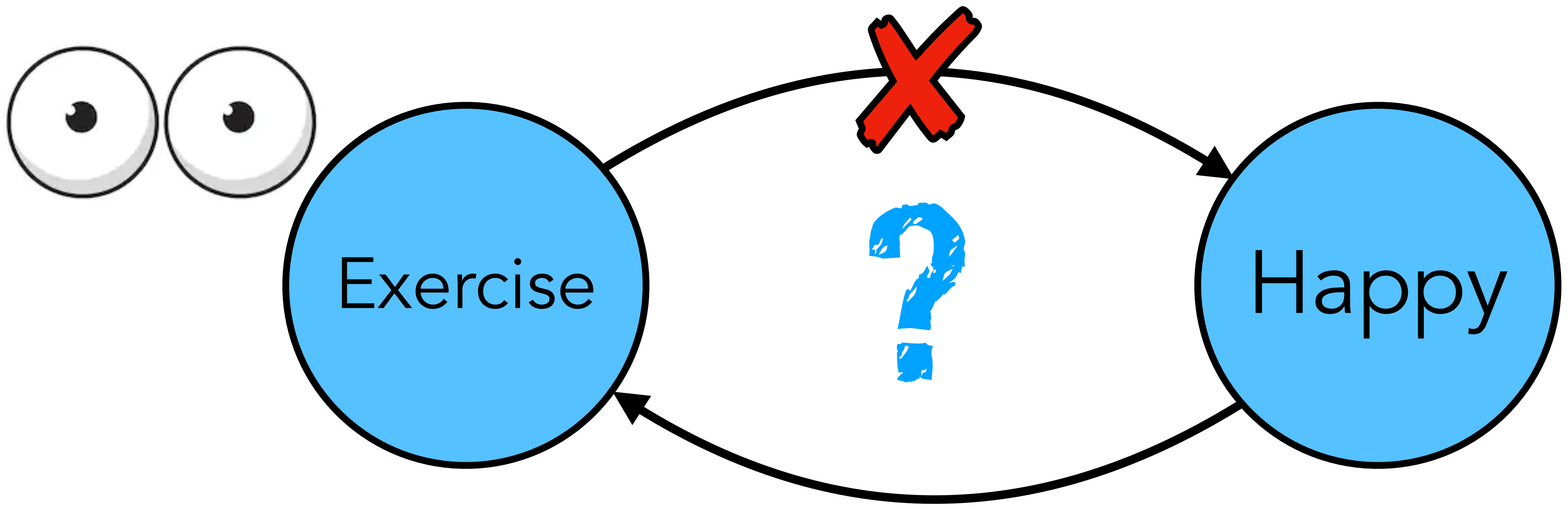
seeing | **doing**



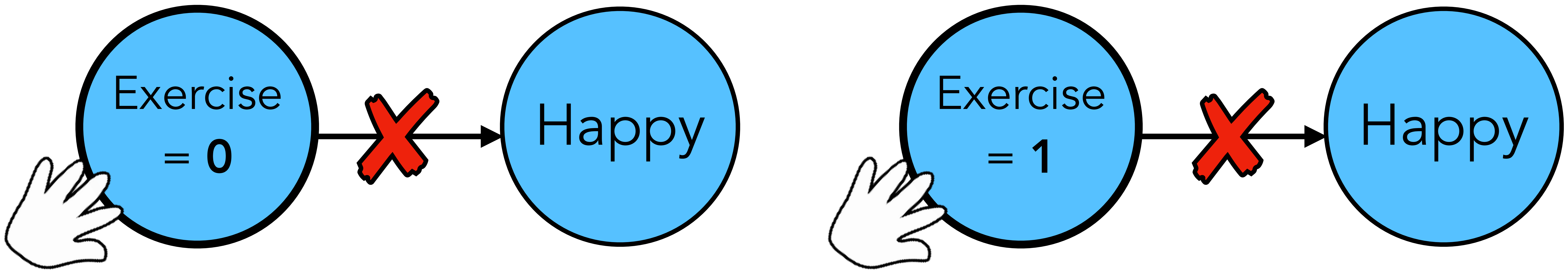
$p(D \mid S) > p(D)$

$p(D \mid \mathrm{do}(S)) = p(D)$

Inferring causal structure through intervention



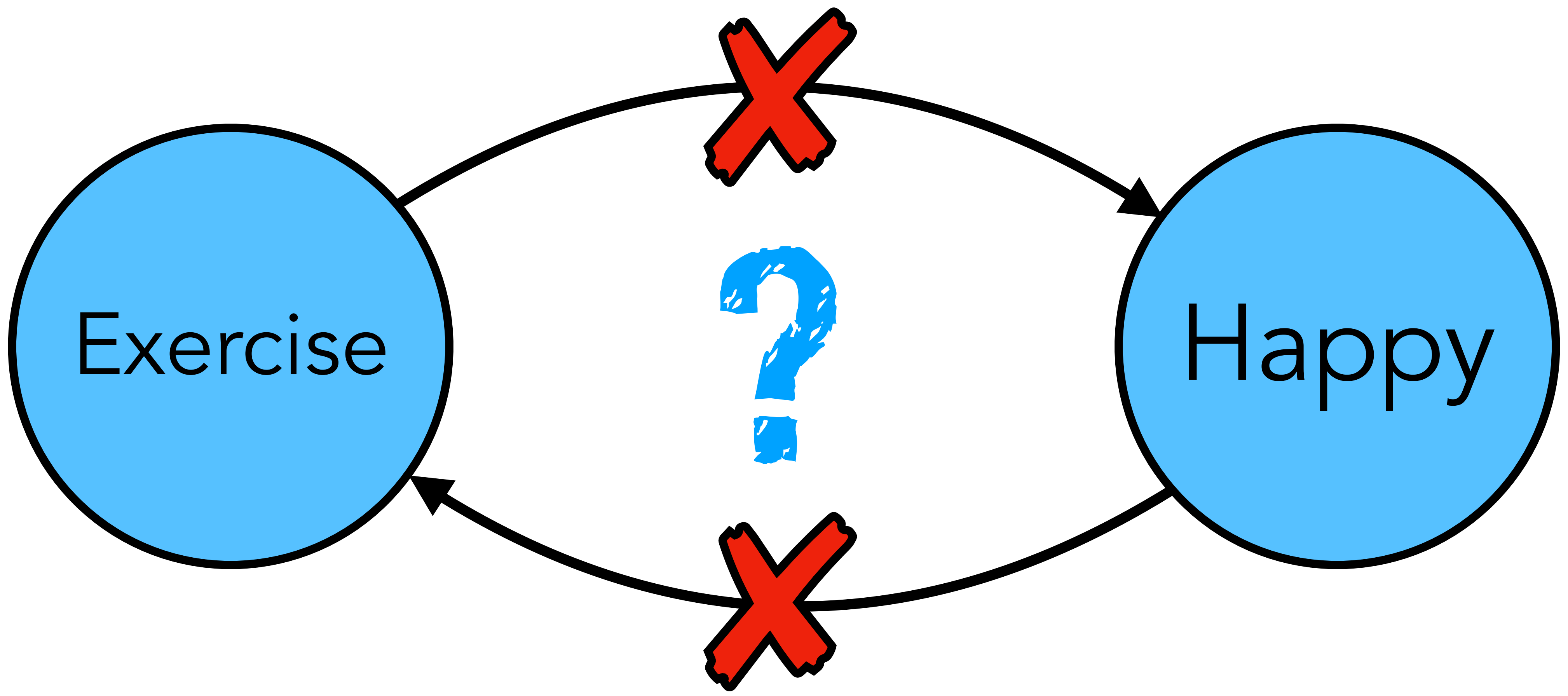
Experiment 1



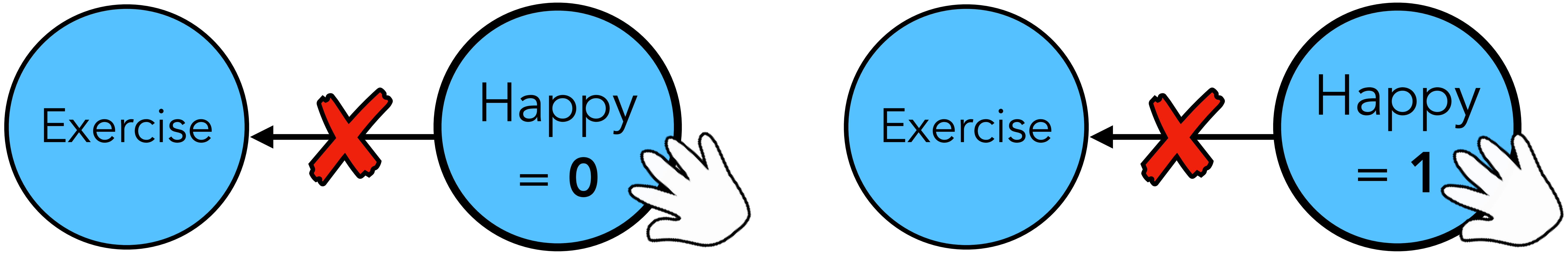
$p(\text{Happy} \mid \text{do}(\text{Exercise} = 0)) = 0.3$

$p(\text{Happy} \mid \text{do}(\text{Exercise} = 1)) = 0.3$

Inferring causal structure through intervention

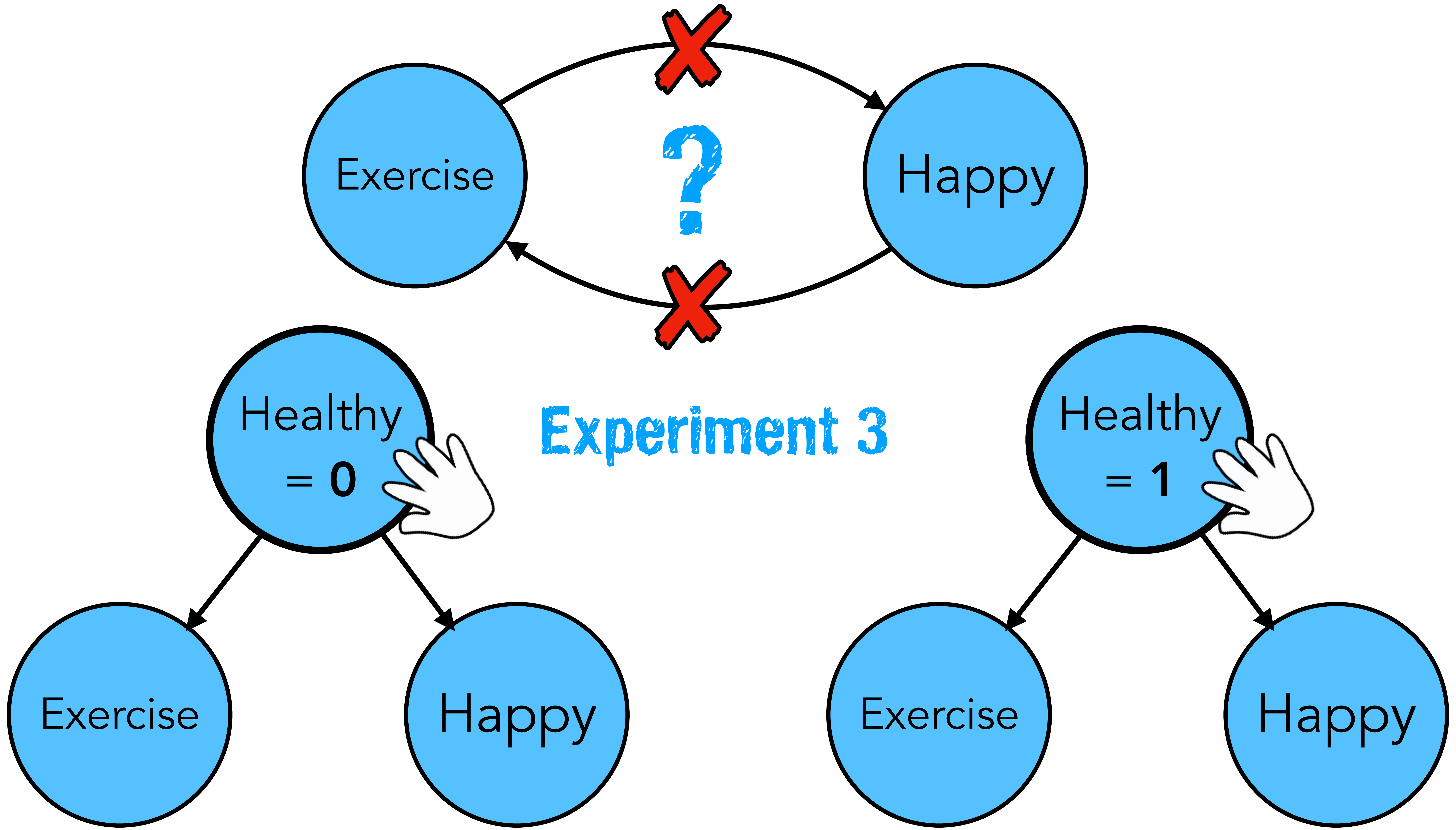


Experiment 2



$p(\text{Exercise} \mid \text{do}(\text{Happy} = 0)) = 0.1$ $p(\text{Exercise} \mid \text{do}(\text{Happy} = 1)) = 0.1$

Inferring causal structure through intervention



$p(\text{Happy} \mid \text{do}(\text{Healthy} = 0)) = 0.05$

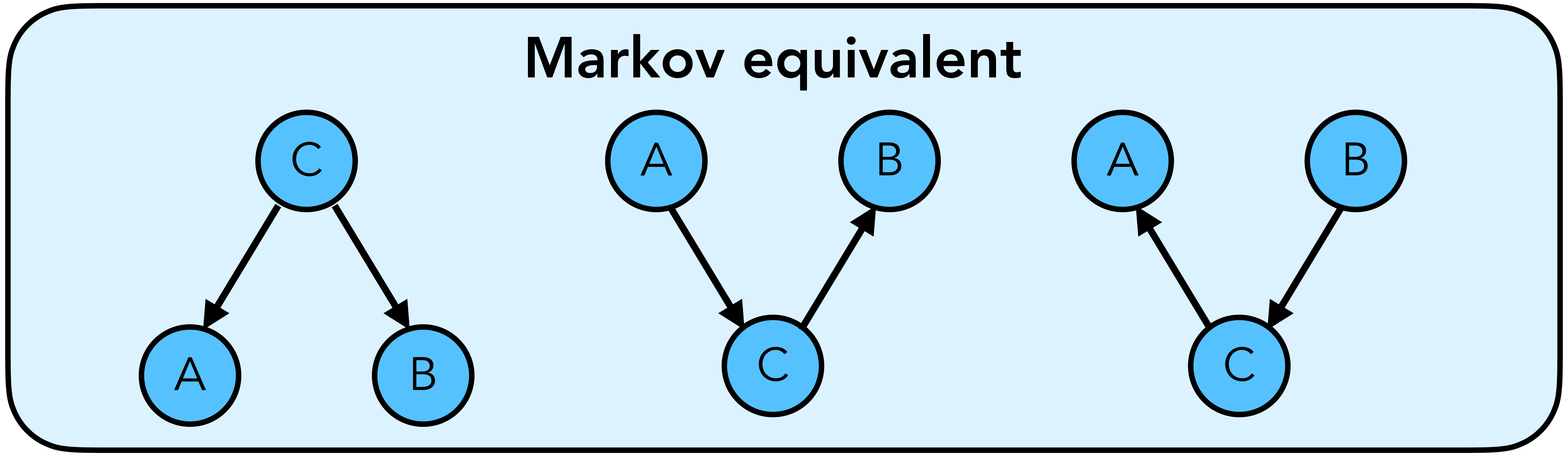
$p(\text{Exercise} \mid \text{do}(\text{Healthy} = 0)) = 0.1$

$p(\text{Happy} \mid \text{do}(\text{Healthy} = 1)) = 0.5$

$p(\text{Exercise} \mid \text{do}(\text{Healthy} = 1)) = 0.75$

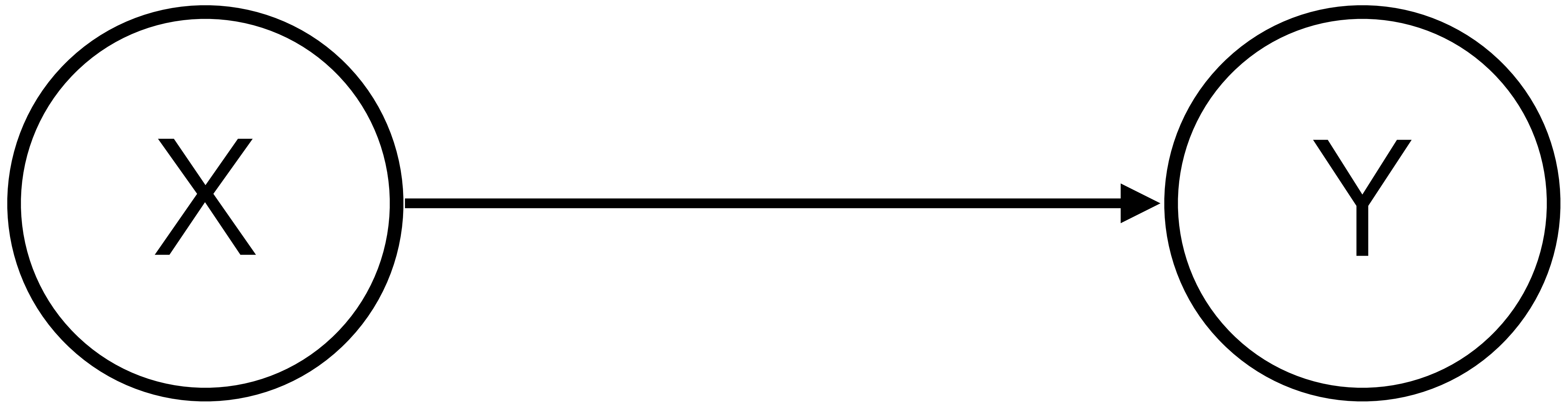
Important take home message

- correlation is not causation
- correlation (= probabilistic dependence) suggests that there is *some* causal relationship
- but we don't know which one it is

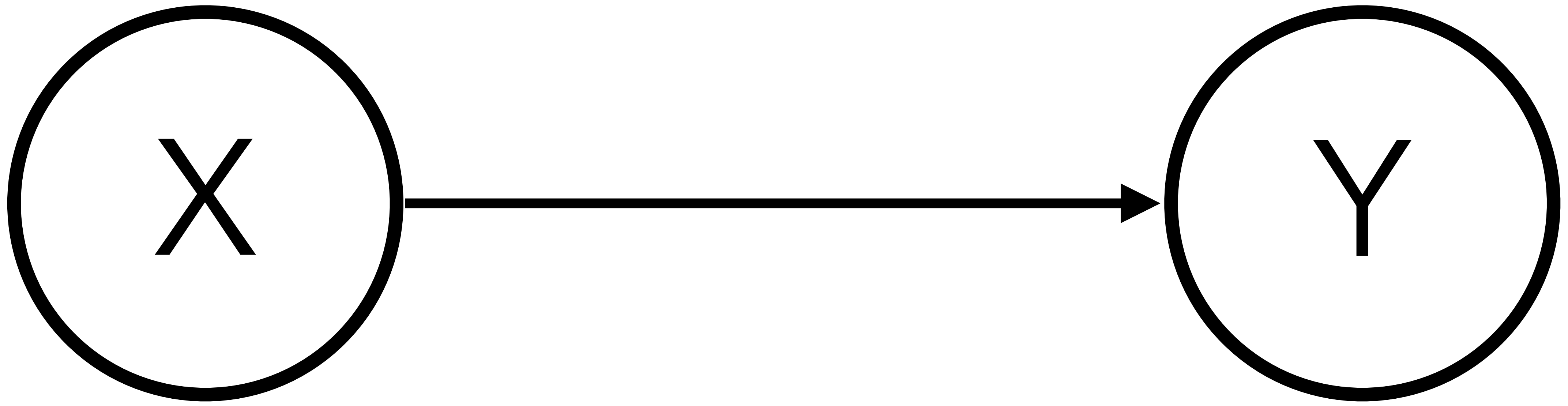


- **causal interventions** / experiments can reveal the underlying causal structure

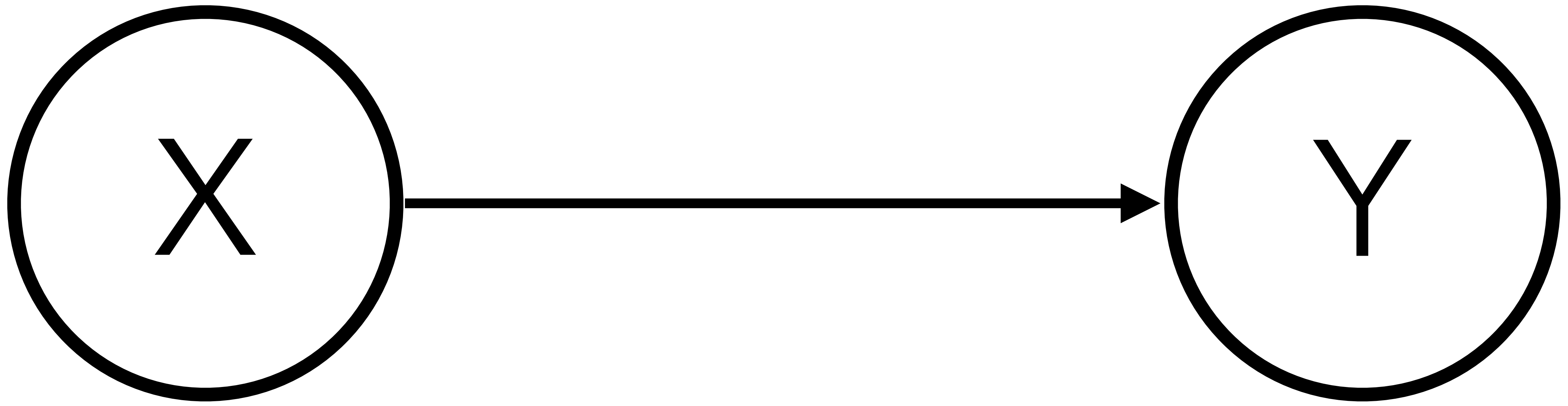
Observation, intervention, counterfactual



Level	Concept	Expression	Activity	Question	Example
I	Observation/ Prediction	$p(y\|x)$	Seeing	How does x change my belief in y?	Would the grass be wet if we *found* the sprinkler off?

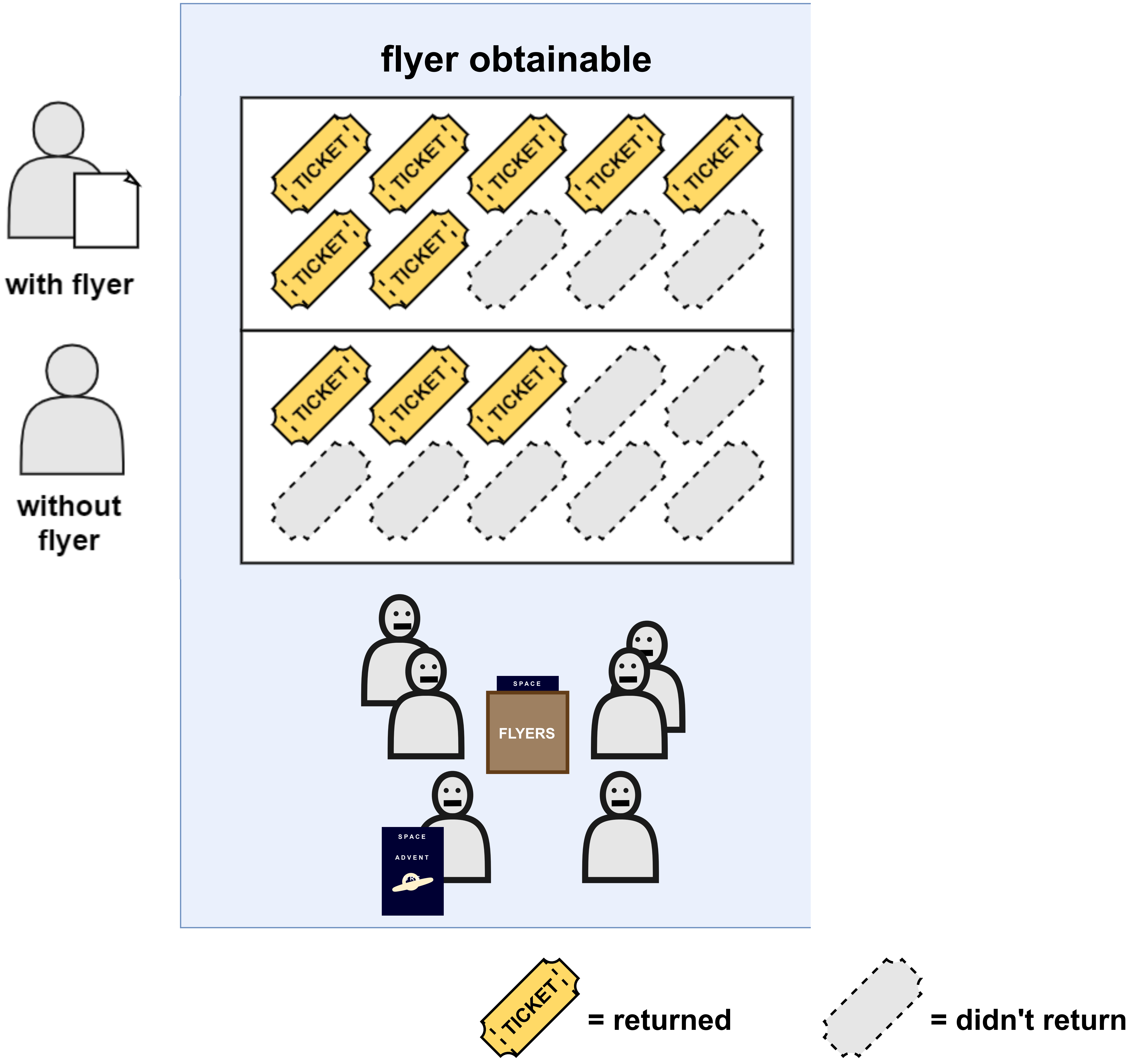


Level	Concept	Expression	Activity	Question	Example
I	Observation/ Prediction	$p(y\|x)$	Seeing	How does x change my belief in y?	Would the grass be wet if we *found* the sprinkler off?
II	Intervention/ Hypothetical	$p(y\|\mathrm{do}(x))$	Doing	Would y happen if I did x?	Would the the grass be wet if *made sure* that the sprinkler was off?

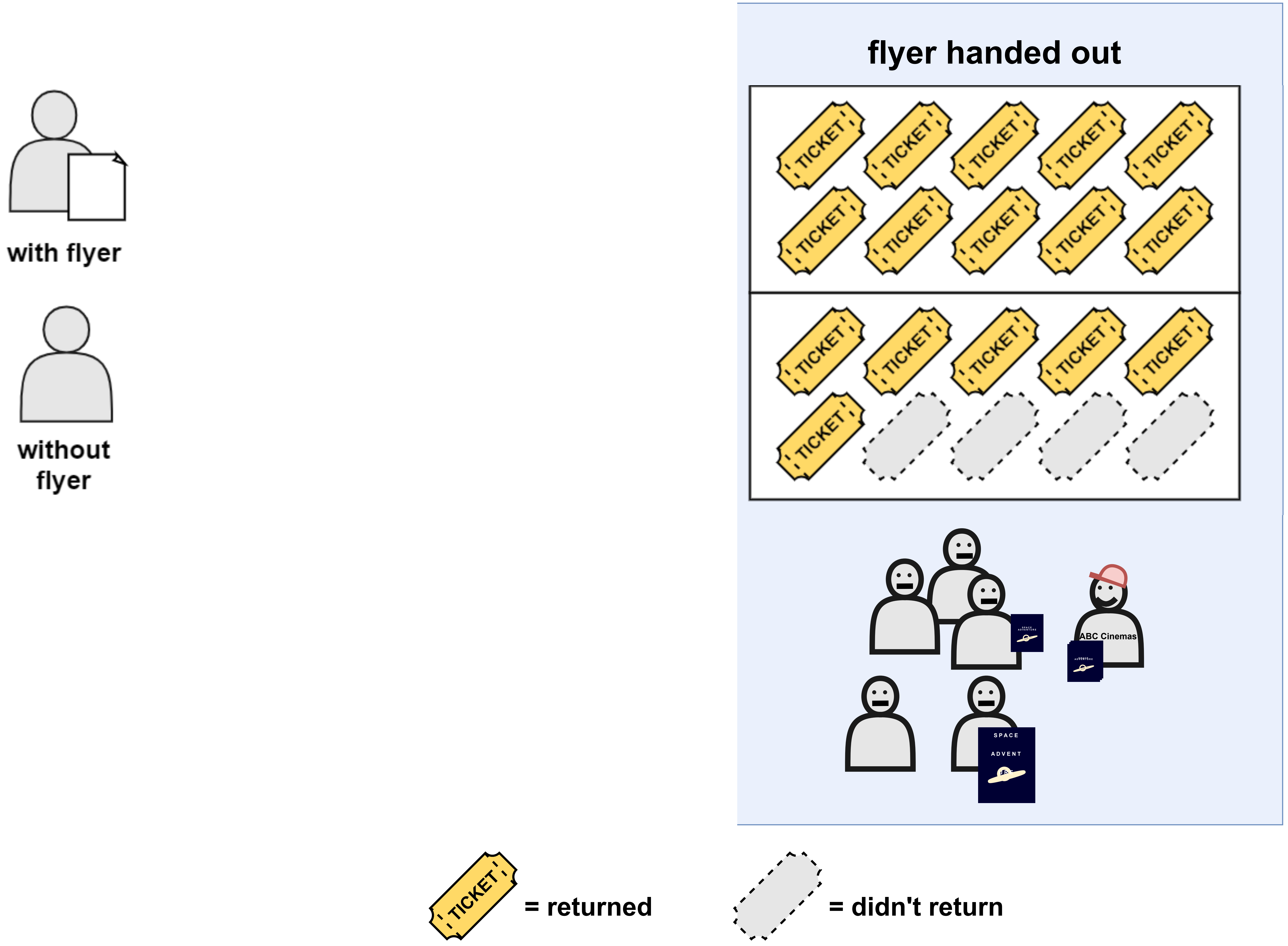


Level	Concept	Expression	Activity	Question	Example
I	Observation/ Prediction	$p(y \mid x)$	Seeing	How does x change my belief in y?	Would the grass be wet if we *found* the sprinkler off?
II	Intervention/ Hypothetical	$p(y \mid \mathrm{do}(x))$	Doing	Would y happen if I did x?	Would the the grass be wet if *made sure* that the sprinkler was off?
III	Counterfactual	$p(y_x \mid x', y')$	Explaining	Would y have happened instead of y', if I had done x instead of x'?	Would the grass have been wet if we *had made sure* that the sprinkler was off, given that the grass is wet and the sprinkler on?

Observation

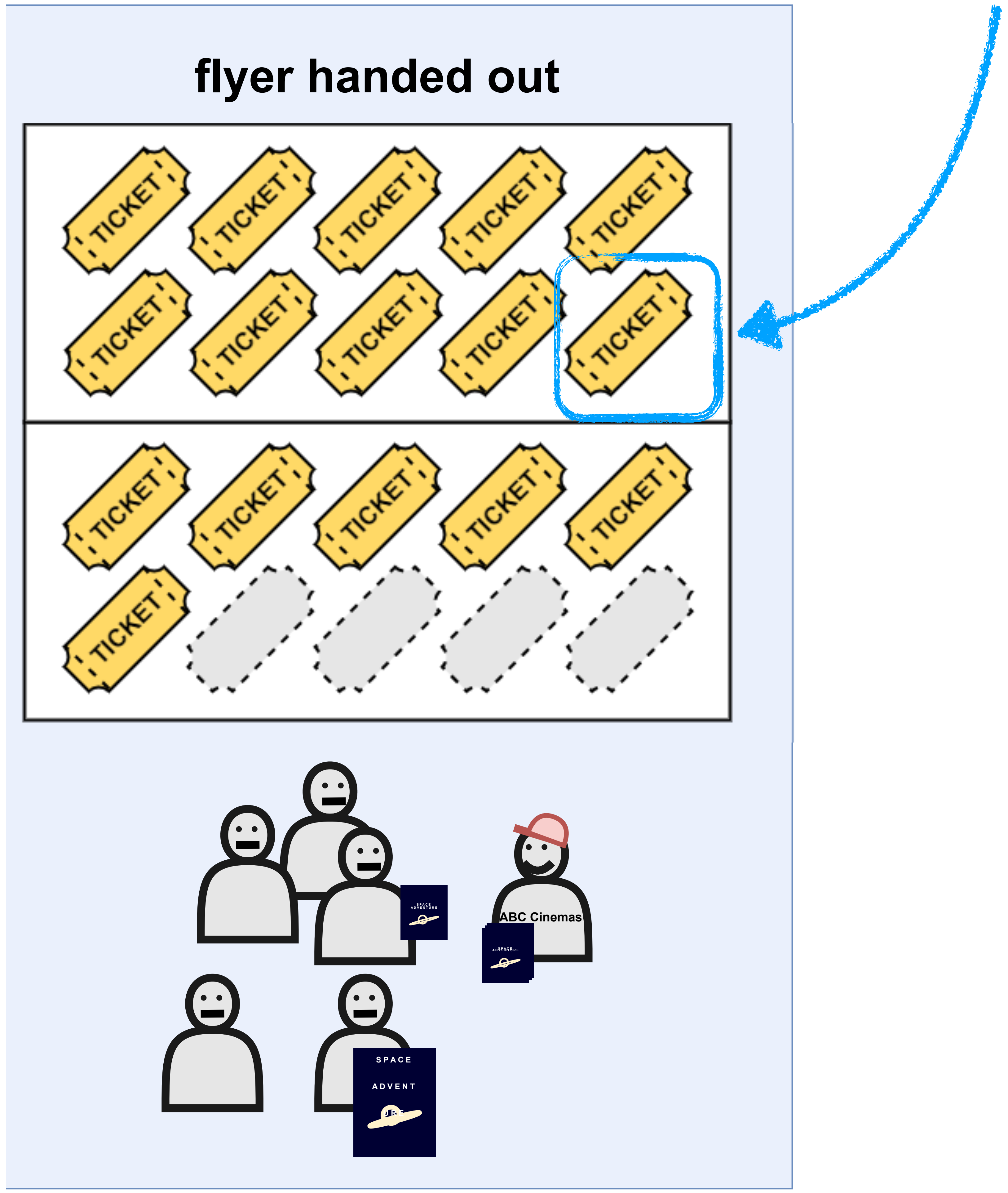


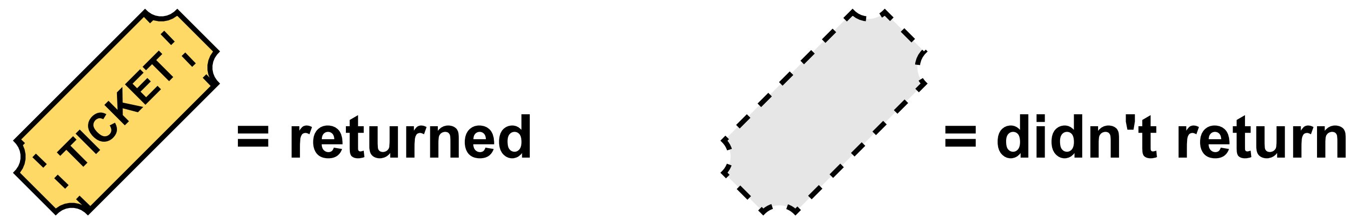
Intervention/Experiment

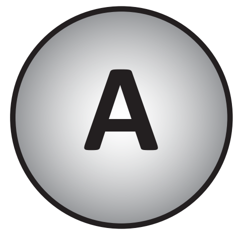
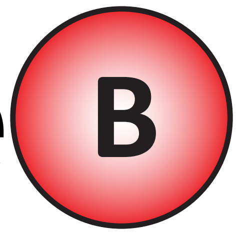


Counterfactual

Did this person go because they were handed a flyer?

