

# Linear mixed effects models 1



02/19/2020

# Your feedback

# Your feedback

The pdf of the model comparison lecture is updated now.

# Your feedback

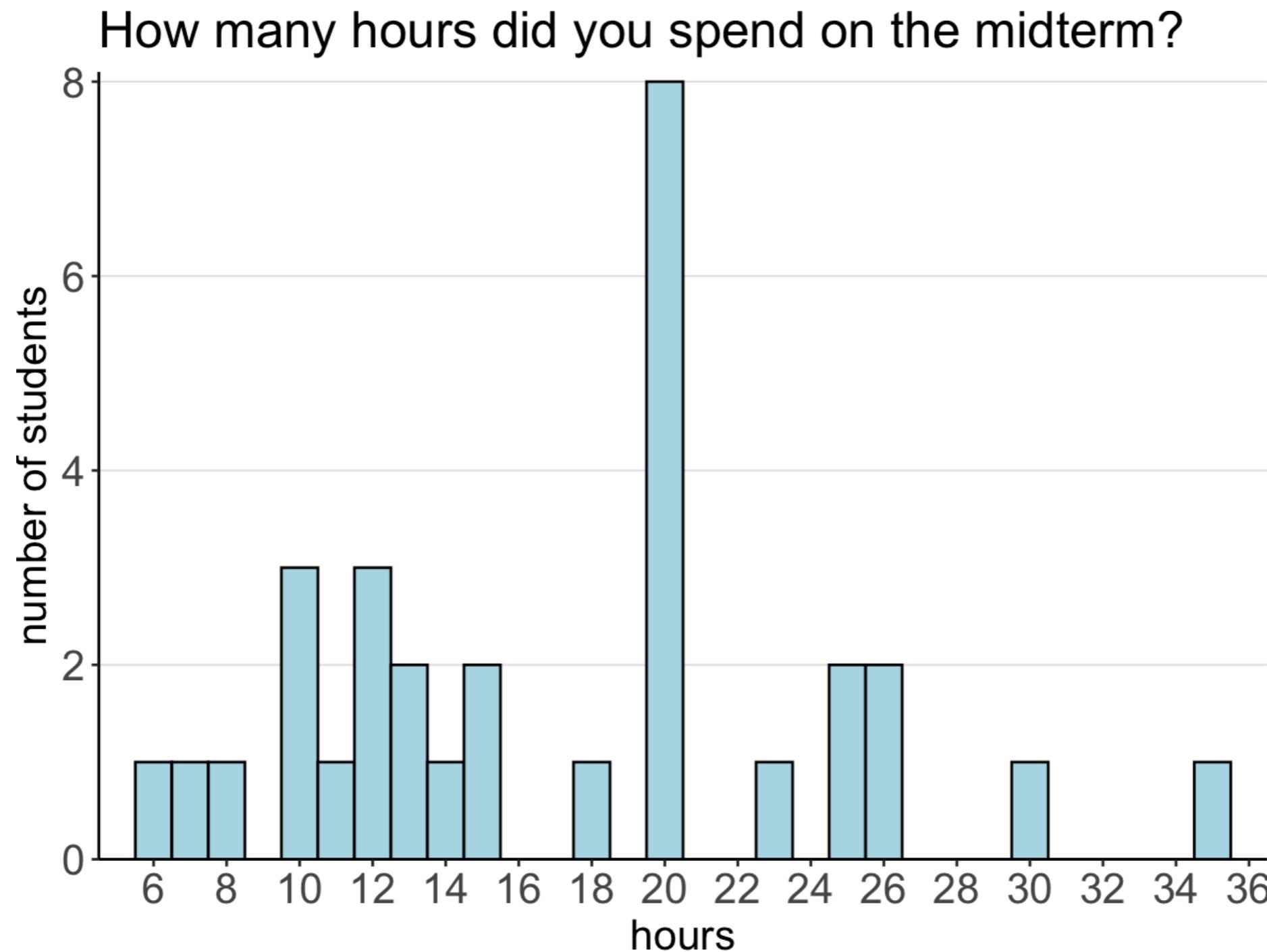
Regarding the midterm, I would like to say that, although it was long (it took me approximately 13 hours), I actually enjoyed doing it! I do think that some of the parts were testing very similar things, so maybe these parts could have been shortened.

**some positive notes**

The last part of the exam was way too hard. I didn't understand the lectures and with the class material I found it impossible to make sense of the simulation. However, for most of the exam, I was just trying to find similar codes on the rmds and didn't really understand the topics.

**many critical notes**

# Your feedback



**we underestimated how long it would take, sorry!!**

# Your feedback

... In terms of the mid-term, the primary tension in the class manifested: is this an R class, or a statistics class? The class, as it's designed, is great at breaking down the bigger statistical concepts, albeit a bit too fast. But the R learning is secondary by a long shot. I spent 90% of the time on the midterm stressing and fretting about code. **I would just prefer to stress over concepts more than knitting.** Analysis and interpretation was pushed aside. If you give me data, and a visualization, then my focus is the interpretation. If you ask me do elaborate code, then my focus is the coding. If I'm supposed to learn coding on my own, then why make it such a integral part of nearly every assignment? **Any ability that I have to interpret data or understand certain methods of data analysis is constantly limited by my R knowledge.** Assessment would be better if you **found a way to separate the two domains of knowledge, at least on occasion.**

 we will try to do  
this more in future  
assessments

this class is trying to do both (teach stats and R), and I know that that's a lot to ask for

I very much hope that you will ultimately find this beneficial

# Your feedback

I think it's great that you changed the late policy for the midterm to be more lenient. However, I strongly disagree with your approach (that you suggested you might follow in future) of keeping the late policy secret until after these test. **Transparency with students and respecting their ability to make decisions after having all the information about the policy seems best.**

**we won't follow this in the future,  
we will be strict on homework assignments**

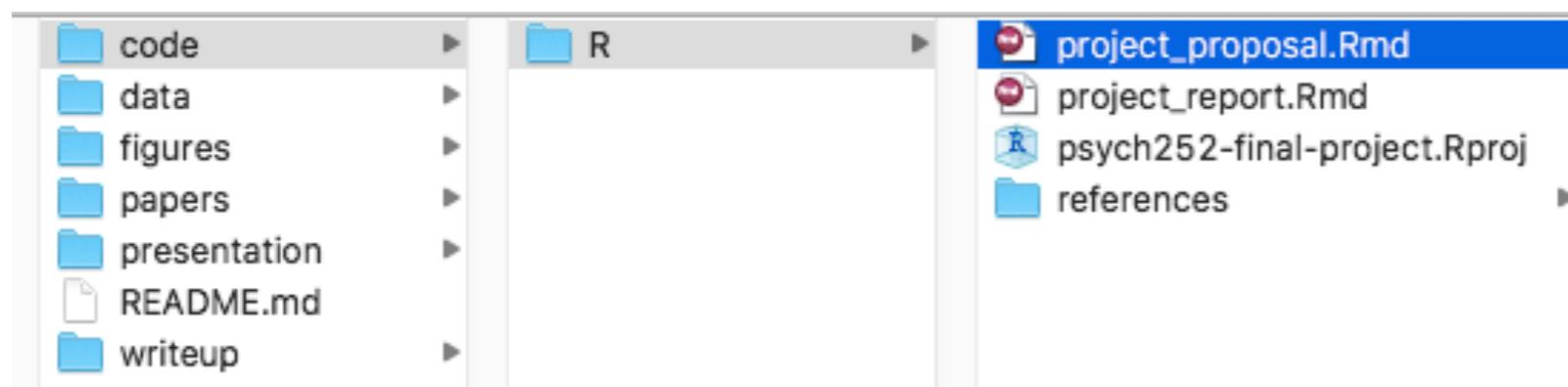
**we will explicitly communicate a late submission policy for the midterm in advance the next time**

# **Logistics**

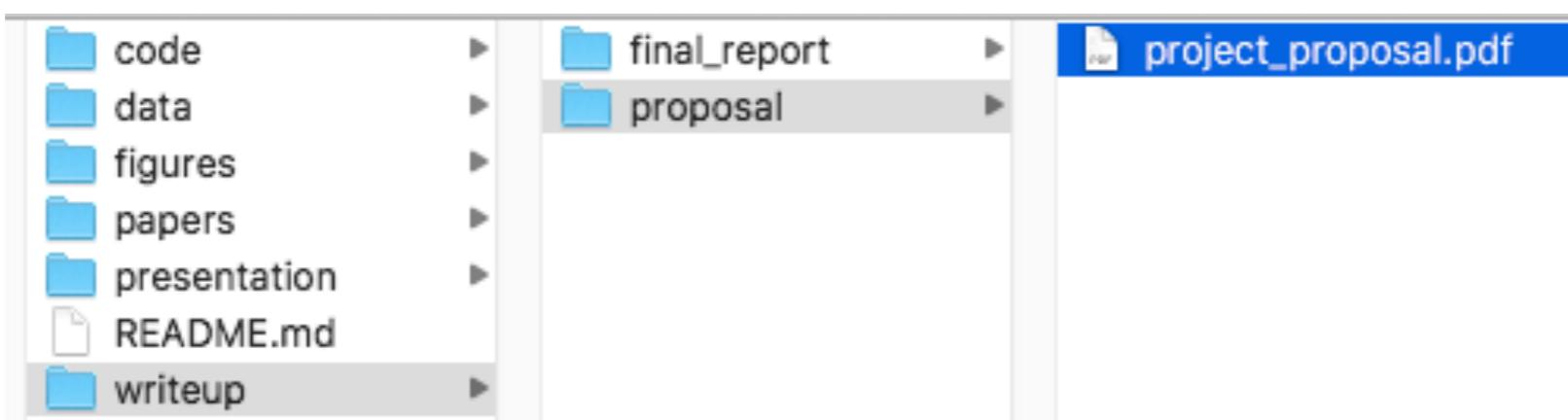
# Project proposal

# Project proposal

the template is here



put the rendered pdf here



... and make sure to upload your pdf on canvas.

# Project proposal

project\_report.Rmd

suggested structure for the final report

## 1 Introduction

### 1.1 Research questions

### 1.2 Hypotheses

## 2 Methods

## 3 Results

### 3.1 Confirmatory analysis

- analyses that you planned to carry out before taking a look at the data go here

### 3.2 Exploratory analysis

- analyses that you carried out after taking a look at the data go here
- don't use null-hypothesis statistical significance testing in this section, you can still report parameter estimates, model fits, etc. (just no significance tests)

## 4 Discussion

analysis you planned (e.g. in your proposal or later, but before you've taken a look at the data)

exploratory analysis after you've taken a look at the data

# **Application section**

# Application section

Thursday, February 20th, 4:30pm - 5:20pm in 160-322

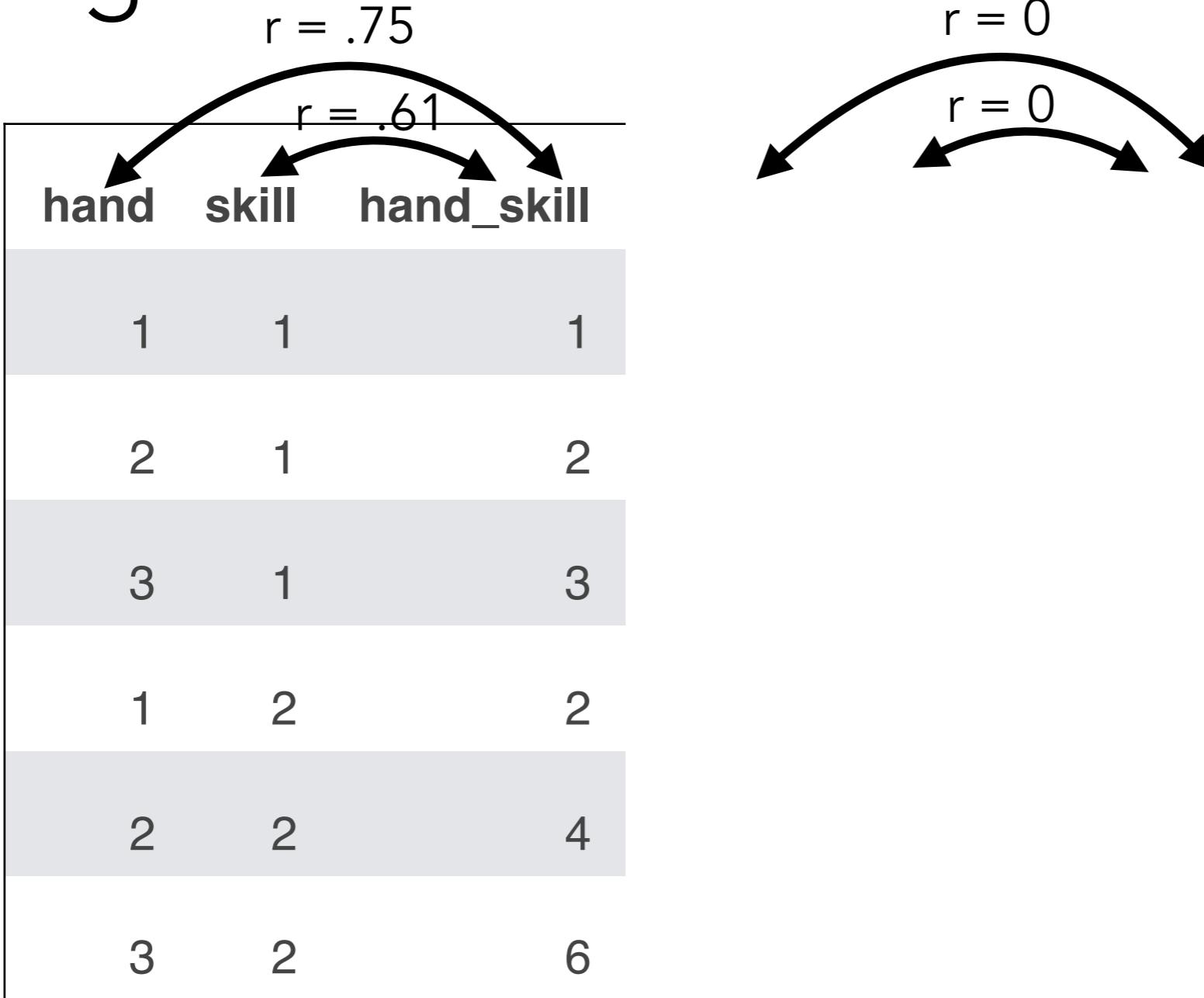
## Applied model comparison



# **Things that came up**

# Centering

# Centering



by centering continuous predictors, individual predictors aren't correlated with their interaction

# **Using pilot studies to determine sample size**

# Using pilot studies to determine sample size

PERSPECTIVES

## Caution Regarding the Use of Pilot Studies to Guide Power Calculations for Study Proposals

Helena Chmura Kraemer, PhD; Jim Mintz, PhD; Art Noda, MS; Jared Tinklenberg, MD; Jerome A. Yesavage, MD



Journal of Clinical Epidemiology 56 (2003) 717–720

Journal of  
Clinical  
Epidemiology

Underpowerin

## ON THE USE OF A PILOT SAMPLE FOR SAMPLE SIZE DETERMINATION

*Integrative Medicine Service*

RICHARD H. BROWNE

Texas Scottish Rite Hospital for Children Clinical Epidemiology

Dovepress

open access to scientific and medical research

Open Access Full Text Article

METHODOLOGY

## Guidance for using pilot studies to inform the design of intervention trials with continuous outcomes

Bell, M. L., Whitehead, A. L., & Julious, S. A. (2018). Guidance for using pilot studies to inform the design of intervention trials with continuous outcomes. *Clinical Epidemiology*, Volume 10, 153–157. <https://doi.org/10.2147/CLEP.S146397>

Browne, R. H. (1995). On the use of a pilot sample for sample size determination. *Statistics in Medicine*, 14(17), 1933–1940. <https://doi.org/10.1002/sim.4780141709>

Kraemer, H. C., Mintz, J., Noda, A., Tinklenberg, J., & Yesavage, J. A. (2006). Caution Regarding the Use of Pilot Studies to Guide Power Calculations for Study Proposals. *Archives of General Psychiatry*, 63(5), 484. <https://doi.org/10.1001/archpsyc.63.5.484>

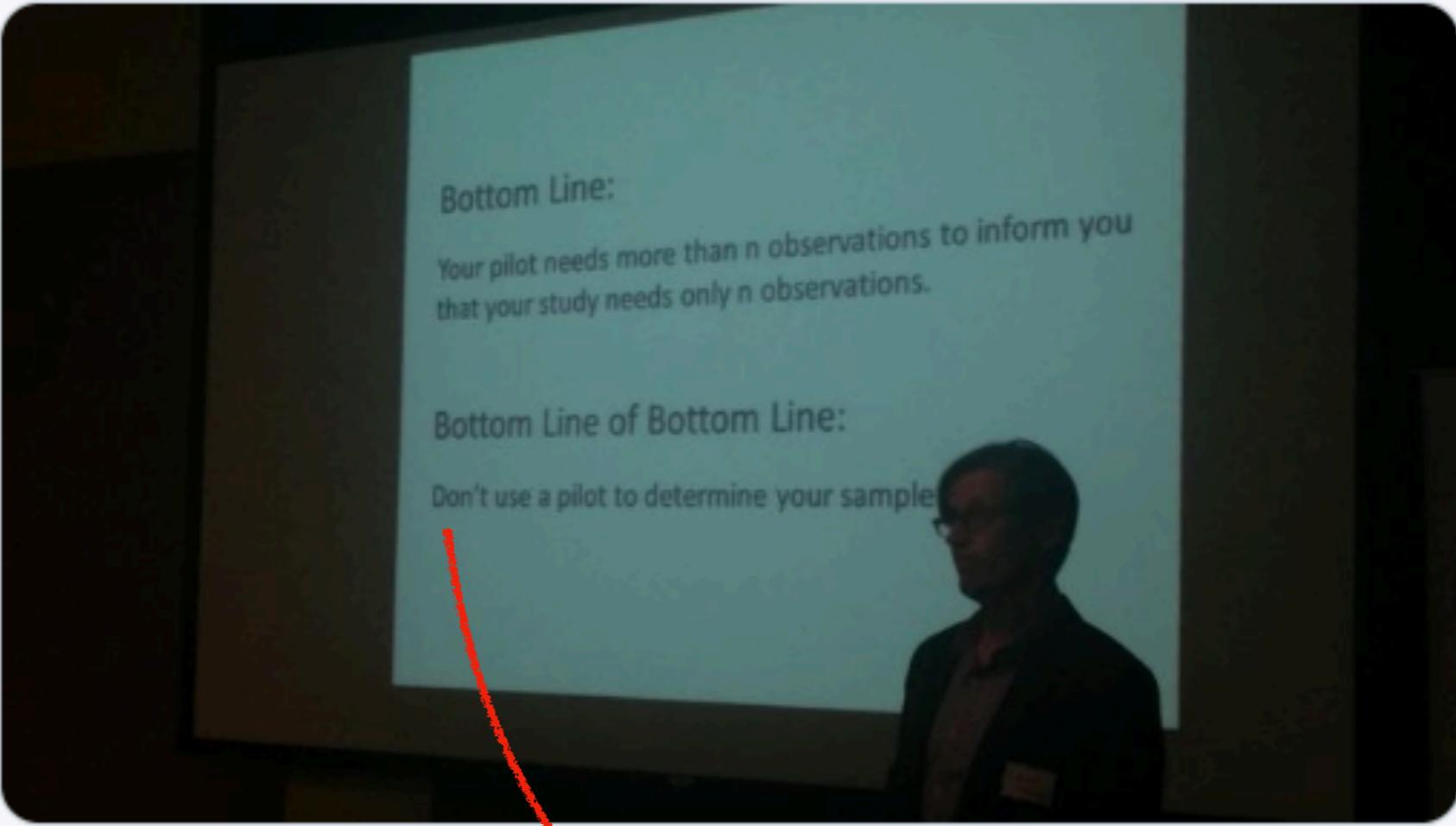
Vickers, A. J. (2003). Underpowering in randomized trials reporting a sample size calculation. *Journal of Clinical Epidemiology*, 56(8), 717–720. [https://doi.org/10.1016/S0895-4356\(03\)00141-0](https://doi.org/10.1016/S0895-4356(03)00141-0)

Whitehead, A. L., Julious, S. A., Cooper, C. L., & Campbell, M. J. (2016). Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. *Statistical Methods in Medical Research*, 25(3), 1057–1073. <https://doi.org/10.1177/096228021558241>

# Using pilot studies to determine sample size

 Research Transparency @UCBITSS · Jun 8, 2016

Statistical bummer re:running a pilot to determine sample size  
#BITSS2016



Bottom Line:  
Your pilot needs more than  $n$  observations to inform you  
that your study needs only  $n$  observations.

Bottom Line of Bottom Line:  
Don't use a pilot to determine your sample size

2 3 6

what else should we do?

# Using pilot studies to determine sample size

- doing pilot studies is (almost) never a bad idea
- using pilot studies to determine sample size needs to be done with caution
  - **key problem:** small studies tend to underestimate the standard deviation
- **solutions:**
  - *meaningful effect:* what would be a meaningful difference?
  - *sensitivity analysis:* determine what  $n$  would be required for larger standard deviations than what was found in the pilot study

# Plan for today

- Model comparison
  - Cross-validation
  - AIC and BIC
- Mediation
- Moderation
- Linear mixed effects model
  - modeling dependence in data

# Plan for today

- Model comparison
  - Cross-validation
  - AIC and BIC
- Mediation
- Moderation
- Linear mixed effects model
  - modeling dependence in data

# Any problems with our approach?

sometimes it doesn't work ...