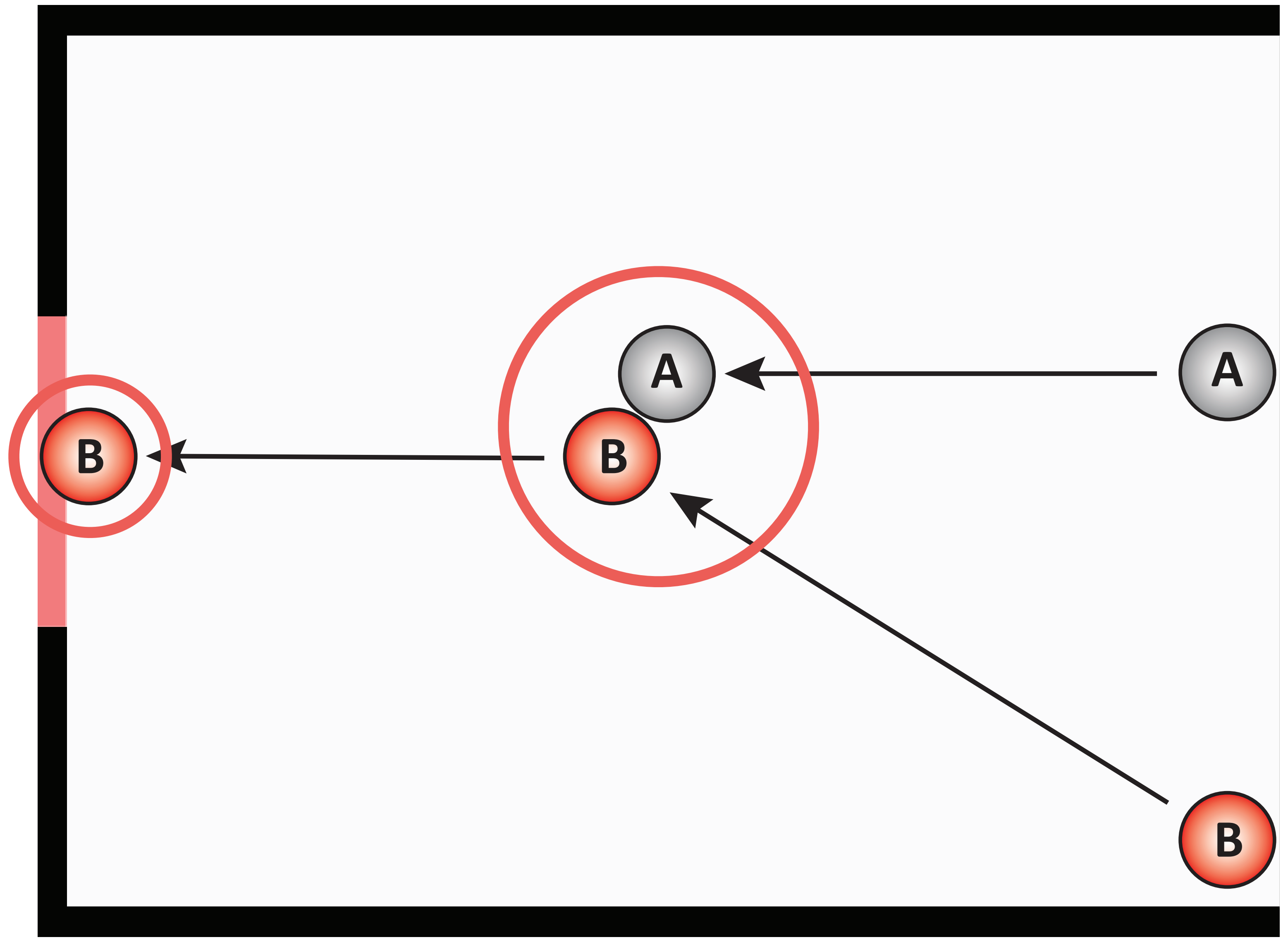




Did  cause  to go through the gate?

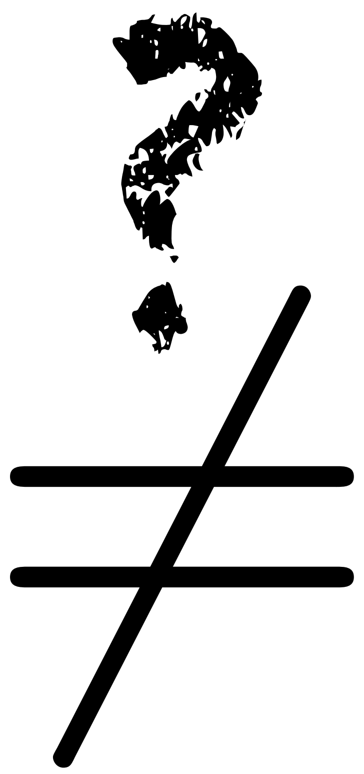
Counterfactual Simulation Model

What happened?

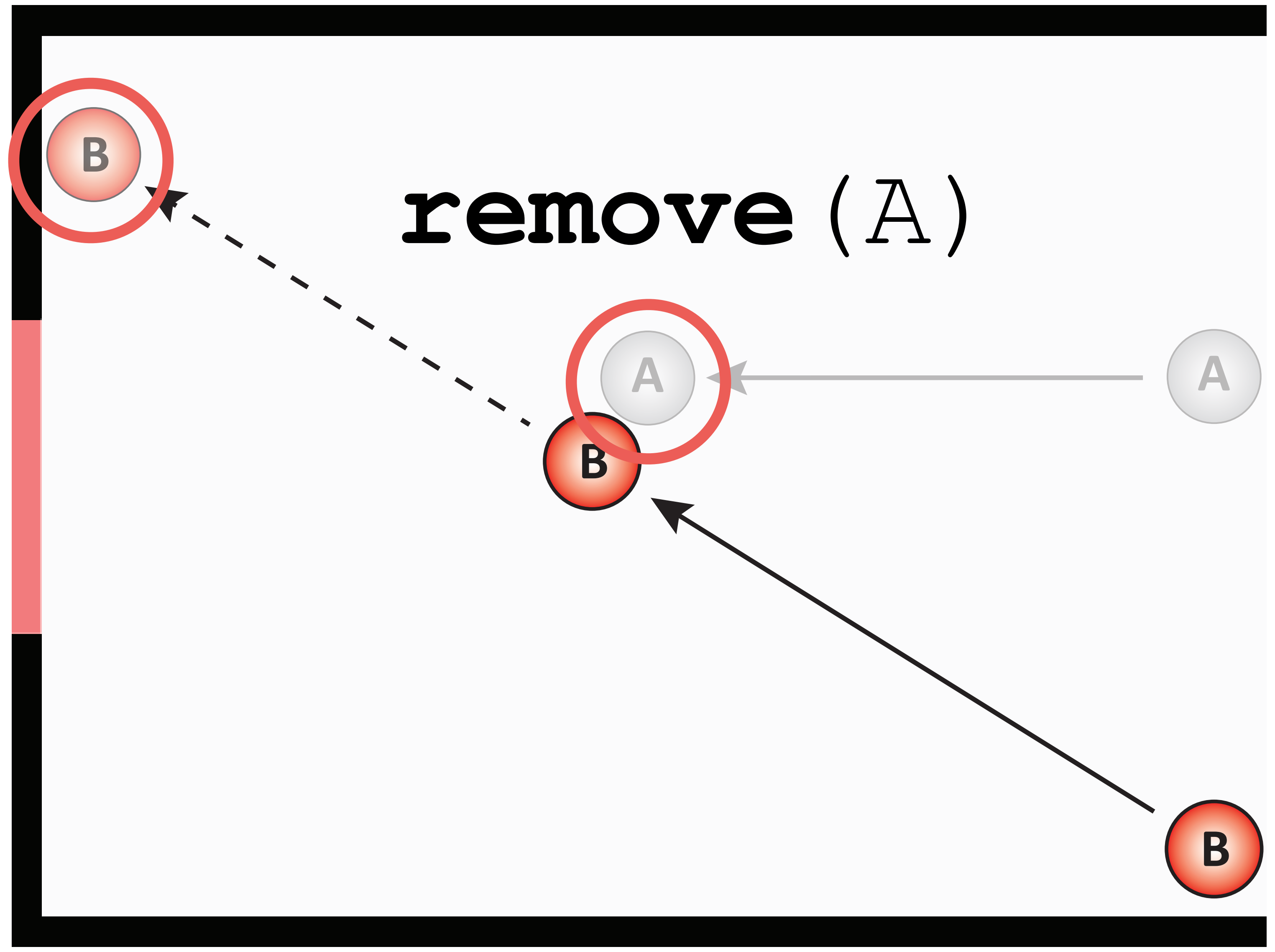


Actual situation

B went through the gate



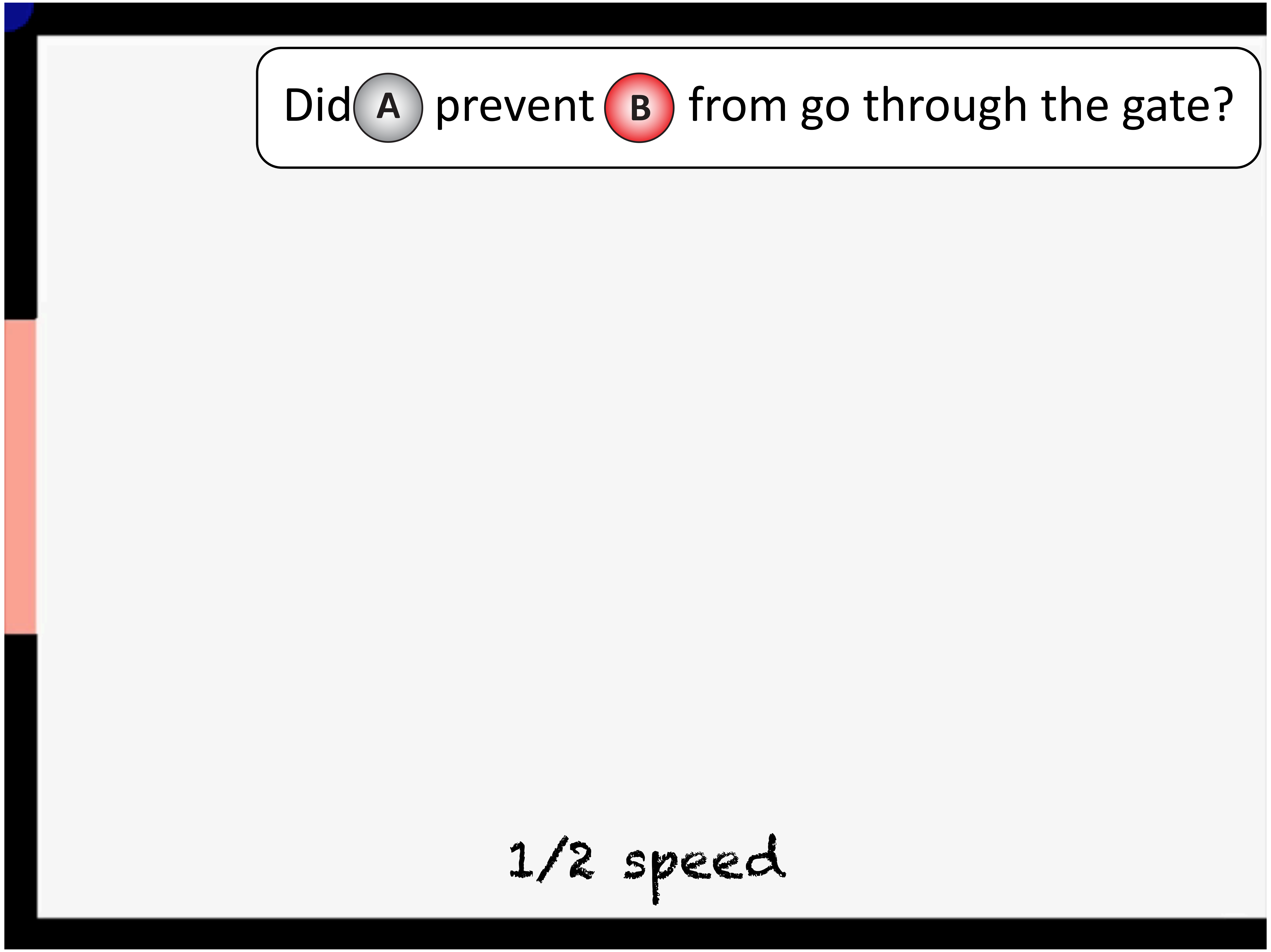
What would have happened?



Counterfactual situation

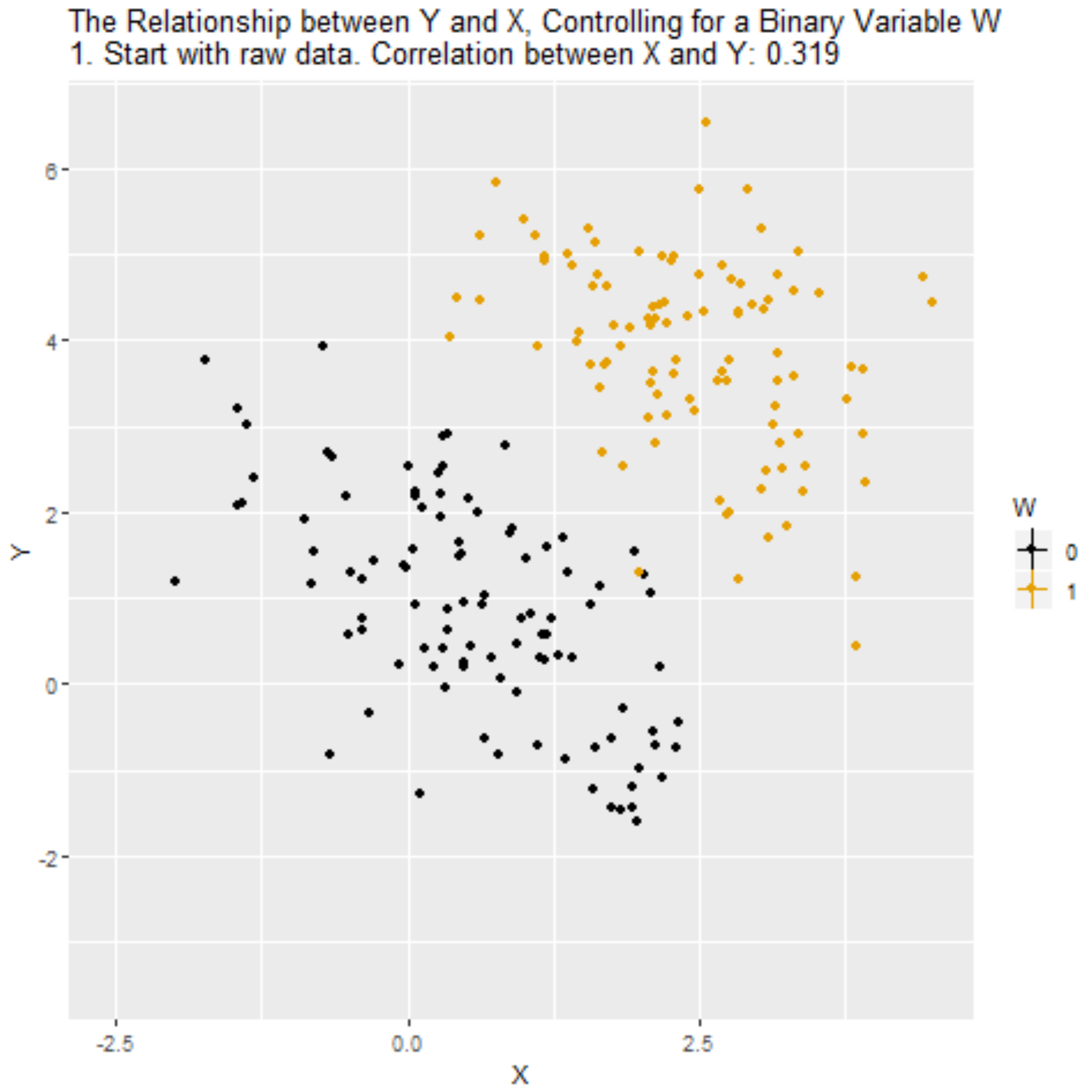
B would have missed the gate

Gerstenberg, Goodman, Lagnado, & Tenenbaum (2021). A counterfactual simulation model of causal judgment for physical events. *Psychological Review*



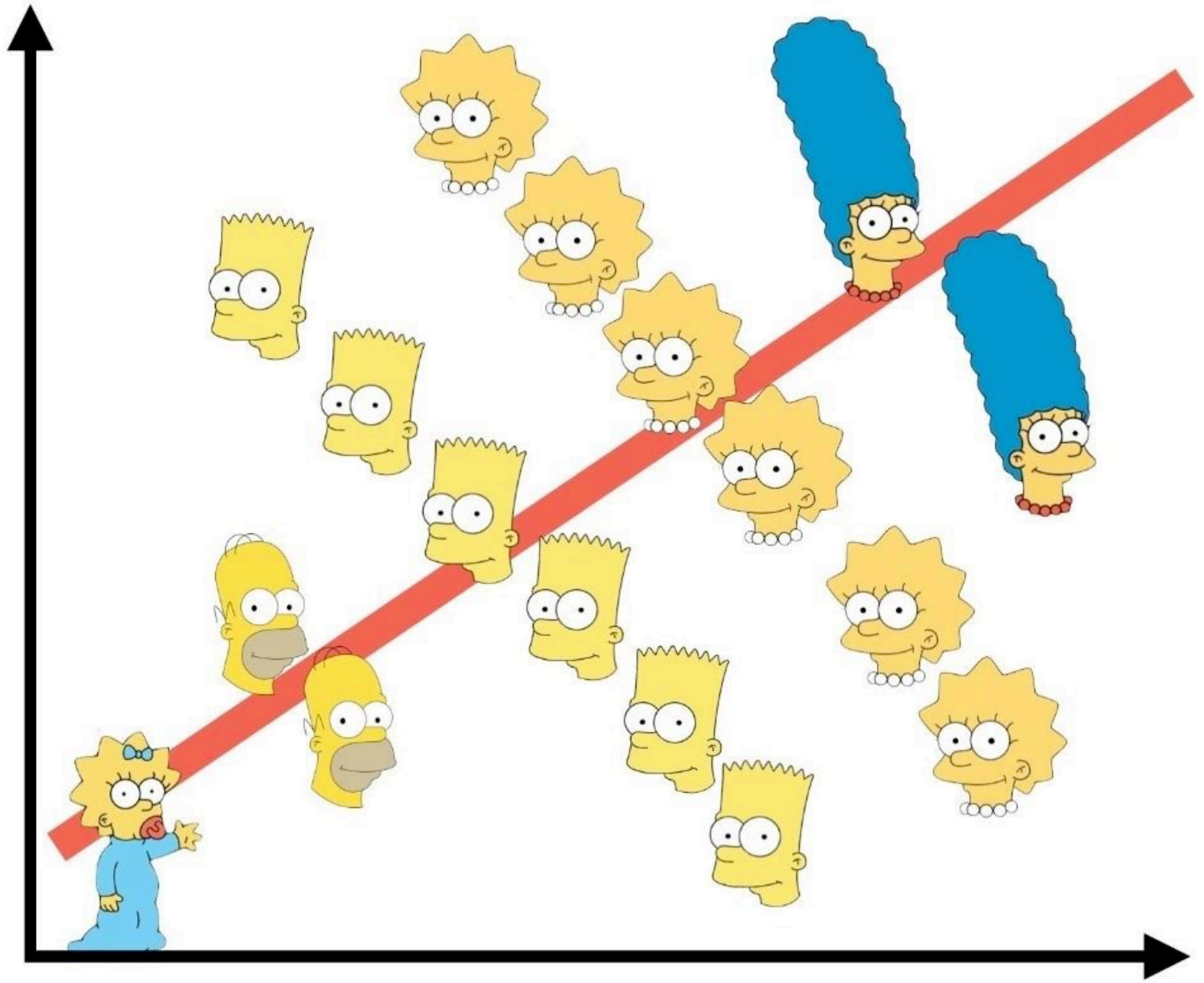
Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum (2017) Eye-tracking causality. *Psychological Science*

Controlling for variables



https://nickchk.com/causalgraphs.html

Simpson's paradox



What does controlling for variables mean?

we are not actually "**controlling**" the variable

instead, we are taking the variable into consideration when making predictions

the hope is that we get a better estimate of the parameter that we are interested in by taking into account other factors

Patterns of inference

Patterns of inference: **Common cause**

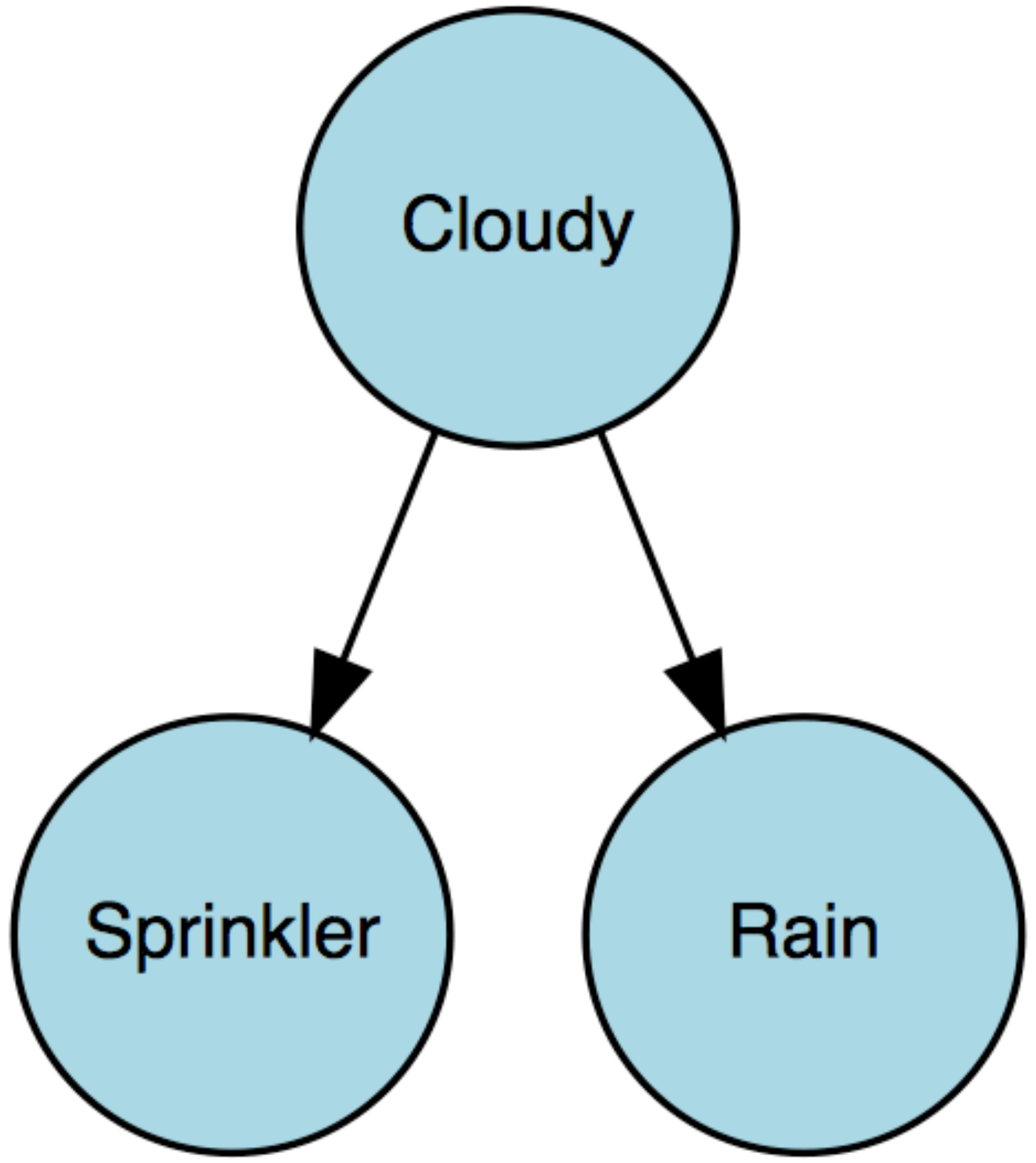
$$p(S \mid R) = p(S)$$

or

$$p(S \mid R) \neq p(S)$$

?

p(C)
0.5



C	p(S)	C	p(R)
F	0.5	F	0
T	0.1	T	0.3

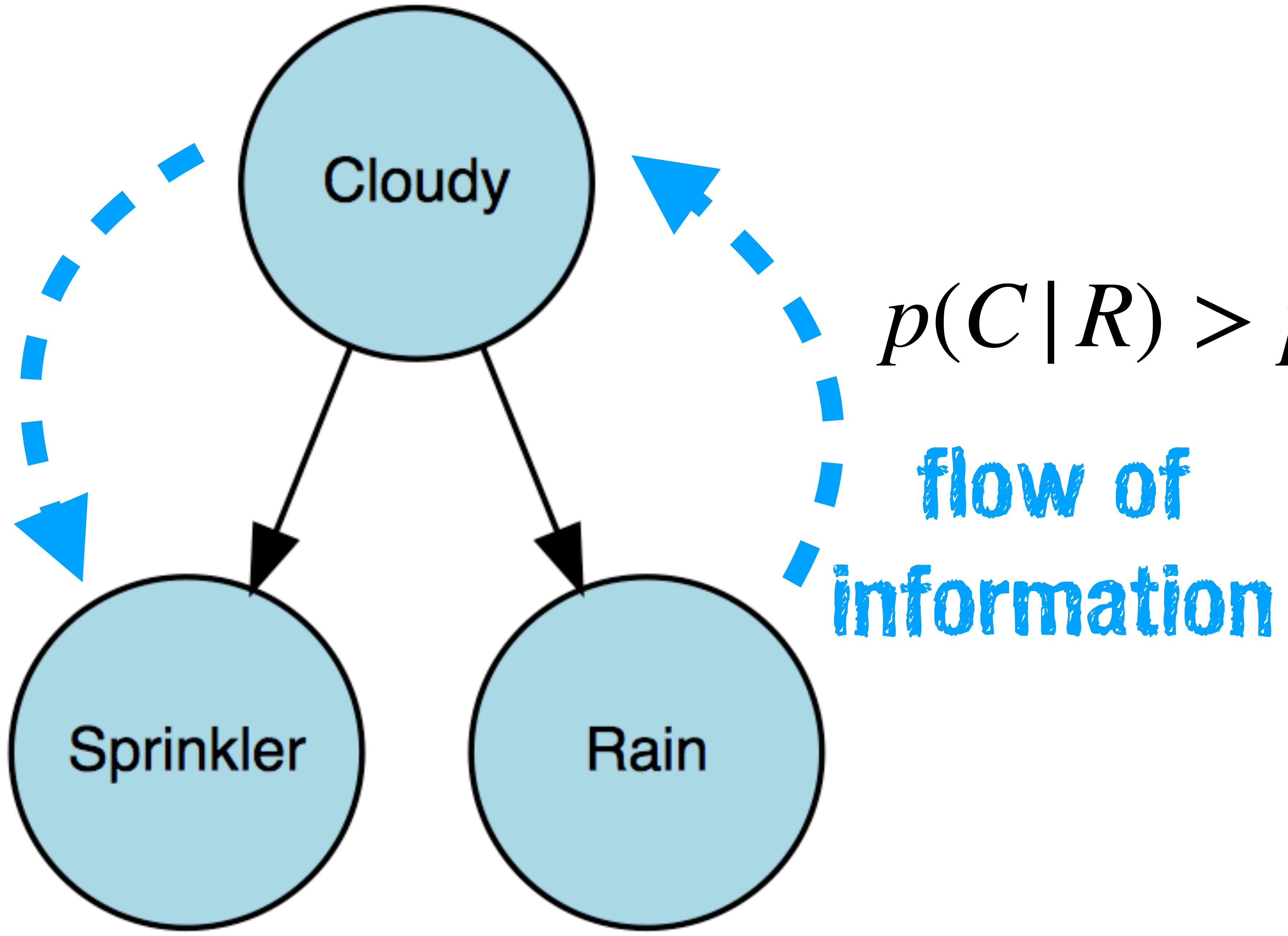
Patterns of inference: **Common cause**

- effects of a common cause are *unconditionally dependent*

$$p(S\,|\,R) \neq p(S)$$

p(C)
0.5

$p(S\,|\,C) < p(S)$



$p(C\,|\,R) > p(C)$

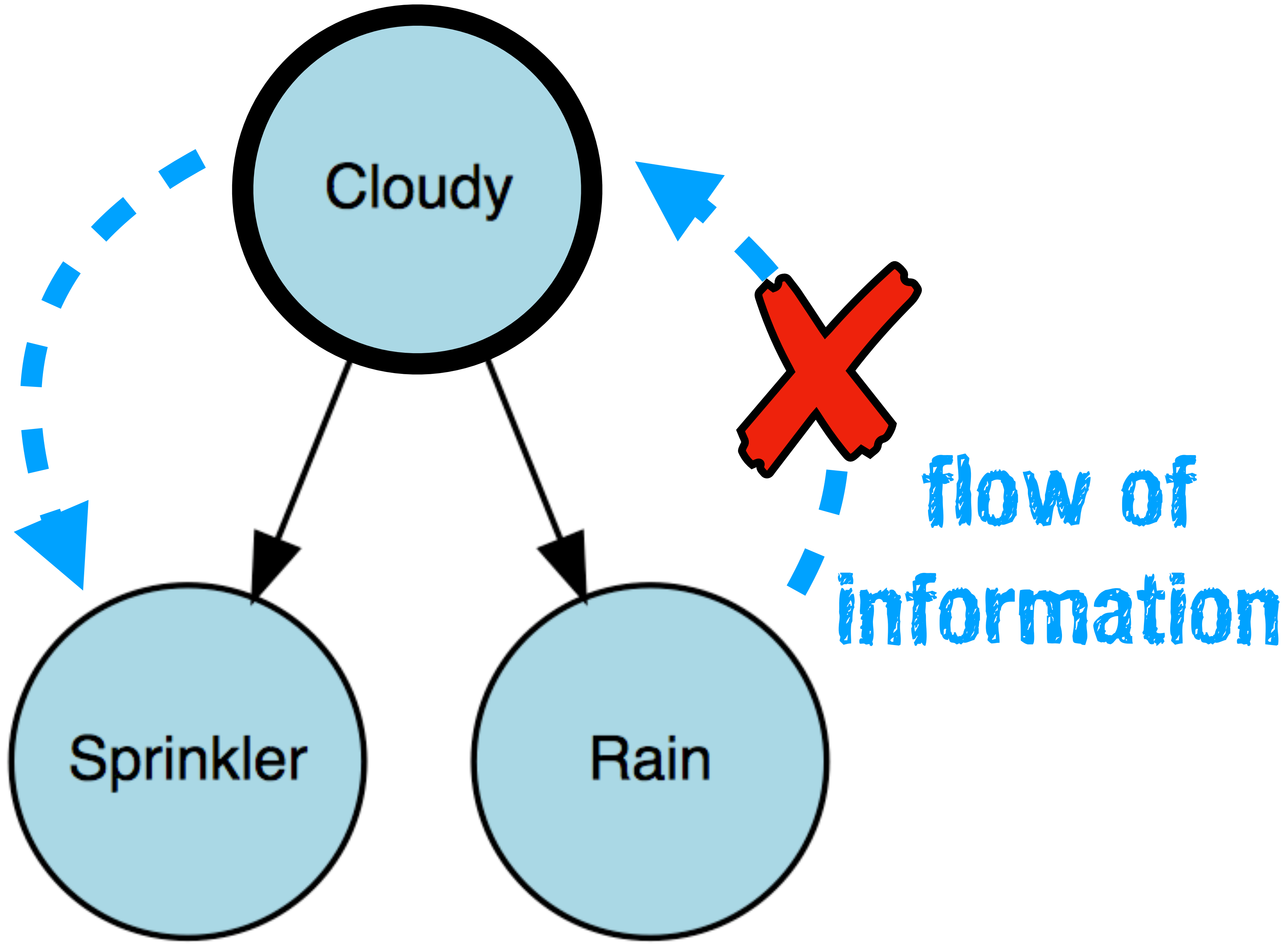
C	p(S)	C	p(R)
F	0.5	F	0
T	0.1	T	0.3

Patterns of inference: **Common cause**

- effects of a common cause are *conditionally independent given the cause*

$$p(S \mid R, C) = p(S \mid C)$$

p(C)
1

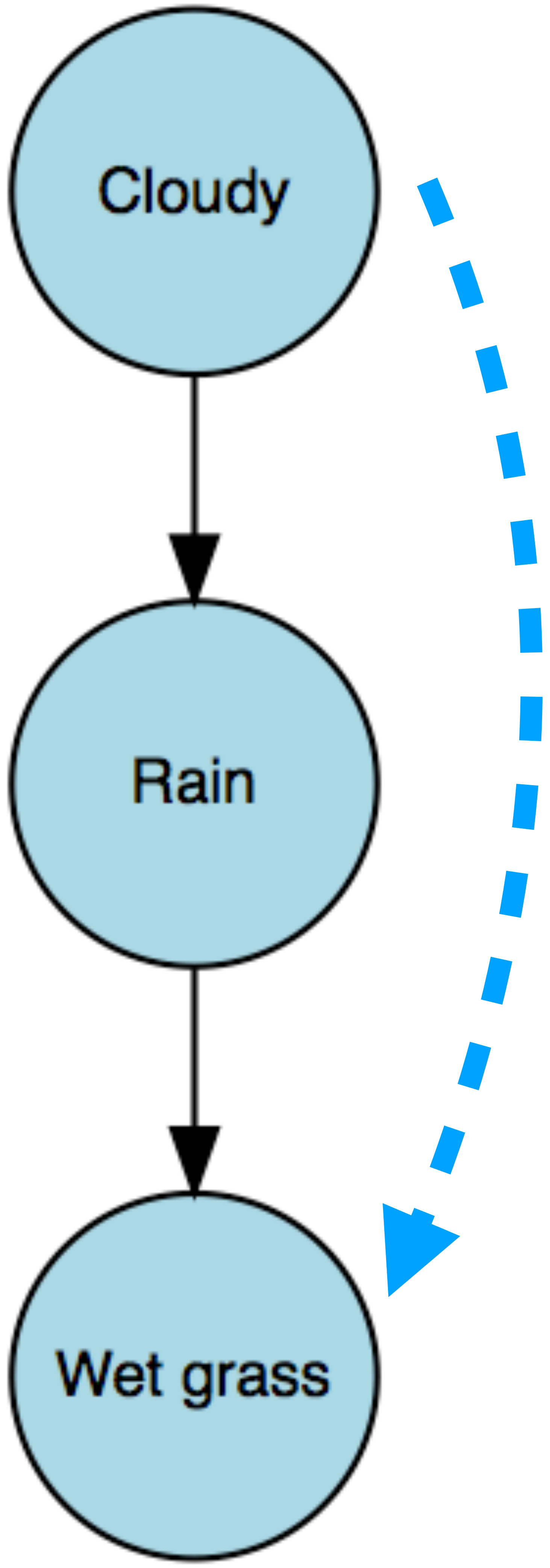


C	p(S)	C	p(R)
F	0.5	F	0
T	0.1	T	0.3

Patterns of inference: **Causal chain**

- cause and effect in a causal chain are *unconditionally dependent*

$$p(W \mid C) \neq p(W)$$

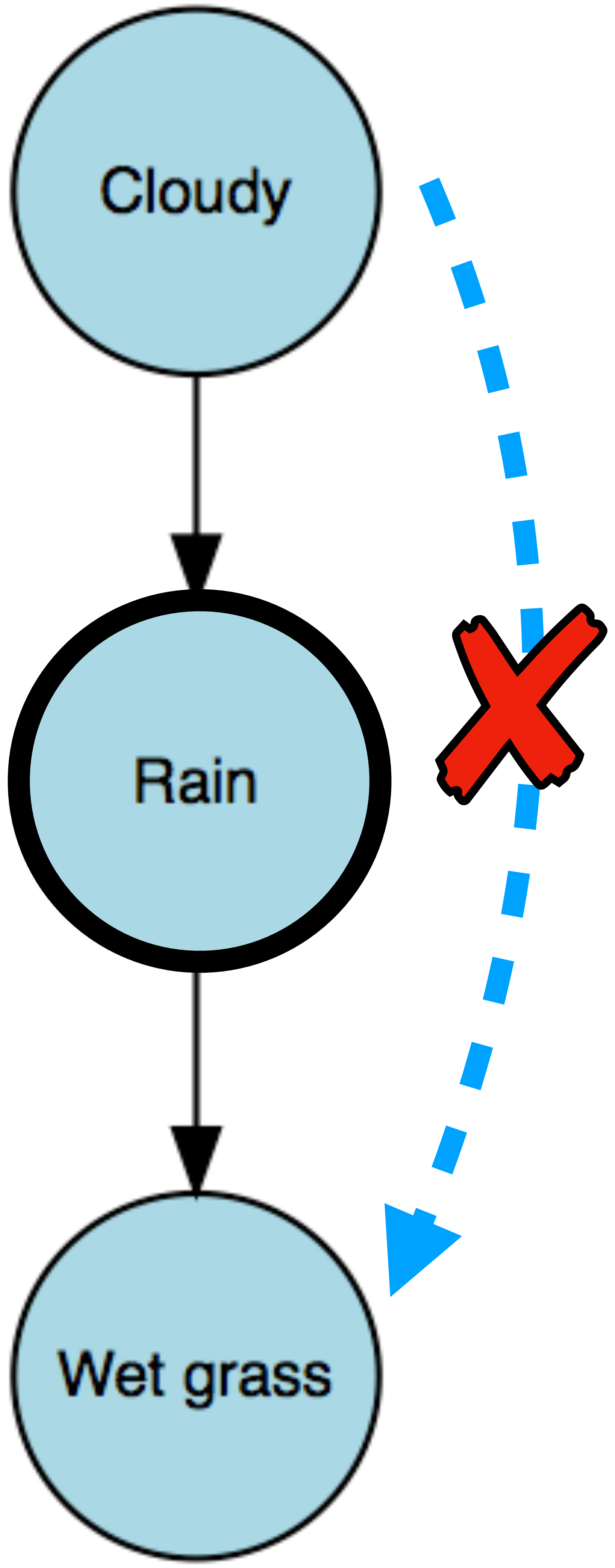


Patterns of inference: **Causal chain**

- cause and effect in a causal chain are *conditionally independent*

$$p(W \mid C, R) = p(W \mid R)$$

screening off

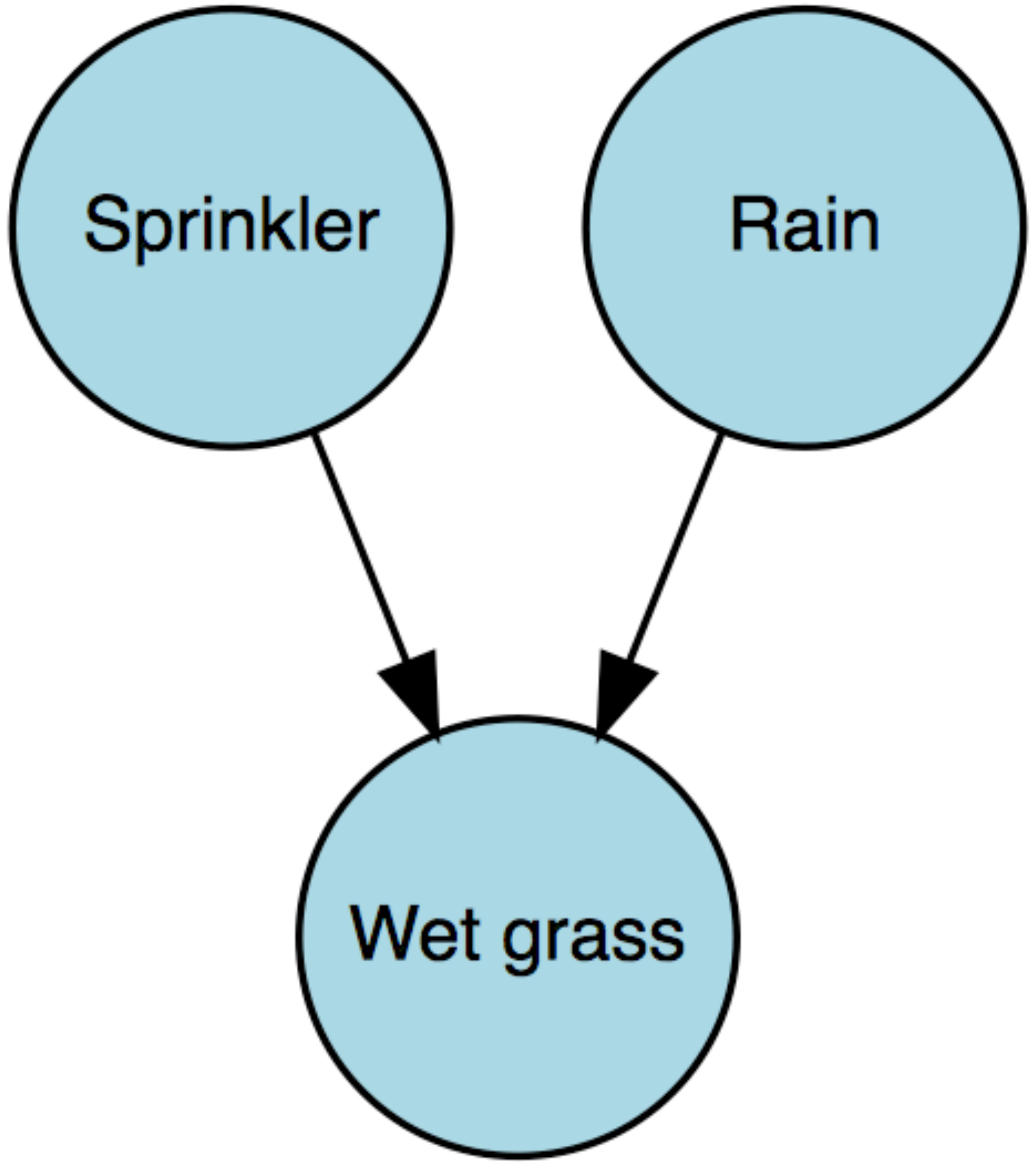


Patterns of inference: **Common effect**

- two causes of a common effect are *unconditionally independent*

$$p(S \mid R) = p(S)$$

(e.g. Sprinkler is set by a timer)

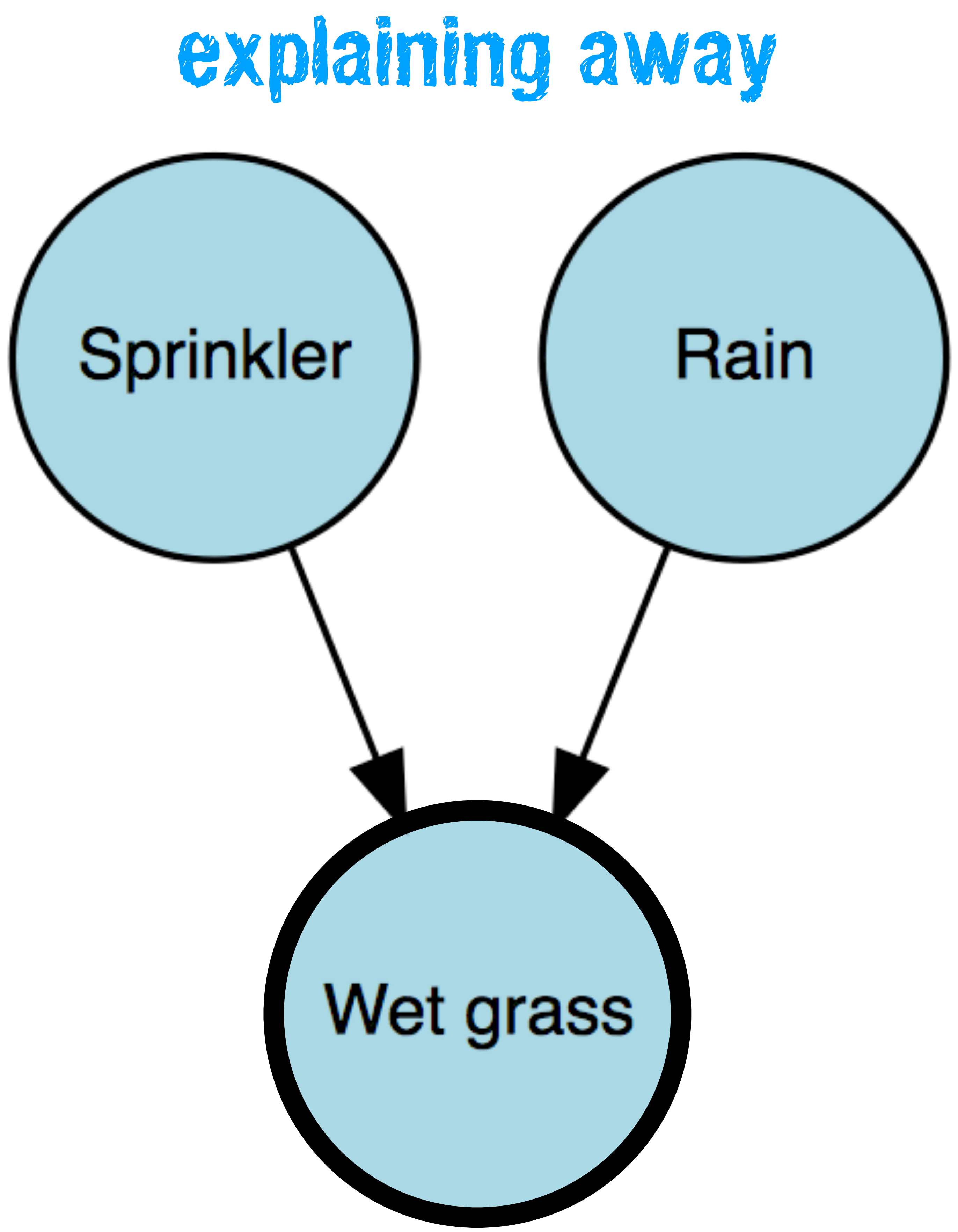


Patterns of inference: **Common effect**

- two causes of a common effect are *conditionally dependent given the effect*

$$p(S \mid R, W) \neq p(S \mid W)$$

- intuitively: both causes compete to explain the effect



Note: The pattern of inference depends on the structural form which captures how `Sprinkler` and `Rain` jointly affect `Wet grass`. Explaining away holds for the commonly used noisy-or integration function.

Should I control?

When should I control for variables?

recent advances in graphical models have produced a way to help distinguish good from bad controls

***d*-separation**

directional

decide from a causal graph whether a set of variables X is independent of another set Y, given a third set Z

Goal: we want a precise (and unbiased) estimate of the predictive relationship between X and Y

we want to block all other paths from X to Y

http://bayes.cs.ucla.edu/BOOK-09/ch11-1-2-final.pdf

When should I control for variables?

How can I tell whether two variables are independent?

Recipe for independence

1. Draw the ancestral graph
2. "Moralize" the graph by "marrying" the parents
3. "Disorient" the graph by replacing arrows with edges
4. Delete the givens and their edges
5. Read the answer off the graph

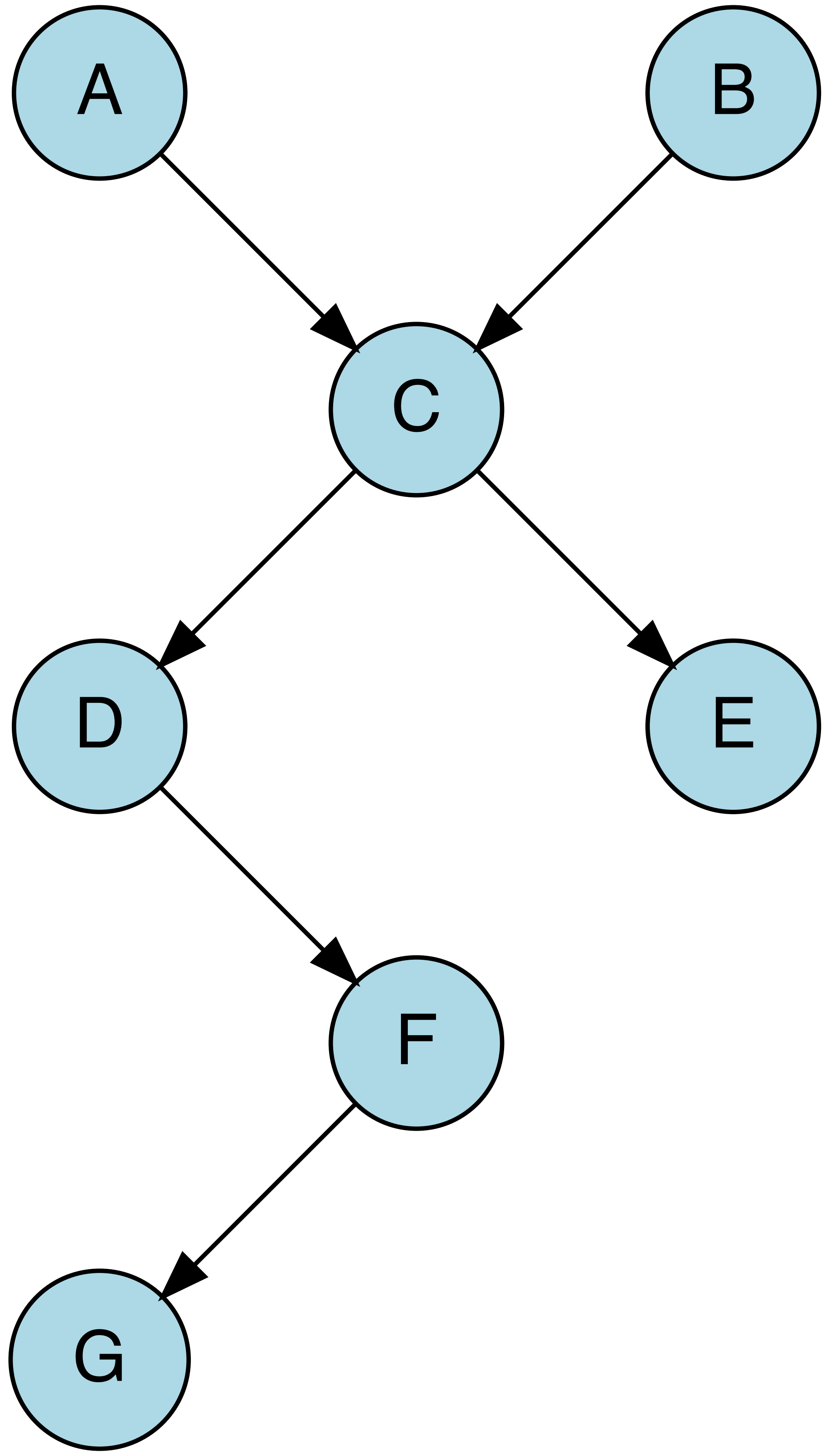
- if variables are **disconnected** they are independent
- if variables are connected (have a path between them) they are not guaranteed to be independent

http://web.mit.edu/jmn/www/6.034/d-separation.pdf

When should I control for variables?

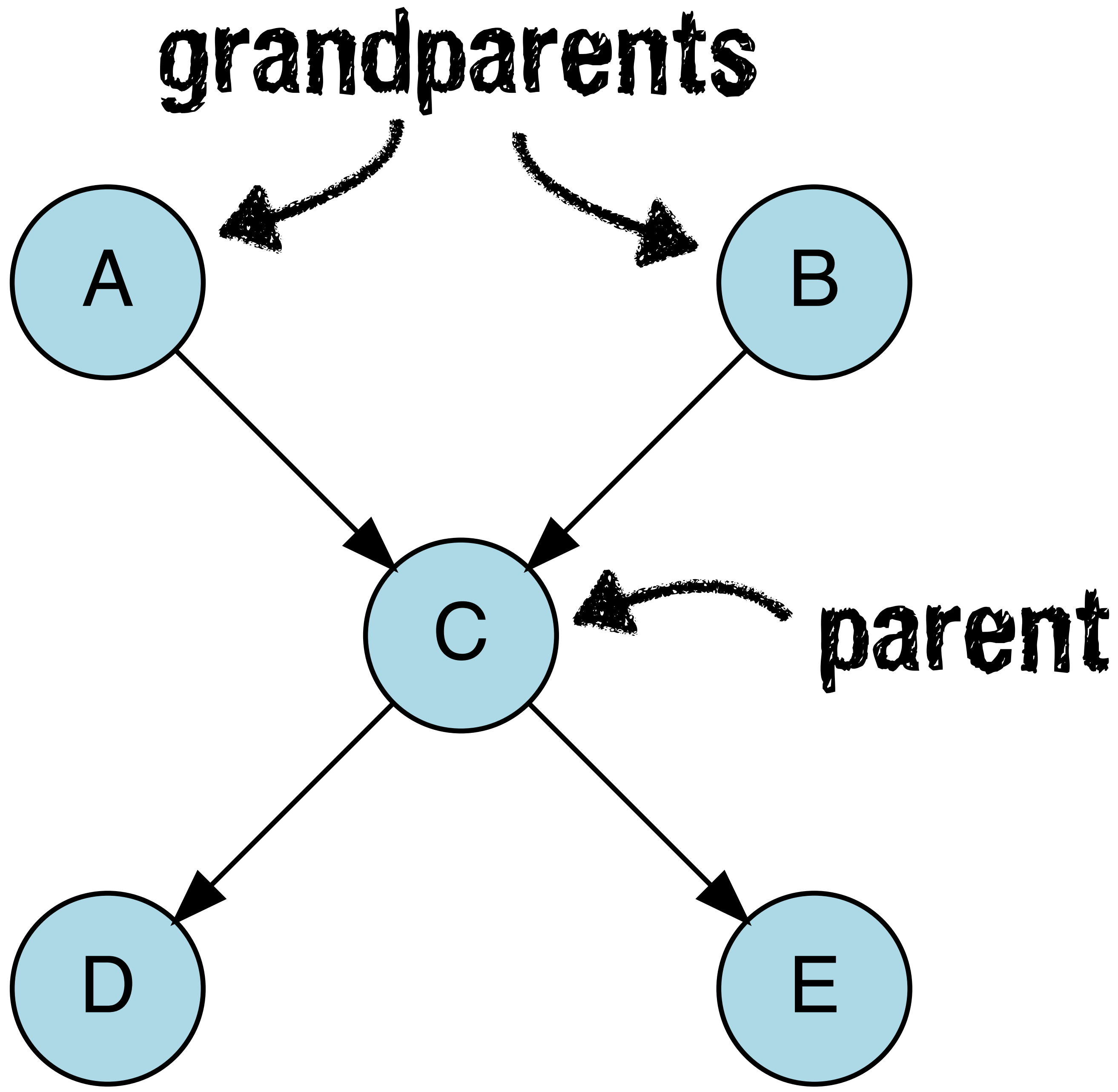
Are D and E independent?

$$p(D \mid E) = p(D) \text{ ?}$$



1. Draw the ancestral graph

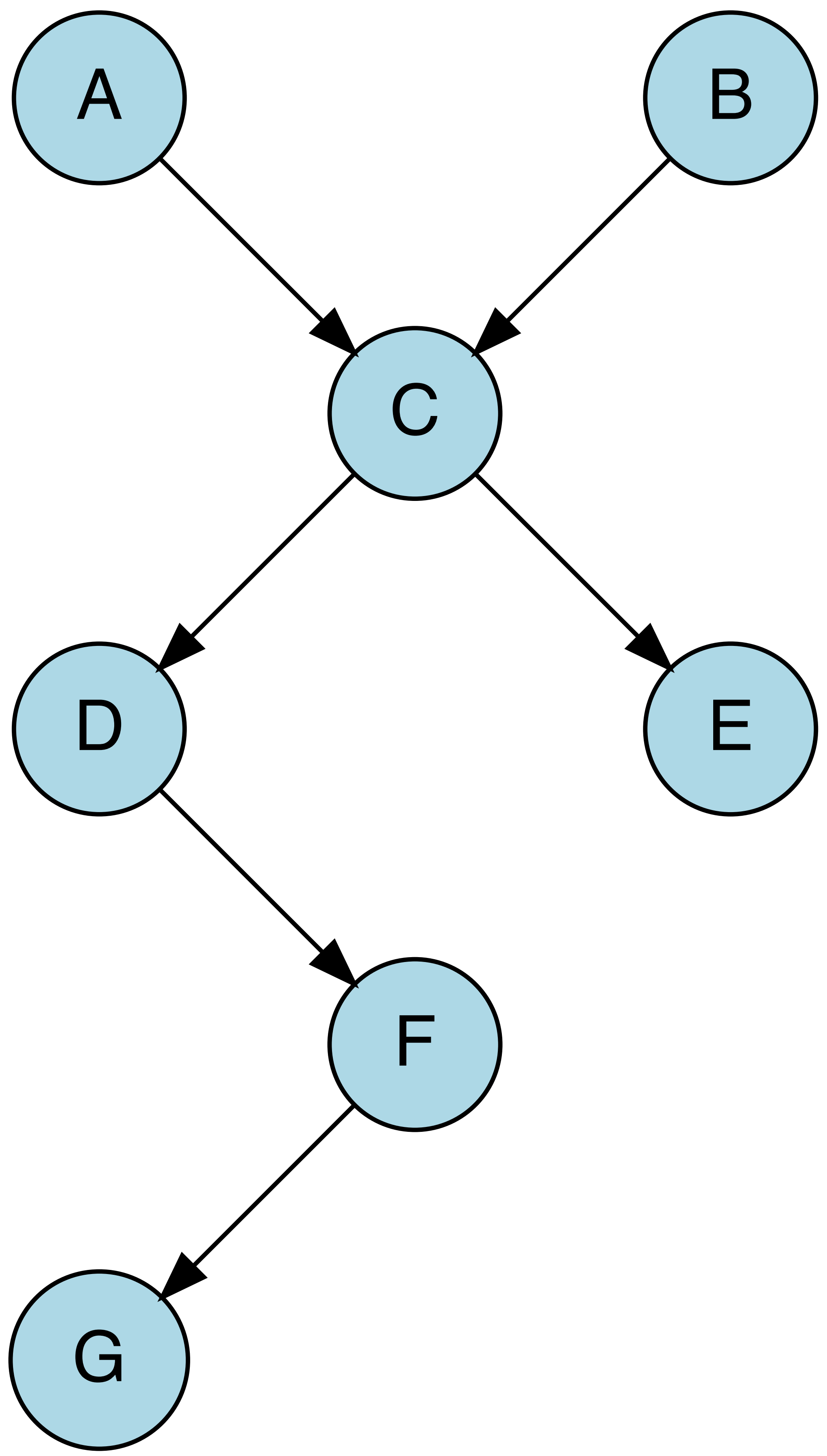
Construct the "ancestral graph" of all variables mentioned in the probability expression. This is a reduced version of the original net, consisting only of the variables mentioned and all of their ancestors (parents, parents' parents, etc.)



When should I control for variables?

Are D and E independent?

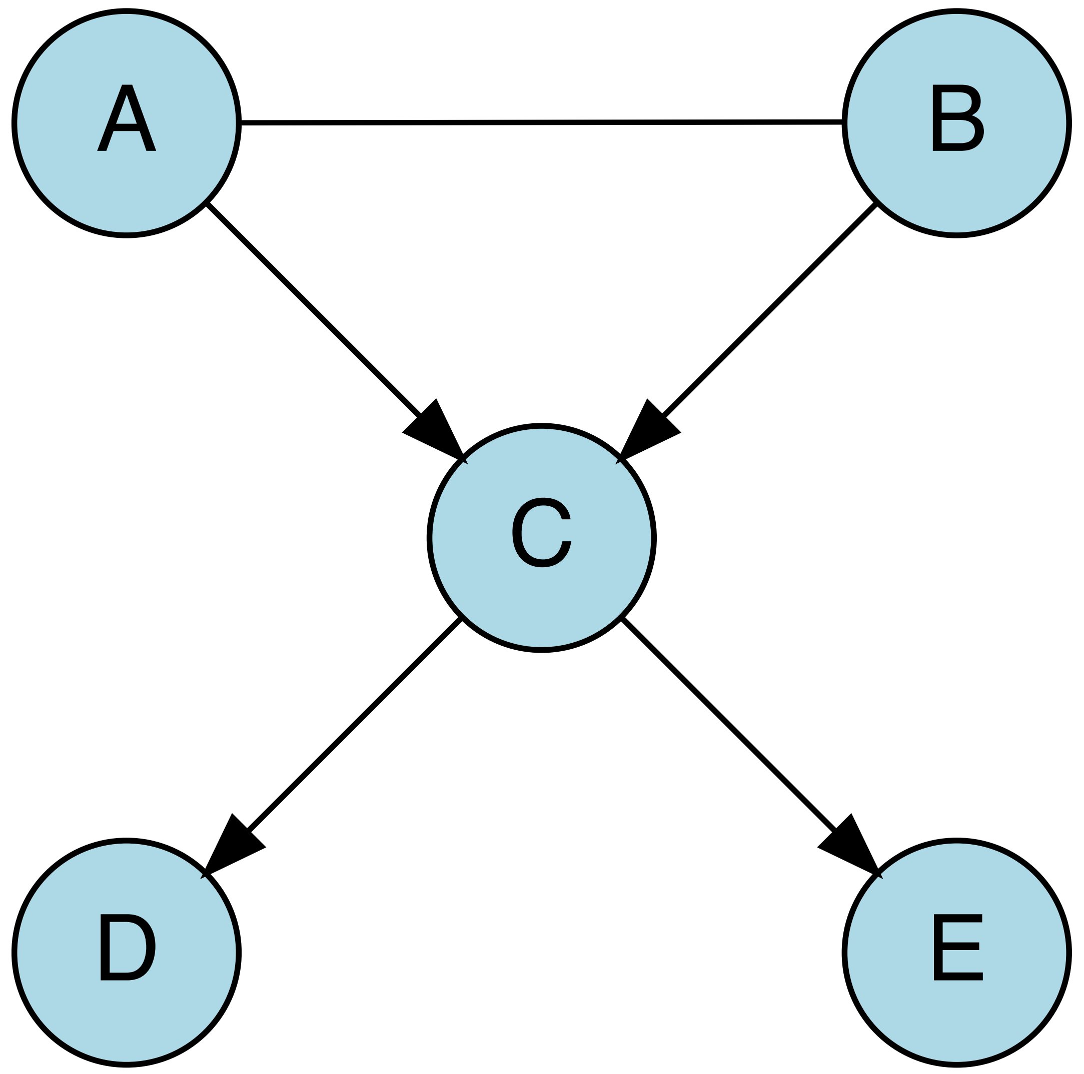
$$p(D \mid E) = p(D) \ ?$$



2. "Moralize" the graph

let's get married!

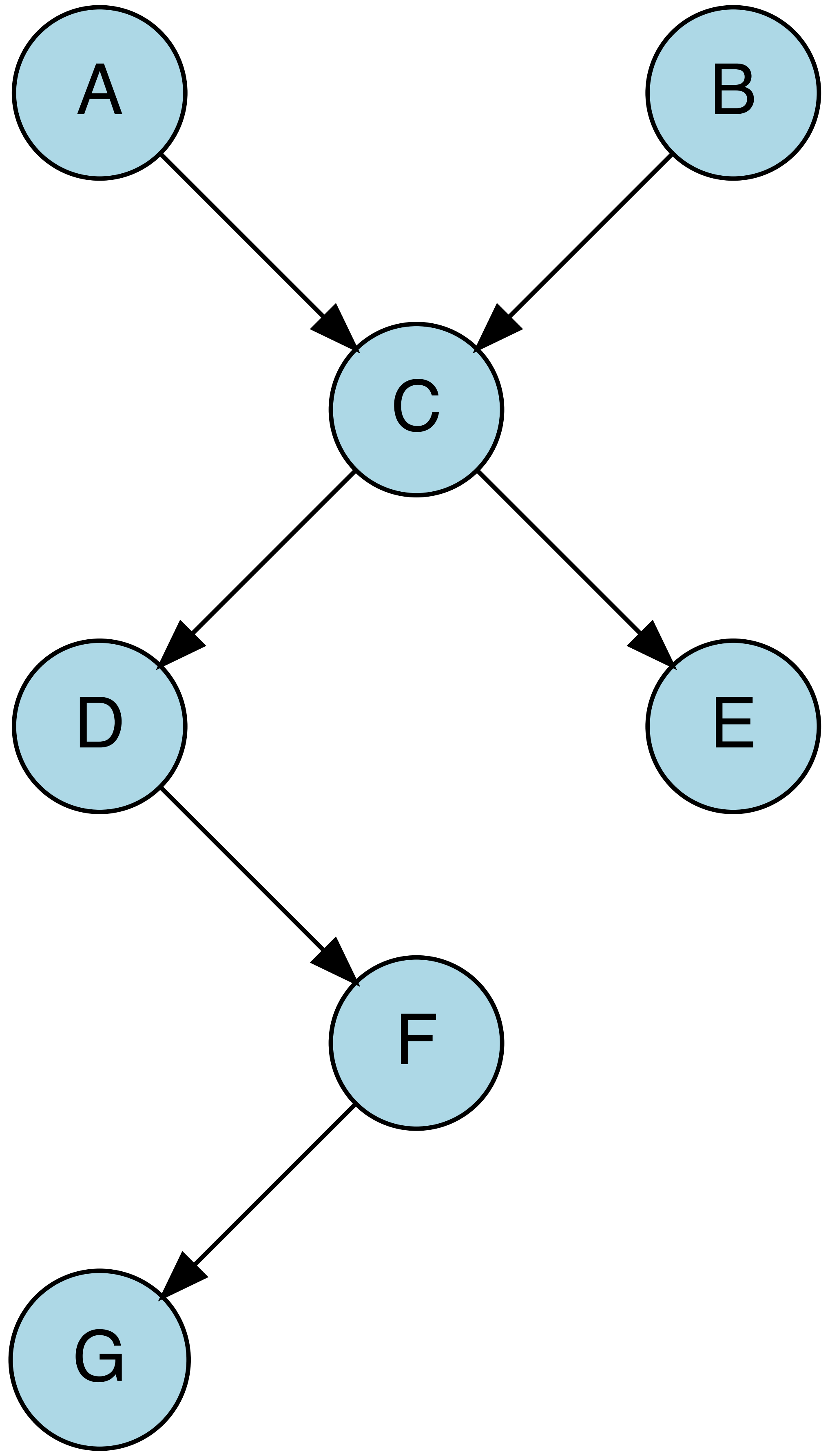
For each pair of variables with a common child, draw an undirected edge (line) between them. (If a variable has more than two parents, draw lines between every pair of parents.)