## Model C

$$\text{balance}_i = \beta_0 + \epsilon_i$$

$$\text{balance}_i = \beta_0 + \beta_1 \cdot \text{student}_i + \epsilon_i$$

$$\text{balance}_i = \beta_0 + \beta_1 \cdot \text{student}_i + \epsilon_i$$

$$\text{balance}_i = \beta_0 + \beta_1 \cdot \text{student}_i + \epsilon_i$$

## Model A

$$\text{balance}_i = \beta_0 + \beta_1 \cdot \text{student}_i + \beta_2 \cdot \text{age}_i + \epsilon_i$$

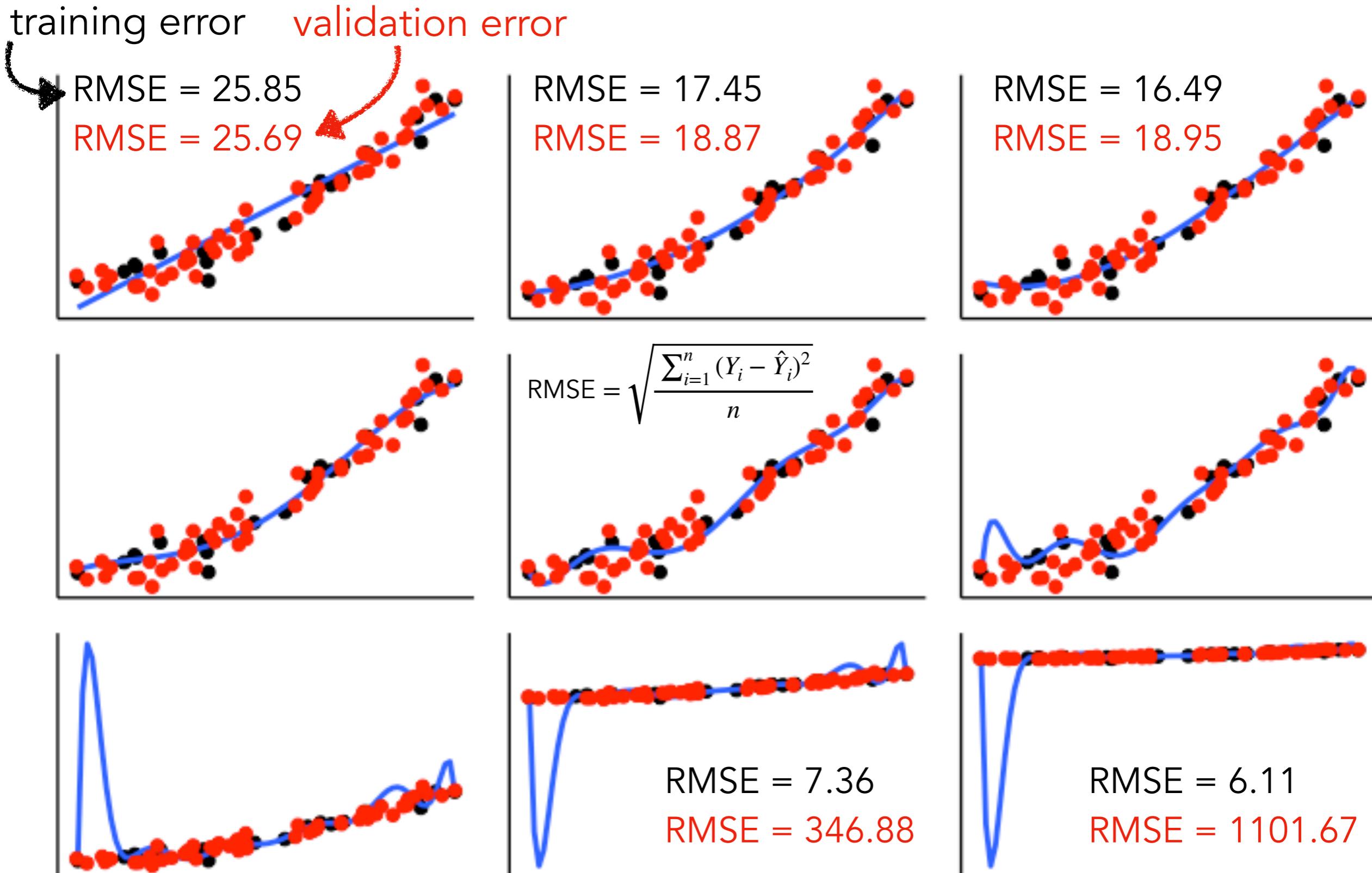
$$\text{balance}_i = \beta_0 + \beta_1 \cdot \text{student}_i + \beta_2 \cdot \text{age}_i + \epsilon_i$$

$$\text{balance}_i = \beta_0 + \beta_1 \cdot \text{age}_i + \epsilon_i$$

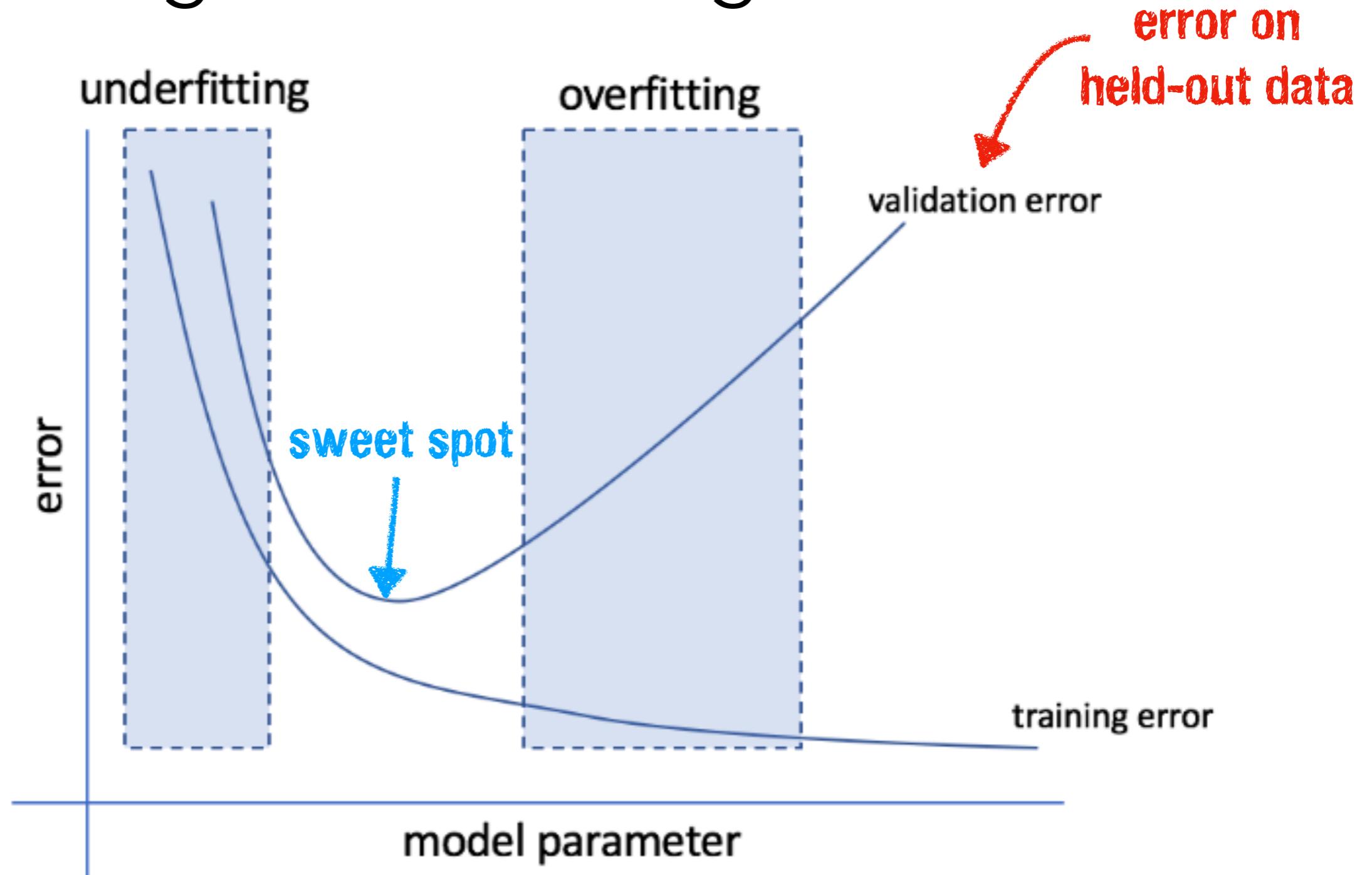
$$\text{balance}_i = \beta_0 + \beta_1 \cdot \text{age}_i + \beta_2 \cdot \text{degree}_i + \epsilon_i$$



# Which model describes the data best?



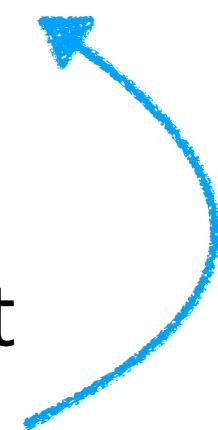
# Underfitting vs. Overfitting



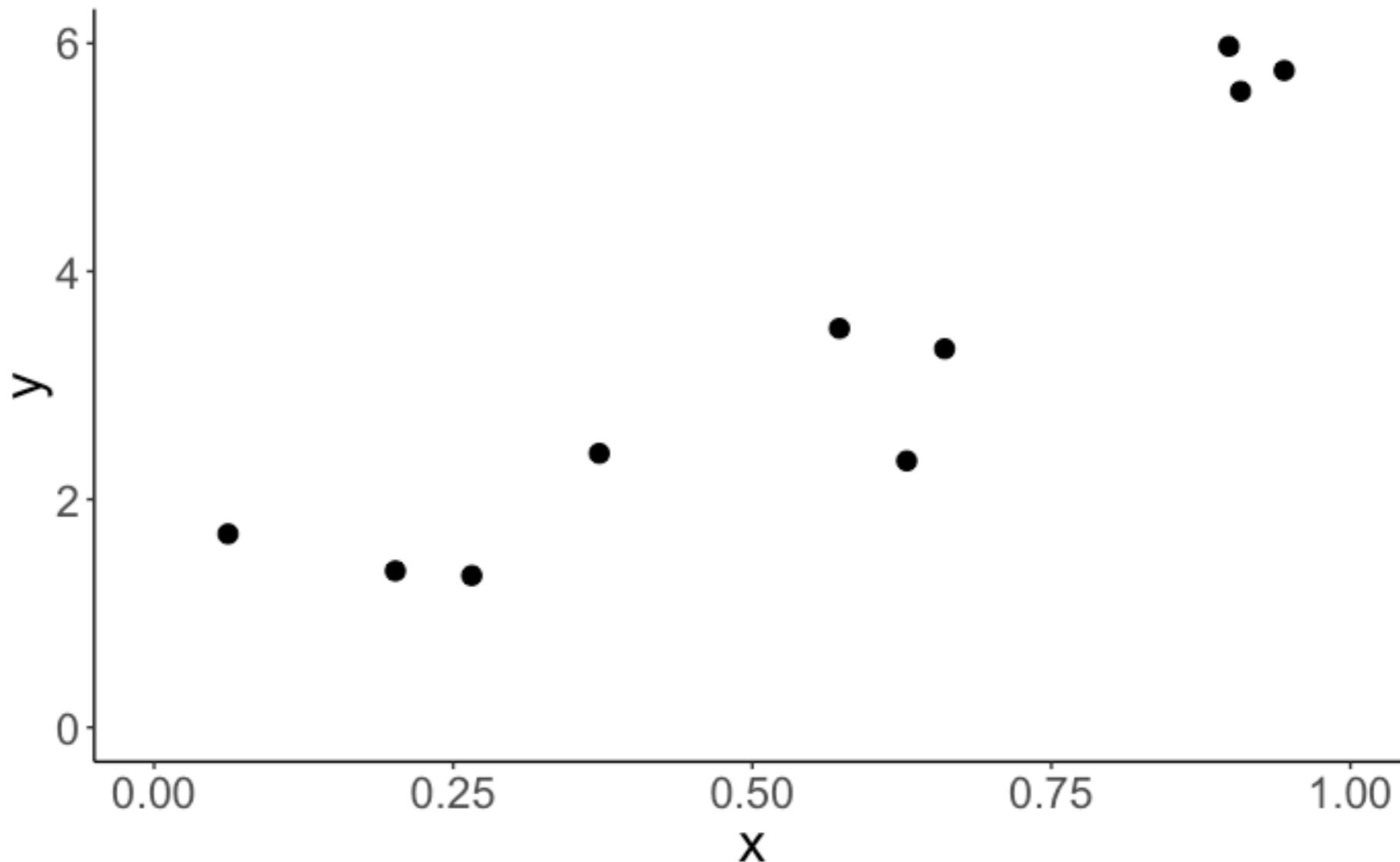
in machine learning, the goal is often to find the sweet spot between underfitting and overfitting

# **Leave-one-out crossvalidation**

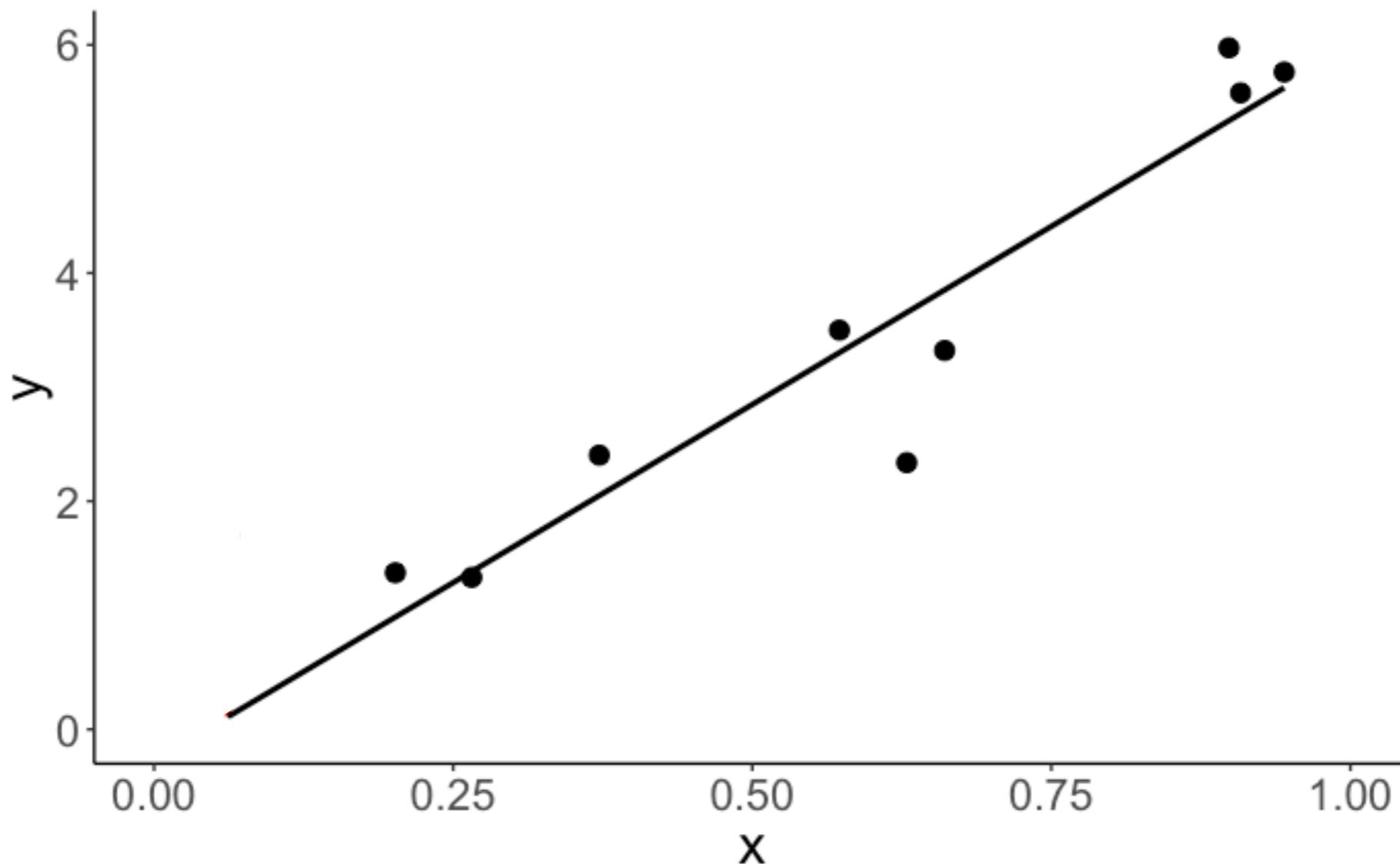
# Leave one out cross-validation

- train the model on all the data points except for one
  - calculate the prediction error for the held-out data point
- repeat for all data points**
- 

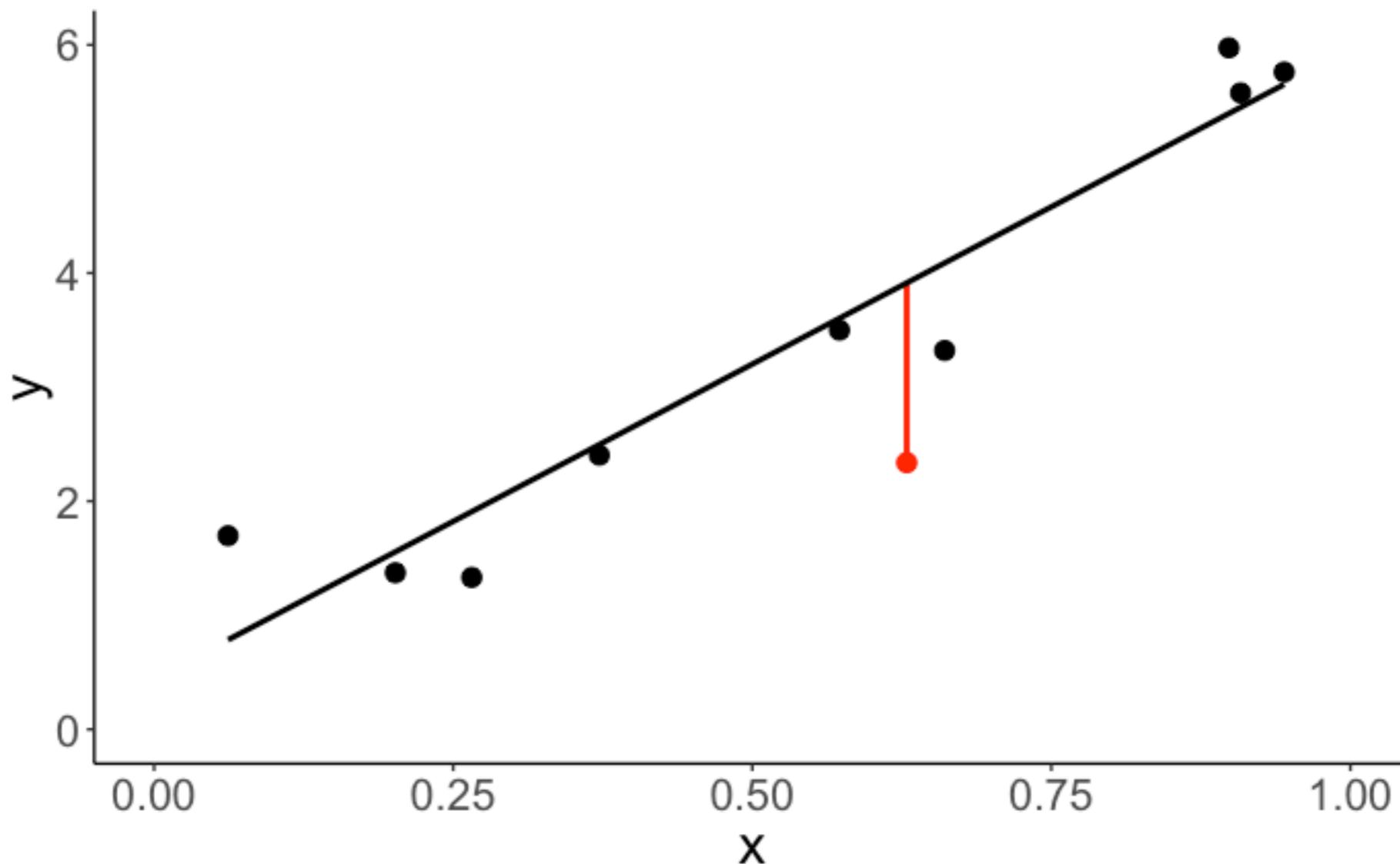
# Leave one out cross-validation



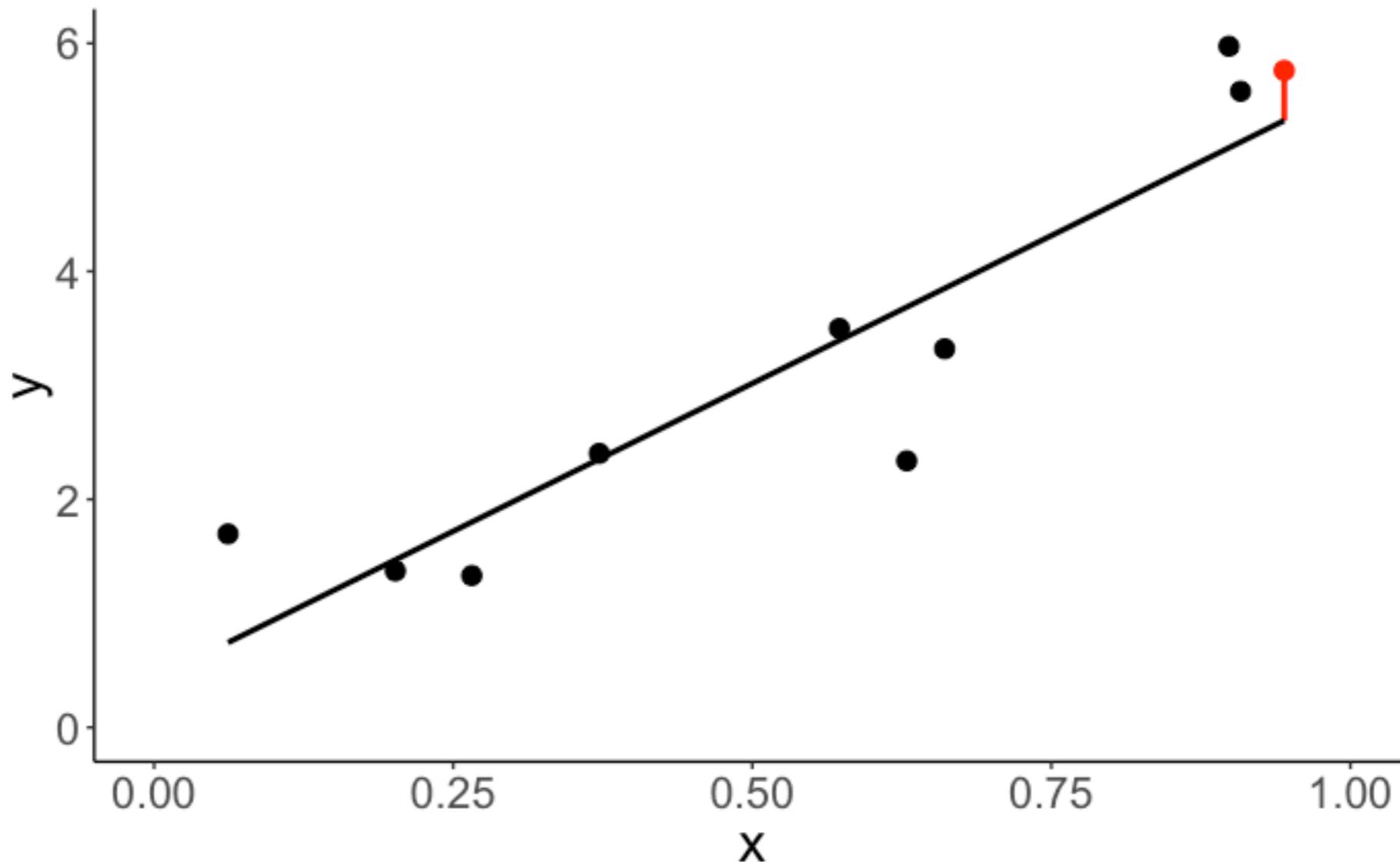
# Leave one out cross-validation



# Leave one out cross-validation



# Leave one out cross-validation



# Leave-one out crossvalidation

```
1 library("modelr")
2
3 df.cross = df.data %>%
4   crossv_loo() %>%
5   mutate(model_simple = map(train, ~ lm(y ~ 1 + x, data = .)),
6         model_correct = map(train, ~ lm(y ~ 1 + x + I(x^2), data = .)),
7         model_complex = map(train, ~ lm(y ~ 1 + x + I(x^2) + I(x^3), data = .))) %>%
8   pivot_longer(cols = contains("model"),
9                 names_to = "model",
10                values_to = "fit") %>%
11   mutate(rmse = map2_dbl(.x = fit, .y = test, ~ rmse(.x, .y)))
```

**calculate the root mean squared error for each model on the test data set**

```
1 df.cross %>%
2   group_by(model) %>%
3   summarize(mean_rmse = mean(rmse))
```

model	mean_rmse
simple	0.65
correct	0.48
complex	0.70

**the correct model has the lowest prediction error**

# Leave-one out crossvalidation

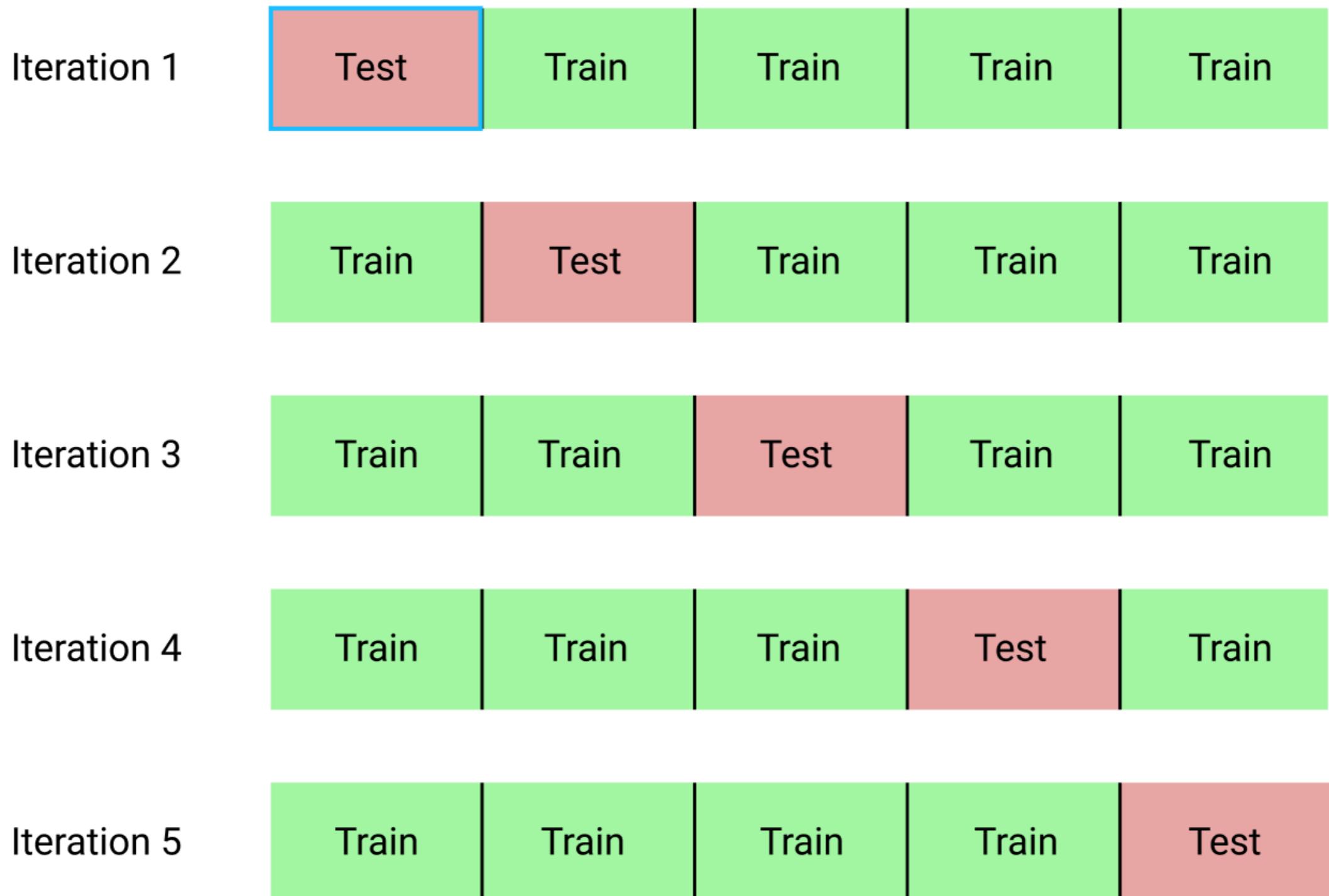
**Any potential problems with LOO?**

can be computationally expensive since it requires fitting the model  $n$  times (once for each data point) ...

# **k-fold cross validation**

# k-fold crossvalidation

Full data set



# k-fold crossvalidation

## k-fold crossvalidation

```
1 df.cross = df.data %>%
2   crossv_kfold(k = 10) %>%
3   mutate(model_simple = map(train, ~ lm(y ~ 1 + x, data = .)),
4         model_correct = map(train, ~ lm(y ~ 1 + x + I(x^2), data = .)),
5         model_complex = map(train, ~ lm(y ~ 1 + x + I(x^2) + I(x^3), data = .))) %>%
6   pivot_longer(cols = contains("model"),
7                 names_to = "model",
8                 values_to = "fit") %>%
9   mutate(rsquare = map2_dbl(.x = fit, .y = test, ~ rsquare(.x, .y))
```

why didn't we use R<sup>2</sup> for LOO?

using R<sup>2</sup> as a measure

this wouldn't work for LOO  
since we only have one data  
point in the test data...

model	median_rsquare
simple	0.839
correct	0.865
complex	0.860

the correct model accounts for  
the most variance in the test data

# k-fold vs. leave-one-out crossvalidation

- LOO:
  - trained on **more** data
  - more variance
  - less bias
- k-fold:
  - trained on **less** data
  - less variance
  - more bias

# Monte Carlo crossvalidation

# Monte Carlo crossvalidation

```
random splits into  
training and test data  
crossv_mc(n = 50, test = 0.5)  
number of  
training-test splits  
proportion of test  
data in each split
```

# Plan for today

- **Model comparison**
  - Cross-validation
  - AIC and BIC
- Mediation
- Moderation
- Linear mixed effects model
  - modeling dependence in data

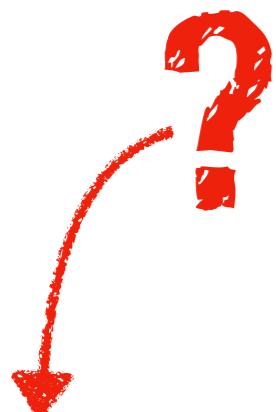
# AIC and BIC

# AIC and BIC

- AIC = Akaike Information Criterion
- BIC = Bayesian Information Criterion

**not that much Bayesian about it ...**

$$\text{AIC} = 2k - 2 \ln(\hat{L})$$



$$\text{BIC} = \ln(n)k - 2 \ln(\hat{L})$$

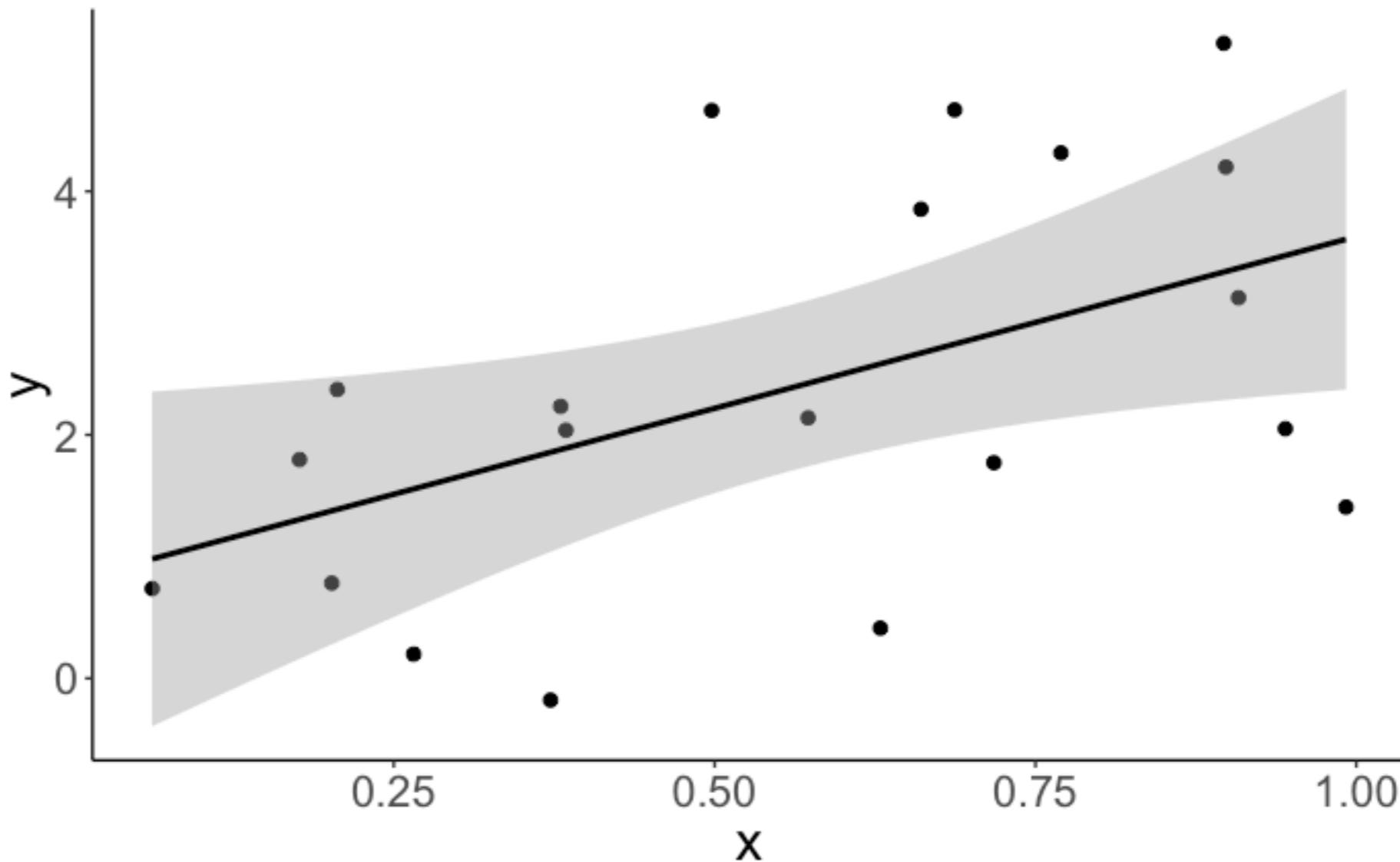
- $\hat{L}$  = maximized value of the likelihood function of the model  
 $k$  = number of parameters in the model  
 $n$  = number of observations

# AIC and BIC

- How do we get the likelihood of our model?
  - in a linear regression, minimizing least squares is equivalent to maximizing the likelihood of the data given the model
- Assumptions of the linear model:
  - residuals are normally distributed with:
    - mean = 0 and sd = sigma
    - calculate overall likelihood by computing the likelihood of each residual, and then multiplying

# AIC and BIC

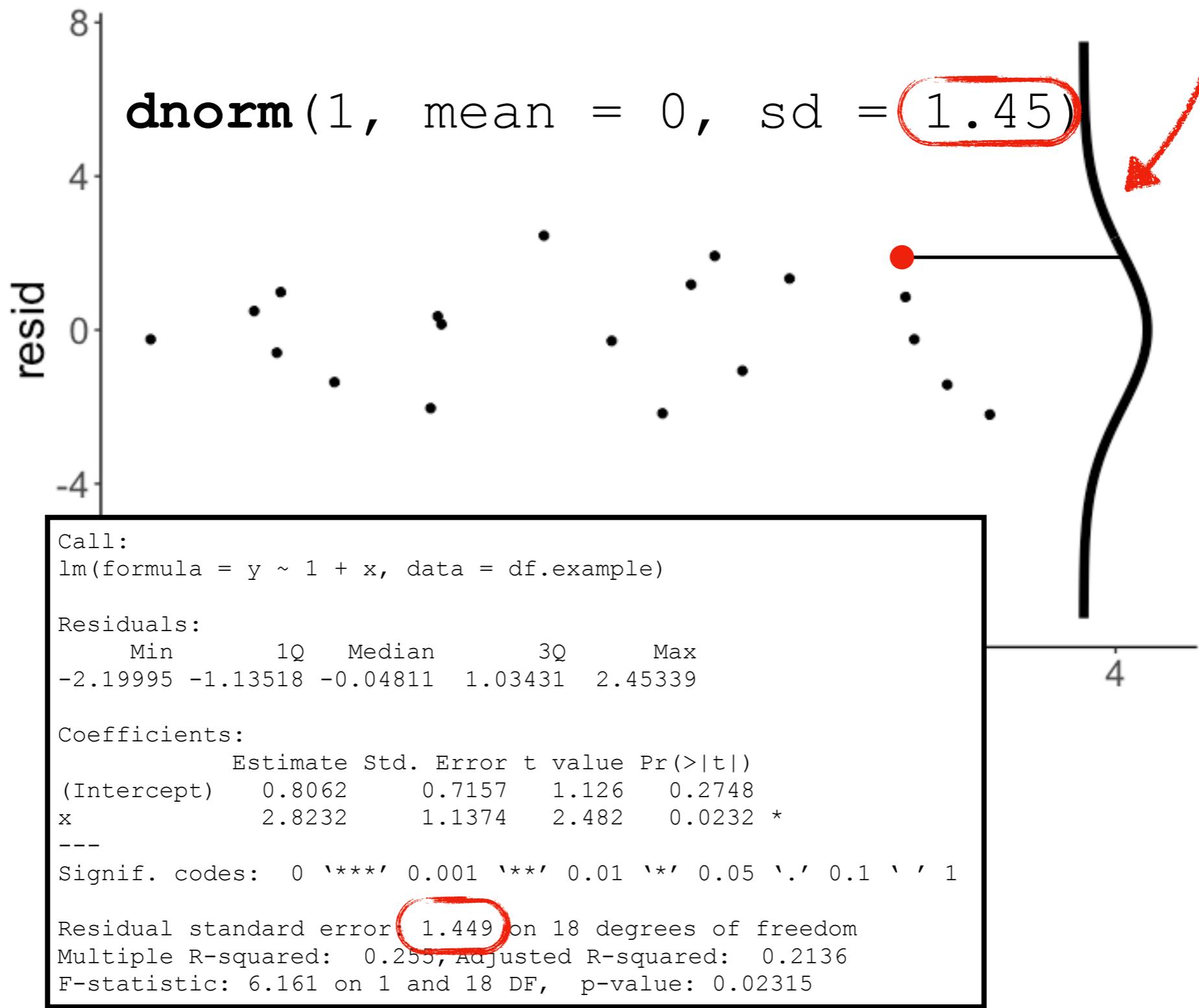
data with linear model fit



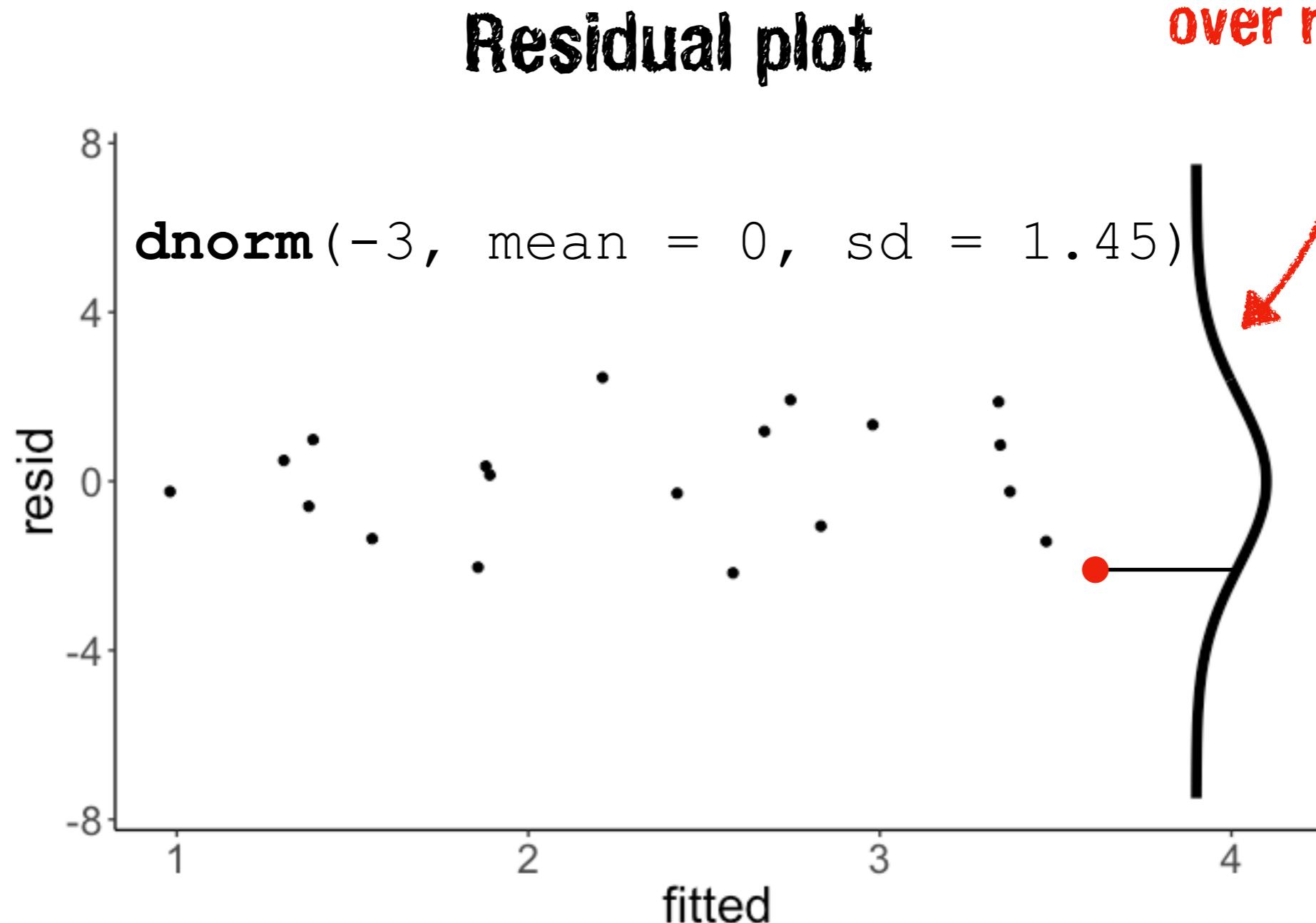
# AIC and BIC

normal distribution  
over residuals

## Residual plot



# AIC and BIC



since the data points are independent, we can calculate the overall likelihood by multiplying the likelihood of each observation