When should I control for variables?

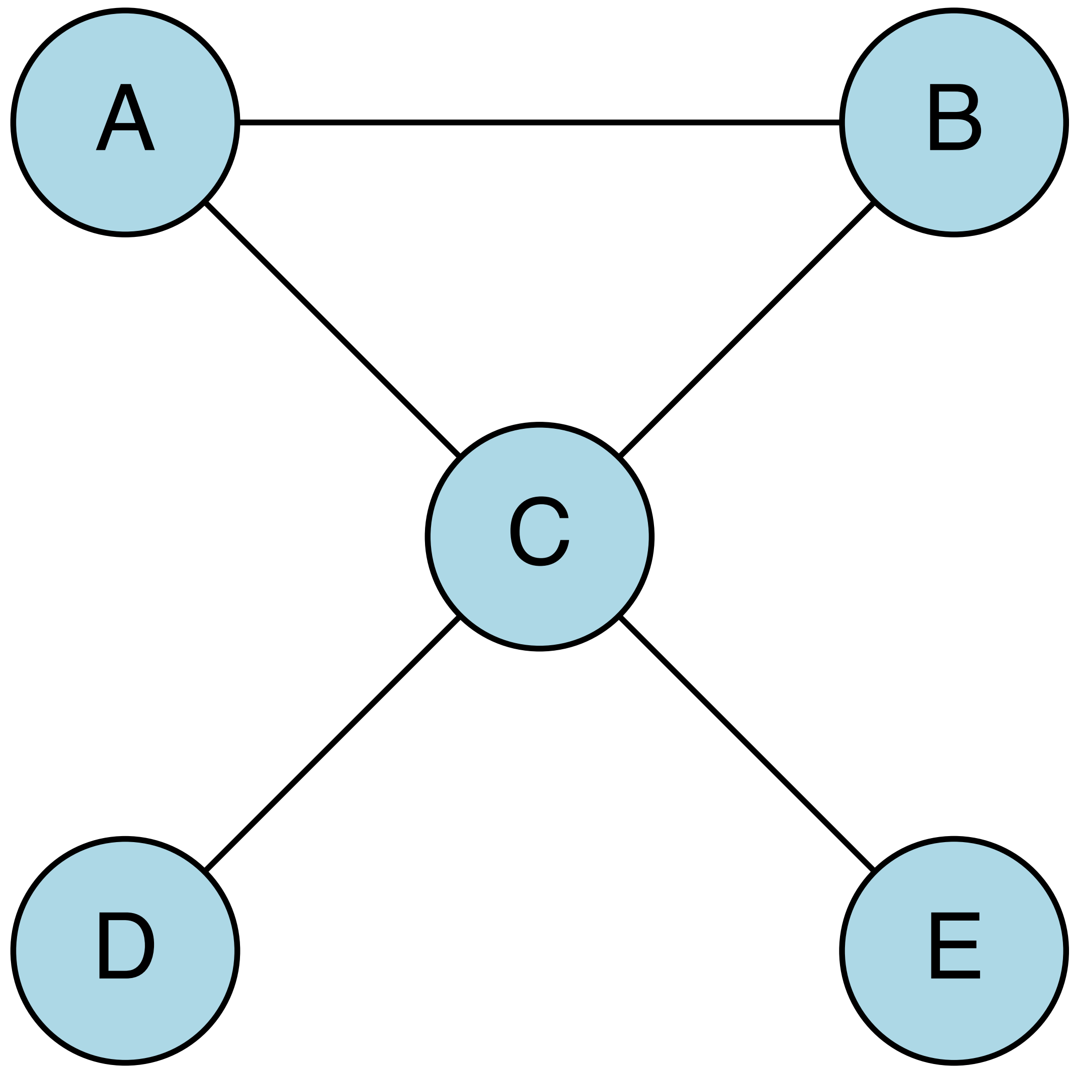
Are D and E independent?

$$p(D \mid E) = p(D) \ ?$$



3. "Disorient" the graph

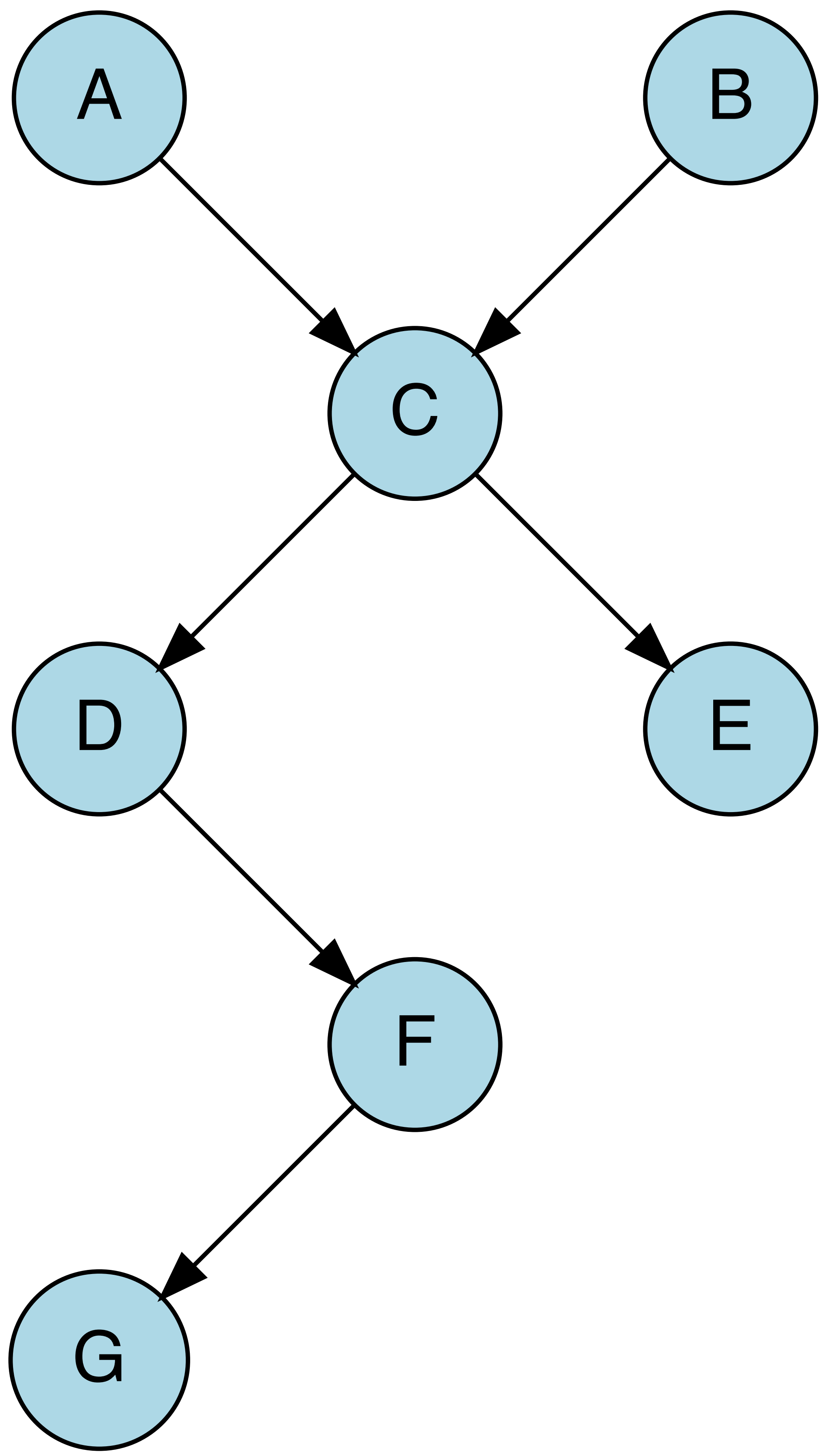
Replace arrows arrows with lines



When should I control for variables?

Are D and E independent?

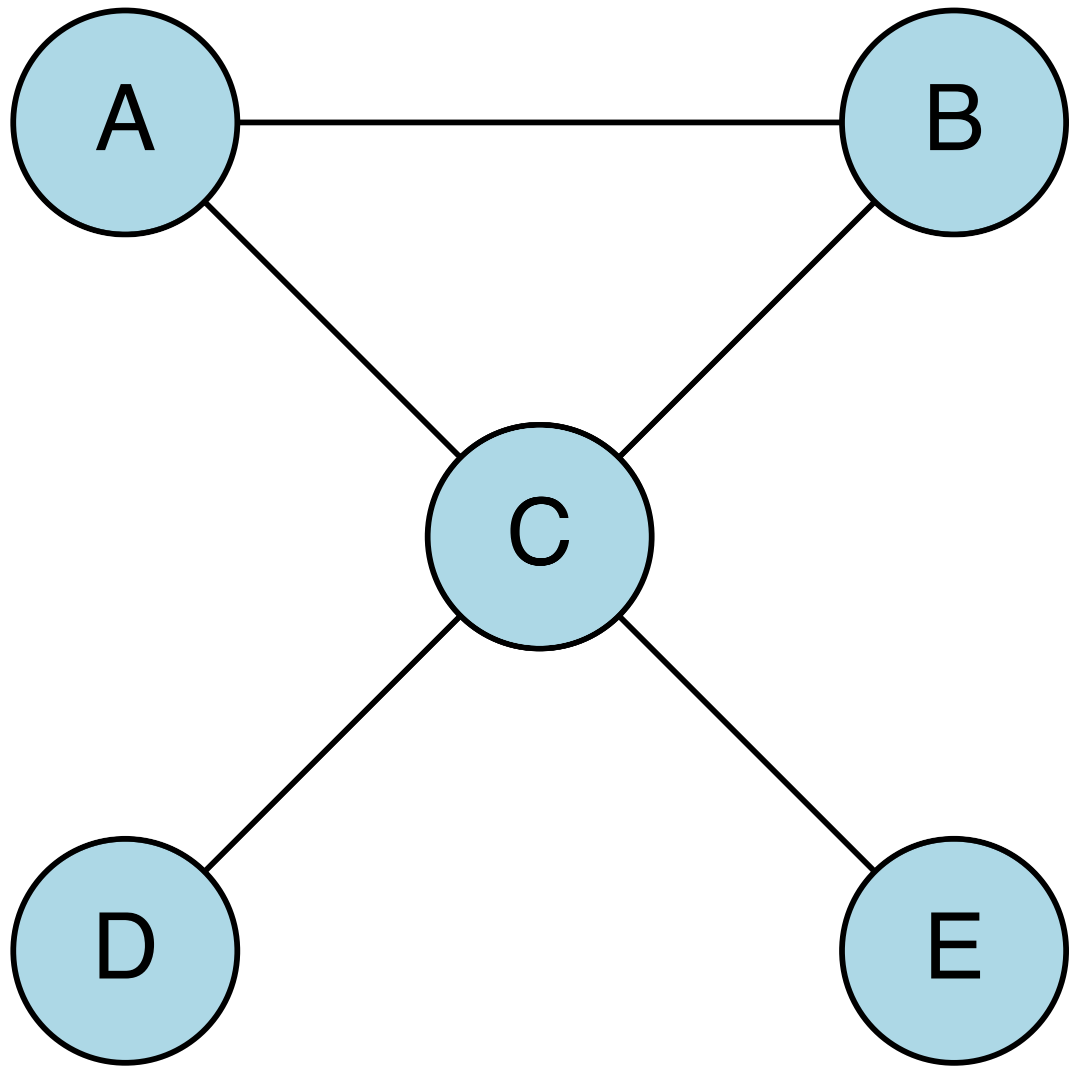
$$p(D \mid E) = p(D) \ ?$$



4. Delete the givens

Remove the variables that we condition on, as well as their edges

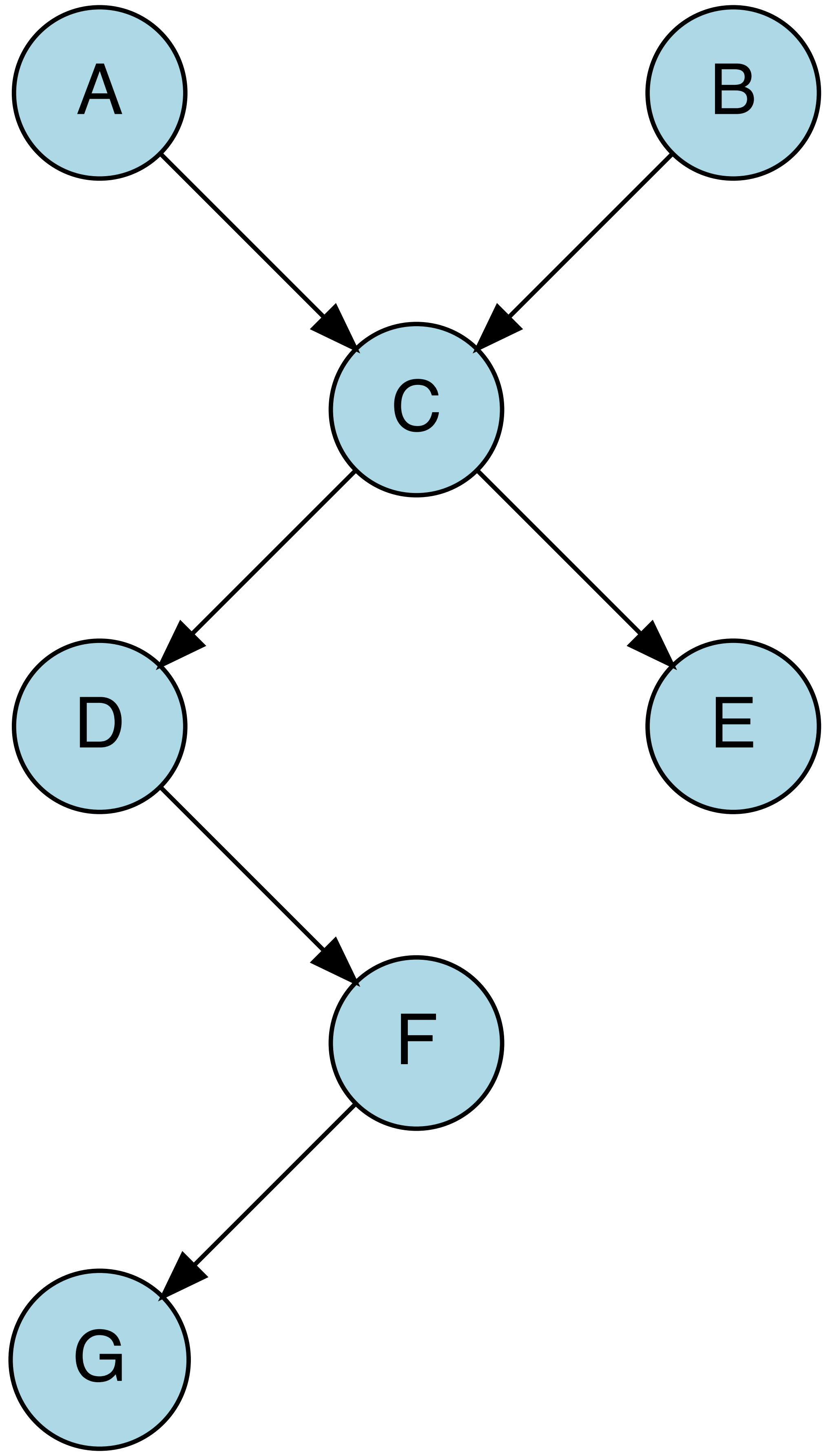
we didn't condition on anything, so there is nothing to delete



When should I control for variables?

Are D and E independent?

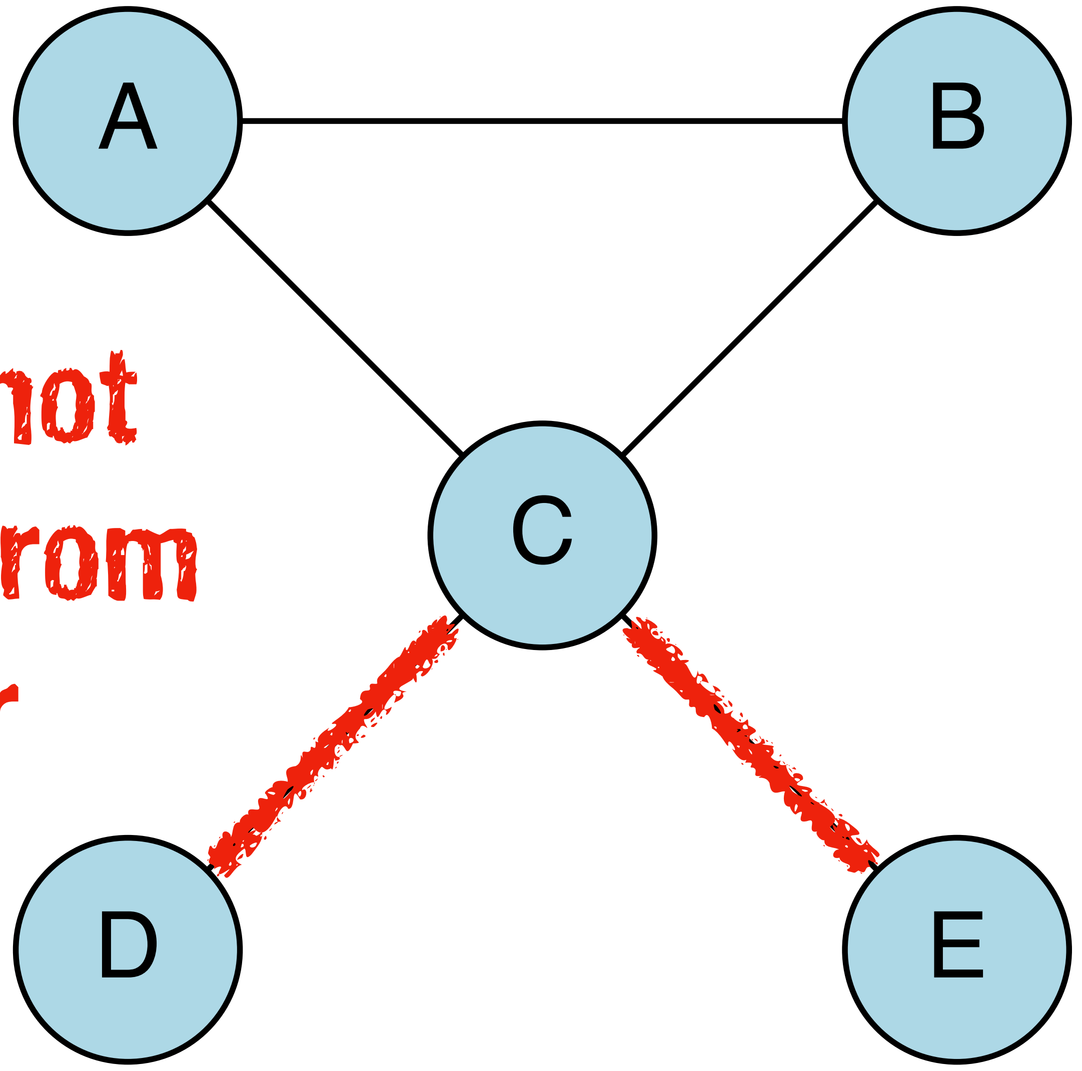
$$p(D\,|\,E) = p(D) \text{ ?}$$



5. Read answer off the graph

- if variables are **disconnected** they are independent
- if variables are connected (have a path between them) they are not guaranteed to be independent

D and E are not independent from each other

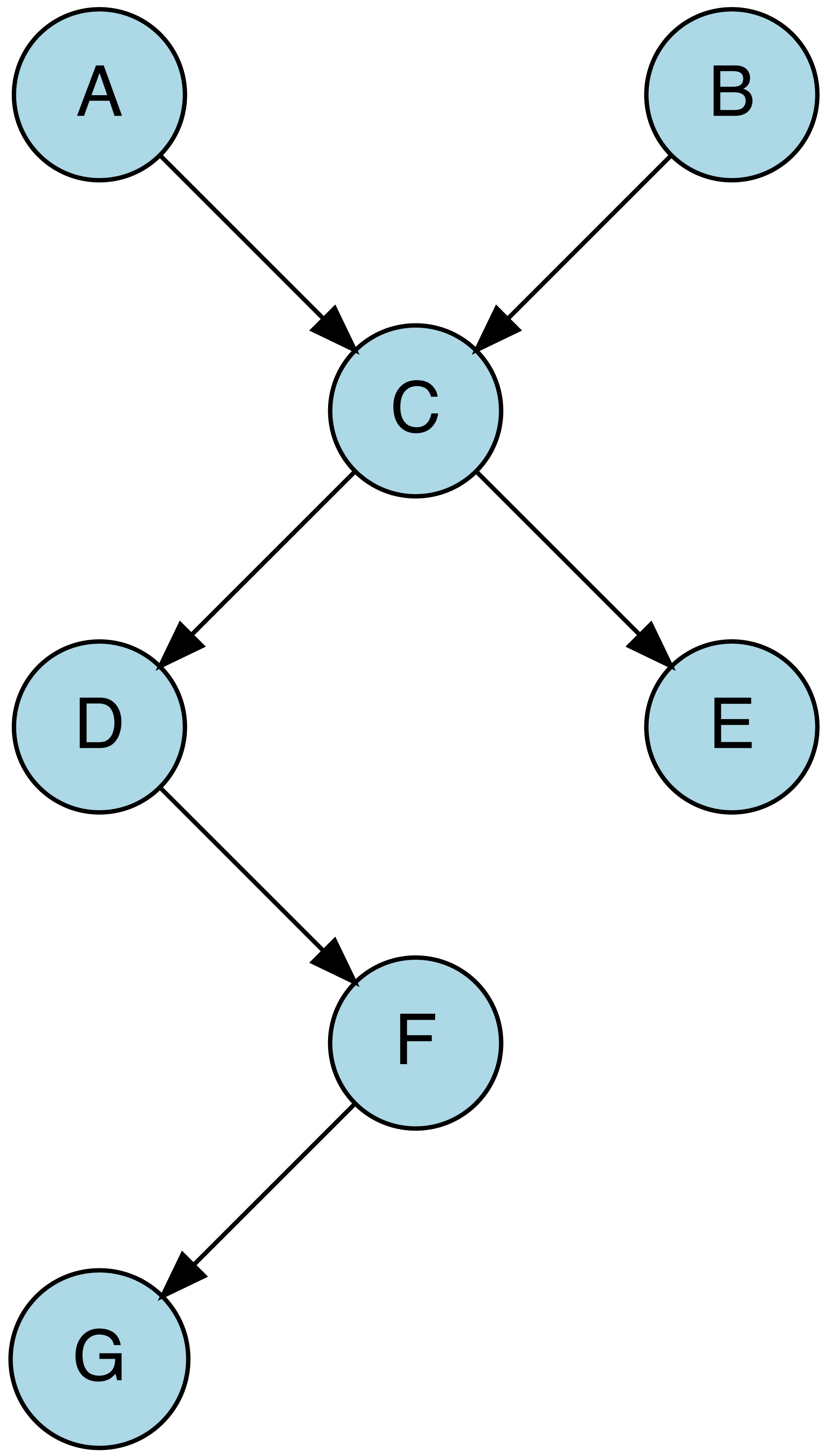


they are connected via at least one path

When should I control for variables?

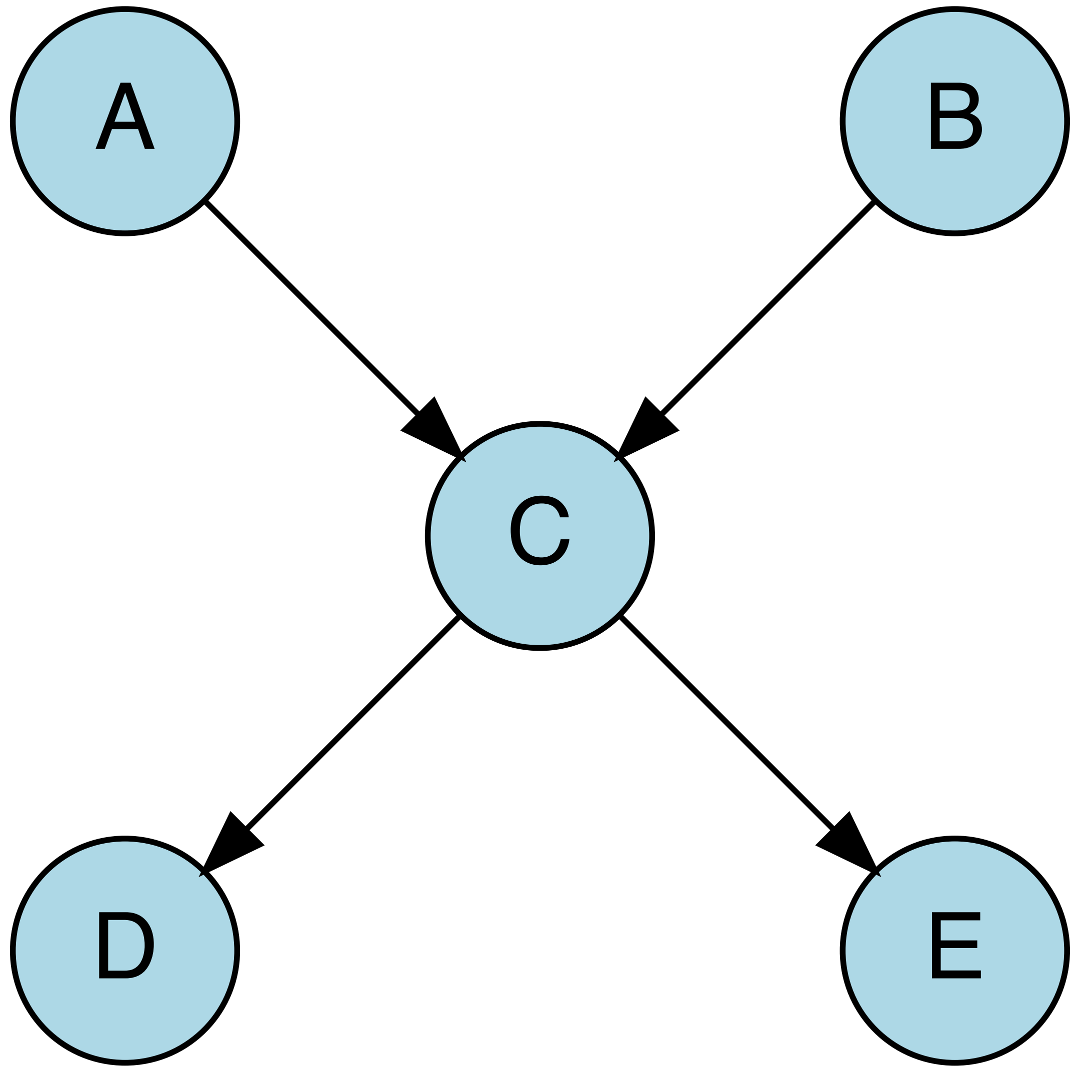
Are D and E independent, given C?

$$p(D \mid E, C) = p(D \mid C) \ ?$$



1. Draw the ancestral graph

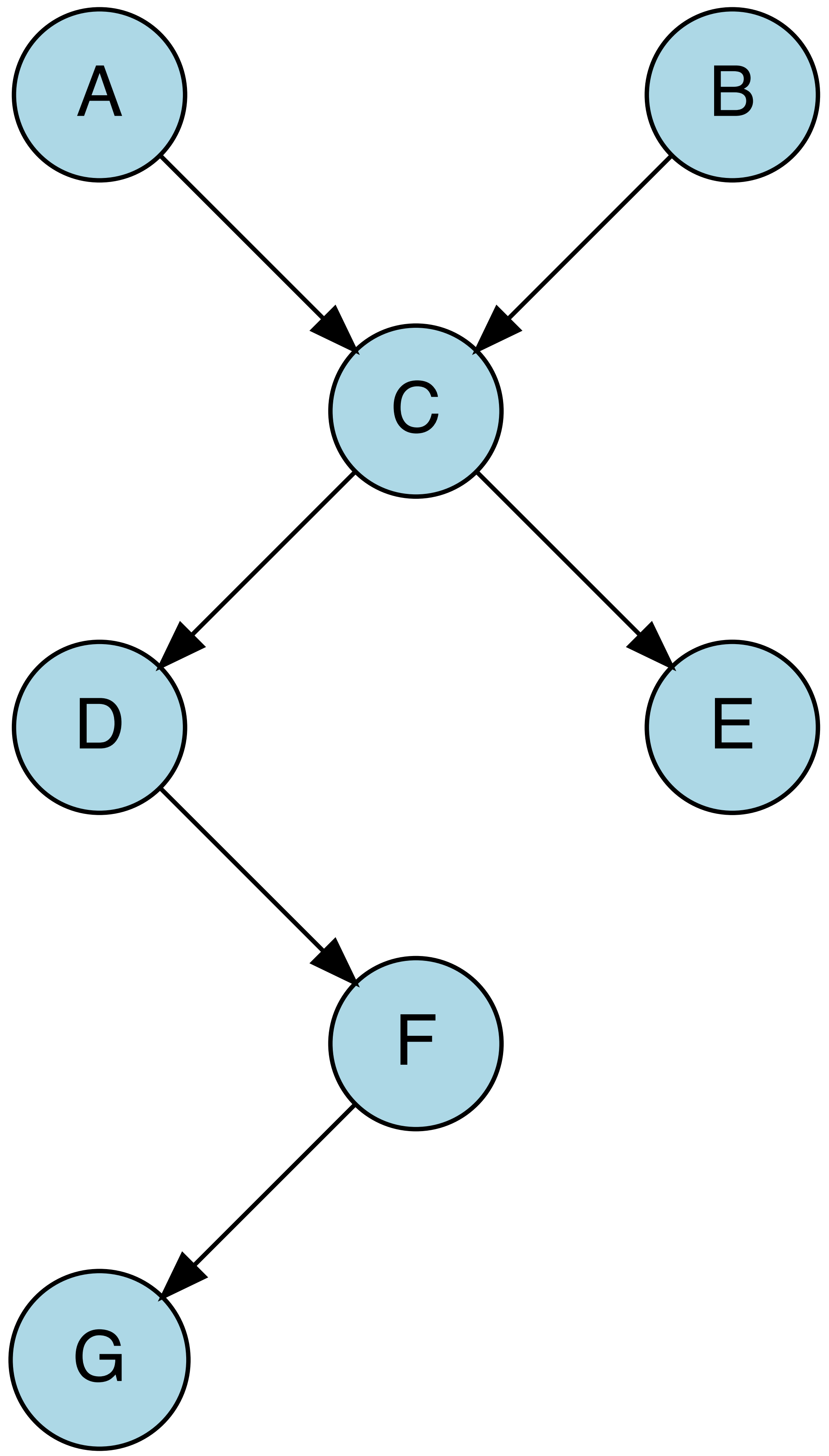
Construct the "ancestral graph" of all variables mentioned in the probability expression. This is a reduced version of the original net, consisting only of the variables mentioned and all of their ancestors (parents, parents' parents, etc.)



When should I control for variables?

Are D and E independent, given C?