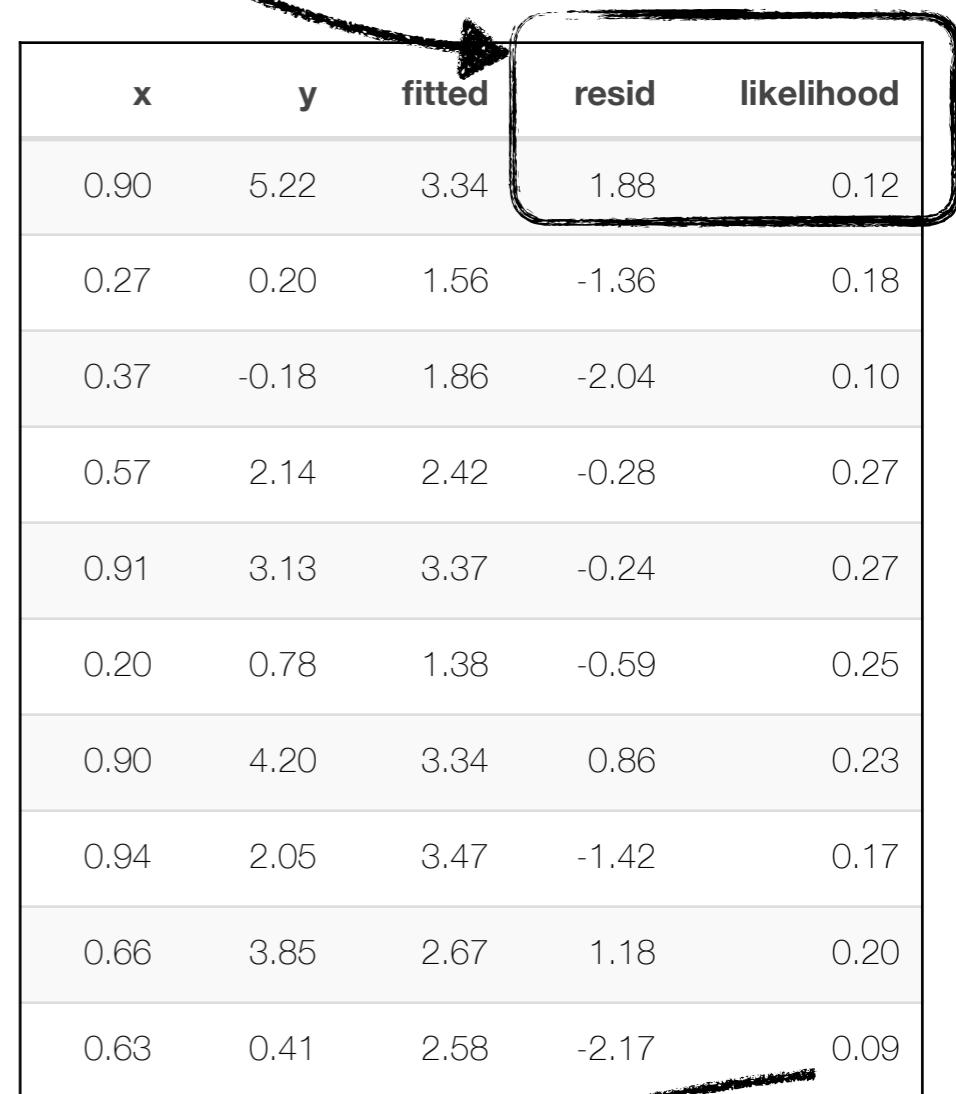
# AIC and BIC

```

1 # generate some data
2 df.like = tibble(
3   x = runif(20, min = 0, max = 1),
4   y = 1 + 3 * x + rnorm(20, sd = 2)
5 )
6
7 # fit the model
8 fit = lm(formula = y ~ x,
9           data = df.like)
10
11 # model summary
12 fit %>%
13   glance()

```

`dnorm(1.88, mean = 0, sd = 1.45) = 0.12`



x	y	fitted	resid	likelihood
0.90	5.22	3.34	1.88	0.12
0.27	0.20	1.56	-1.36	0.18
0.37	-0.18	1.86	-2.04	0.10
0.57	2.14	2.42	-0.28	0.27
0.91	3.13	3.37	-0.24	0.27
0.20	0.78	1.38	-0.59	0.25
0.90	4.20	3.34	0.86	0.23
0.94	2.05	3.47	-1.42	0.17
0.66	3.85	2.67	1.18	0.20
0.63	0.41	2.58	-2.17	0.09

**inferred standard deviation of the error**

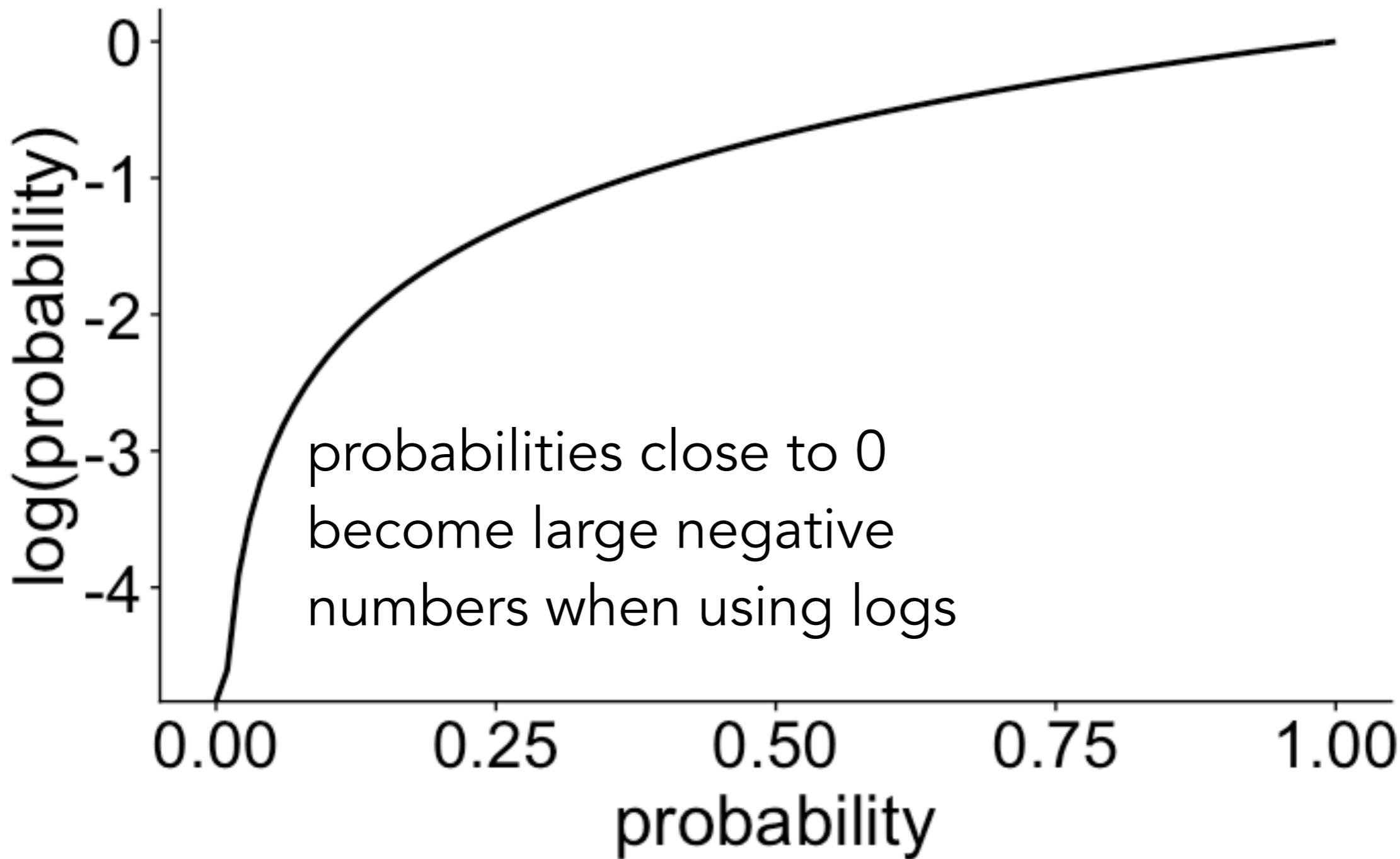
$$\sum_{i=1}^n \ln(\text{likelihood})$$



r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual
0.25	0.21	1.45	6.16	0.02	2	-34.74	75.47	78.46	37.77	18

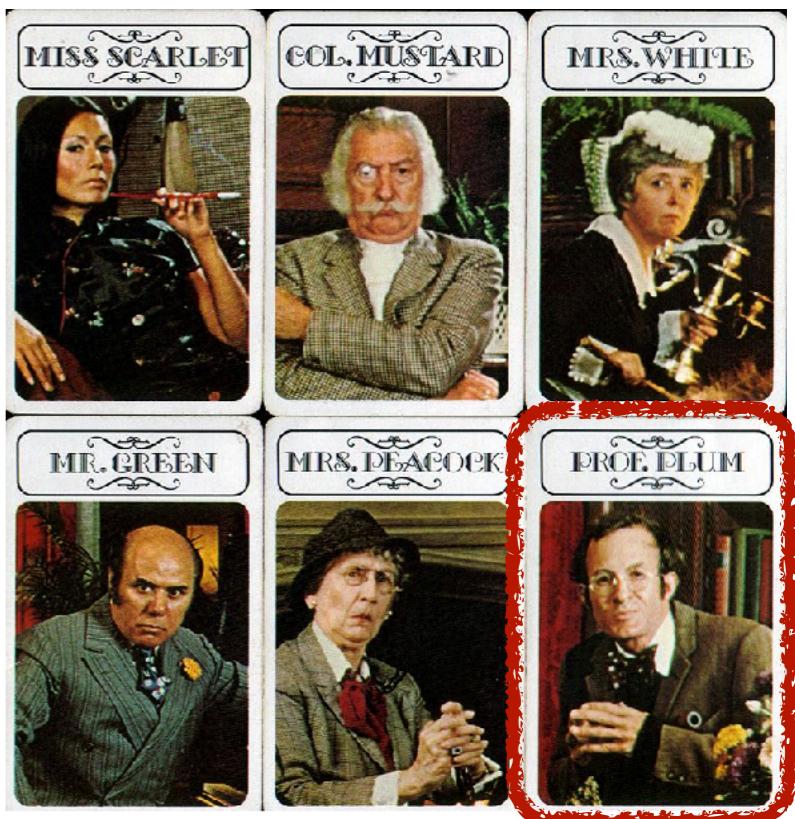
$e \sim \mathcal{N}(\text{mean} = 0, \text{sd} = 1.45)$

# `log()` is your friend!



# Clue guide to probability

Who?



- joint probability:

- if A and B are independent then

- Definition:  $p(A, B) = p(A) \cdot p(B)$

- $p(\text{Prof Plum, candle stick}) =$

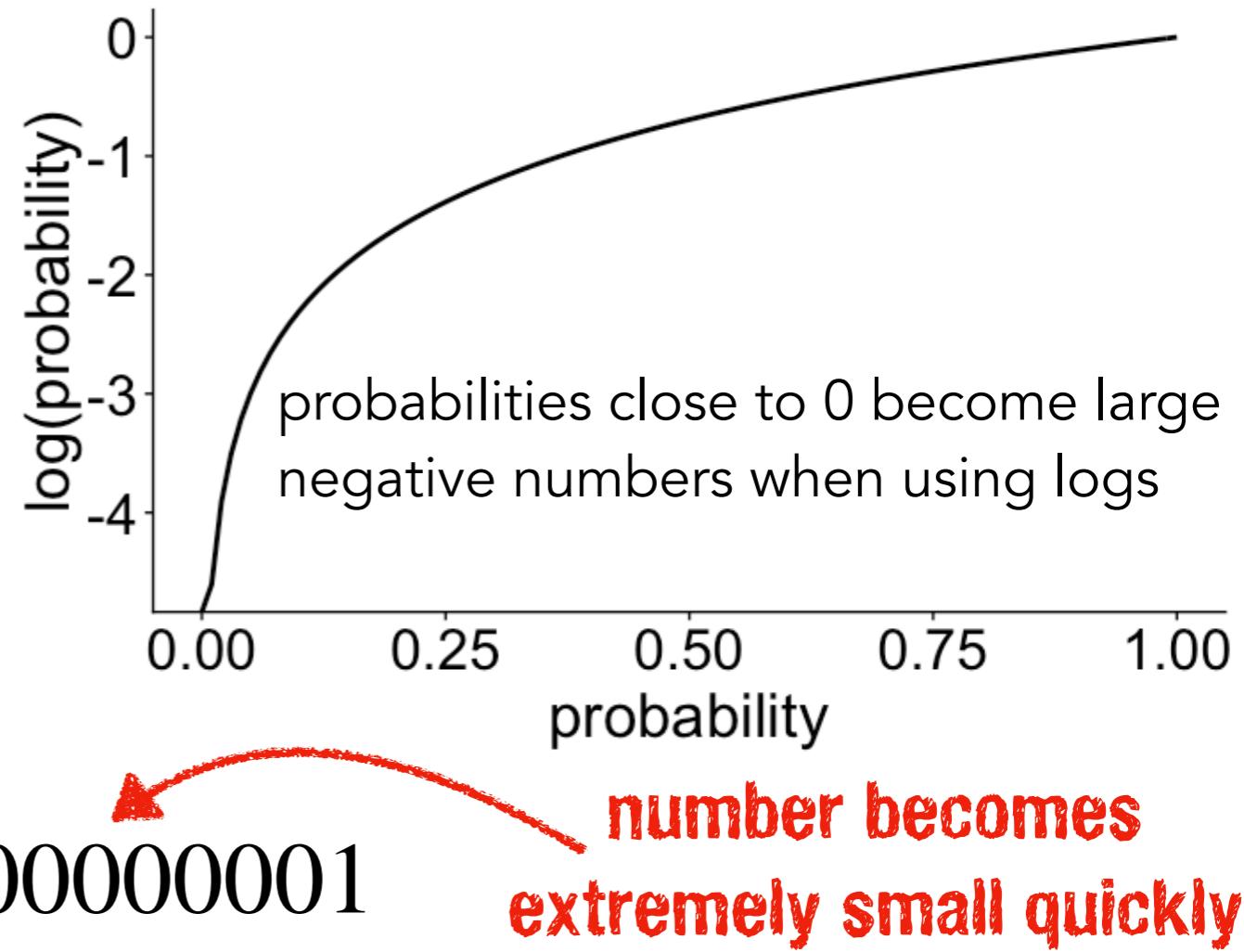
$$p(\text{Prof Plum}) \cdot p(\text{candle stick}) =$$

$$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

What?



# `log()` is your friend!



## multiplying probabilities

$$0.01 \cdot 0.01 \cdot 0.01 \cdot 0.01 = 0.00000001$$

## take `log()`

$$\log(0.01) = -4.60517$$

## summing logs

$$(-4.60517) + (-4.60517) + (-4.60517) + (-4.60517) = -18.42068$$

## transform back into probability

$$\exp(-18.42068) = 0.00000001$$

often not necessary since we just use `logLikelihood`

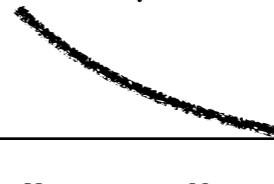
# AIC and BIC

```

1 # generate some data
2 df.like = tibble(
3   x = runif(20, min = 0, max = 1),
4   y = 1 + 3 * x + rnorm(20, sd = 2)
5 )
6
7 # fit the model
8 fit = lm(formula = y ~ x,
9           data = df.like)
10
11 # model summary
12 fit %>%
13   glance()

```

`dnorm(1.88, mean = 0, sd = 1.45) = 0.12`



x	y	fitted	resid	likelihood
0.90	5.22	3.34	1.88	0.12
0.27	0.20	1.56	-1.36	0.18
0.37	-0.18	1.86	-2.04	0.10
0.57	2.14	2.42	-0.28	0.27
0.91	3.13	3.37	-0.24	0.27
0.20	0.78	1.38	-0.59	0.25
0.90	4.20	3.34	0.86	0.23
0.94	2.05	3.47	-1.42	0.17
0.66	3.85	2.67	1.18	0.20
0.63	0.41	2.58	-2.17	0.09

**inferred standard deviation of the error**




r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual
0.25	0.21	1.45	6.16	0.02	2	-34.74	75.47	78.46	37.77	18

$e \sim \mathcal{N}(\text{mean} = 0, \text{sd} = 1.45)$

# AIC and BIC

$$\text{AIC} = 2k - 2 \ln(\hat{L})$$

$$\text{BIC} = \ln(n)k - 2 \ln(\hat{L})$$

$\ln(\hat{L})$  = maximized value of the likelihood function of the model **-34.74**

$k$  = number of parameters in the model **3**

$n$  = number of observations **20**

the sd of the normal distribution modeling the residuals counts as a parameter

```
lm(formula = y ~ 1 + x, data = df.example)
```

r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual
0.25	0.21	1.45	6.16	0.02	2	-34.74	75.47	78.46	37.77	18

# AIC and BIC

$$\text{AIC} = 2k - 2 \ln(\hat{L}) = 2 \cdot 3 - 2 \cdot (-34.74) = 75.47$$

$$\text{BIC} = \ln(n)k - 2 \ln(\hat{L}) = \ln(20) \cdot 3 - 2 \cdot (-34.74) = 78.46$$

$\ln(\hat{L})$  = maximized value of the likelihood function of the model **-34.74**

$k$  = number of parameters in the model **3**

$n$  = number of observations **20**

the sd of the normal distribution modeling the residuals counts as a parameter

```
lm(formula = y ~ 1 + x, data = df.example)
```

r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance	df.residual
0.25	0.21	1.45	6.16	0.02	2	-34.74	75.47	78.46	37.77	18

# AIC and BIC

$$\text{AIC} = 2k - 2 \ln(\hat{L})$$

$$\text{BIC} = \ln(n)k - 2 \ln(\hat{L})$$

- for both AIC and BIC, *lower* is better!
- neither provide a test of a model in the sense of testing a null hypothesis
  - AIC or BIC tell us nothing about the absolute quality of a model, only the quality relative to other models
- BIC generally penalizes free parameters more strongly than AIC (though it depends on the size of  $n$ )

$\Delta\text{BIC}$	Evidence against higher BIC
0 to 2	Not worth more than a bare mention
2 to 6	Positive
6 to 10	Strong
>10	Very Strong

# What shall I use when?

- Use it all!
- ideally, the different measures provide converging evidence

**Table 2**

Summary of the model results. Values for  $r$  and RMSE indicate means (with 5% and 95% quantiles in parentheses) based on 100 split-half cross-validation runs. BIC scores are based on running the models on the full data set.

Model	$r$	RMSE	BIC
Difference & pivotality	.86 (.66, .95)	10.56 (6.17, 17.21)	158.59
Difference	.70 (.30, .90)	26.92 (16.4, 40.6)	209.74
Pivotality	.63 (.41, .77)	14.23 (11.39, 17.54)	199.53
Optimality	.66 (.42, .84)	14.55 (10.54, 17.91)	199.47

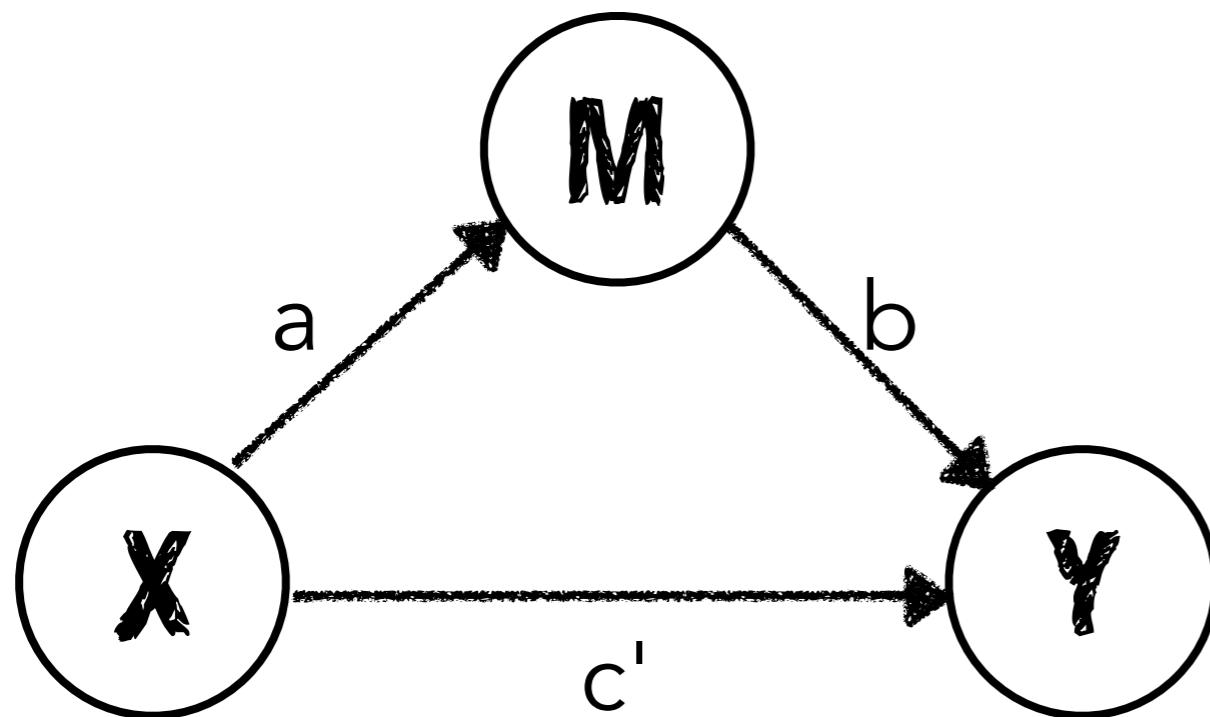
Note: BIC = Bayesian Information Criterion (lower values indicate better model performance).

# Plan for today

- Model comparison
  - Cross-validation
  - AIC and BIC
- **Mediation**
- Moderation
- Linear mixed effects model
  - modeling dependence in data

# **Mediation**

# Definition

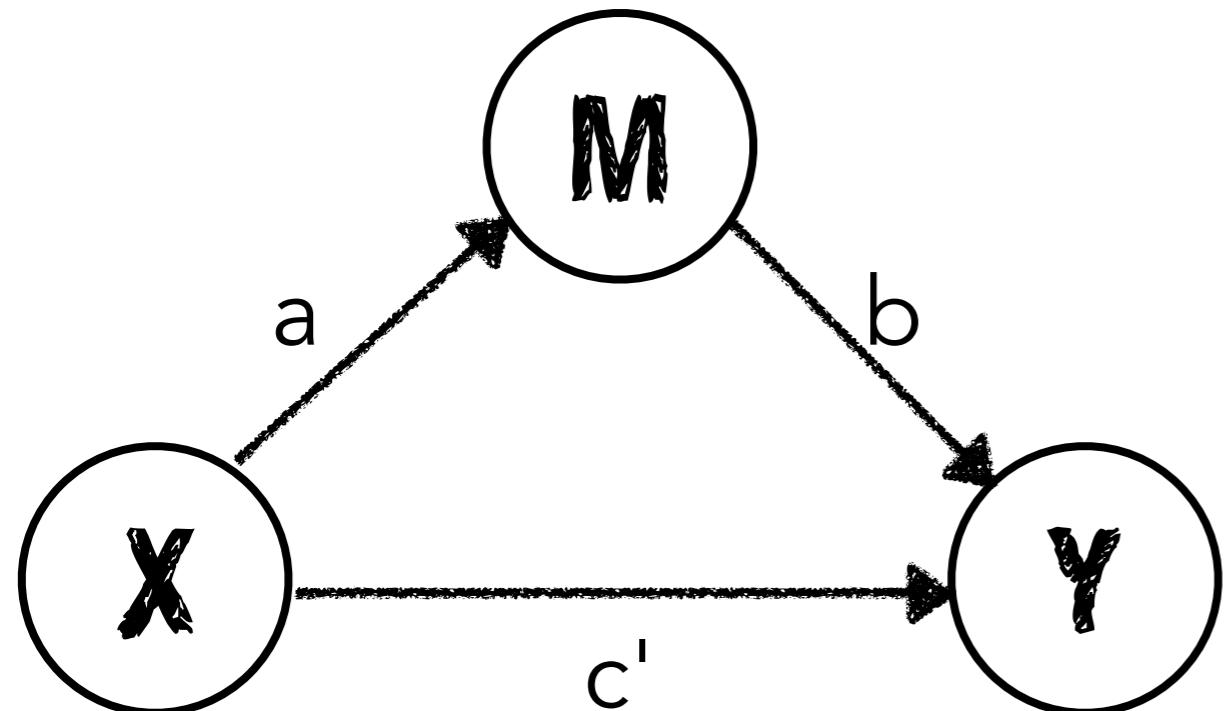


Rather than a direct causal relationship between **X** and **Y**, a mediation model proposes that **X** influences the mediator variable **M**, which in turn influences **Y**. Thus, the mediator variable serves to clarify the nature of the relationship between **X** and **Y**.