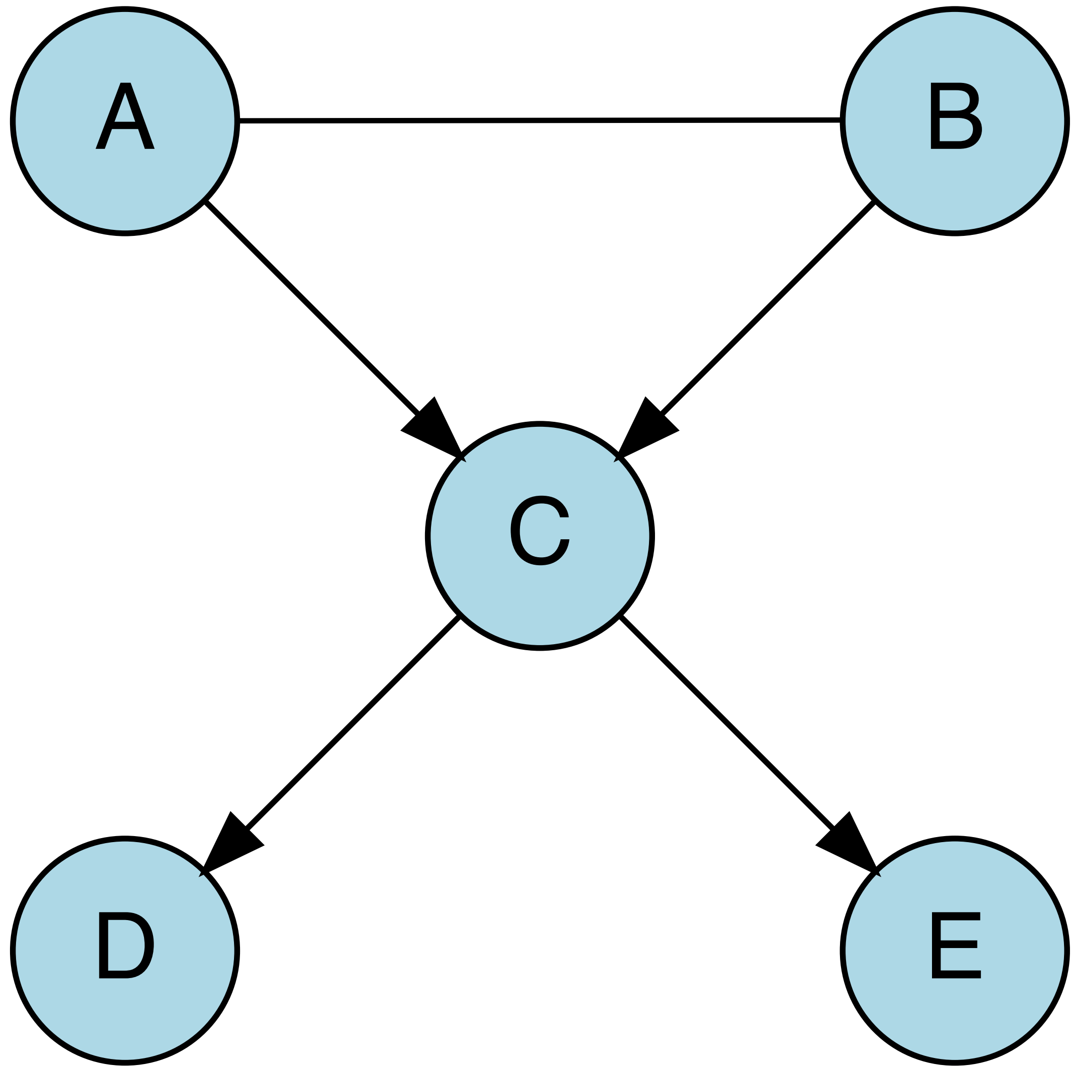
$$p(D \mid E, C) = p(D \mid C) \ ?$$



2. "Moralize" the graph

let's get married!

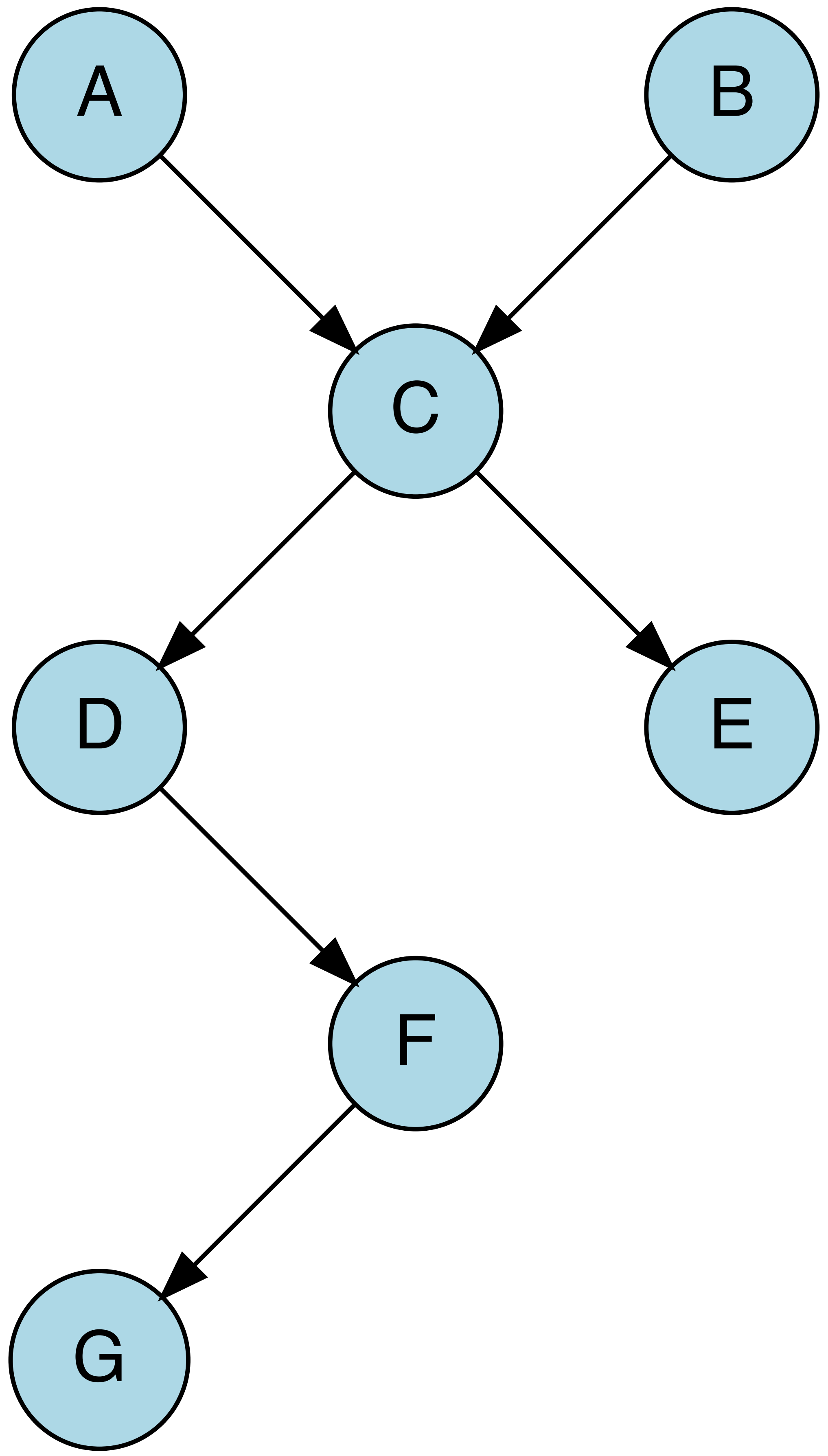
For each pair of variables with a common child, draw an undirected edge (line) between them. (If a variable has more than two parents, draw lines between every pair of parents.)



When should I control for variables?

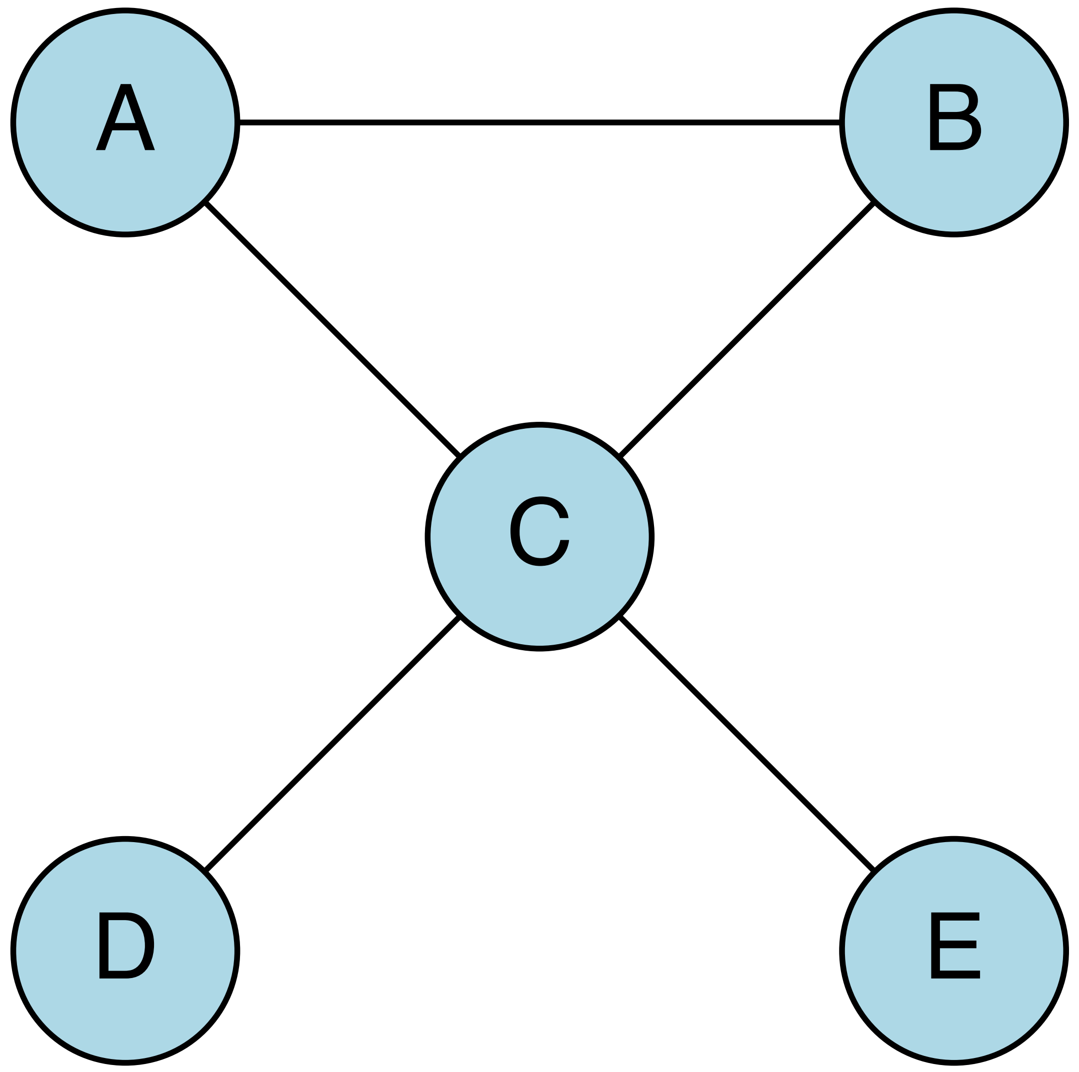
Are D and E independent, given C?

$$p(D \mid E, C) = p(D \mid C) \text{ ?}$$



3. "Disorient" the graph

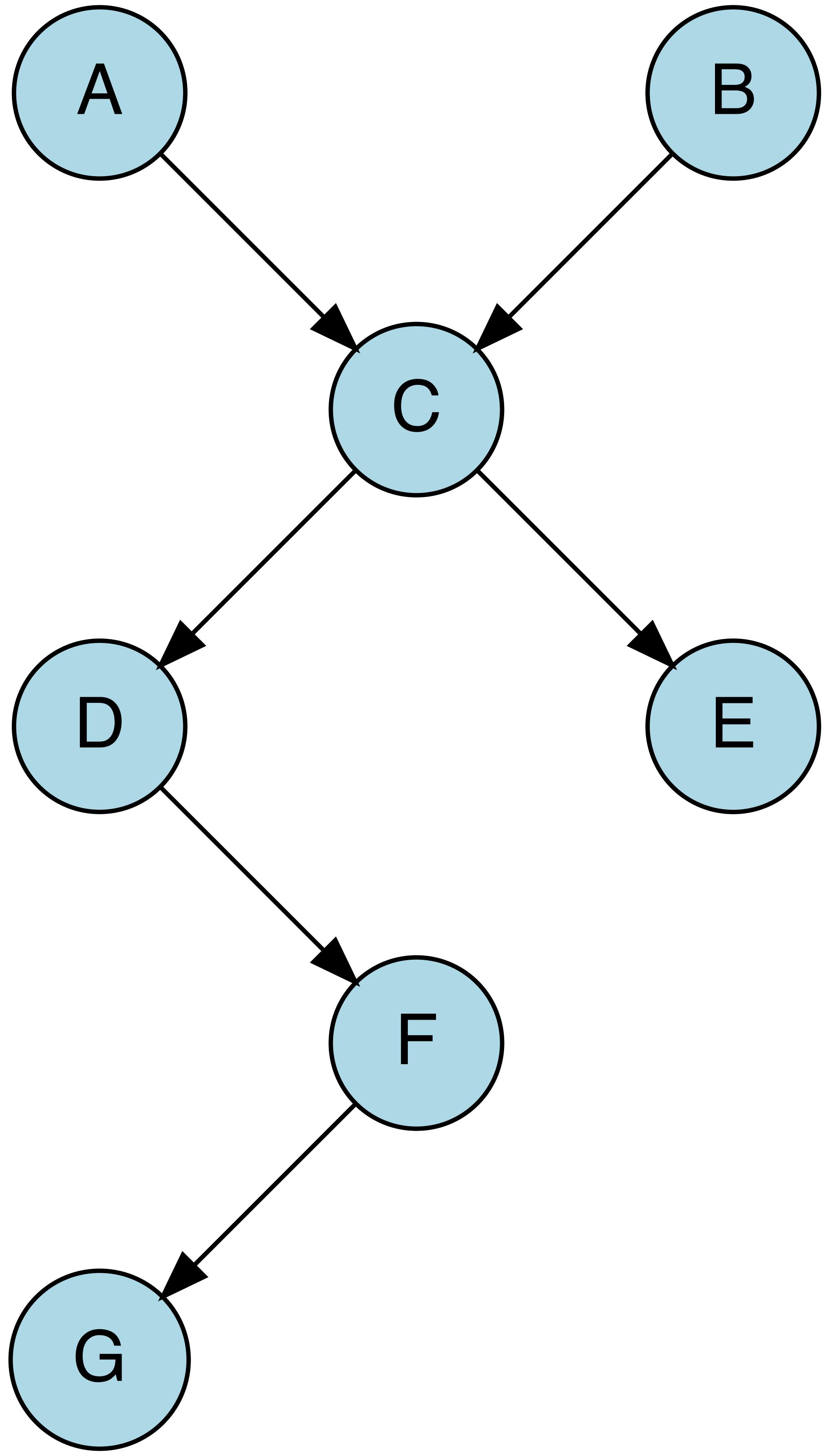
Replace arrows arrows with lines



When should I control for variables?

Are D and E independent, given C?

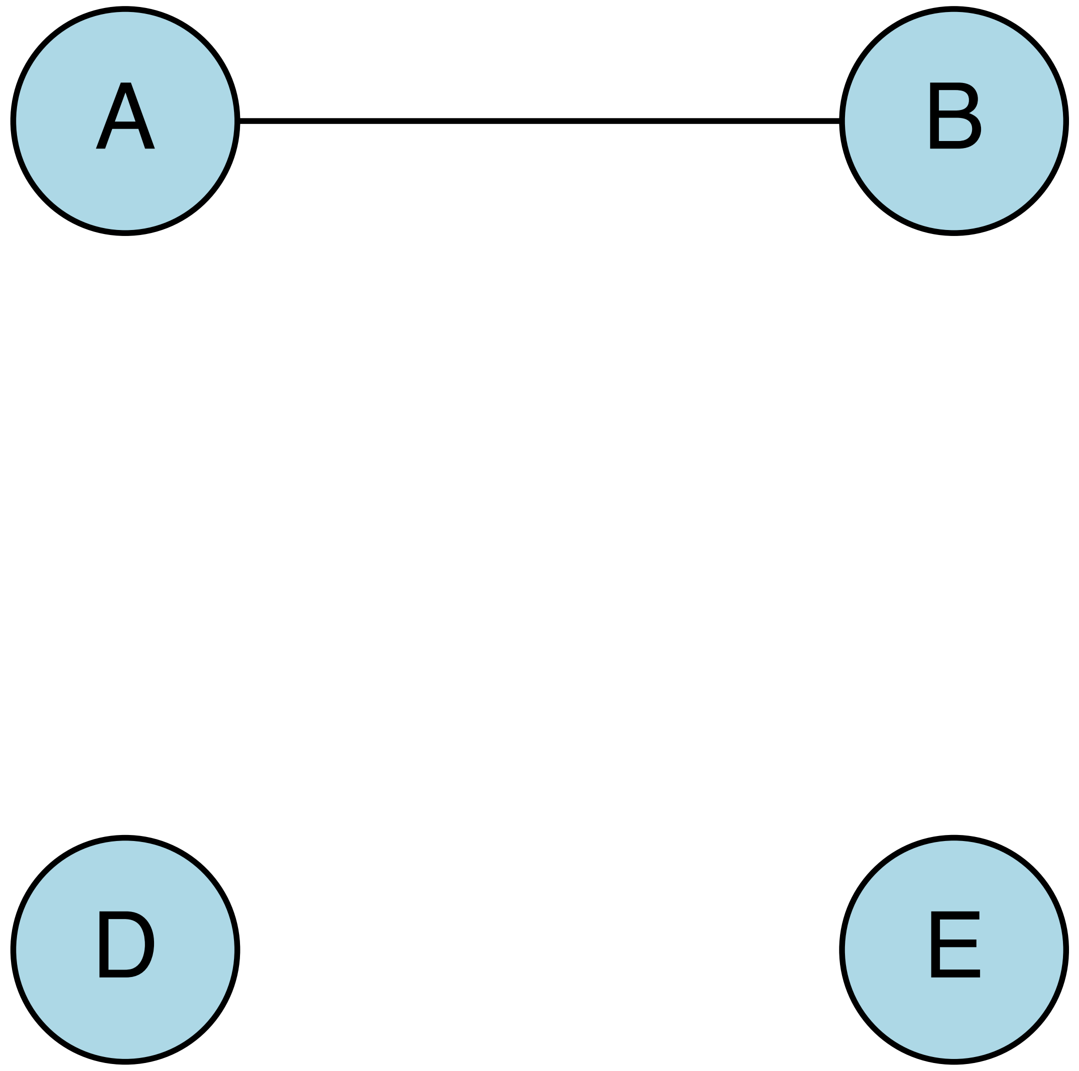
$$p(D \mid E, C) = p(D \mid C) \ ?$$



4. Delete the givens

Remove the variables that we condition on, as well as their edges

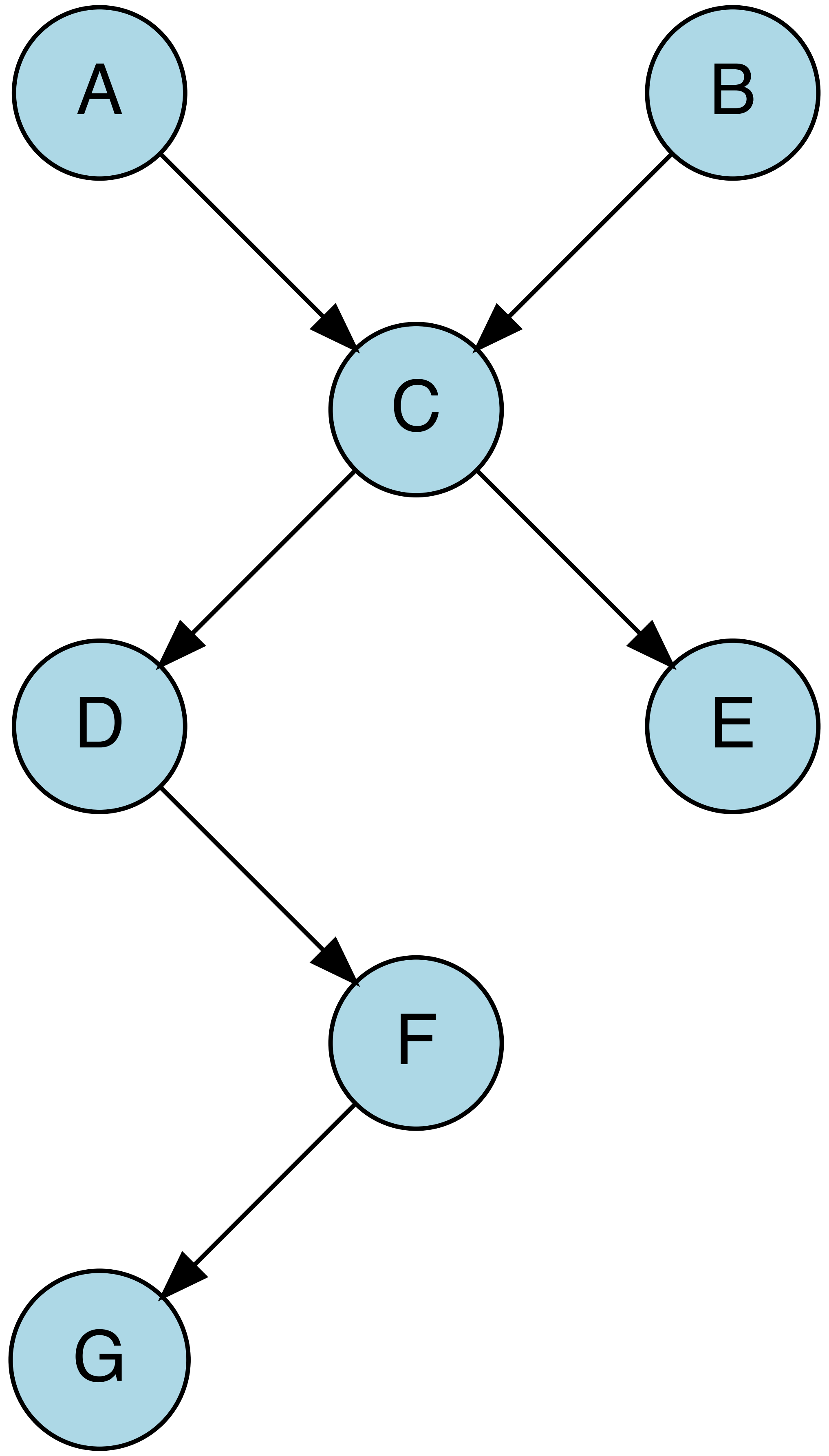
we conditioned on C!



When should I control for variables?

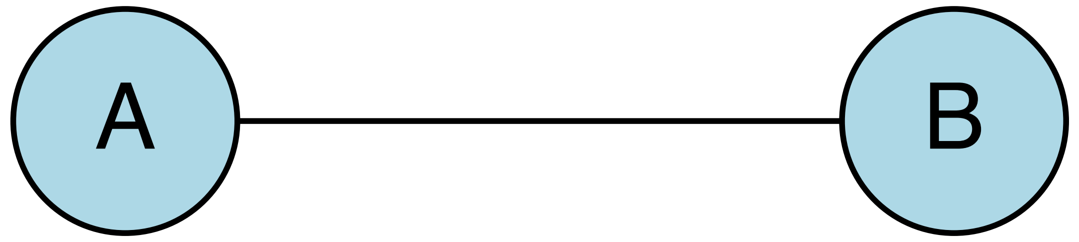
Are D and E independent, given C?

$$p(D \mid E, C) = p(D \mid C) \; ?$$



5. Read answer off the graph

- if variables are **disconnected** they are independent
- if variables are connected (have a path between them) they are not guaranteed to be independent



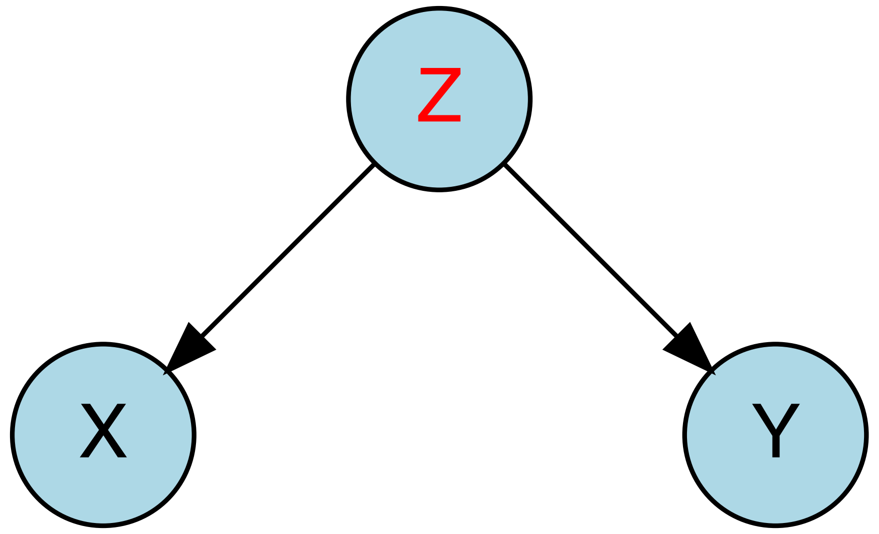
D and E are independent from each other conditioned on C

they aren't connected via a path

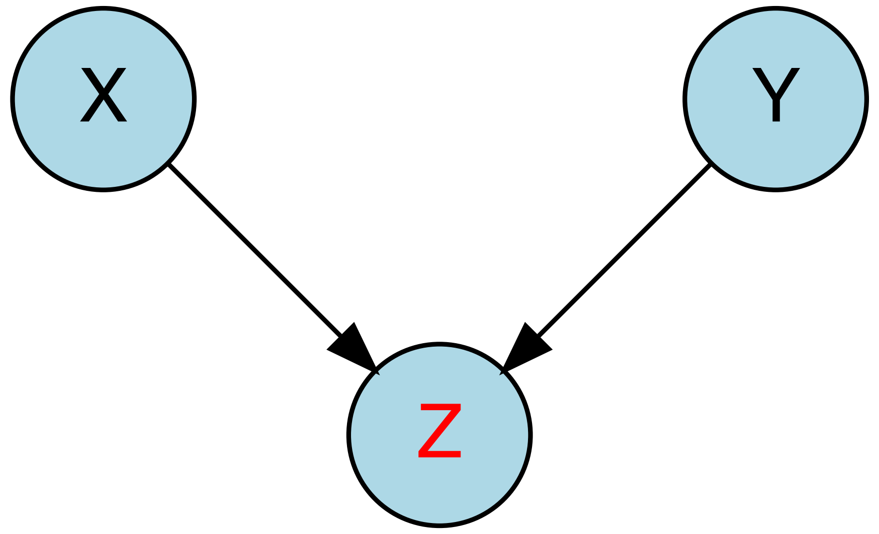
So what?

Patterns of inference

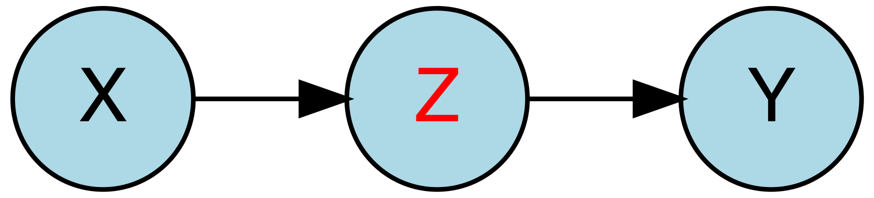
We want to estimate the (causal) relationship between `X` and `Y`



common cause



common effect



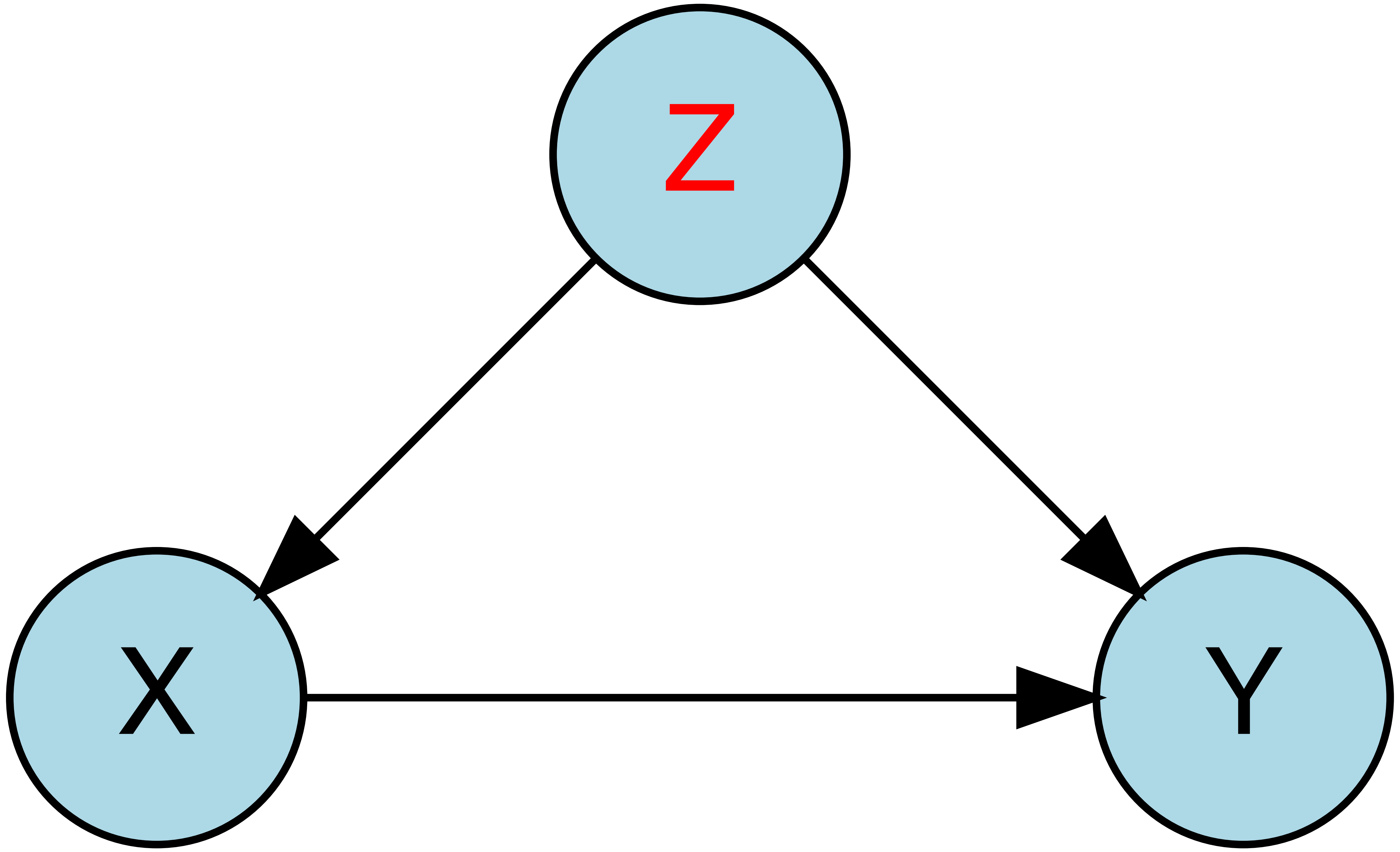
causal chain

by controlling for `Z` we hope to get a better estimate of the relationship between `X` and `Y`

d-separation helps us tell apart **good controls** from **bad controls**

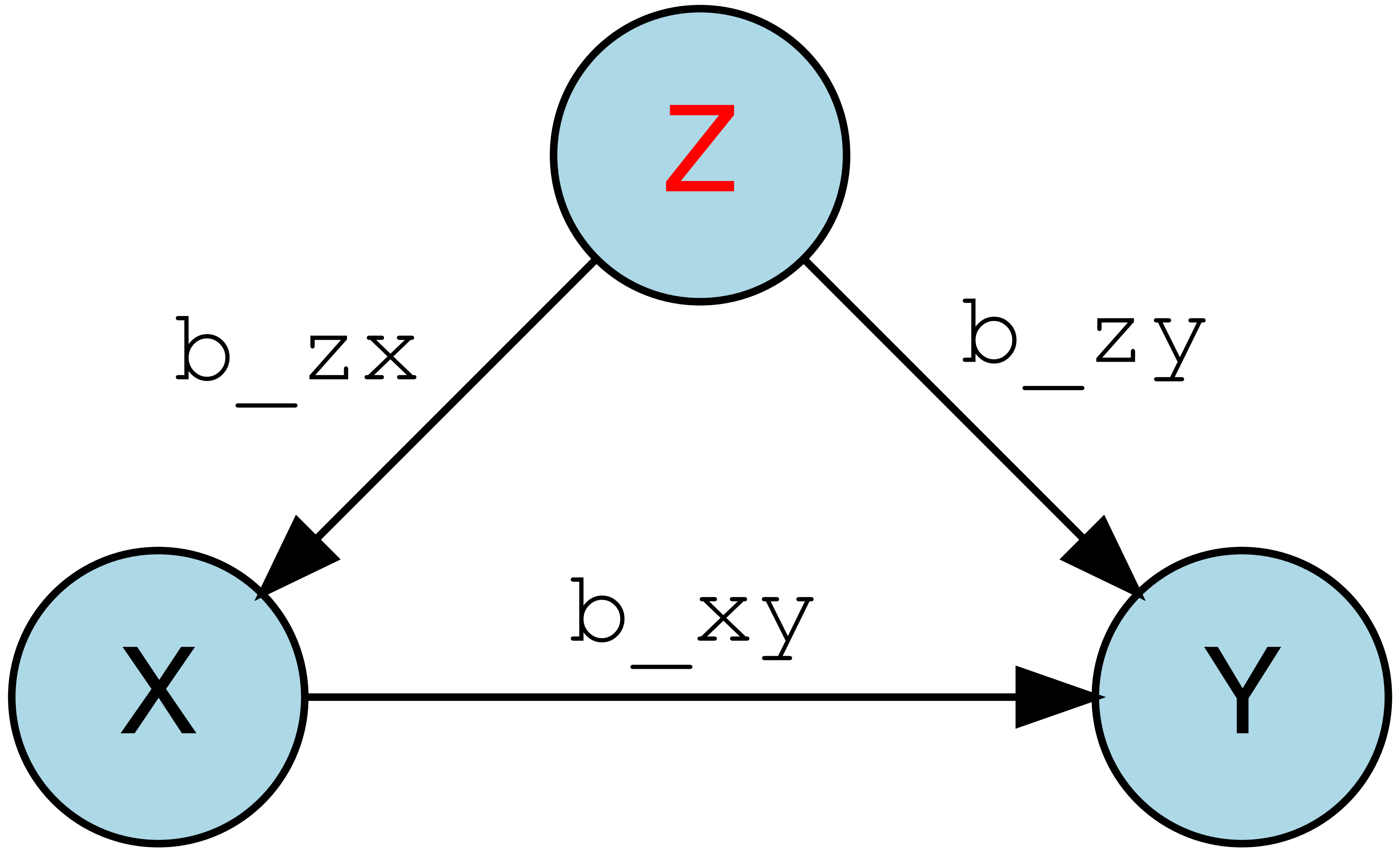
When should I control for variables?