**Adapted from Wikipedia**

[https://en.wikipedia.org/wiki/Mediation\\_\(statistics\)](https://en.wikipedia.org/wiki/Mediation_(statistics))

# Example



**X** = grades in Psych 252

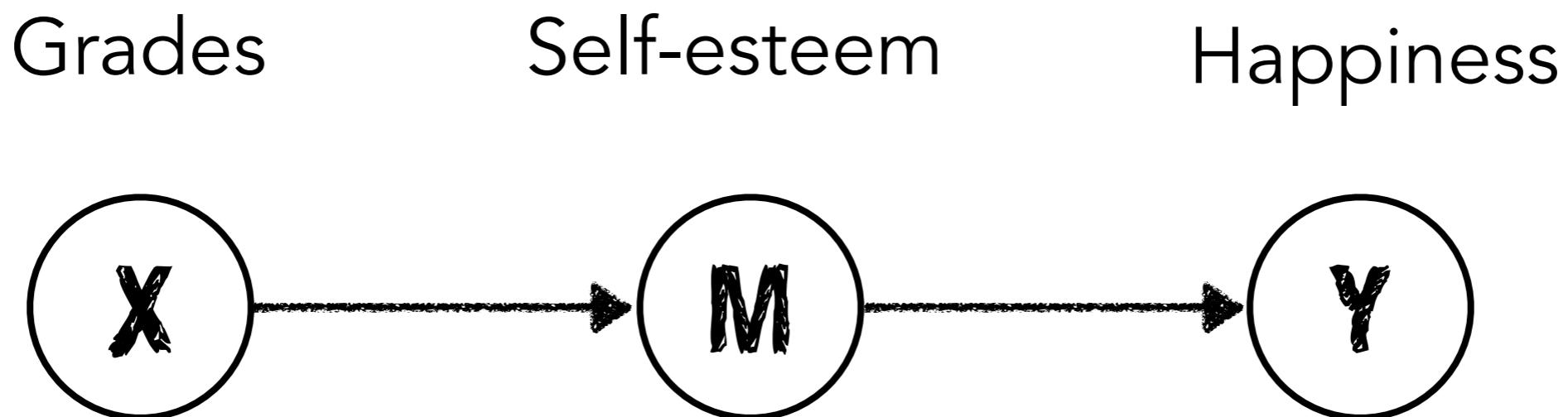
**M** = feelings of self-esteem

**Y** = happiness

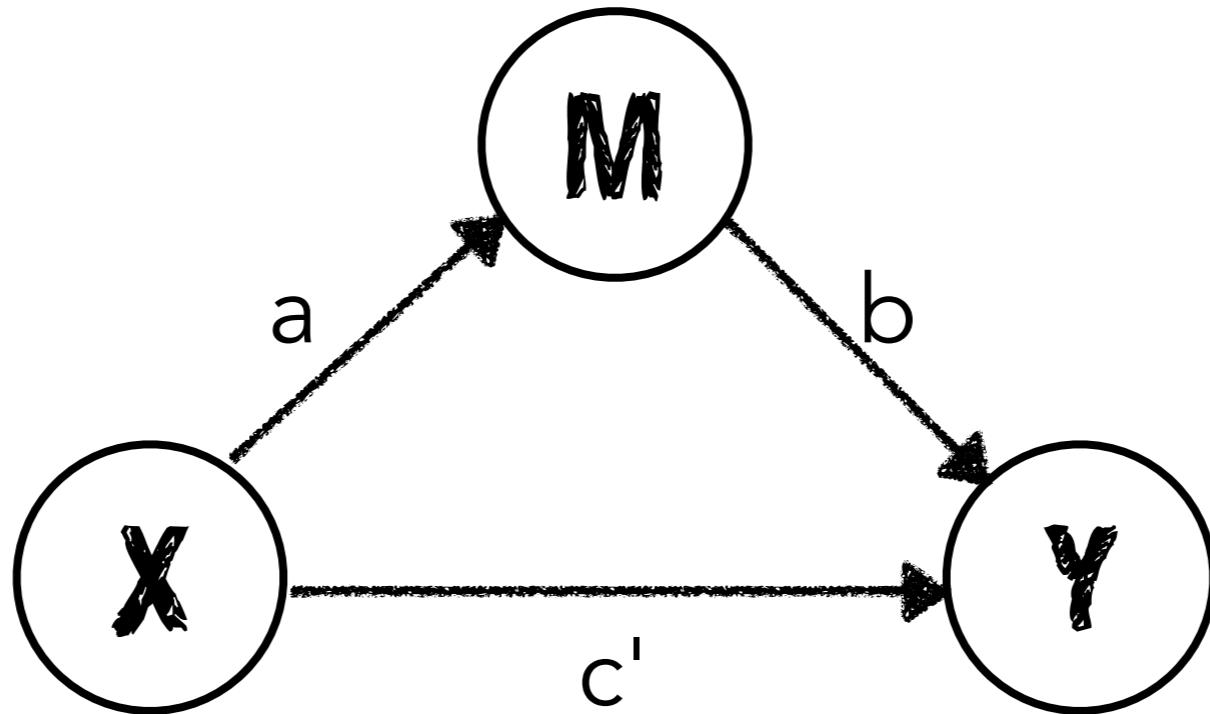
Is the relationship between grades in Psych 252 and happiness mediated by feelings of self-esteem?

# Simulate a mediation analysis

```
1 # number of participants
2 n = 100
3
4 # generate data
5 df.mediation = tibble(
6   x = rnorm(n, 75, 7),           # grades
7   m = 0.7 * x + rnorm(n, 0, 5), # self-esteem
8   y = 0.4 * m + rnorm(n, 0, 5) # happiness
9 )
```



# Baron and Kenny's (1986) steps for mediation



## Sequence of regression models

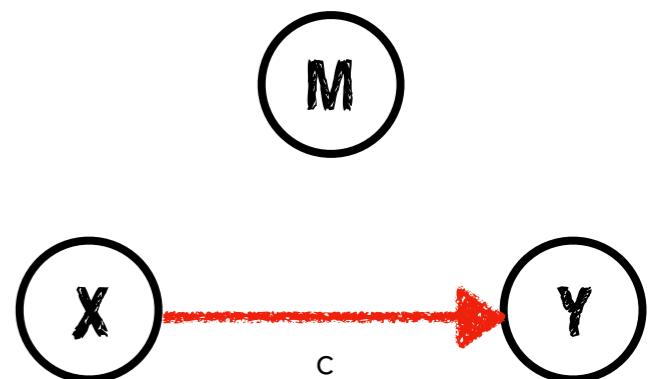
1. Is there a relationship between **X** and **Y**?
2. Is there a relationship between **X** and **M**?
3. Does the relationship between **X** and **Y** change, once we control for **M**?

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?

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

```
1 # fit the model
2 fit.y_x = lm(formula = y ~ 1 + x,
3               data = df.mediation)
4
5 # summarize the results
6 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

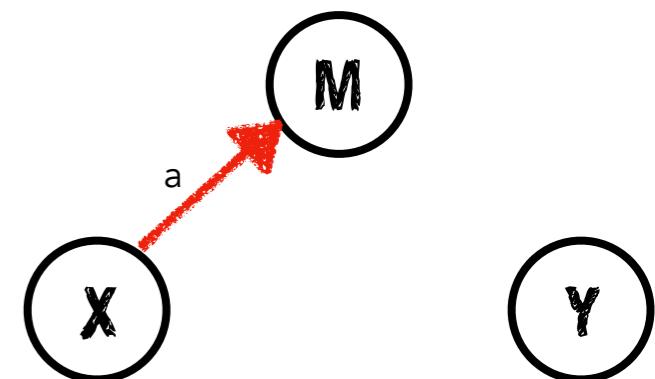
Residual standard error: 5.16 on 98 degrees of freedom
Multiple R-squared:  0.04237, Adjusted R-squared:  0.0326 
F-statistic: 4.336 on 1 and 98 DF,  p-value: 0.03993
```

significant  
relationship

# Is there a relationship between X and M?

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

```
1 # fit the model
2 fit.m_x = lm(formula = m ~ 1 + x,
3               data = df.mediation)
4
5 # summarize the results
6 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 

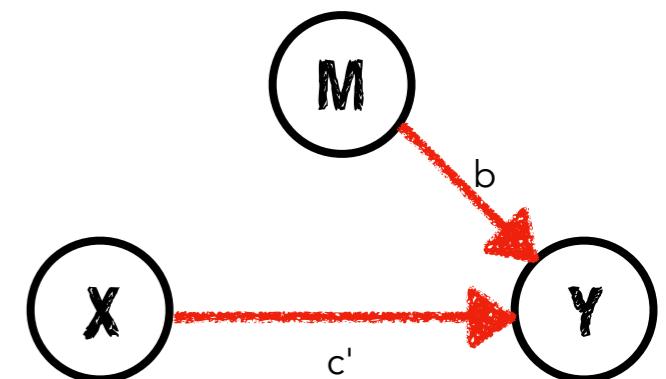
Residual standard error: 4.854 on 98 degrees of freedom
Multiple R-squared:  0.4346,    Adjusted R-squared:  0.4288 
F-statistic: 75.31 on 1 and 98 DF,  p-value: 8.872e-14
```

**significant  
relationship**

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

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

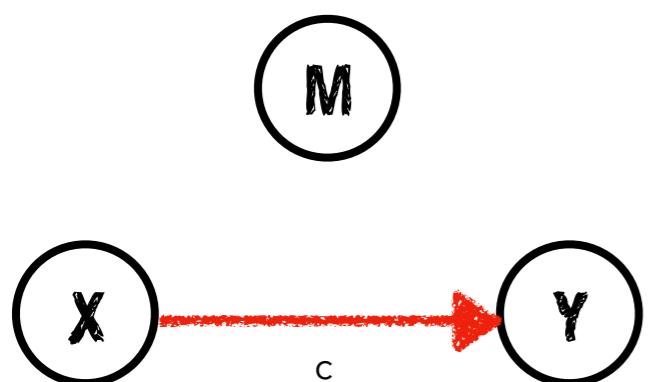
```
1 # fit the model  
2 fit.y_mx = lm(formula = y ~ 1 + m + x,  
3                 data = df.mediation)  
4  
5 # summarize the results  
6 fit.y_mx %>% summary()
```



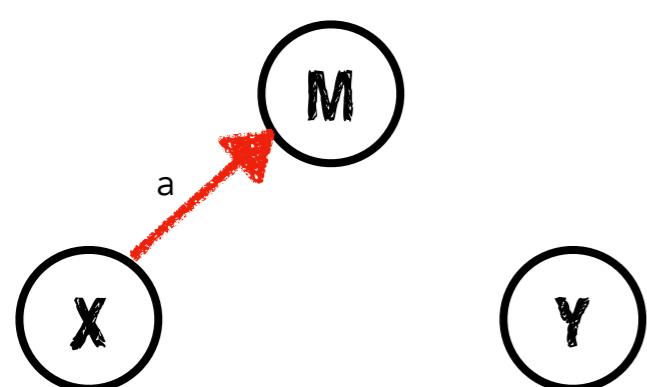
```
Call:  
lm(formula = y ~ 1 + m + x, data = df.mediation)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-9.3651 -3.3037 -0.6222  3.1068 10.3991  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 7.80952   5.68297   1.374   0.173  
m            0.42381   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**

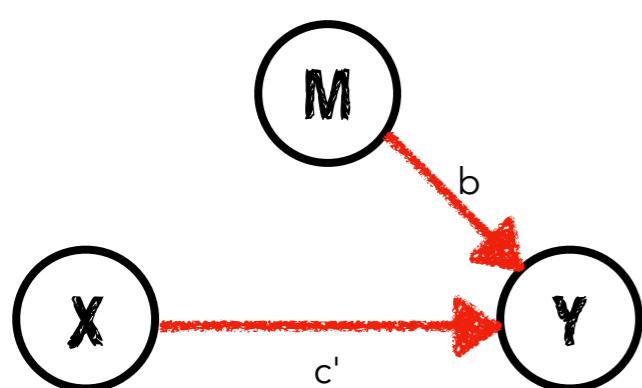
# 3 Step procedure



Relationship between X and Y?  
 $\hat{y} = b_0 + b_1 \cdot x$



Relationship between X and M?  
 $\hat{m} = b_0 + b_1 \cdot x$



Relationship between X and Y,  
controlling for M?

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

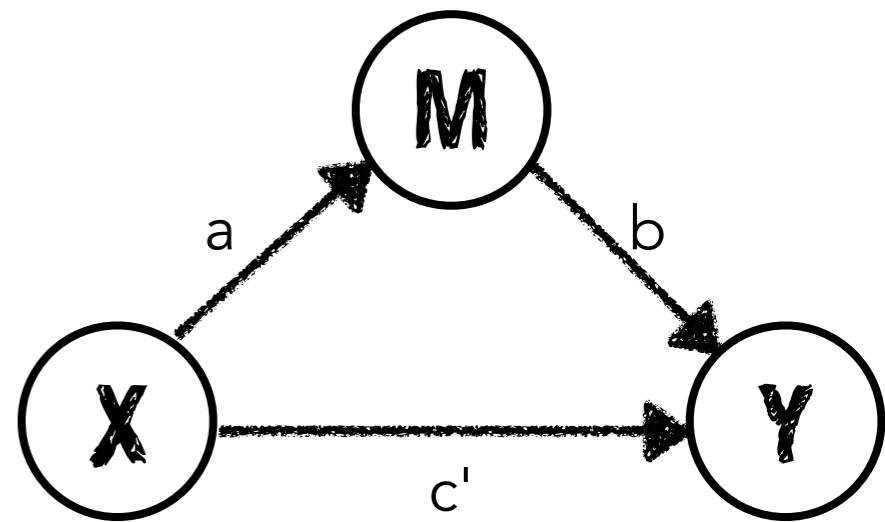
## 3. Bayesian mediation analysis



later in class!

# 1. Sobel test

```
1 library("multilevel")
2
3 # run the sobel test
4 fit.sobel = sobel(pred = df.mediation$x,
5                      med = df.mediation$m,
6                      out = df.mediation$y)
7
8 # calculate the p-value
9 (1 - pnorm(fit.sobel$z.value))2
```

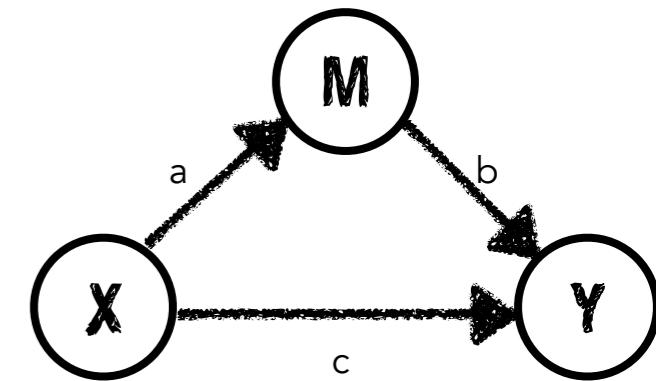


"It is becoming increasingly more difficult to publish tests of mediation based purely on the Baron and Kenny method or tests that make distributional assumptions such as the Sobel test."

A Wikipedia person

## 2. Bootstrapping

```
1 library("mediation")
2
3 # bootstrapped mediation
4 fit.mediation = mediate(model.m = fit.m_x, ←  $\hat{m} = b_0 + b_1 \cdot x$ 
5 model.y = fit.y_mx, ←  $\hat{y} = b_0 + b_1 \cdot m + b_2 \cdot x$ 
6 treat = "x",
7 mediator = "m",
8 boot = T)
9
10 # summarize results
11 fit.mediation %>% summary()
```



```
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
```

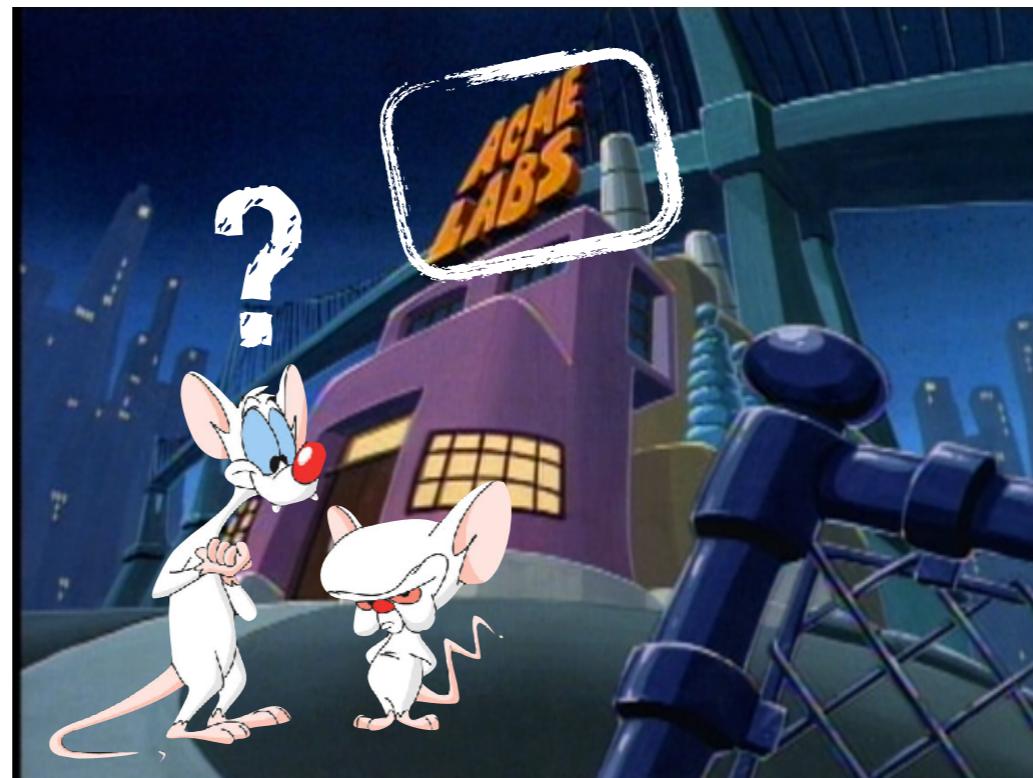
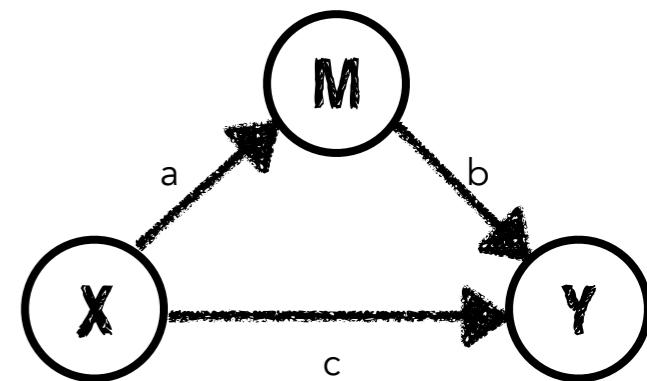
# 2. Bootstrapping

Causal Mediation Analysis

Nonparametric Bootstrap Confidence Intervals with the Percentile Method

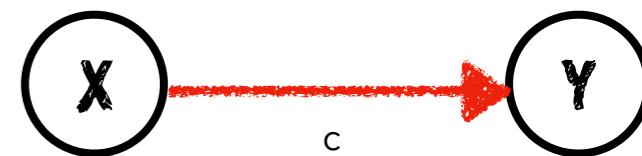
	Estimate	95% CI Lower	95% CI Upper	p-value	
ACME	0.28078	0.14059	0.42	<2e-16	***
ADE	-0.11179	-0.29276	0.10	0.272	
Total Effect	0.16899	-0.00415	0.34	0.064	.
Prop. Mediated	1.66151	-3.22476	11.46	0.064	.
---					
Signif. codes:	0 '***'	0.001 '**'	0.01 '*'	0.05 '.'	0.1 ' '
Sample Size Used:	100				

Simulations: 1000



# 2. Bootstrapping

M



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 ' '
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 *

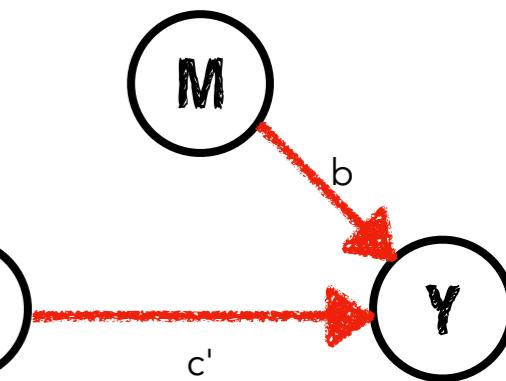
## 2. Bootstrapping

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 ' '
Sample Size Used:	100				

Simulations: 1000



$$\hat{y} = b_0 + b_1 \cdot m + b_2 \cdot x \quad \text{ADE: Average direct effect}$$

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

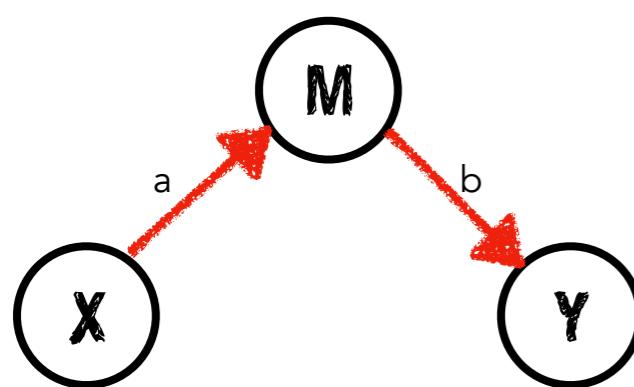
Causal Mediation Analysis

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 ' '
Sample Size Used:	100				

Simulations: 1000

**ACME**



indirect effect

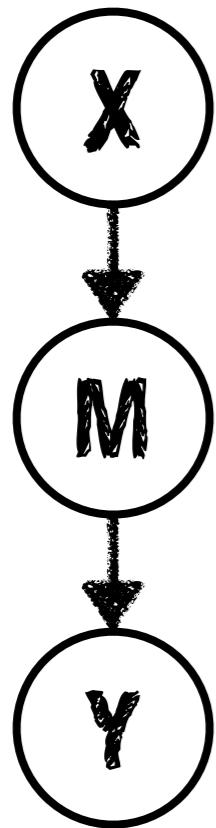
**ADE: Average direct effect**

**ACME: Average causal mediation effect**

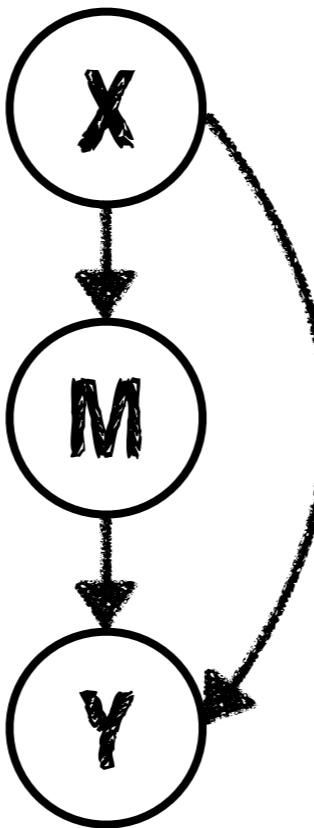
**ACME = Total effect - ADE**

# Underlying causal model

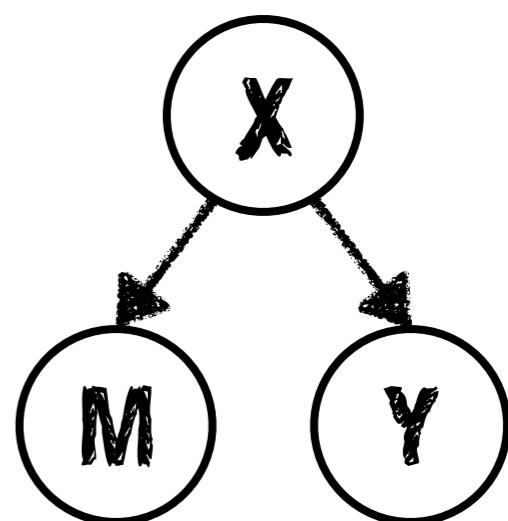
**Full mediation**



**Partial mediation**



**No mediation**

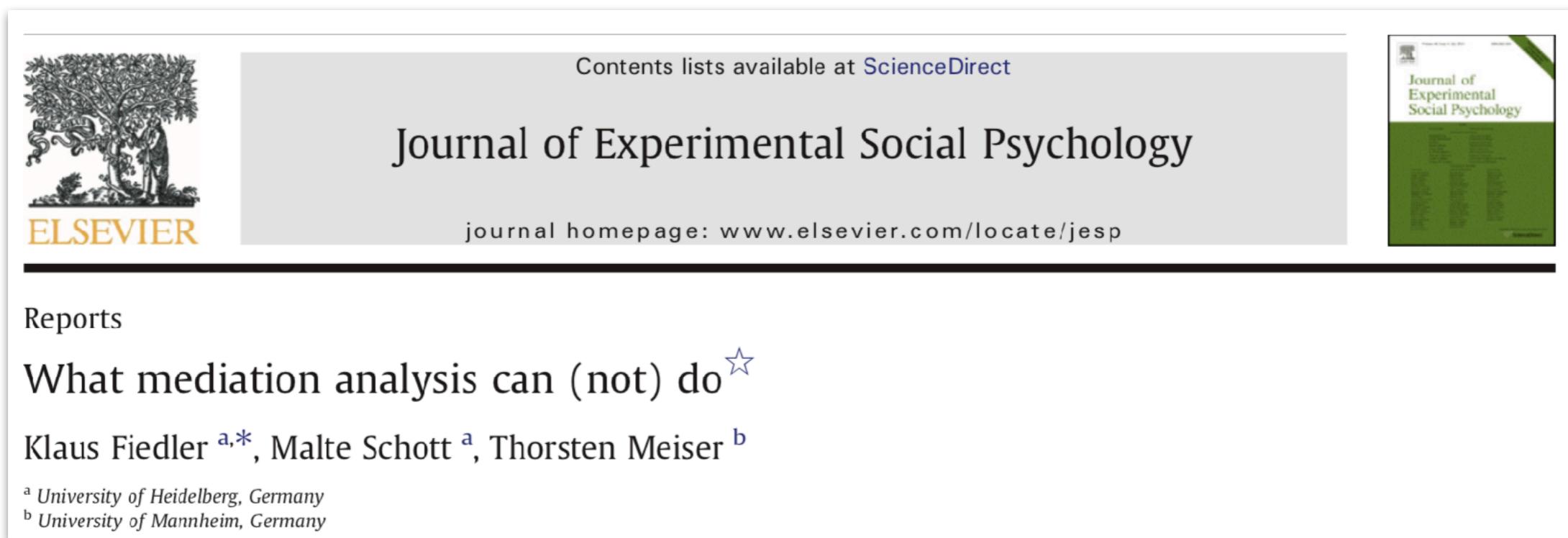


**Full mediation:** When the effect of **X** on **Y** completely disappears, **M** fully mediates between **X** and **Y**.

**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

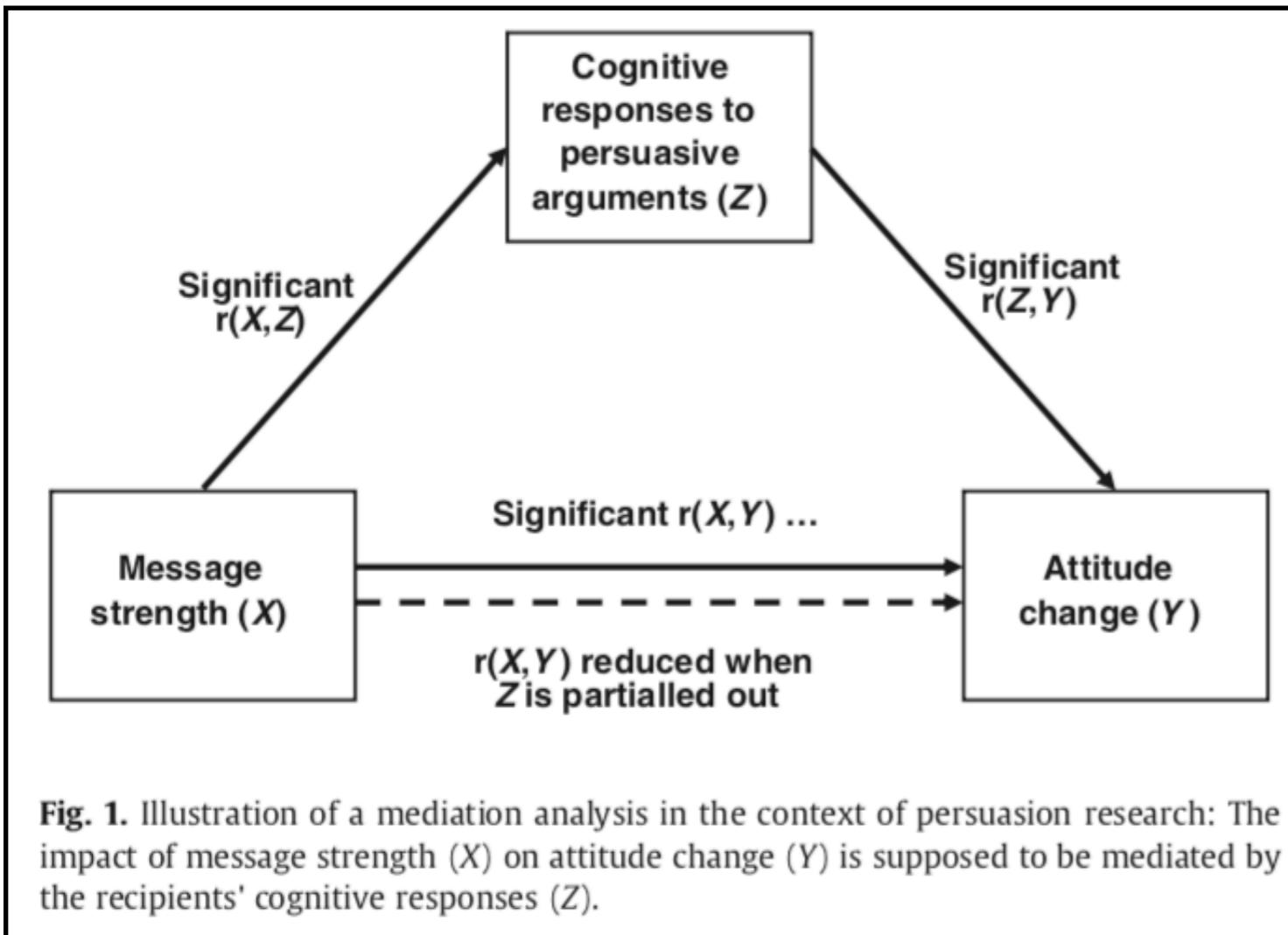


The image shows a screenshot of a journal article from the Journal of Experimental Social Psychology. At the top left is the Elsevier logo, which includes a tree and the word 'ELSEVIER'. To the right of the logo is the journal title 'Journal of Experimental Social Psychology' and its website 'journal homepage: www.elsevier.com/locate/jesp'. Above the journal title is a link 'Contents lists available at ScienceDirect'. To the right of the journal title is a small thumbnail image of the journal cover. Below the header, the word 'Reports' is visible. The main title of the article is 'What mediation analysis can (not) do<sup>☆</sup>', with authors listed as Klaus Fiedler <sup>a,\*</sup>, Malte Schott <sup>a</sup>, and Thorsten Meiser <sup>b</sup>. Footnotes indicate <sup>a</sup> University of Heidelberg, Germany and <sup>b</sup> University of Mannheim, Germany.

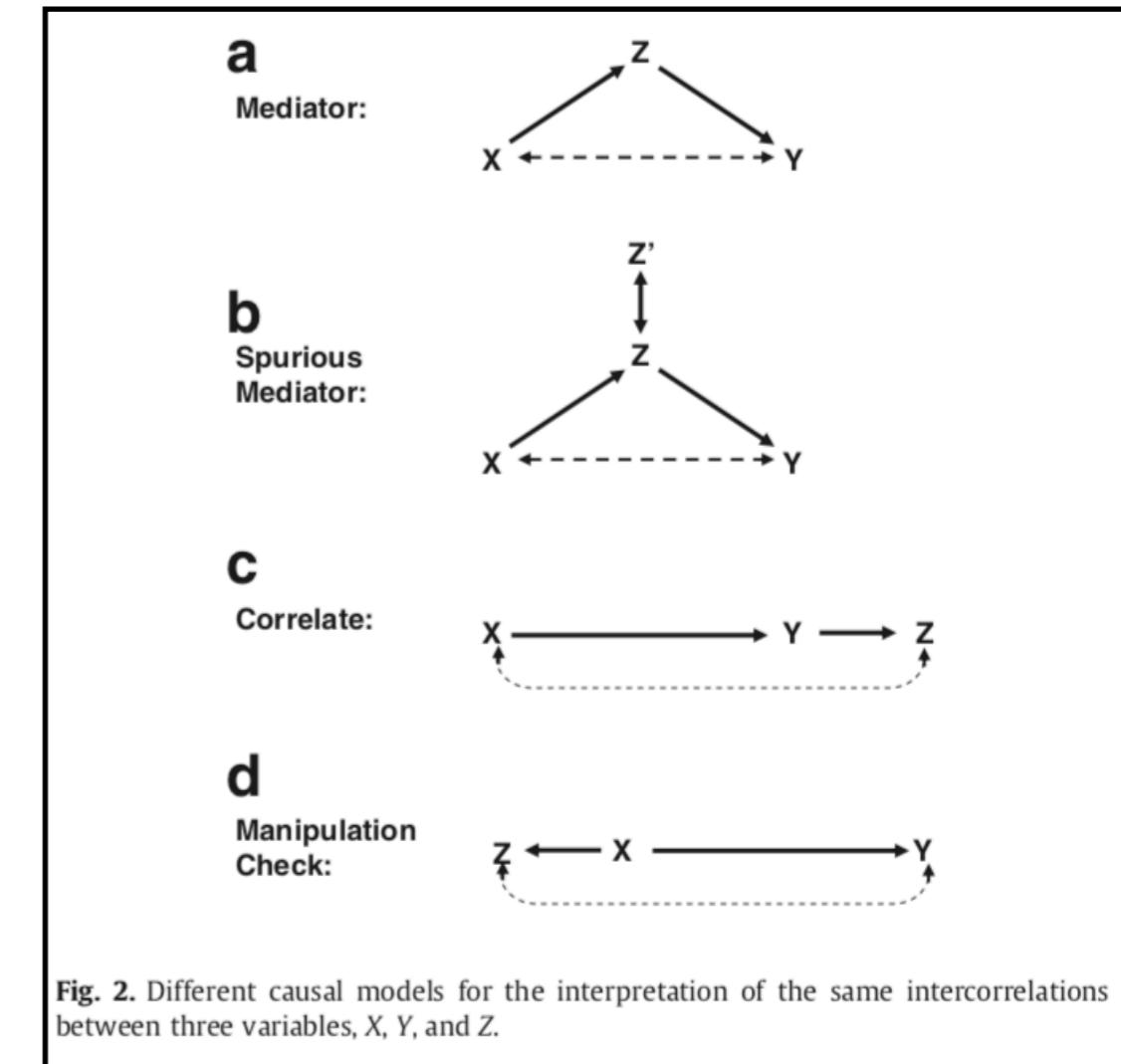
Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, 47(6), 1231-1236. 74