I want to estimate the effect that X has on Y



Is Z a **good** or a **bad** control here?

When should I control for variables?




```r
set.seed(1)

n = 1000
b_zx = 2
b_xy = 2
b_zy = 2
sd = 1

df = tibble(z = rnorm(n = n, sd = sd),
            x = b_zx * z + rnorm(n = n, sd = sd),
            y = b_zy * z + b_xy * x + rnorm(n = n, sd = sd))
```

overestimating X's effect on Y

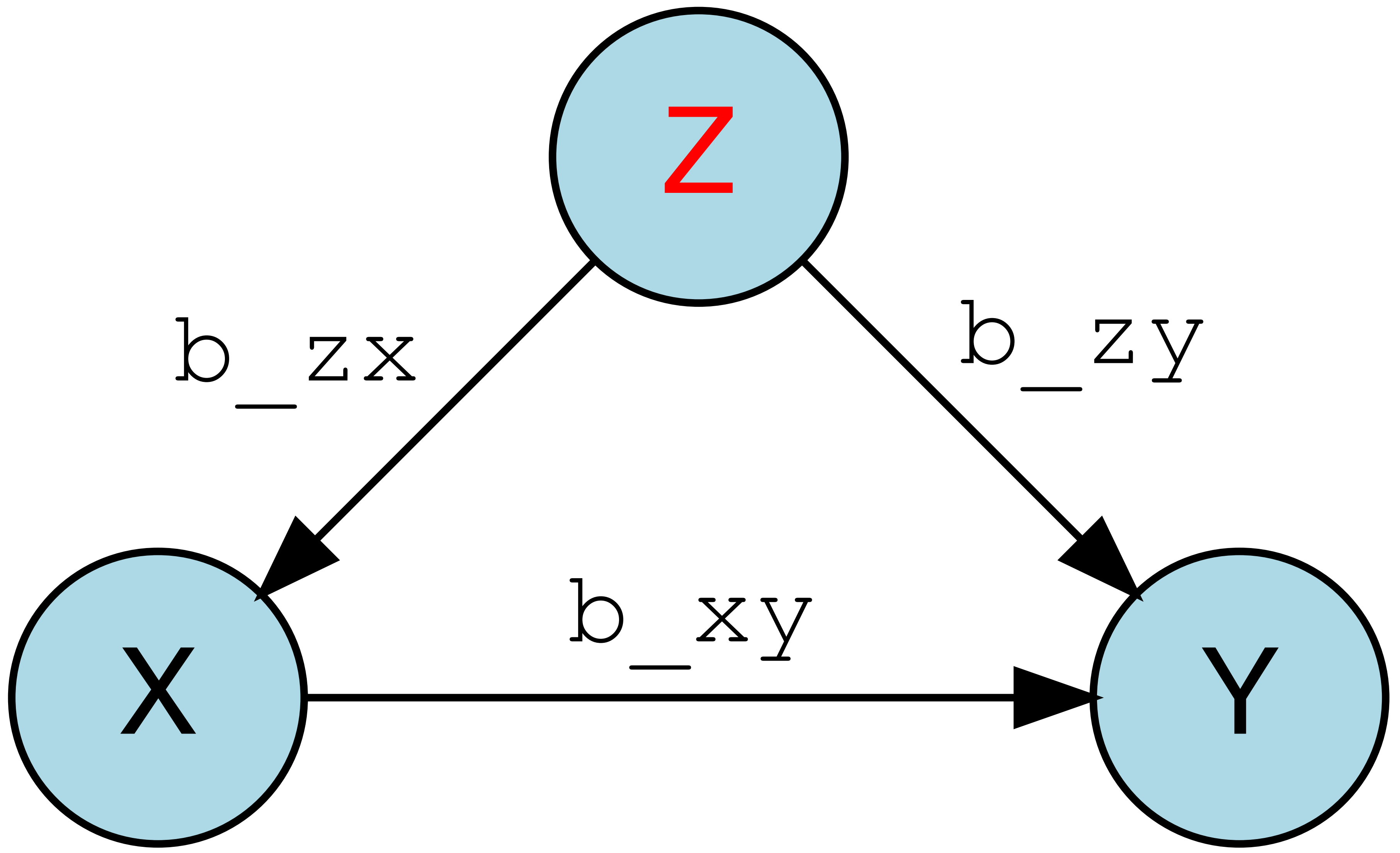
$$Y = b_0 + b_1 \cdot X + e$$

```r
# without control
lm(formula = y ~ 1 + x,
   data = df) %>%
  summary()
```

```
Call:
lm(formula = y ~ x, data = df)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.6011 -0.9270 -0.0506  0.9711  4.0454 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.02449    0.04389   0.558    0.577    
x            2.82092    0.01890 149.225   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.388 on 998 degrees of freedom
Multiple R-squared:  0.9571,	Adjusted R-squared:  0.9571 
F-statistic: 2.227e+04 on 1 and 998 DF,  p-value: < 2.2e-16
```

When should I control for variables?

```r
set.seed(1)

n = 1000
b_zx = 2
b_xy = 2
b_zy = 2
sd = 1

df = tibble(z = rnorm(n = n, sd = sd),
            x = b_zx * z + rnorm(n = n, sd = sd),
            y = b_zy * z + b_xy * x + rnorm(n = n, sd = sd))
```



$$Y = b_0 + b_1 \cdot X + b_2 \cdot Z + e$$

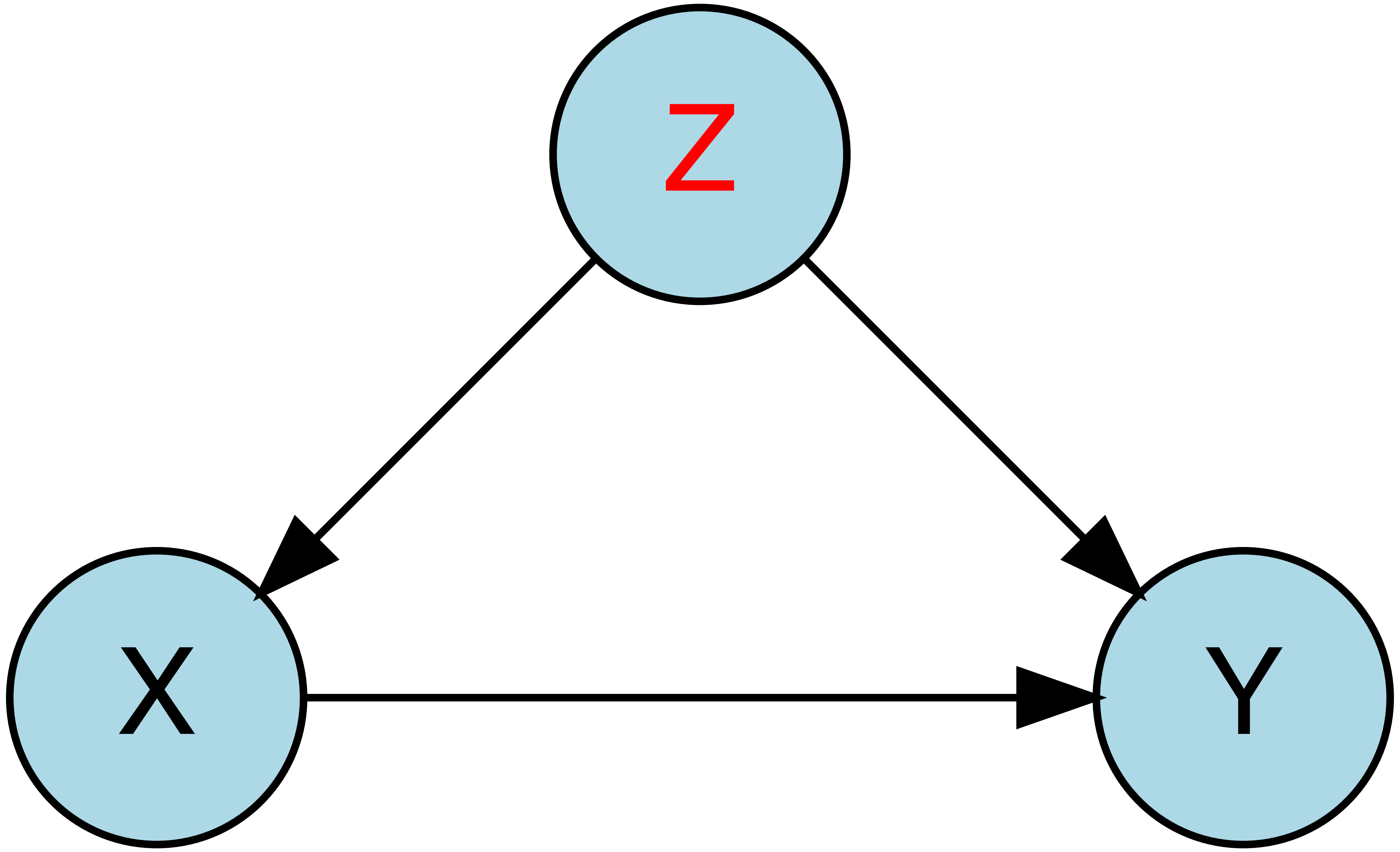
```r
# with control
lm(formula = y ~ 1 + x + z,
   data = df) %>%
  summary()
```

accurate estimate of X's effect on Y

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

Residuals:
    Min      1Q  Median      3Q     Max
-3.6151 -0.6564 -0.0223  0.6815  2.8132

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.01624    0.03260   0.498    0.618
x            2.02202    0.03135  64.489   <2e-16 ***
z            2.00501    0.07036  28.497   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.031 on 997 degrees of freedom
Multiple R-squared:  0.9764,	Adjusted R-squared:  0.9763
F-statistic: 2.059e+04 on 2 and 997 DF,  p-value: < 2.2e-16
```

When should I control for variables?

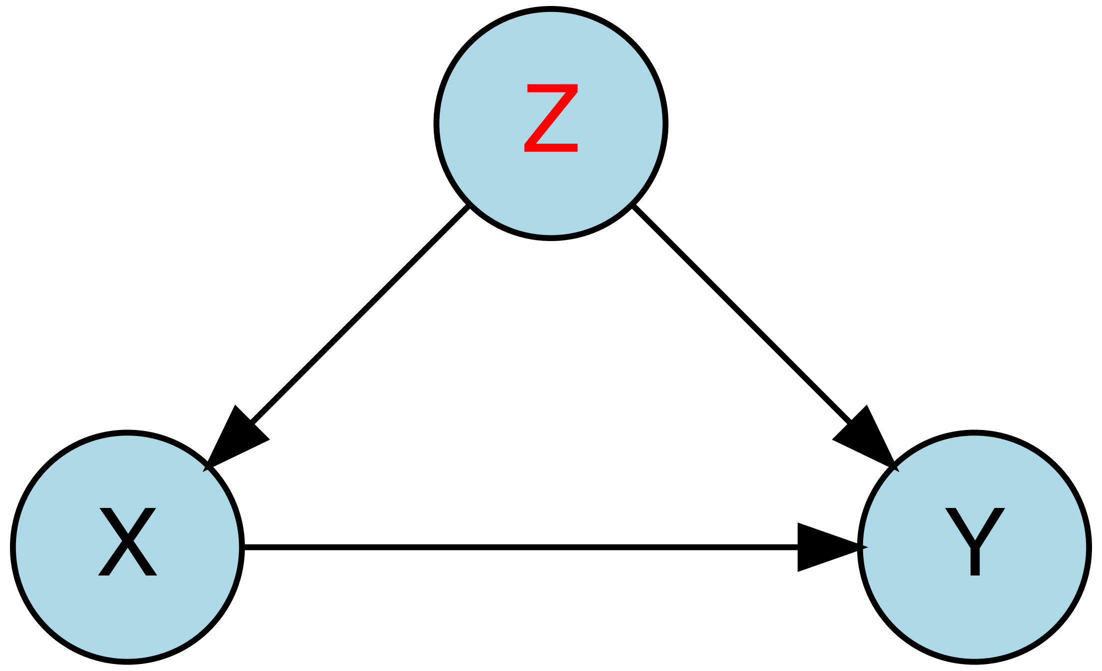
I want to estimate the effect that X has on Y

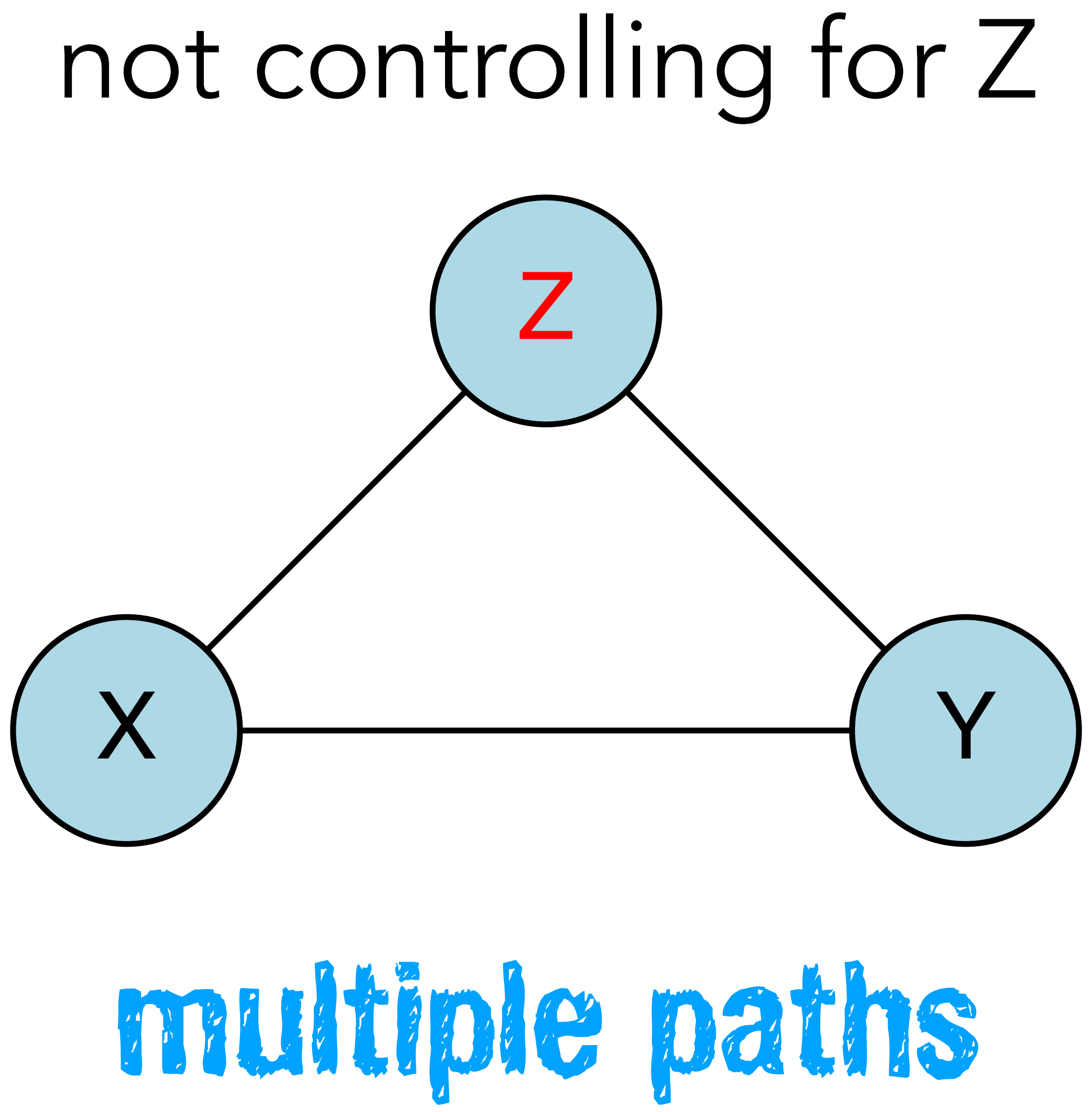


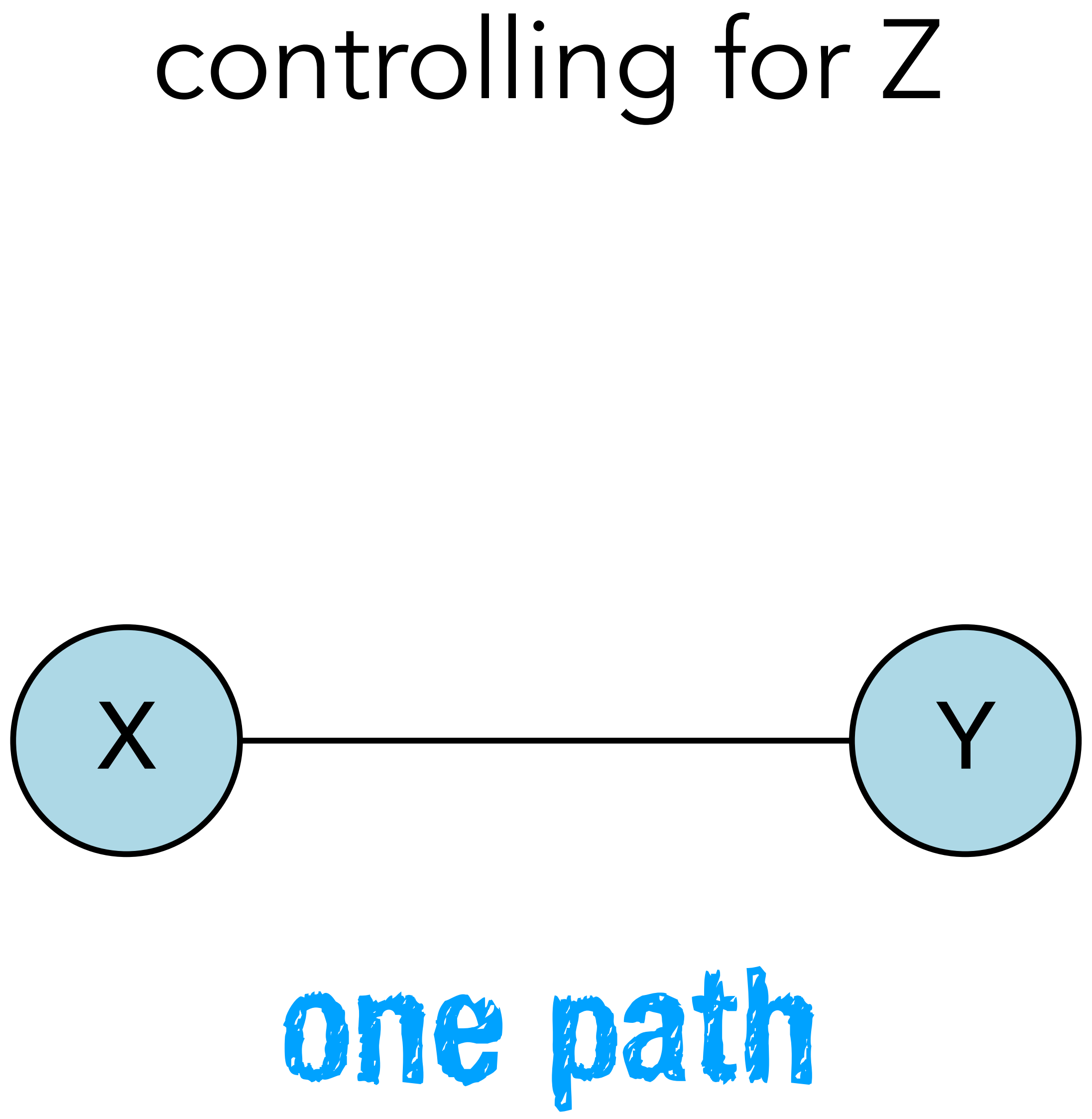
Z is a **good** control here!

When should I control for variables?

I want to estimate the effect that X has on Y



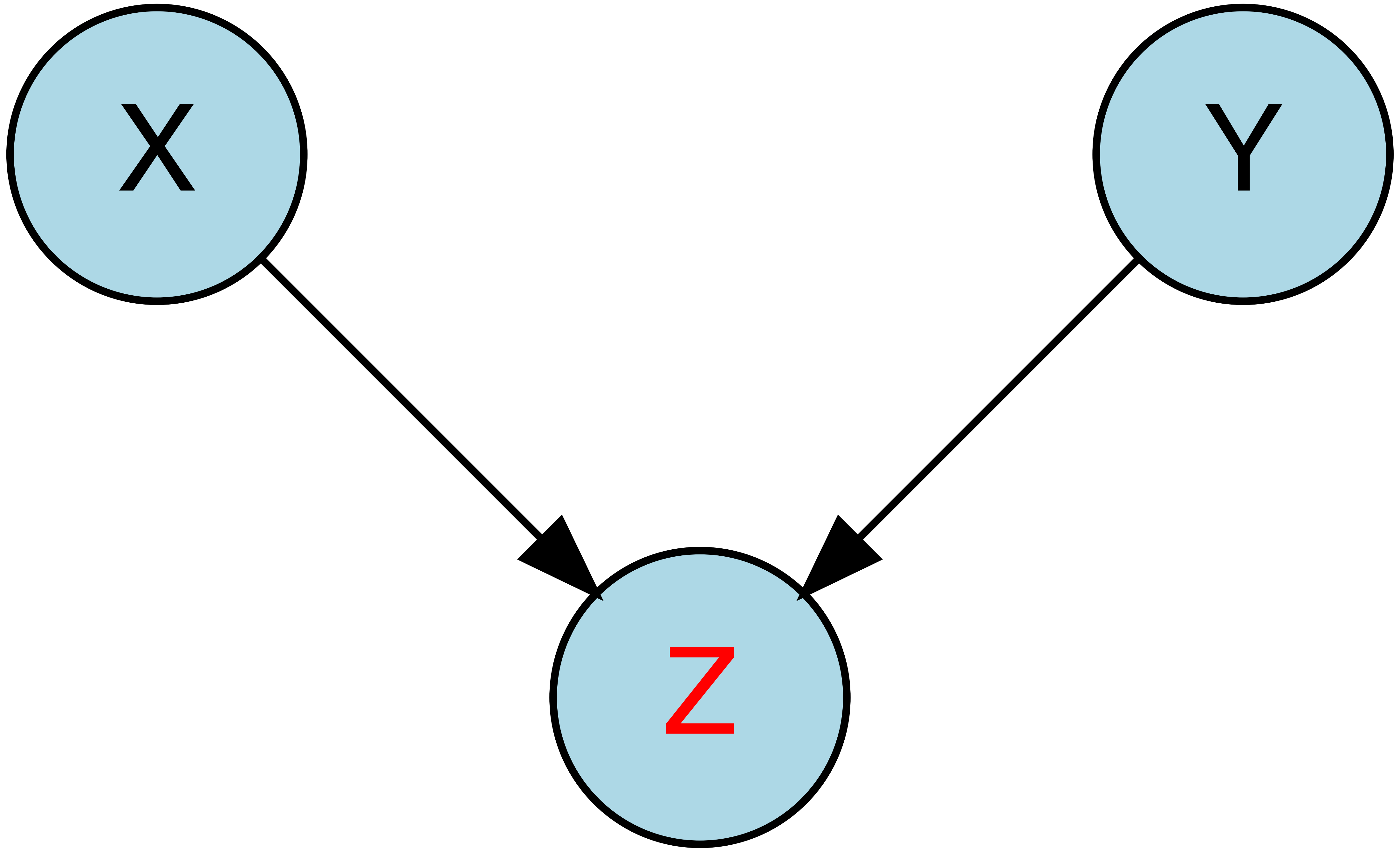




Z is a **good** control here!

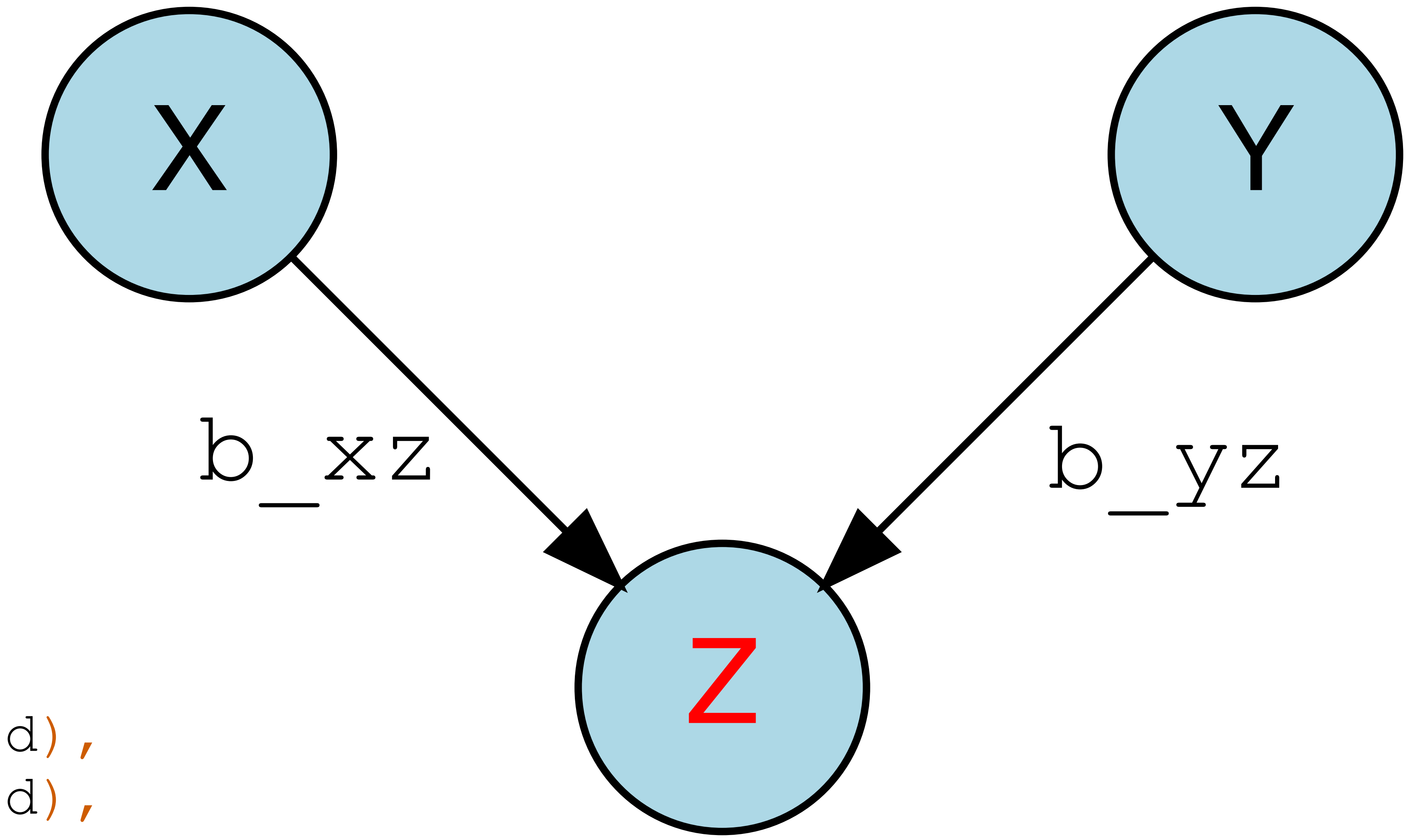
When should I control for variables?

I want to estimate the effect that X has on Y



Is Z a **good** or a **bad** control here?

When should I control for variables?




```r
set.seed(1)

n = 1000
b_xz = 2
b_yz = 2
sd = 1

df = tibble(x = rnorm(n = n, sd = sd),
            y = rnorm(n = n, sd = sd),
            z = x * b_xz + y * b_yz + rnorm(n = n, sd = sd))
```

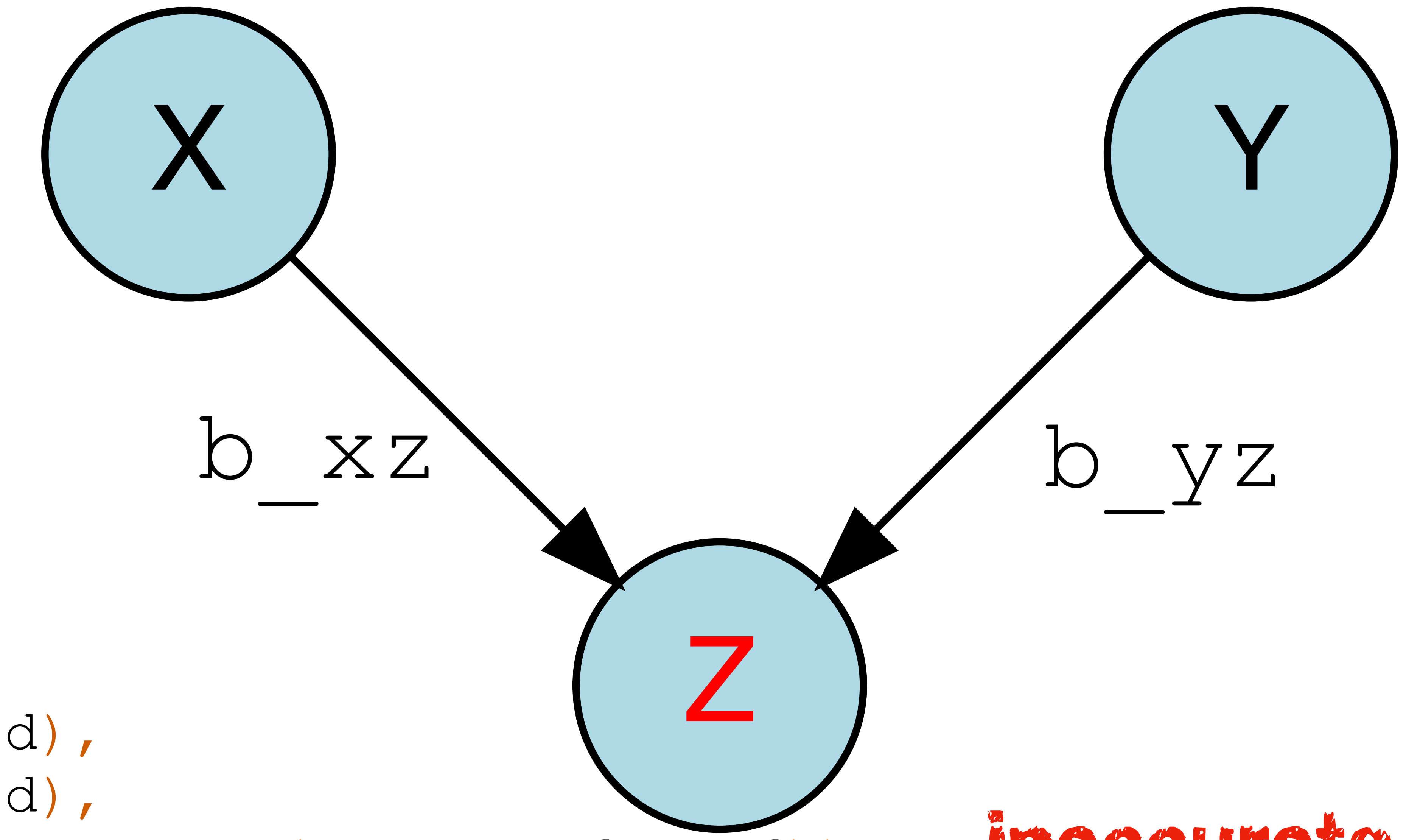
accurate estimate
of X's effect on Y

$$Y = b_0 + b_1 \cdot X + e$$

```r
# without control
lm(formula = y ~ 1 + x,
   data = df) %>%
  summary()
```

```
Call:
lm(formula = y ~ x, data = df)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.2484 -0.6720 -0.0138  0.7554  3.6443 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.016187   0.032905  -0.492    0.623
x            0.006433   0.031809   0.202    0.840

Residual standard error: 1.04 on 998 degrees of freedom
Multiple R-squared:  4.098e-05,	Adjusted R-squared:  -0.000961 
F-statistic: 0.0409 on 1 and 998 DF,  p-value: 0.8398
```

When should I control for variables?