# 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  $Z$ .

**we need experiments to tell apart possible causal structures**

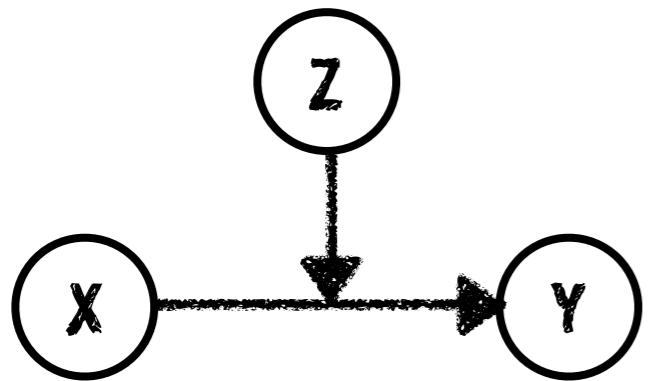
Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology*, 47(6), 1231-1236.

# Plan for today

- Model comparison
  - Cross-validation
  - AIC and BIC
- Mediation
- **Moderation**
- Linear mixed effects model
  - modeling dependence in data

# Moderation

# Definition

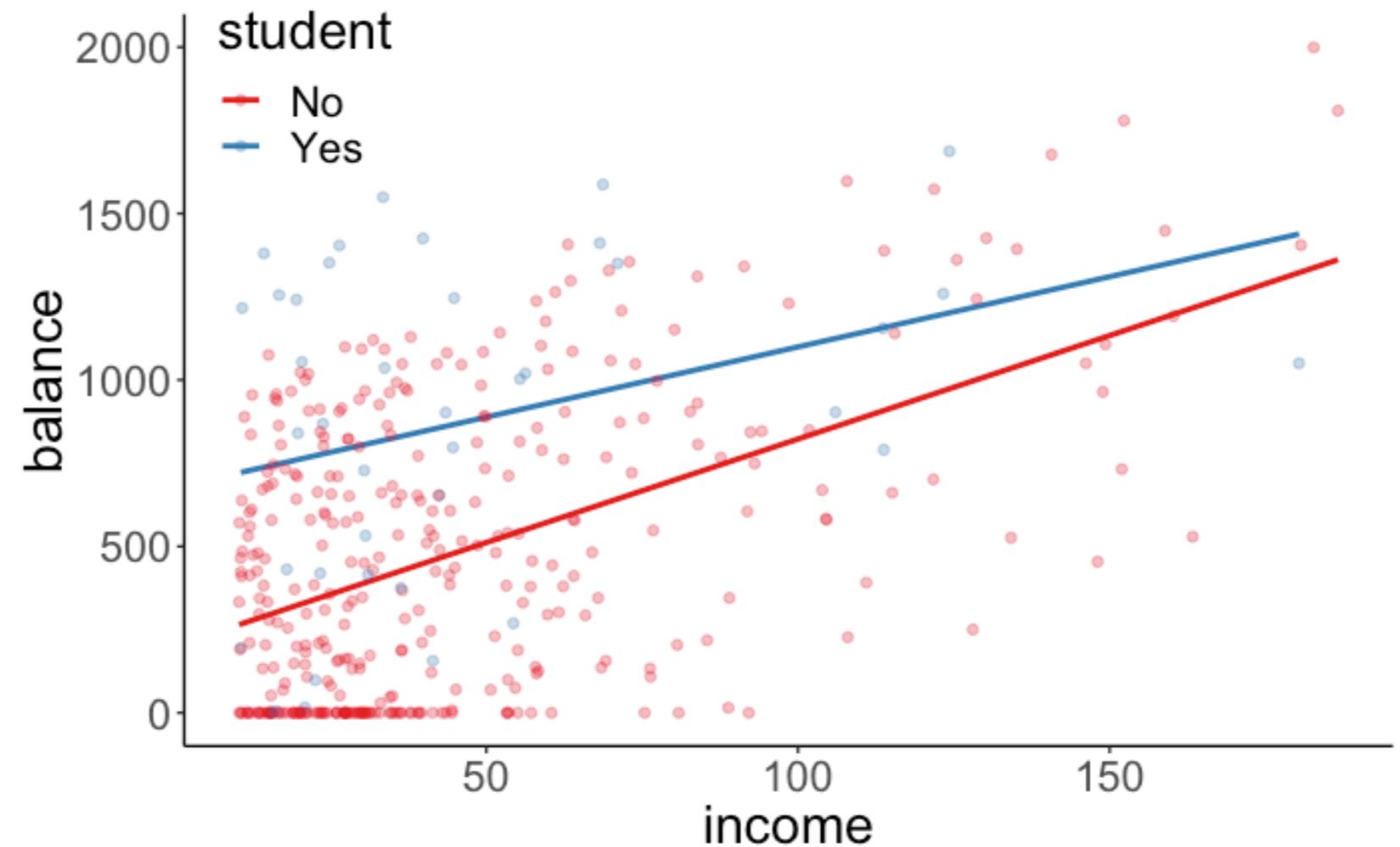


**Moderation** means that the effect of a predictor depends on the value of another.

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.



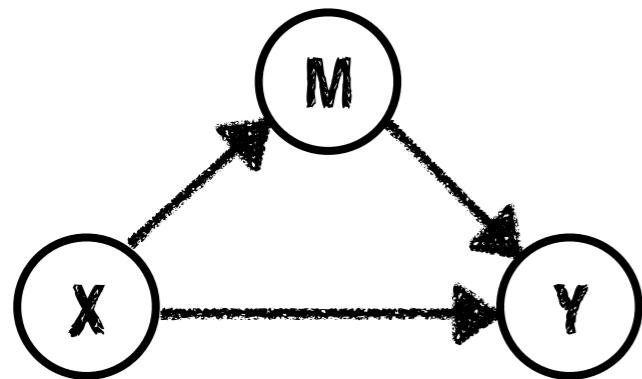
$$\widehat{\text{balance}}_i = 200.62 + 6.22 \cdot \text{income}_i + 476.68 \cdot \text{student}_i - 2.00 \cdot (\text{income}_i \times \text{student}_i)$$

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

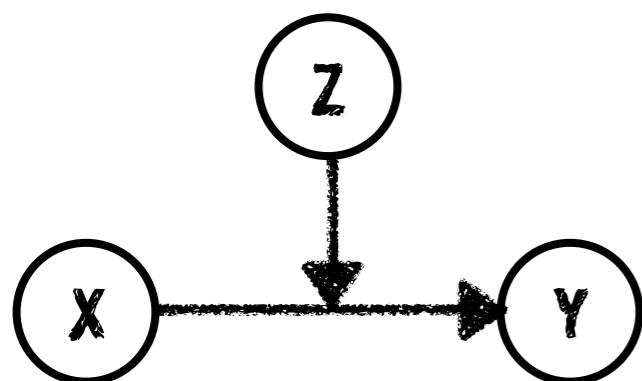
**if student = "Yes"**

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

# Mediation and moderation

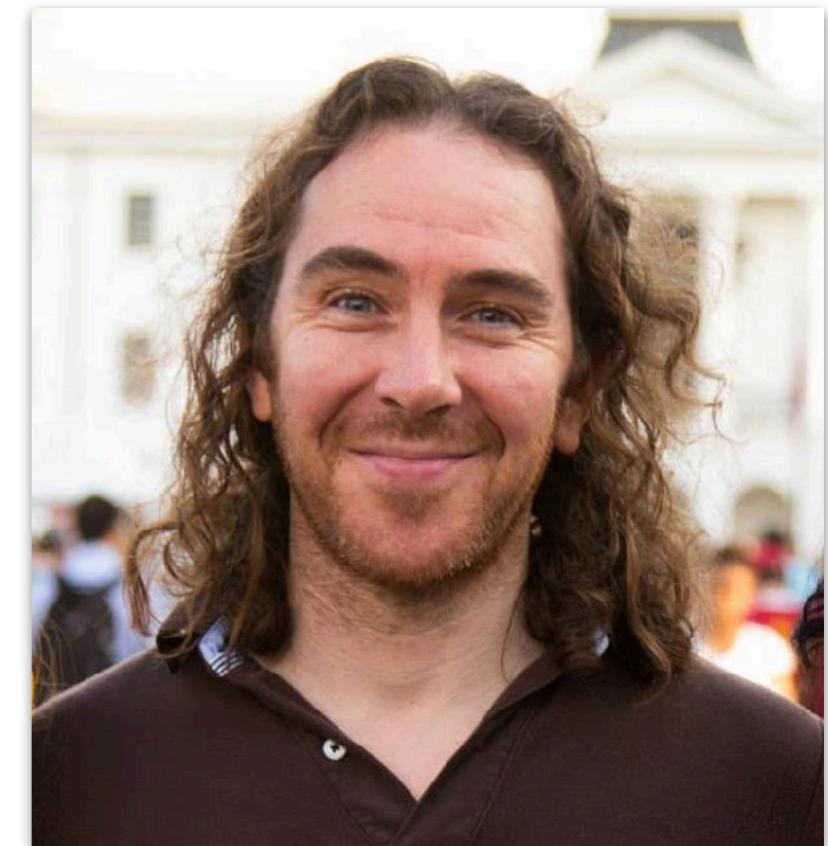
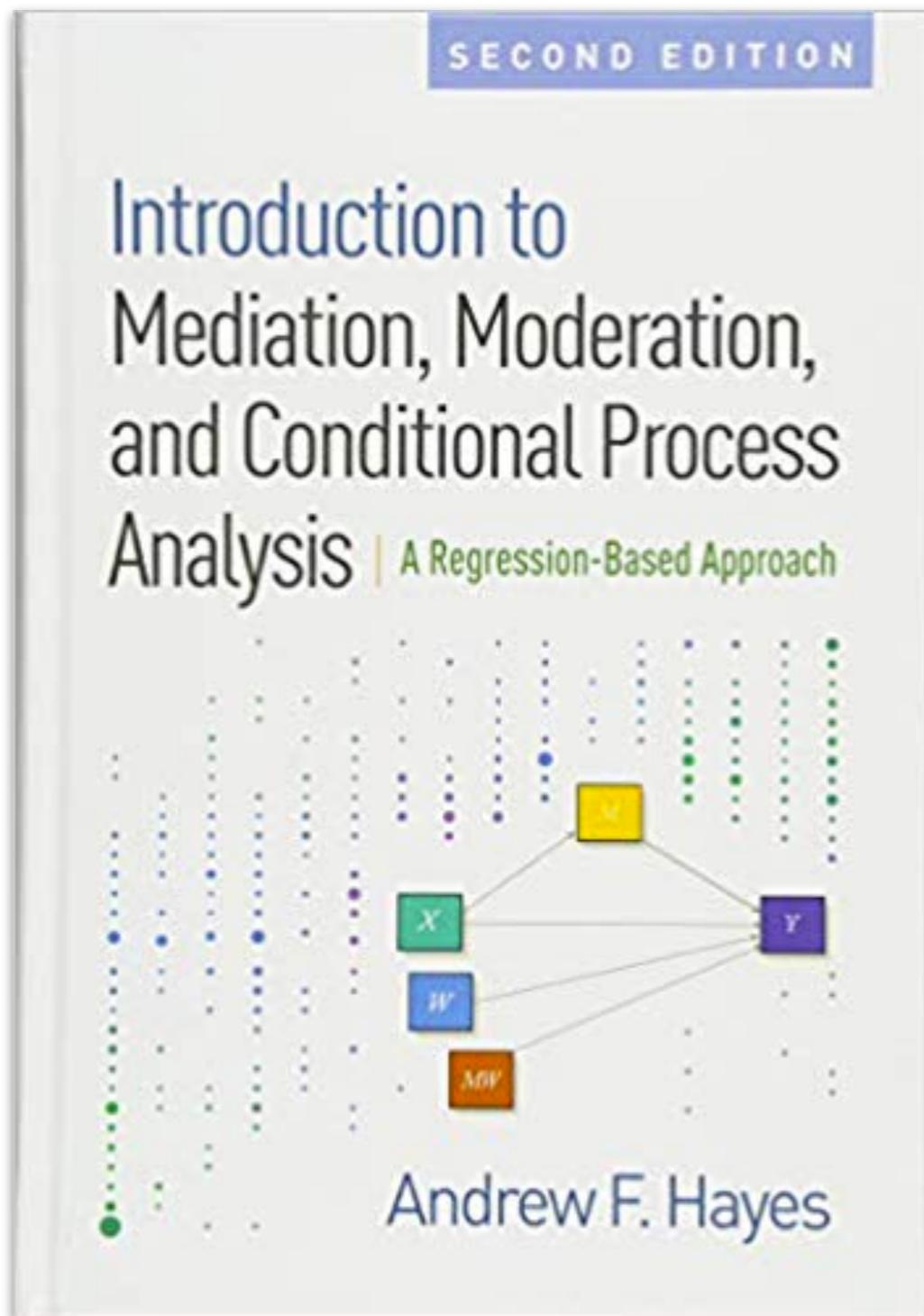


**Mediation** means that the effect of X on Y is (partially) indirect via another variable M.



**Moderation** means that the effect of a X on Y depends on the value of another variable Z.

# Learn more about mediation and moderation



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

# Plan for today

- Model comparison
  - Cross-validation
  - AIC and BIC
- Mediation
- Moderation
- **Linear mixed effects model**
  - **modeling dependence in data**

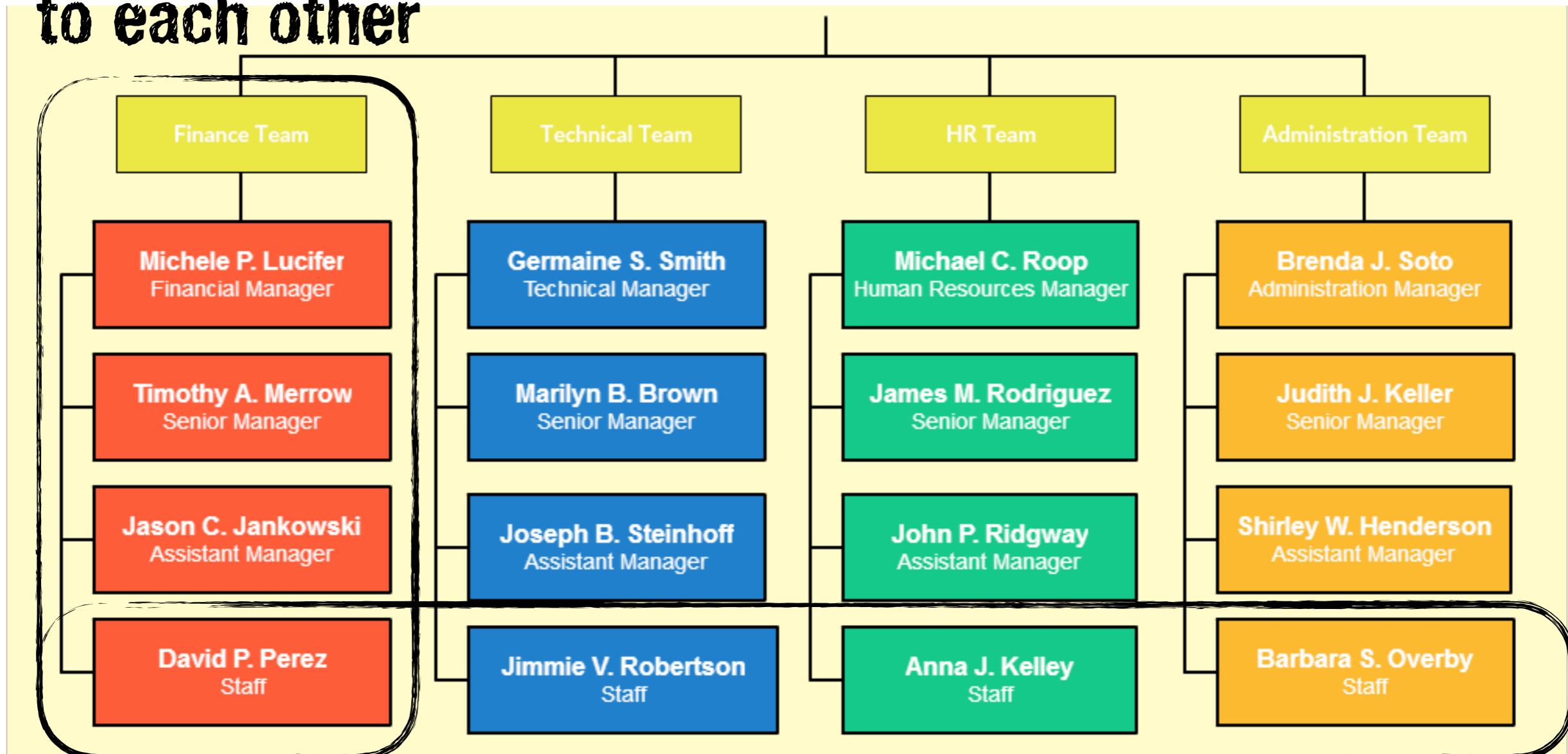
# **Linear mixed effects models**

# Dependence

- so far, all the models that we've discussed (linear model with different kinds of predictors and contrasts) make the assumption that the data are **iid** (independent, and identically distributed)
- often this assumption is violated
  - **psychology experiments**: many observations from the same participants
  - **survey data**: different populations between different states in the US
  - **time series**: distribution at  $t + 1$  depends on  $t$

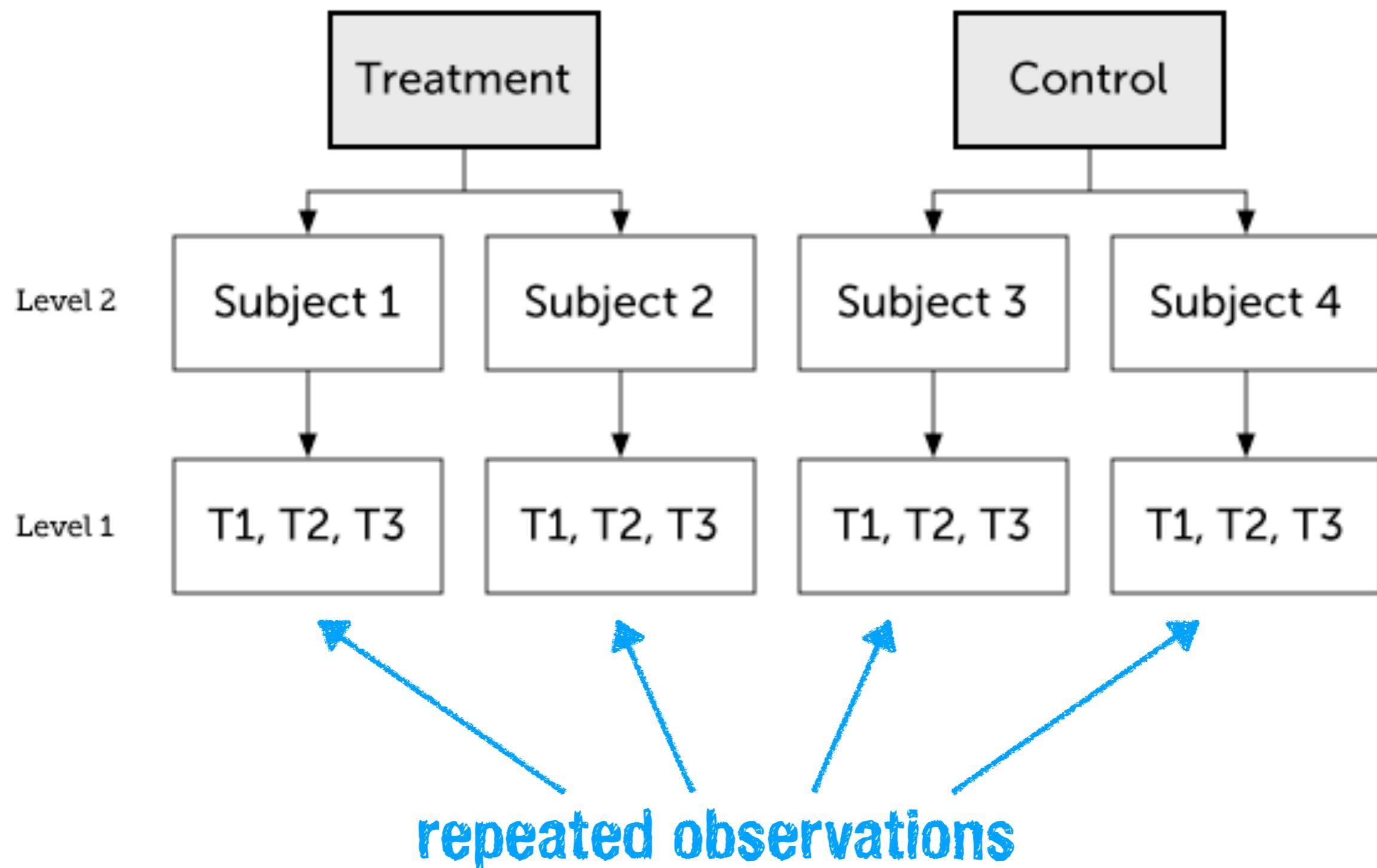
# Main use cases: Hierarchical models

more similar  
to each other



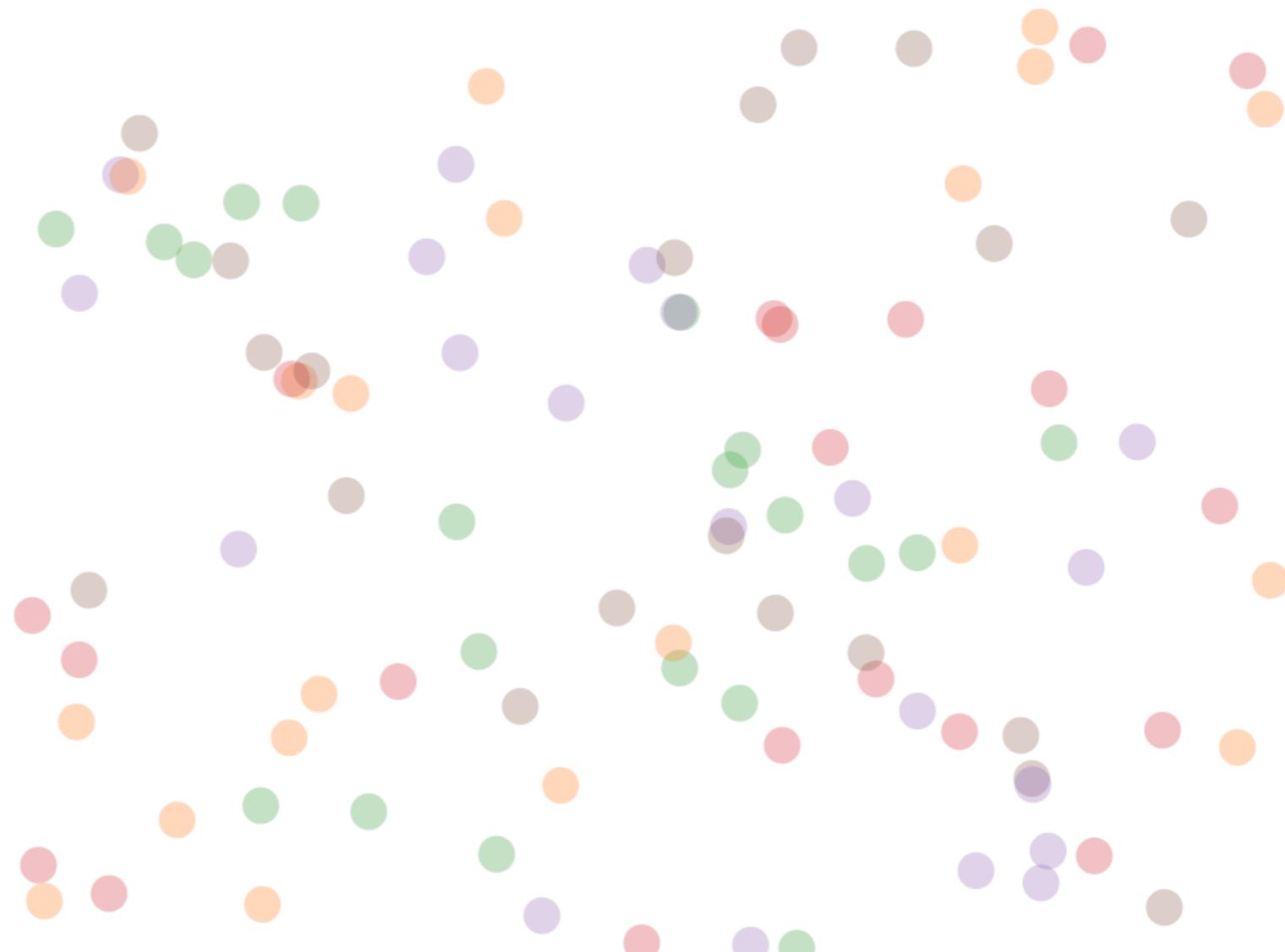
less similar  
to each other

# Main use cases: Longitudinal models



# An Introduction to Hierarchical Modeling

This visual explanation introduces the statistical concept of **Hierarchical Modeling**, also known as *Mixed Effects Modeling* or by [these other terms](#). This is an approach for modeling **nested data**. Keep reading to learn how to translate an understanding of your data into a hierarchical model specification.

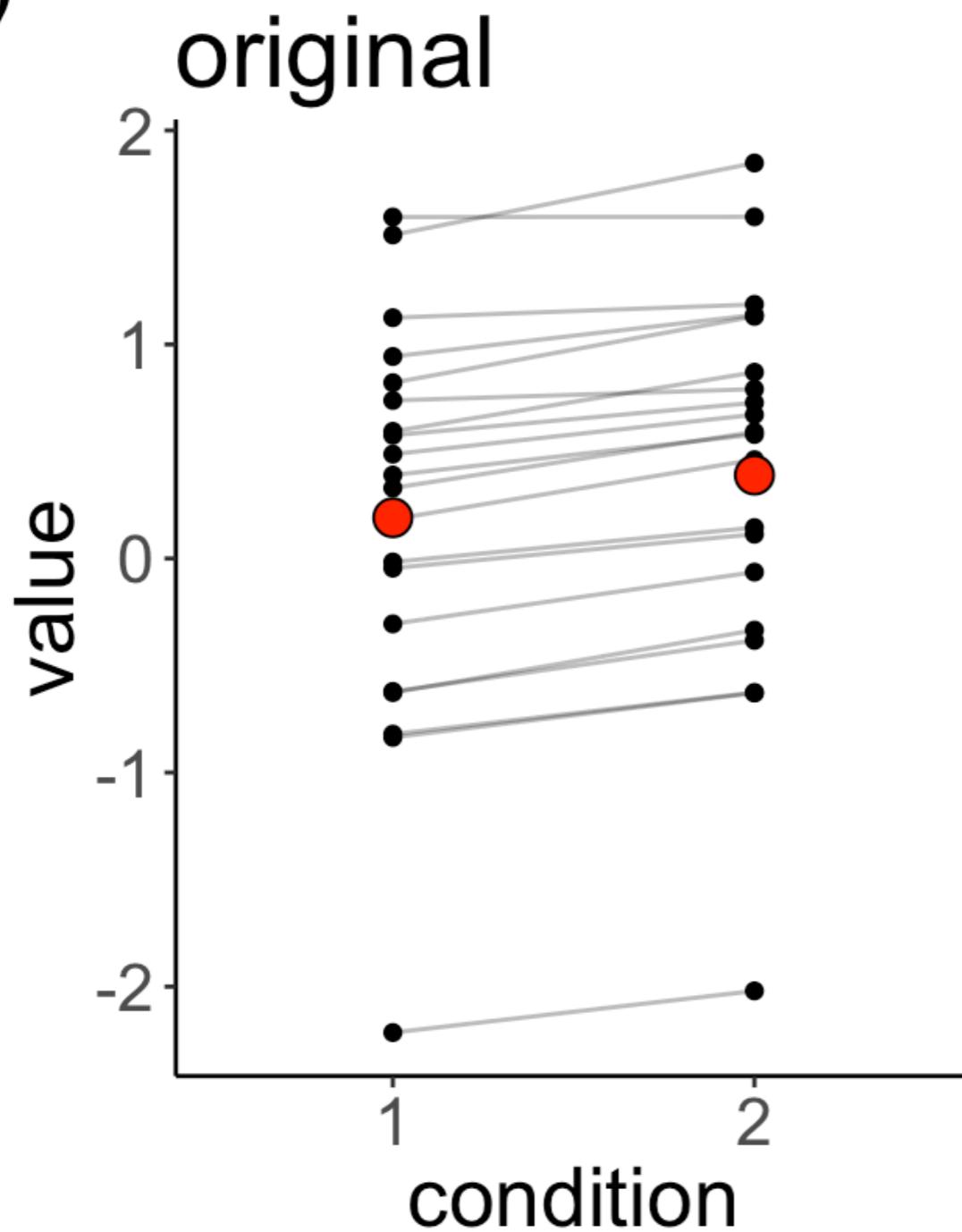


# Dependence

Does it really matter?

Is there a significant difference  
between conditions 1 and 2?

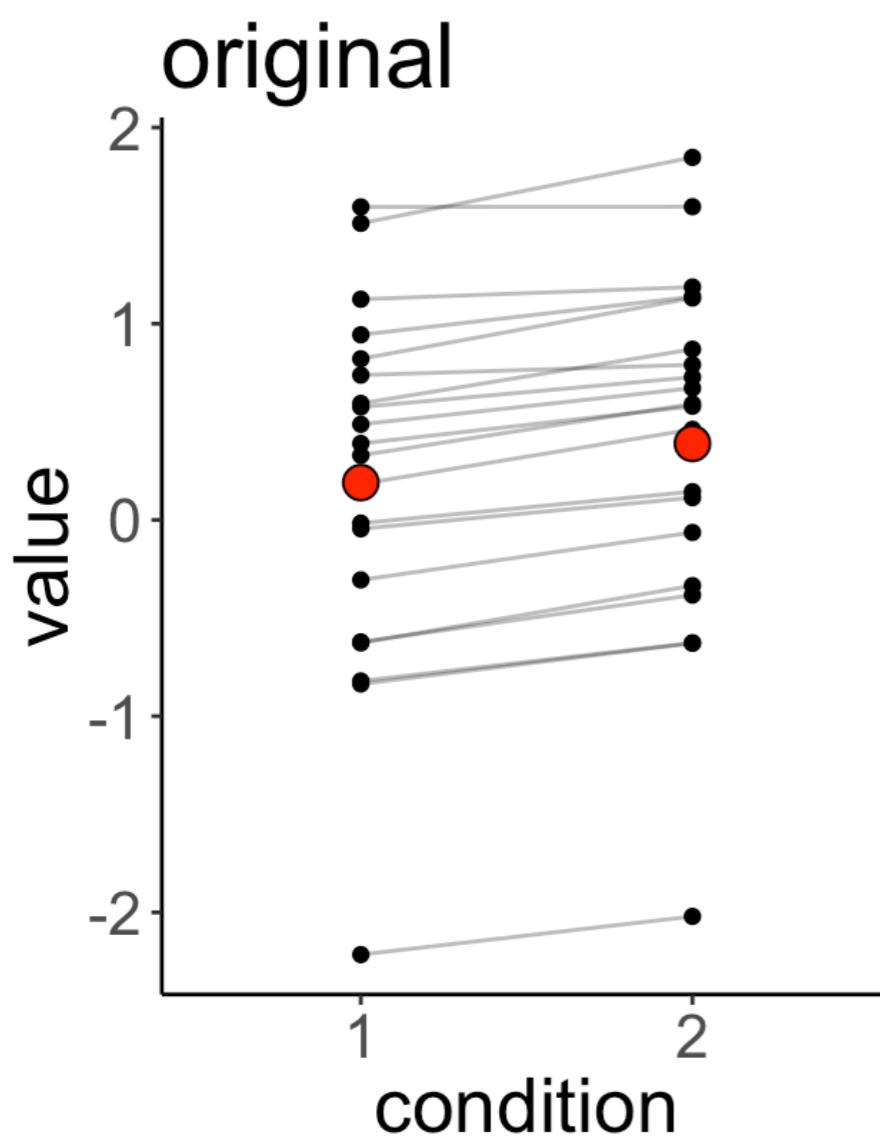
a)



# Dependence

assuming independence!