```r
set.seed(1)

n = 1000
b_xz = 2
b_yz = 2
sd = 1

df = tibble(x = rnorm(n = n, sd = sd),
            y = rnorm(n = n, sd = sd),
            z = x * b_xz + y * b_yz + rnorm(n = n, sd = sd))
```

inaccurate estimate of X's effect on Y

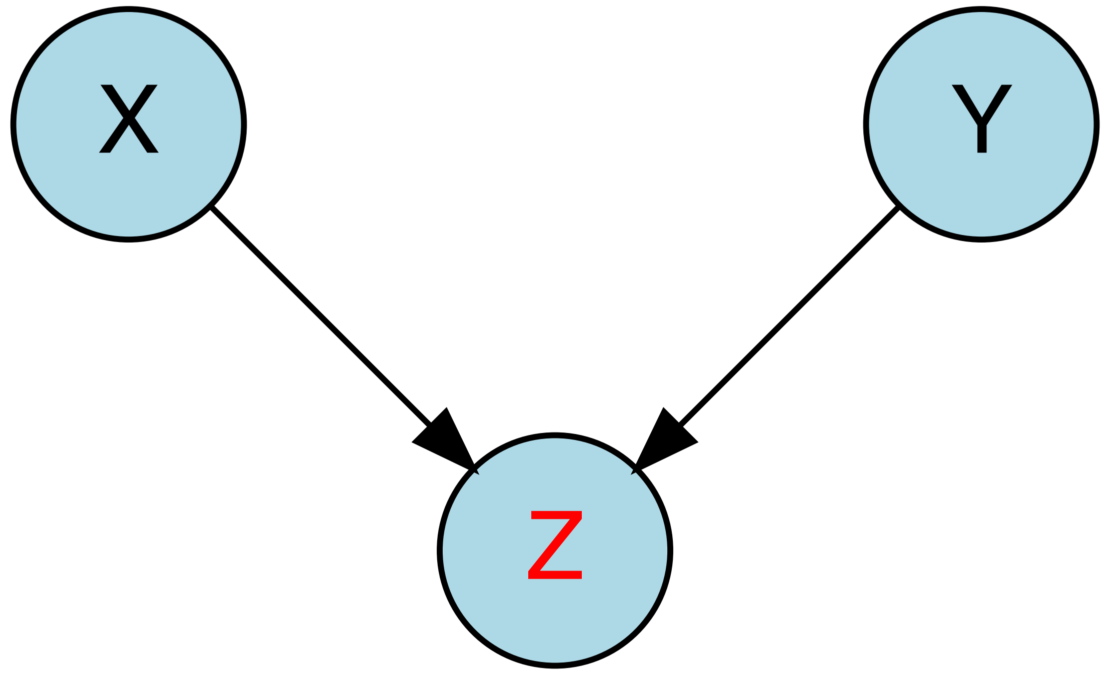
$$Y = b_0 + b_1 \cdot X + b_2 \cdot Z + e$$

```r
# with control
lm(formula = y ~ 1 + x + z,
   data = df) %>%
  summary()
```

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

Residuals:
     Min       1Q   Median       3Q      Max 
-1.35547 -0.30016  0.00298  0.31119  1.73408 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.009608   0.014477  -0.664    0.507    
x           -0.816164   0.018936 -43.102   <2e-16 ***
z            0.398921   0.006186  64.489   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4578 on 997 degrees of freedom
Multiple R-squared:  0.8066,	Adjusted R-squared:  0.8062 
F-statistic:  2079 on 2 and 997 DF,  p-value: < 2.2e-16
```

When should I control for variables?

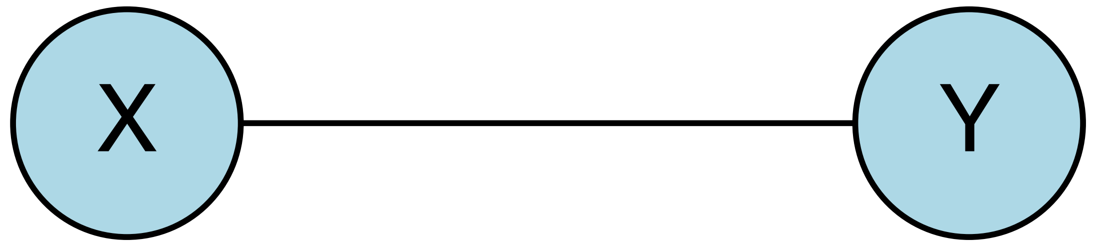
I want to estimate the effect that X has on Y



not controlling for Z

no path

controlling for Z



one path

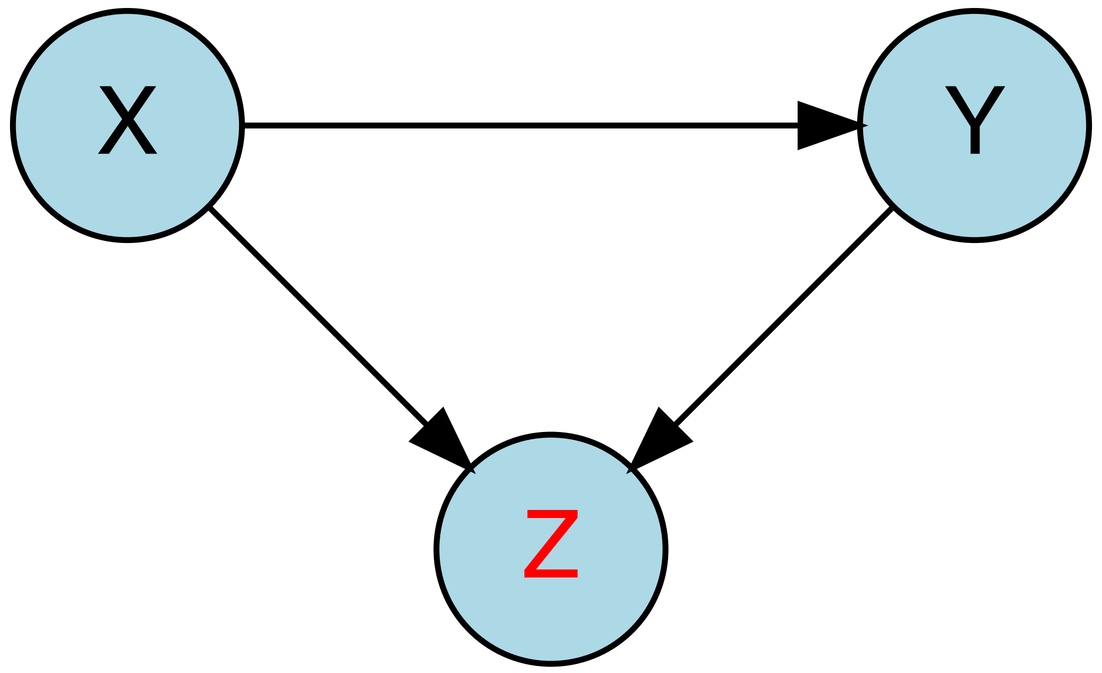
Z is a **bad** control here!

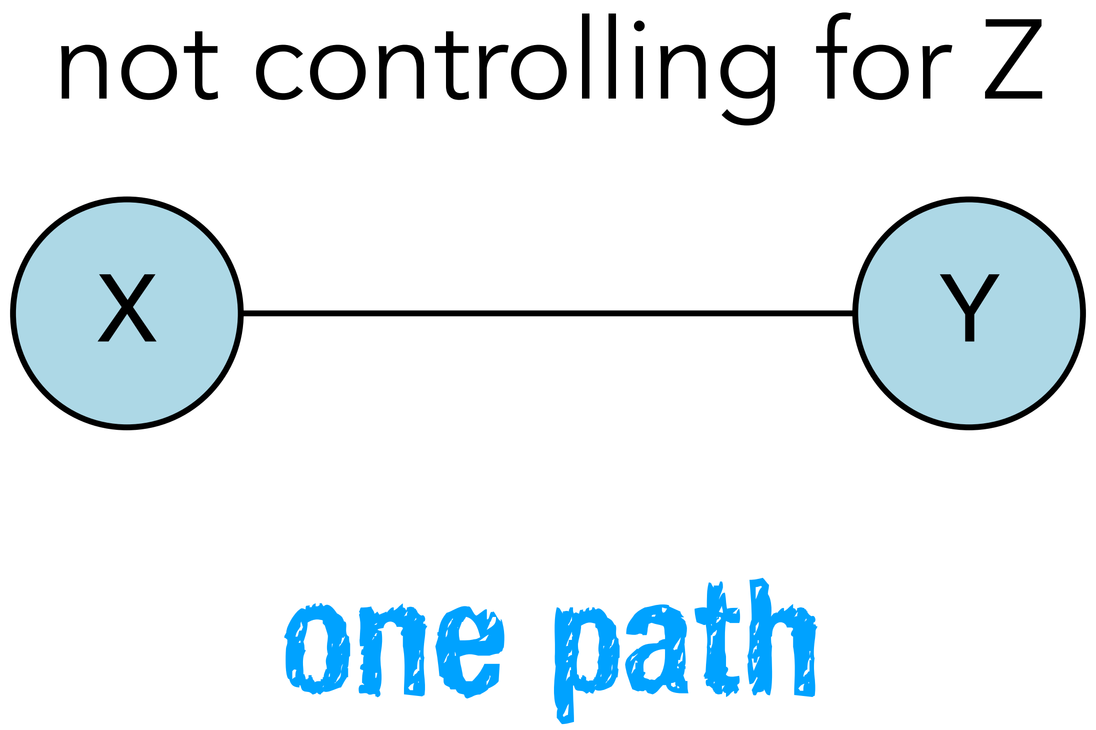
When should I control for variables?

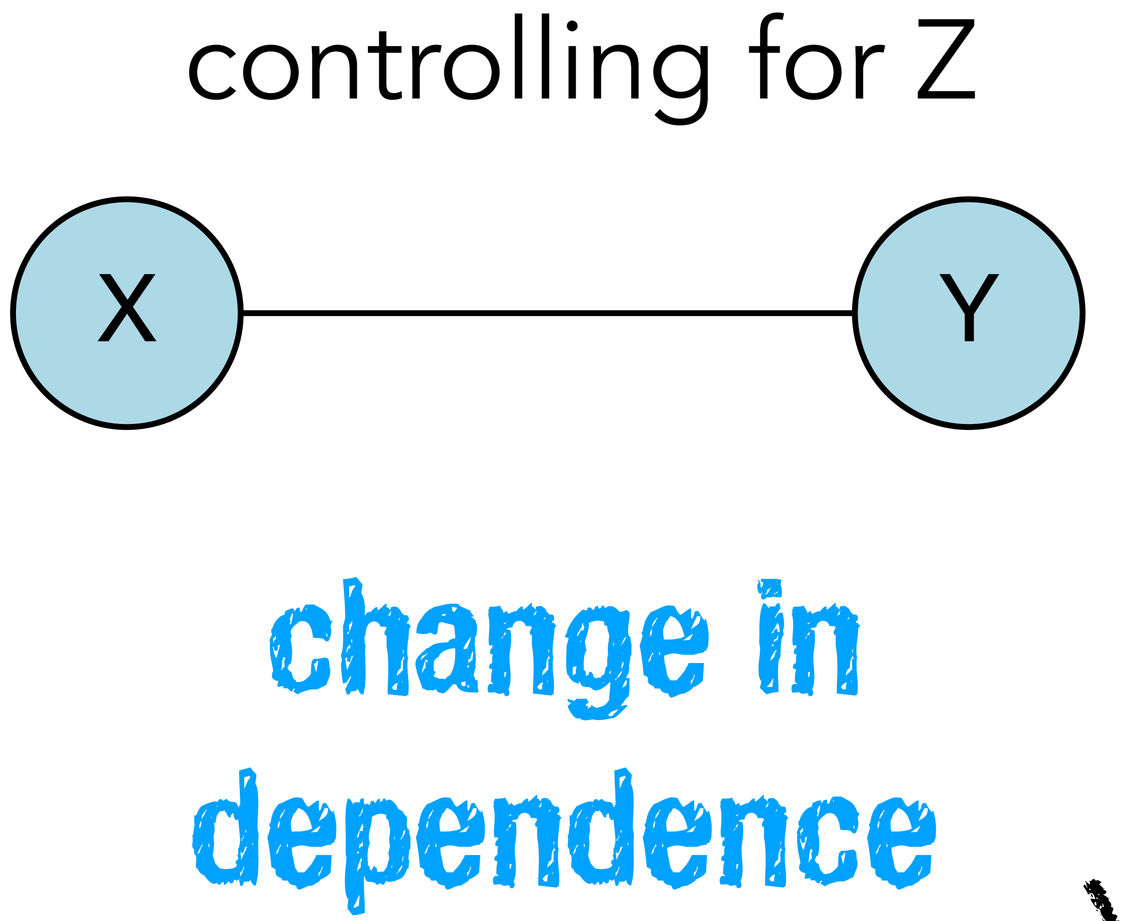
- checking for **d-separation** tells us whether or not variables are (conditionally) independent
- it also tells us whether paths of dependence "open up", or get "closed down"
- the graphical procedure doesn't necessarily reveal whether the dependence between variables changes: it reveals the **structure** of dependence but not the **strength**
- you can always double check via running simulations in R

When should I control for variables?

I want to estimate the effect that X has on Y







not revealed in the undirected graph

Z is a **bad** control here!

When should I control for variables?

- **good controls** reduce additional paths from `X` to `Y` apart from the direct path we are interested in estimating
- **bad controls** introduce additional paths (or change existing ones) that lead to a biased estimate of the direct path between `X` and `Y`

When should I control for variables?

Problem: We don't know the ground truth ...

$X \rightarrow Z \rightarrow Y$ $\qquad$ $X \leftarrow Z \rightarrow Y$ $\qquad$ $X \leftarrow Z \leftarrow Y$

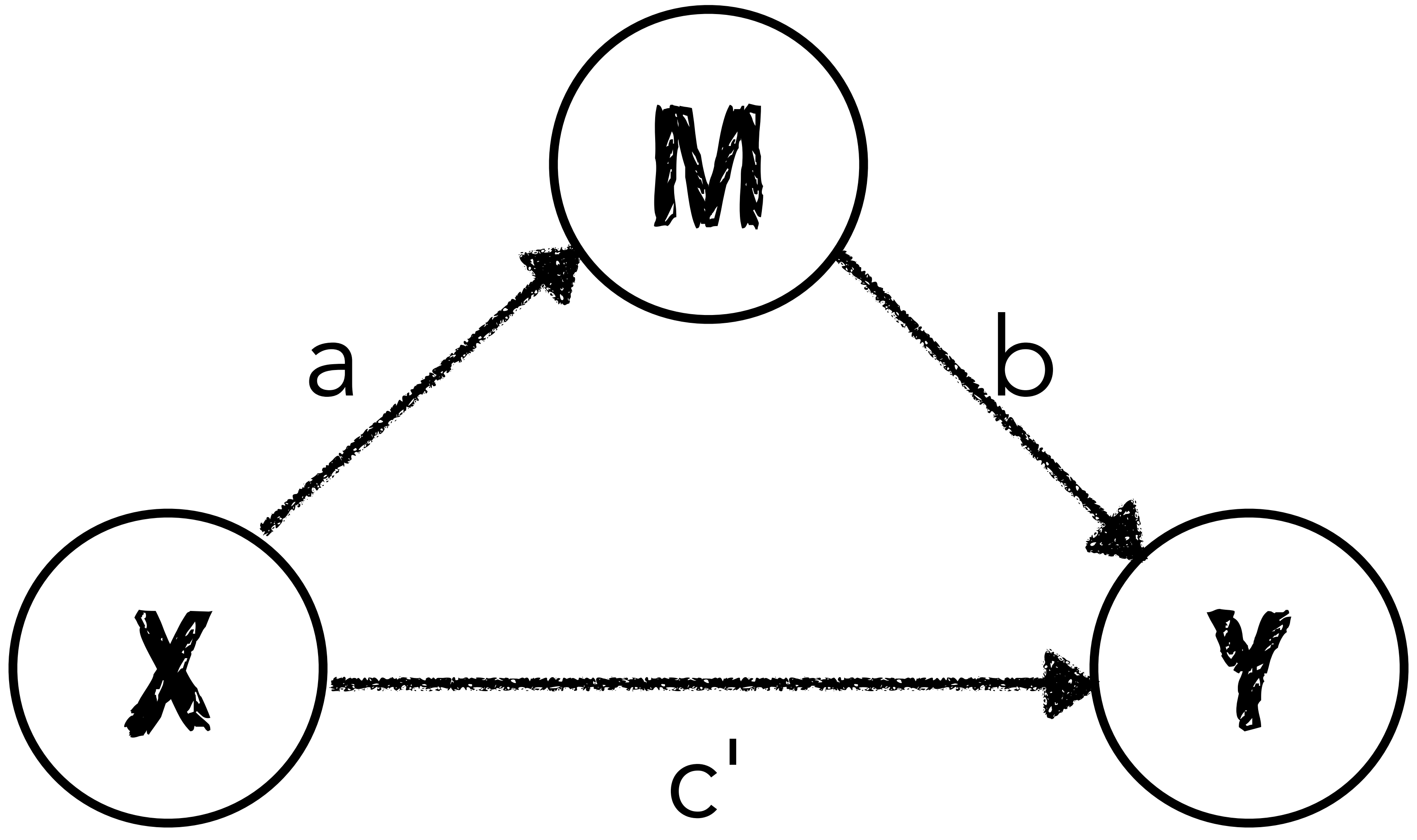
controlling for Z

$X \qquad Y$

we need to manipulate X experimentally to tell these apart

Mediation

Definition

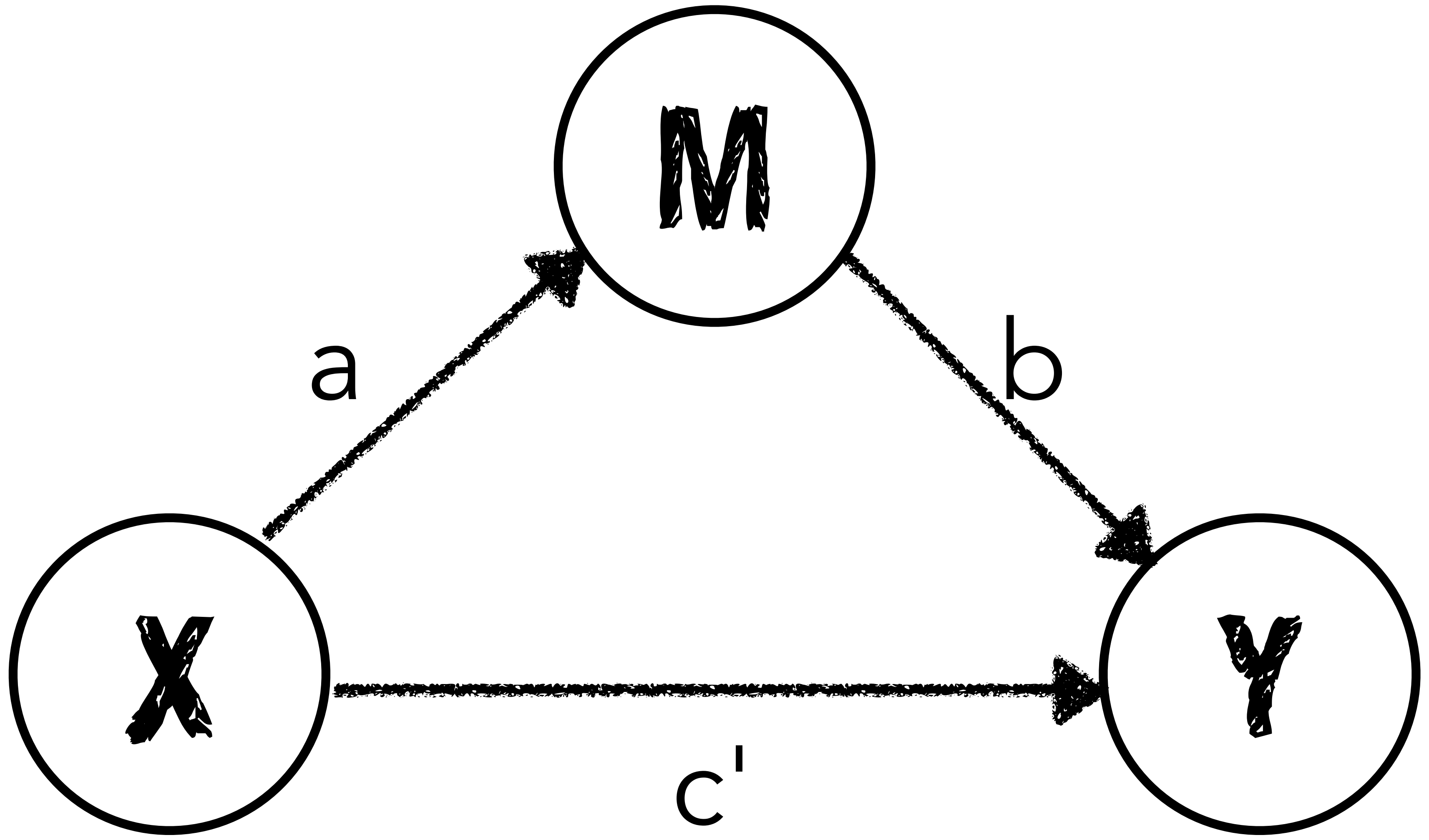


Rather than a direct causal relationship between **X** and **Y**, a mediation model proposes that **X** influences the mediator variable **M**, which in turn influences **Y**. Thus, the mediator variable serves to clarify the nature of the relationship between **X** and **Y**.

Adapted from Wikipedia

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

Example



X = grades in Psych 252

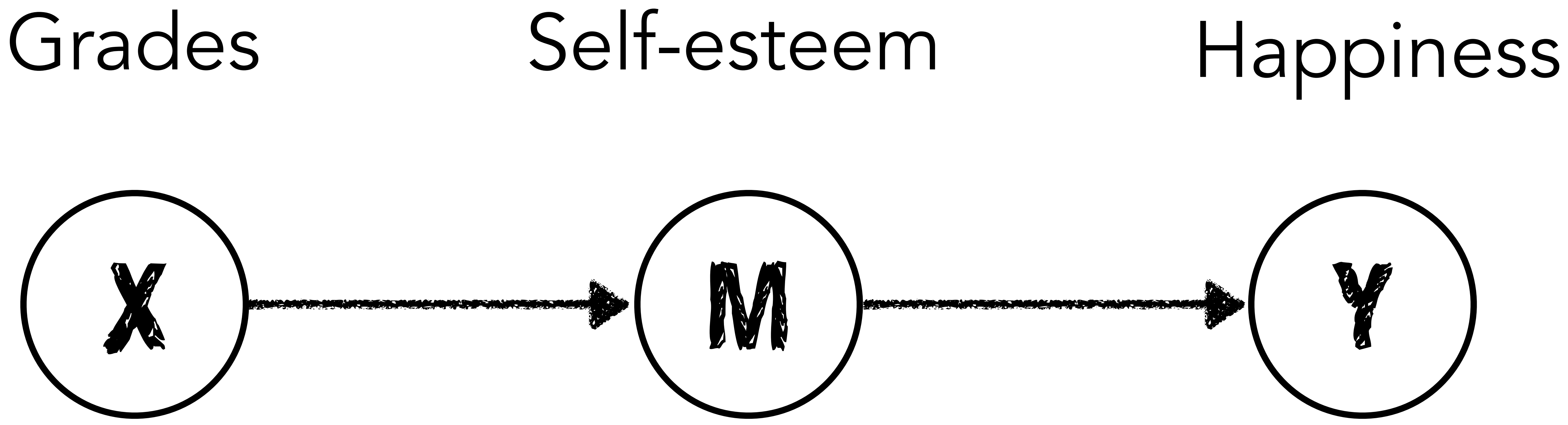
M = feelings of self-esteem

Y = happiness

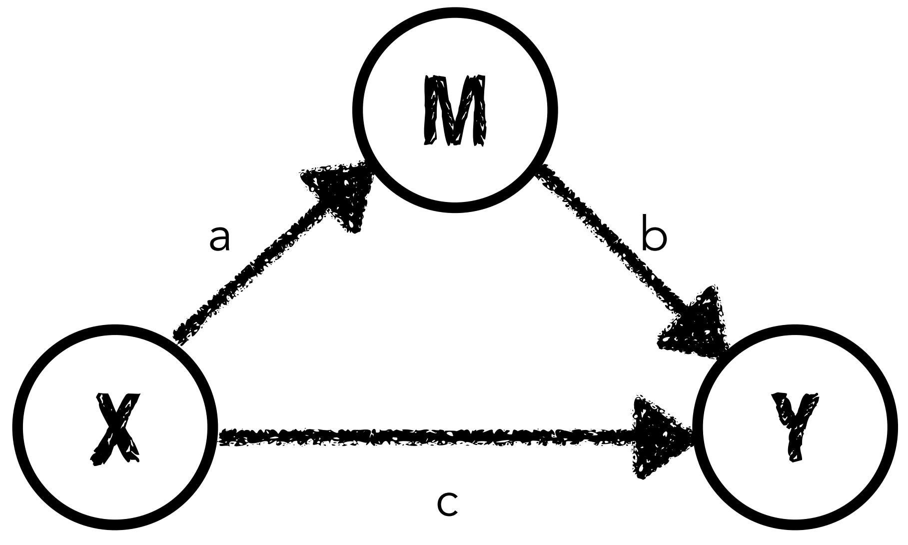
Is the relationship between grades in Psych 252 and happiness mediated by feelings of self-esteem?

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
)
```



Bootstrapping



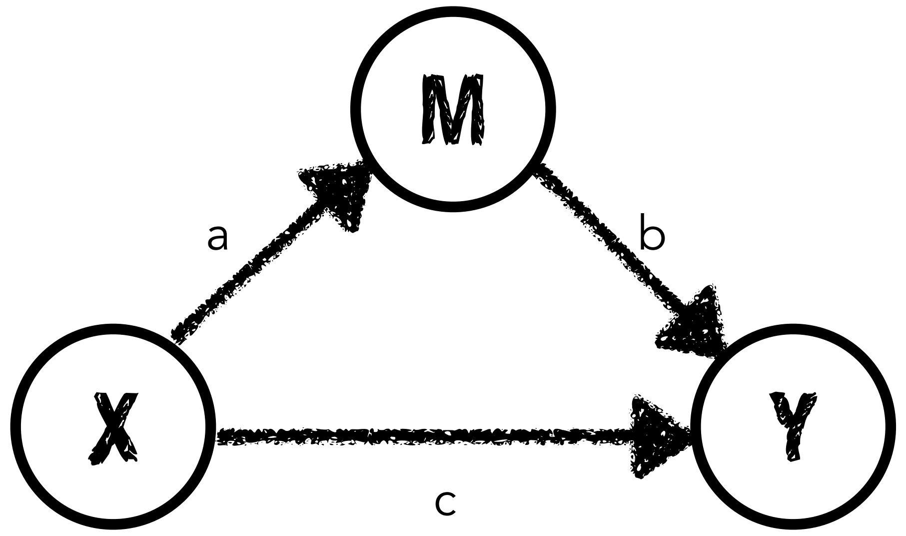

```r
library("mediation")

# 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()
```

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

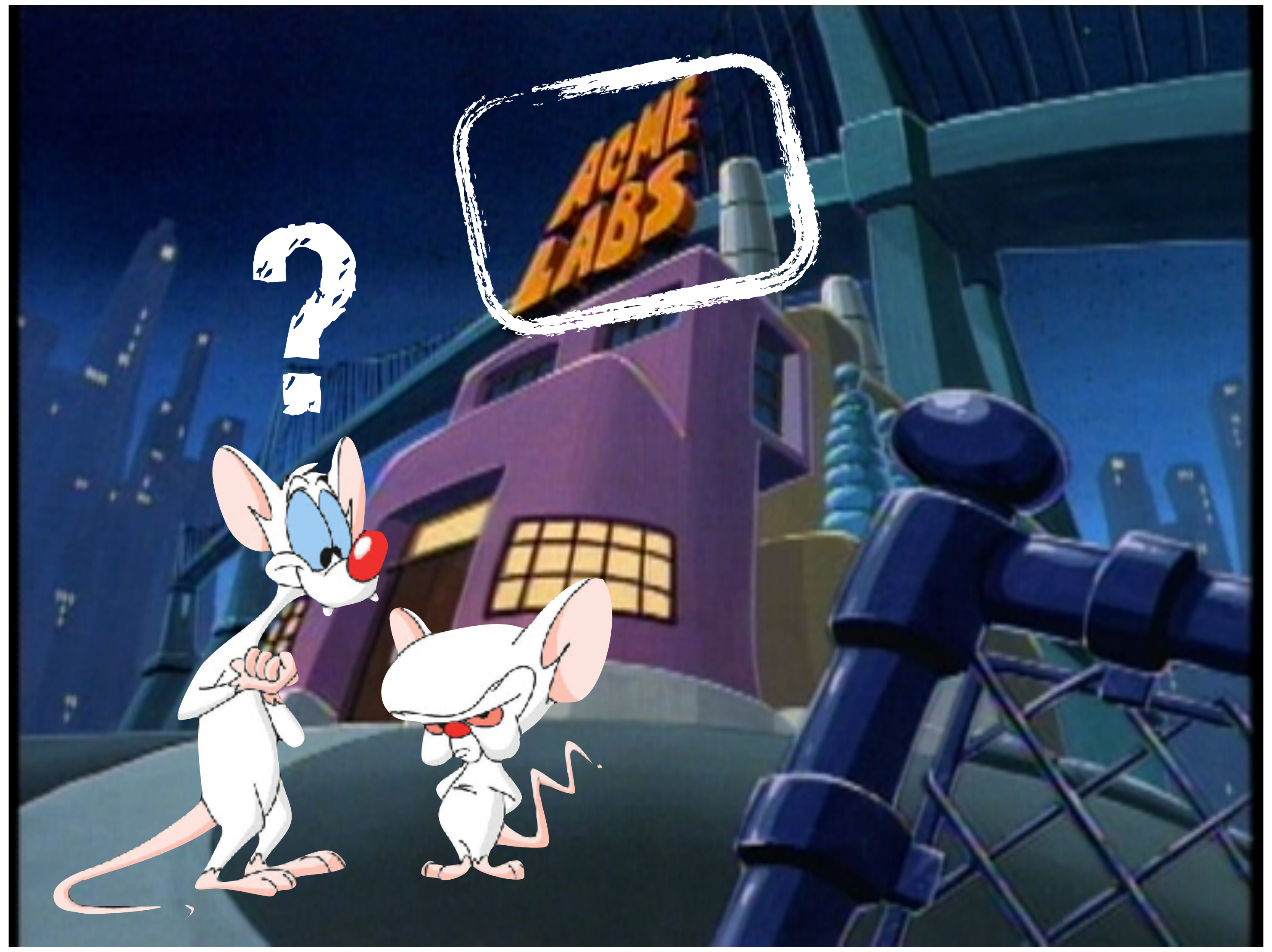
$$\hat{y} = b_0 + b_1 \cdot m + b_2 \cdot x$$

```
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
```

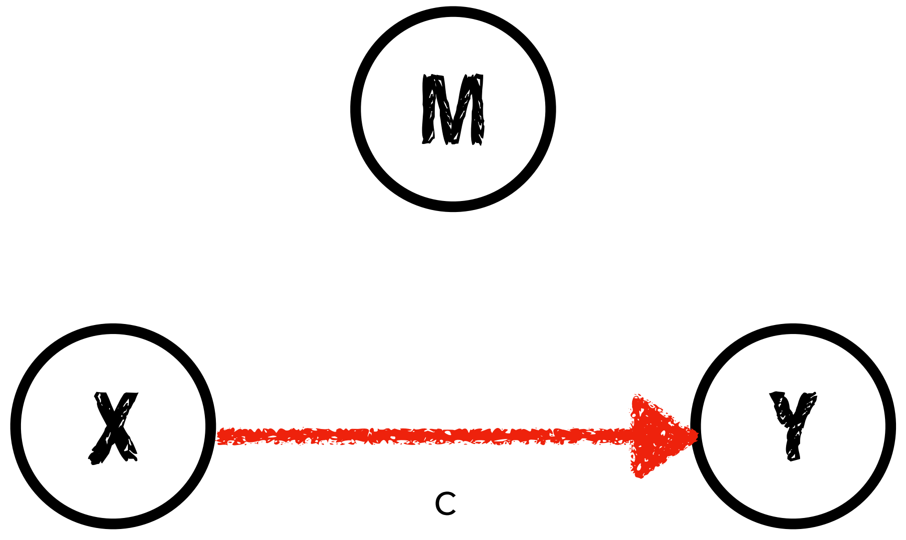
2. Bootstrapping




```
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

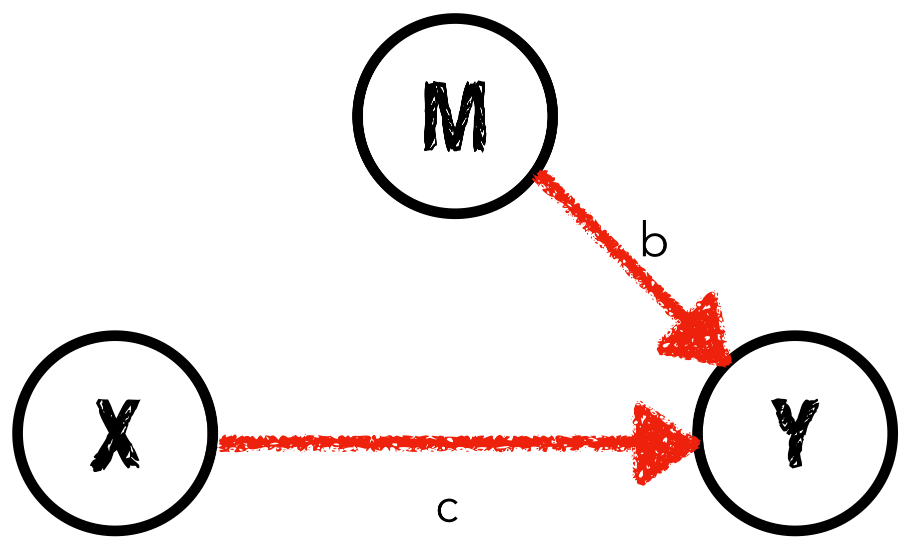



```
Causal Mediation Analysis

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

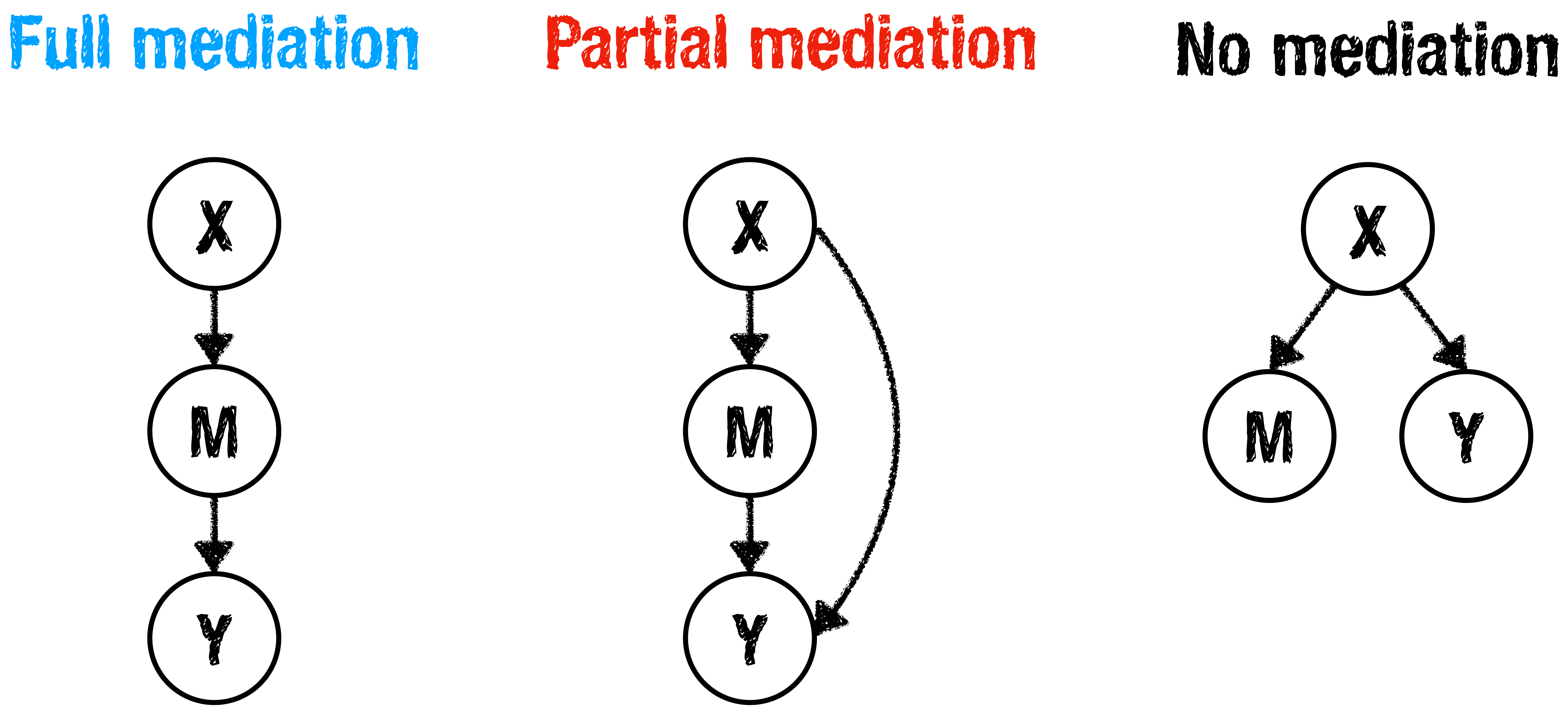
```
Causal Mediation Analysis

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: Average causal mediation effect

ACME = Total effect - ADE

ADE: Average direct effect

Underlying causal model

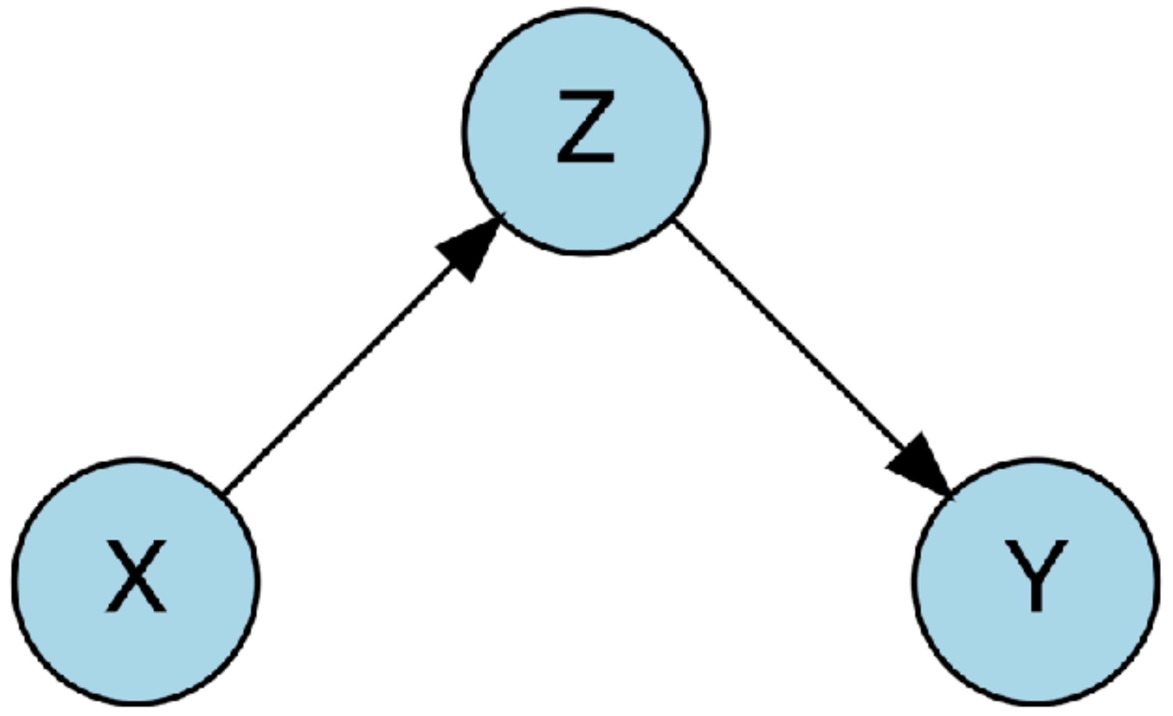


Full mediation: When the effect of **X** on **Y** completely disappears, **M** fully mediates between **X** and **Y**.

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

Beware of mediation analyses!!!

```r
set.seed(1)

n = 100 # number of observations

# causal chain
df.causal_chain = tibble(x = rnorm(n, 0, 1),
                         z = 2 * x + rnorm(n, 0, 1),
                         y = 2 * z + rnorm(n, 0, 1))
```

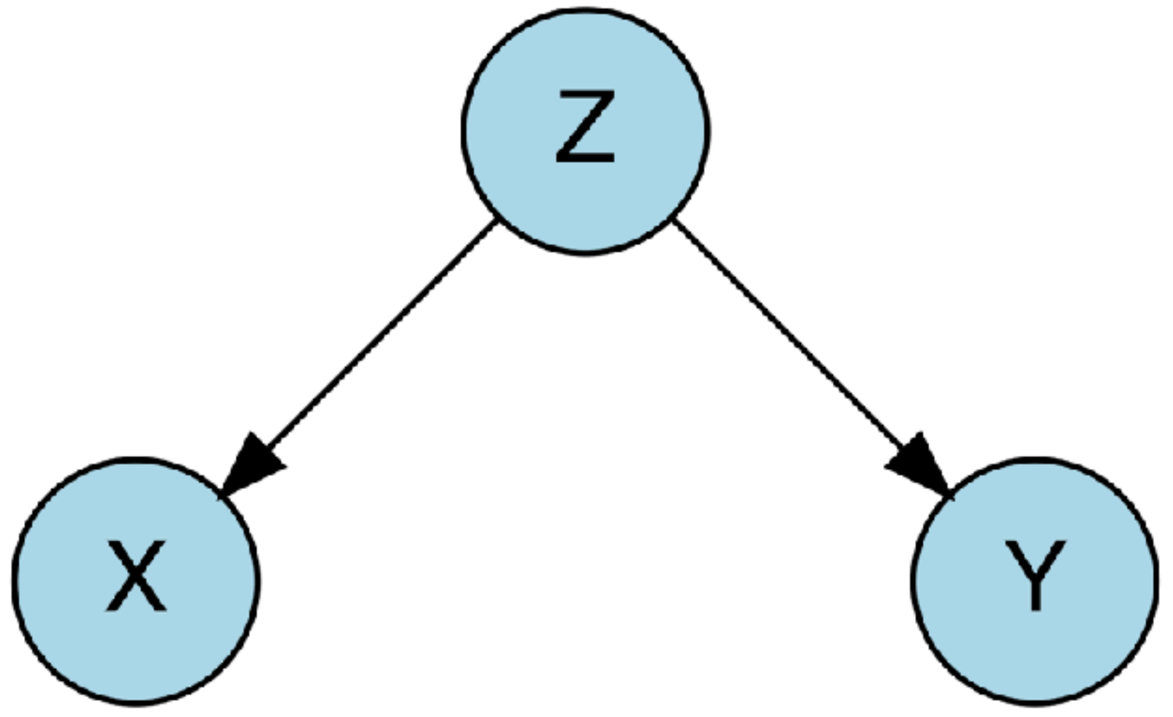



```
Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

               Estimate 95% CI Lower 95% CI Upper p-value    
ACME             0.8287       0.6234         1.05  <2e-16 ***
ADE             -0.0535      -0.2548         0.15    0.55    
Total Effect     0.7752       0.6391         0.90  <2e-16 ***
Prop. Mediated   1.0690       0.8131         1.35  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 100


Simulations: 1000
```

nice mediation result!

Beware of mediation analyses!!!

```r
set.seed(1)

n = 100 # number of observations

# common cause
df.common_cause = tibble(z = rnorm(n, 0, 1),
                         x = 2 * z + rnorm(n, 0, 1),
                         y = 2 * z + rnorm(n, 0, 1))
```

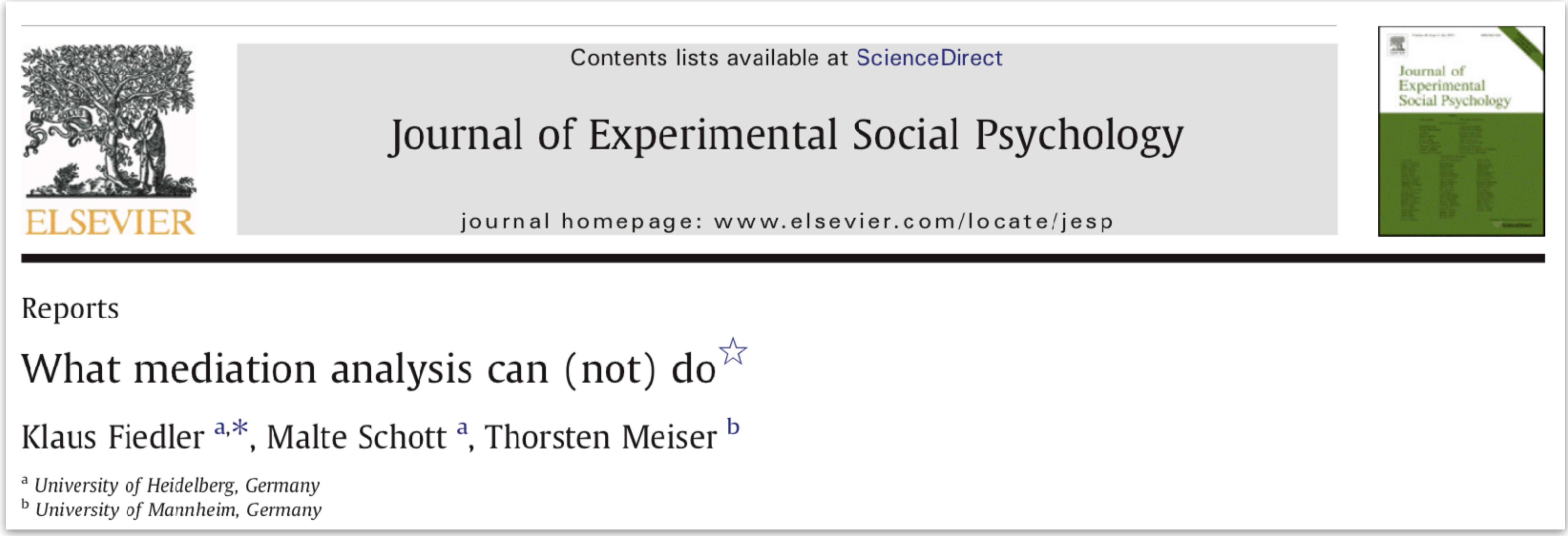



```
Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

               Estimate 95% CI Lower 95% CI Upper p-value    
ACME             0.8287       0.6065         1.04  <2e-16 ***
ADE             -0.0535      -0.2675         0.16    0.56    
Total Effect     0.7752       0.6353         0.90  <2e-16 ***
Prop. Mediated   1.0690       0.8134         1.37  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Sample Size Used: 100


Simulations: 1000
```

(not) nice mediation result!

Limitations

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



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.

Limitations

many-to-one mapping

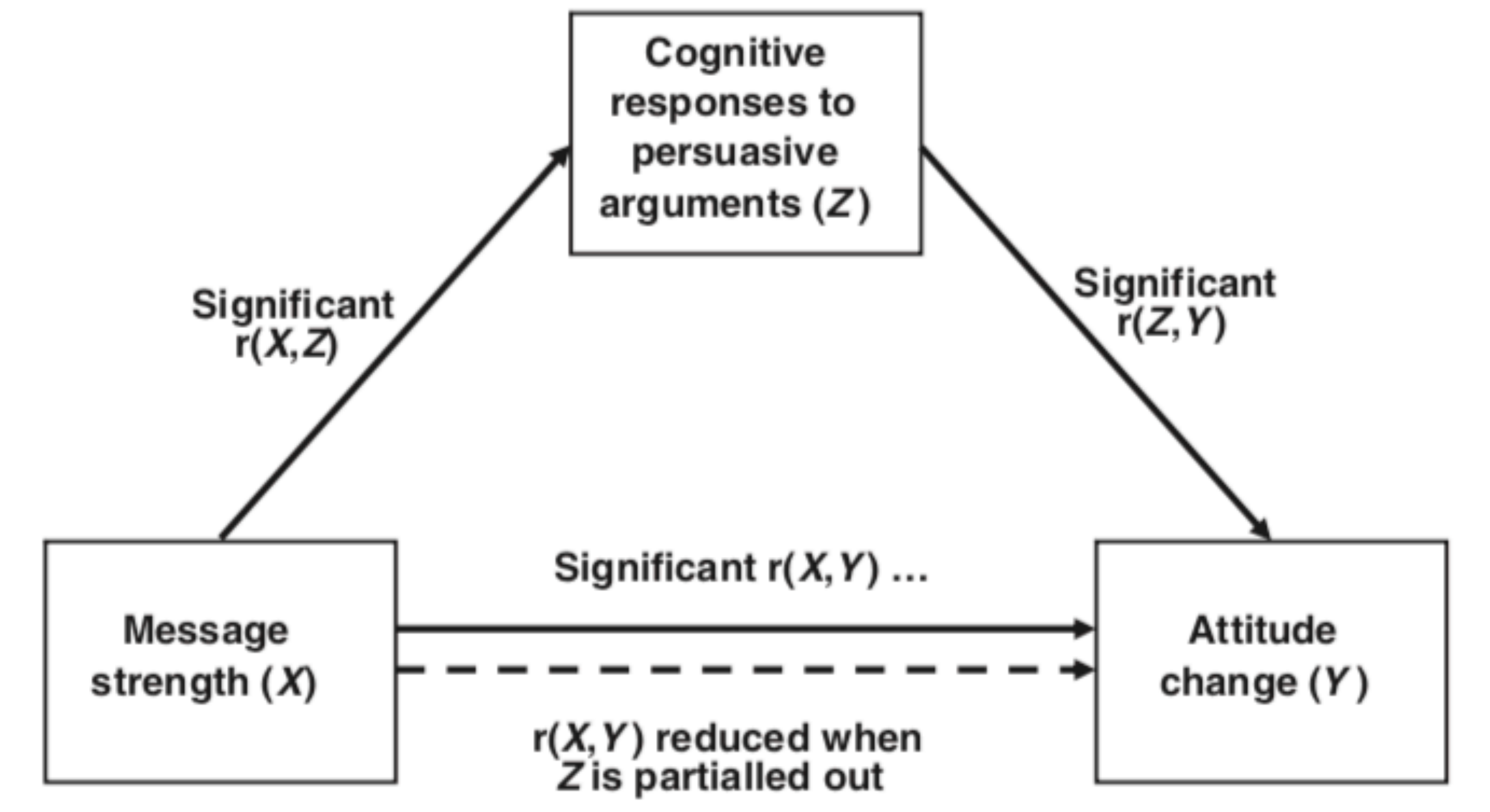


Fig. 1. Illustration of a mediation analysis in the context of persuasion research: The impact of message strength (*X*) on attitude change (*Y*) is supposed to be mediated by the recipients' cognitive responses (*Z*).

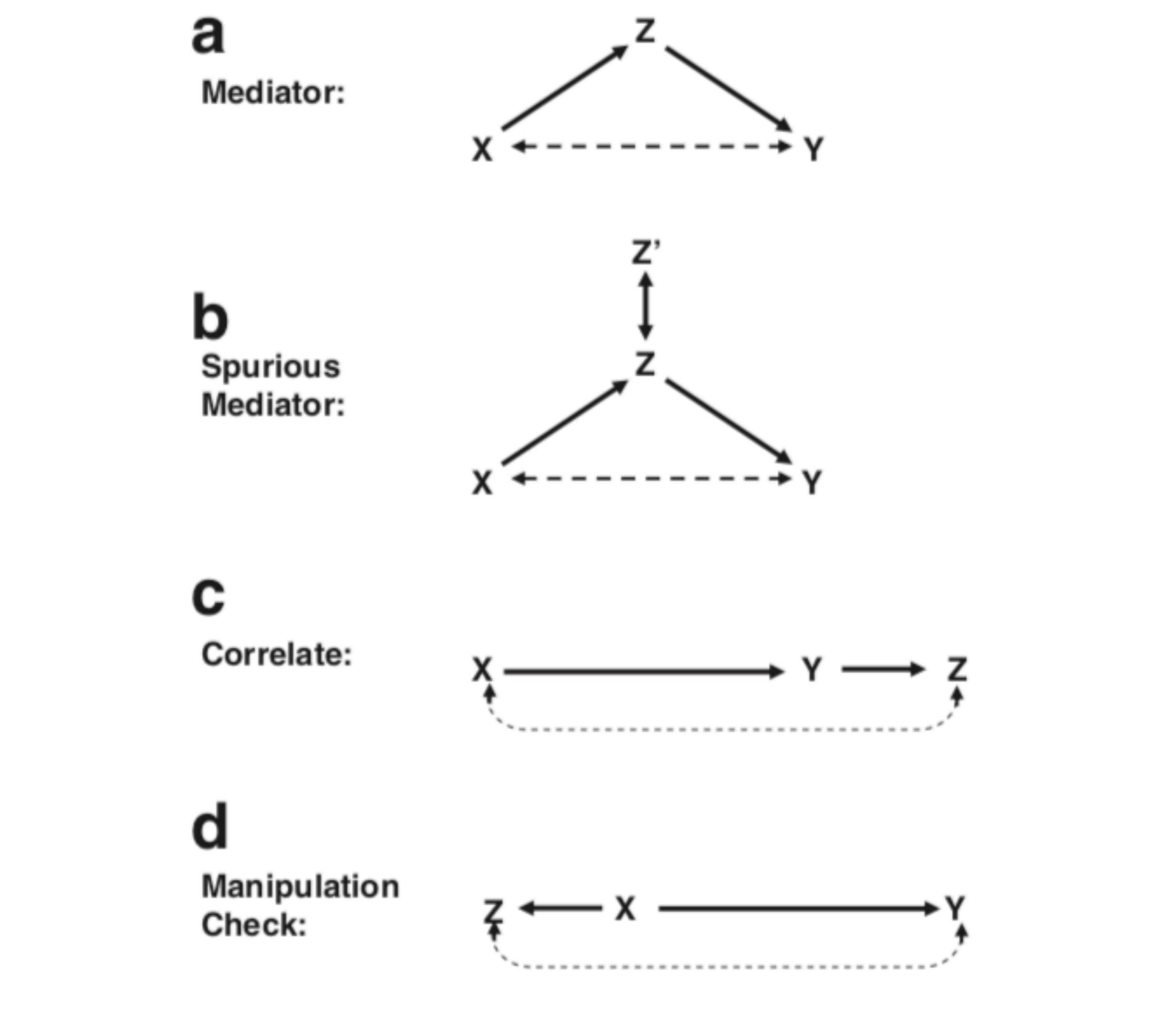


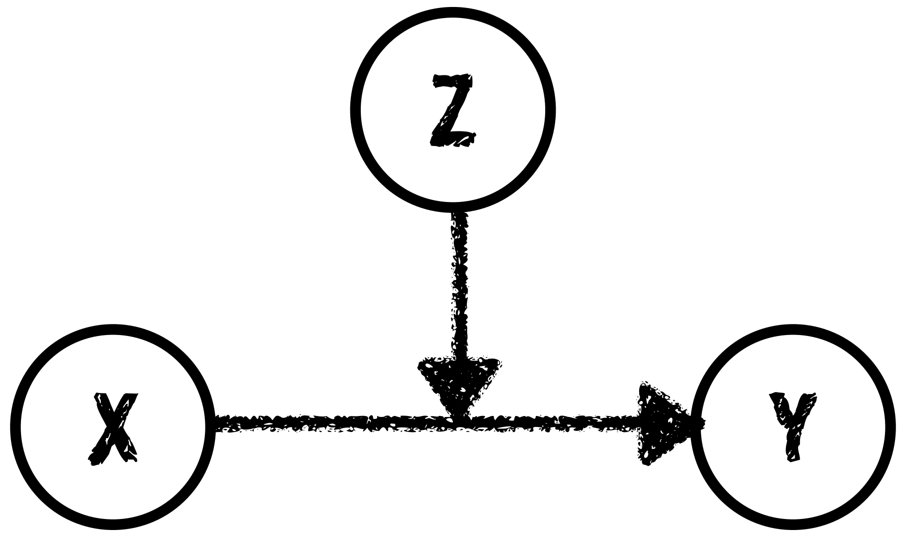
Fig. 2. Different causal models for the interpretation of the same intercorrelations between three variables, *X*, *Y*, and *Z*.

only experiments allow us to tell apart possible causal structures

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

Moderation

Definition

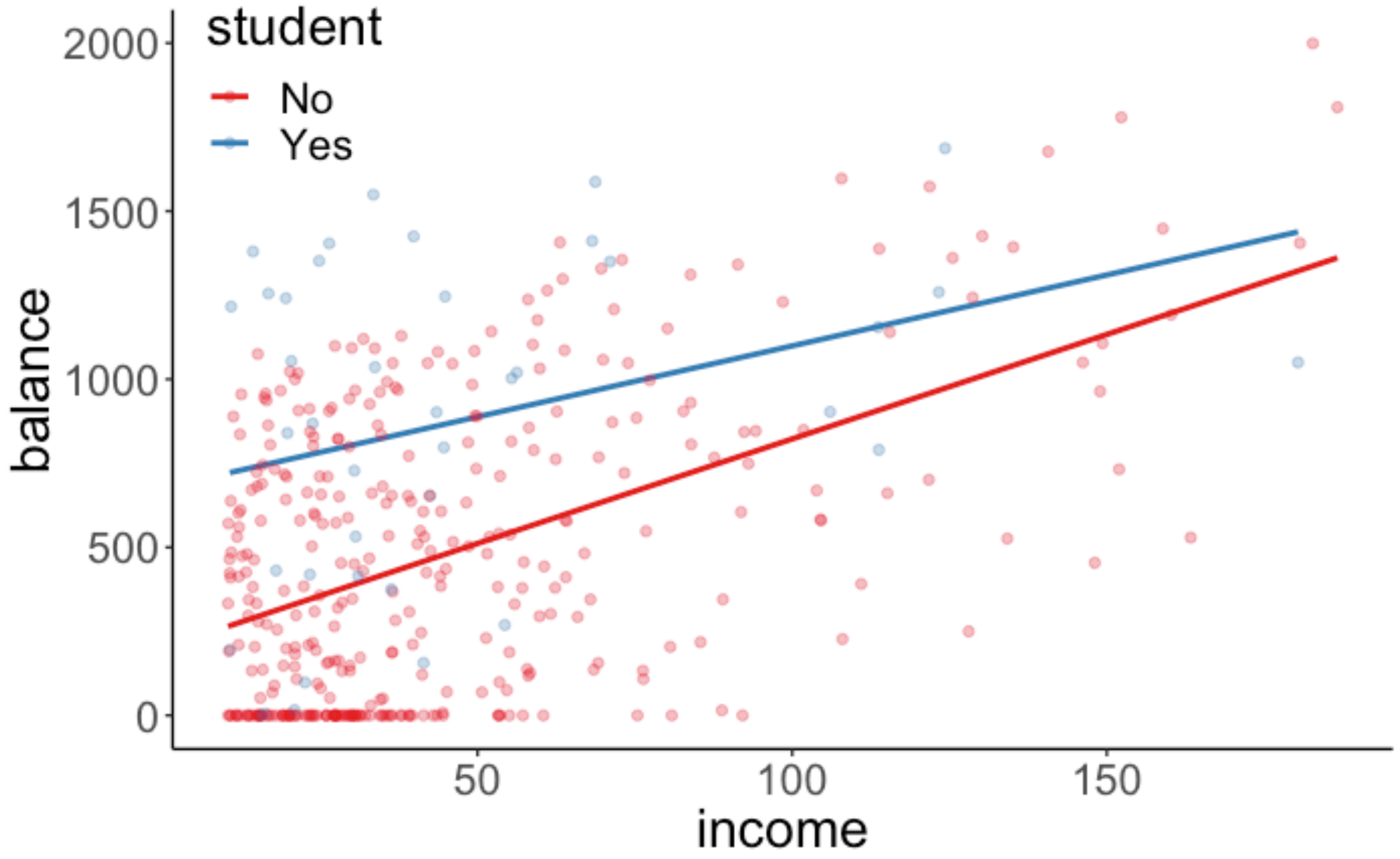


Moderation means that the effect of a predictor depends on the value of another.

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.



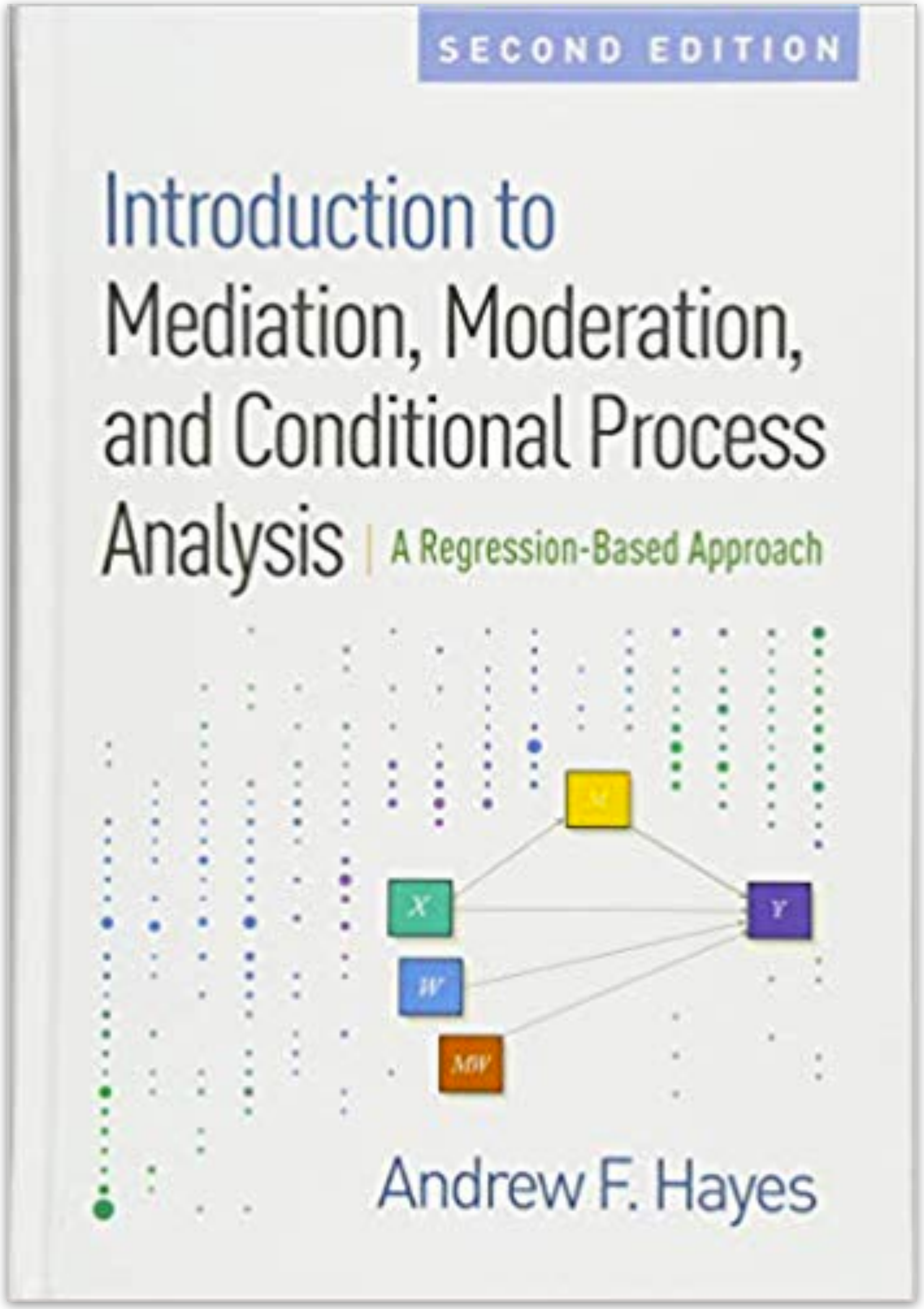
$$\widehat{\text{balance}}_i = 200.62 + 6.22 \cdot \text{income}_i + 476.68 \cdot \text{student}_i - 2.00 \cdot (\text{income}_i \times \text{student}_i)$$

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

if student = "Yes"

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

Learn more about mediation and moderation



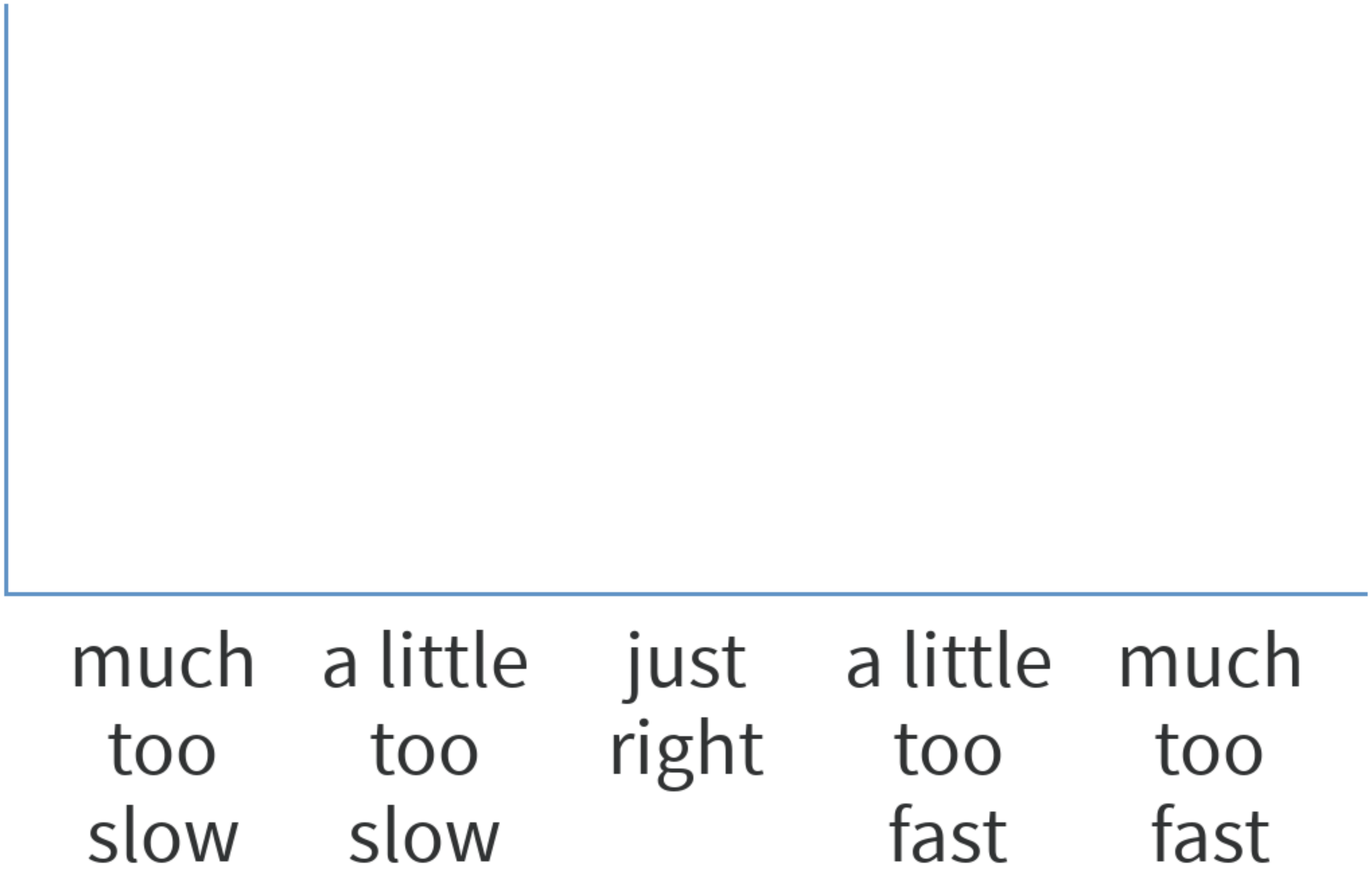
Recoded with `brms` by Solomon Kurz here: https://bookdown.org/connect/#/apps/1523/access

Plan for today

- Quick recap
- Linear Mixed Model
  - Different random effect structures
  - lmer() standard operating procedures
- Causation
  - Causation vs. correlation
  - Controlling for variables
  - Mediation
  - Moderation

Feedback

How was the pace of today's class?



How happy were you with today's class overall?

What did you like about today's class? What could be improved next time?

Answers to this poll are anonymous

Start the presentation to see live content. Still no live content? Install the app or get help at **PollEv.com/app**

Thank you!