```
1 # linear model
2 lm(formula = value ~ condition,
3     data = df.original) %>%
4 summary()
```



```
Call:
lm(formula = value ~ condition, data = df.original)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.4100 -0.5530  0.1945  0.5685  1.4578 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.1905    0.2025   0.941   0.353    
condition2  0.1994    0.2864   0.696   0.491    
                                                        
Residual standard error: 0.9058 on 38 degrees of freedom
Multiple R-squared:  0.01259,    Adjusted R-squared:  -0.0134 
F-statistic: 0.4843 on 1 and 38 DF,  p-value: 0.4907
```

- we ignore the fact that we have repeated observations from the same participants
- in the data it looks like there is a small but consistent effect of condition

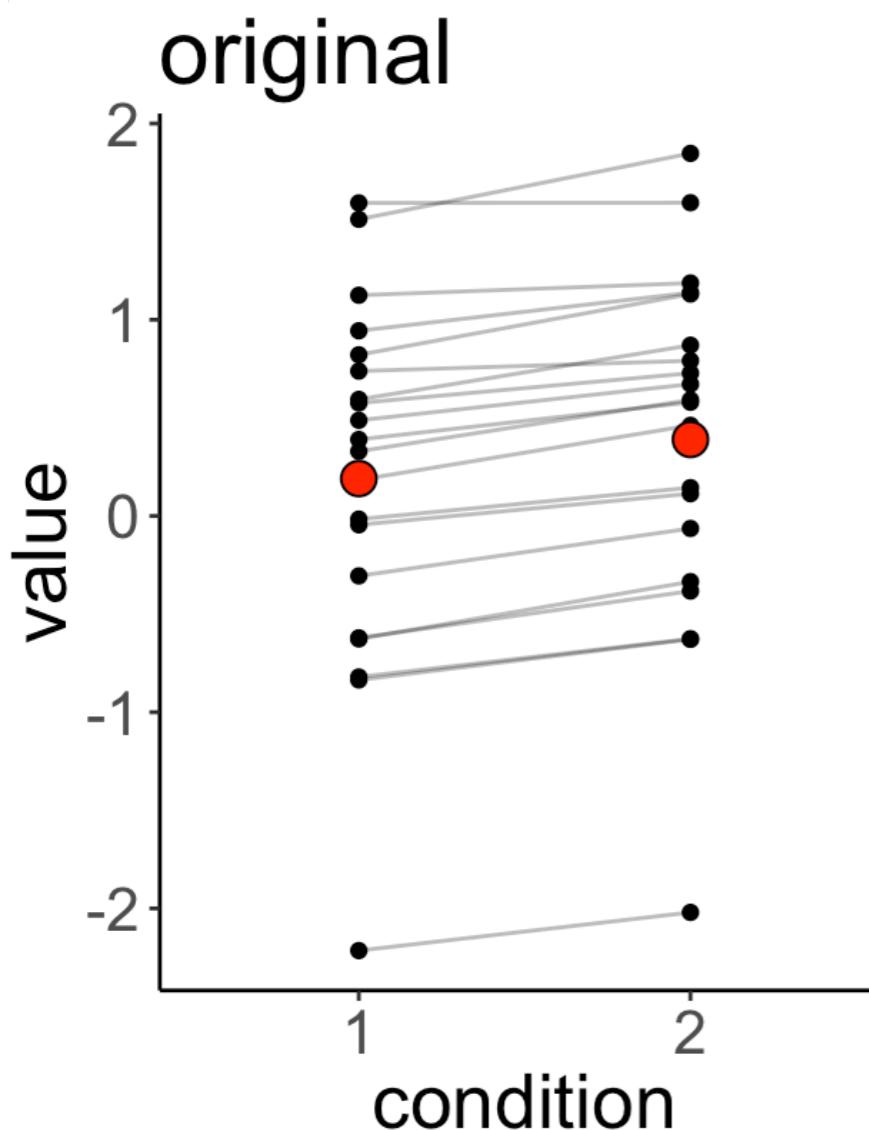
# meet lmer()



# Dependence

new syntax

```
1 # fit a linear mixed effects model
2 lmer(formula = value ~ condition + (1 | participant),
3       data = df.original) %>%
4 summary()
```



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

REML criterion at convergence: 17.3

Scaled residuals:
    Min     1Q Median     3Q    Max 
-1.55996 -0.36399 -0.03341  0.34400  1.65823 

Random effects:
 Groups   Name        Variance Std.Dev. 
 participant (Intercept) 0.816722 0.90373
 Residual           0.003796 0.06161 
Number of obs: 40, groups: participant, 20

Fixed effects:
            Estimate Std. Error t value
(Intercept)  0.19052   0.20255  0.941
condition2   0.19935   0.01948 10.231

Correlation of Fixed Effects:
              (Intr) condition2
condition2 -0.048
```

no p-value!

NO P-VALUE



# Dependence

we can still do our good ol' model comparison trick

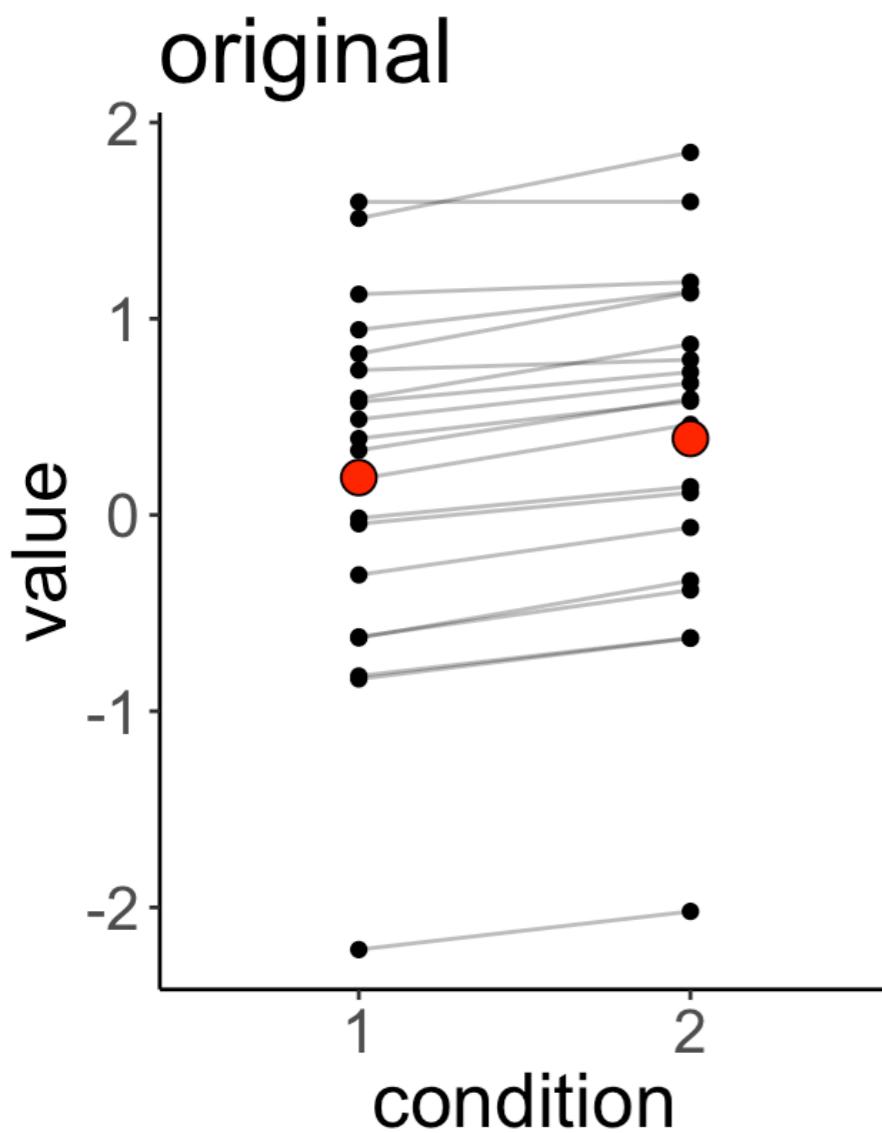
```
1 # fit models
2 fit.compact = lmer(formula = value ~ 1 + (1 | participant),
3                     data = df.original)

4 fit.augmented = lmer(formula = value ~ 1 + condition + (1 | participant),
5                     data = df.original)
6
7 # compare via Chisq-test
8 anova(fit.compact, fit.augmented)
```

```
refitting model(s) with ML (instead of REML)
Data: df.original
Models:
fit.compact: value ~ 1 + (1 | participant)
fit.augmented: value ~ 1 + condition + (1 | participant)
              Df     AIC     BIC   logLik deviance    Chisq Chi Df Pr(>Chisq)
fit.compact     3 53.315 58.382 -23.6575     47.315
fit.augmented   4 17.849 24.605  -4.9247      9.849 37.466          1 9.304e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# Dependence

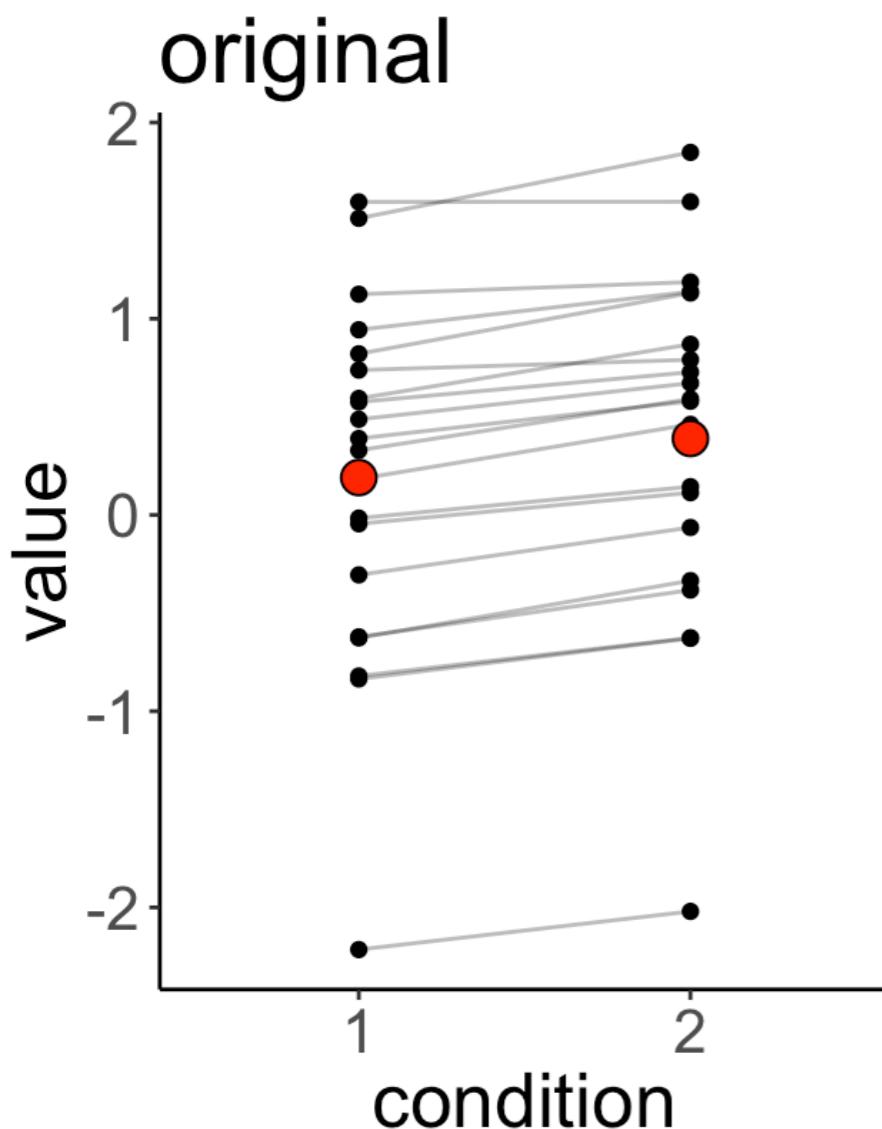
**Why is the effect of condition significant when we account for the dependence in the data?**



- there are large interindividual differences in the baseline
- the variance explained by the effect of condition is (much) smaller than the interindividual variance
- **but:** the effect of condition is highly consistent

# Dependence

**Why is the effect of condition significant when we account for the dependence in the data?**



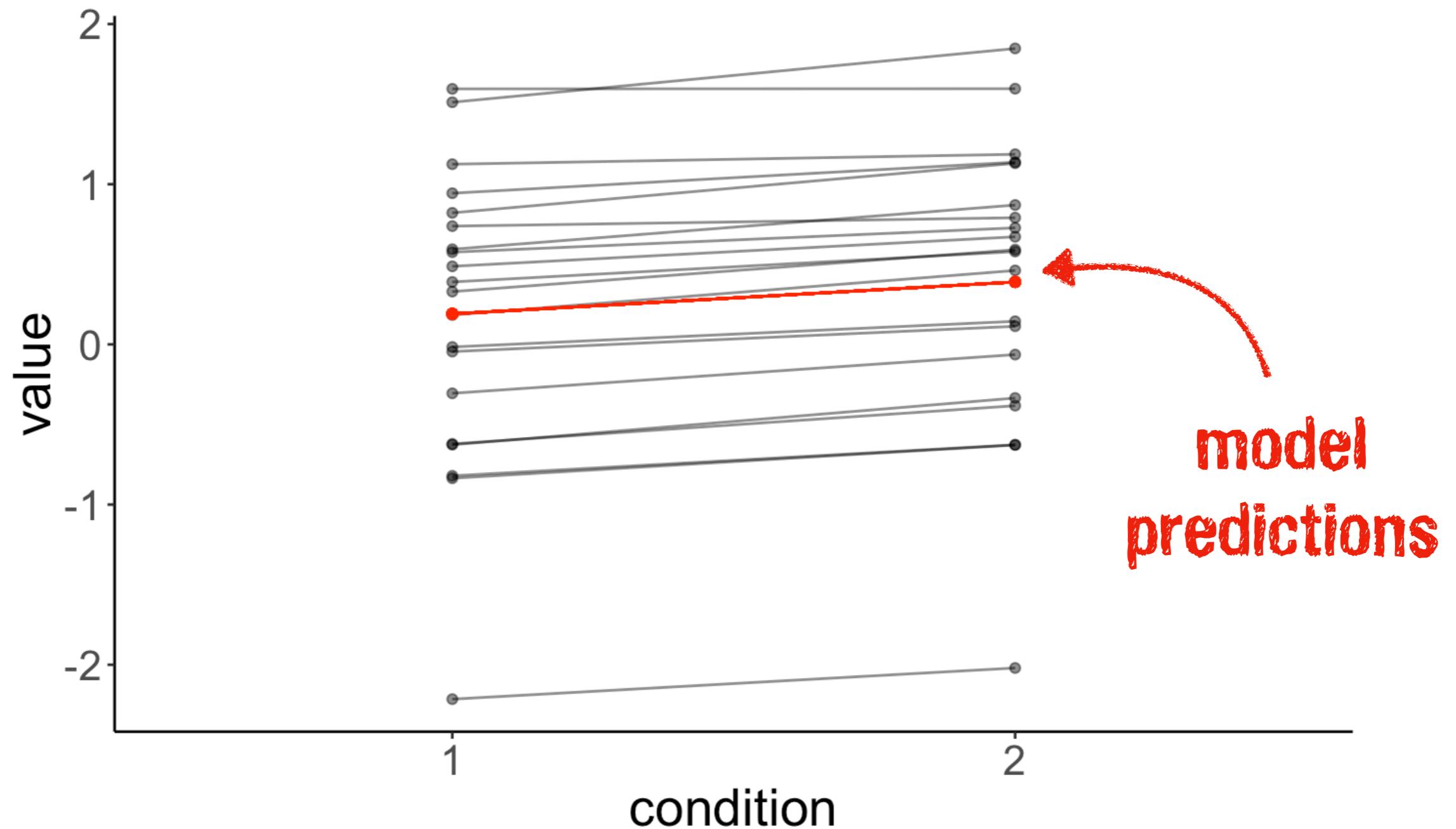
- by explicitly modeling the dependence in the data, we account for the interindividual differences

**let's visualize the model predictions!**

# Linear model (assuming independence)

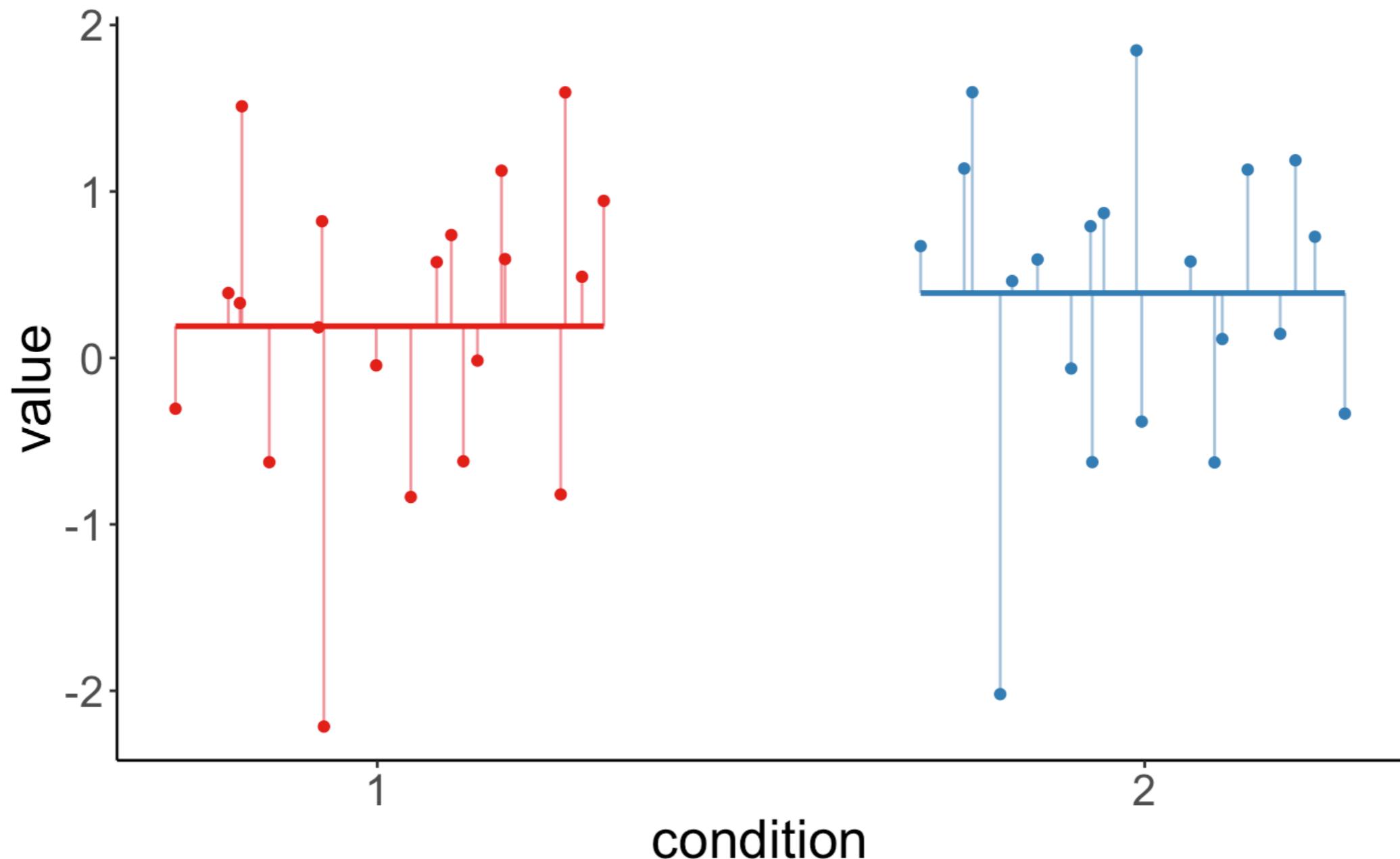
Predictions by the linear model which assumes independence

```
lm (formula = value ~ condition,  
    data = df.original)
```



# Linear model (assuming independence)

## Residuals of the model

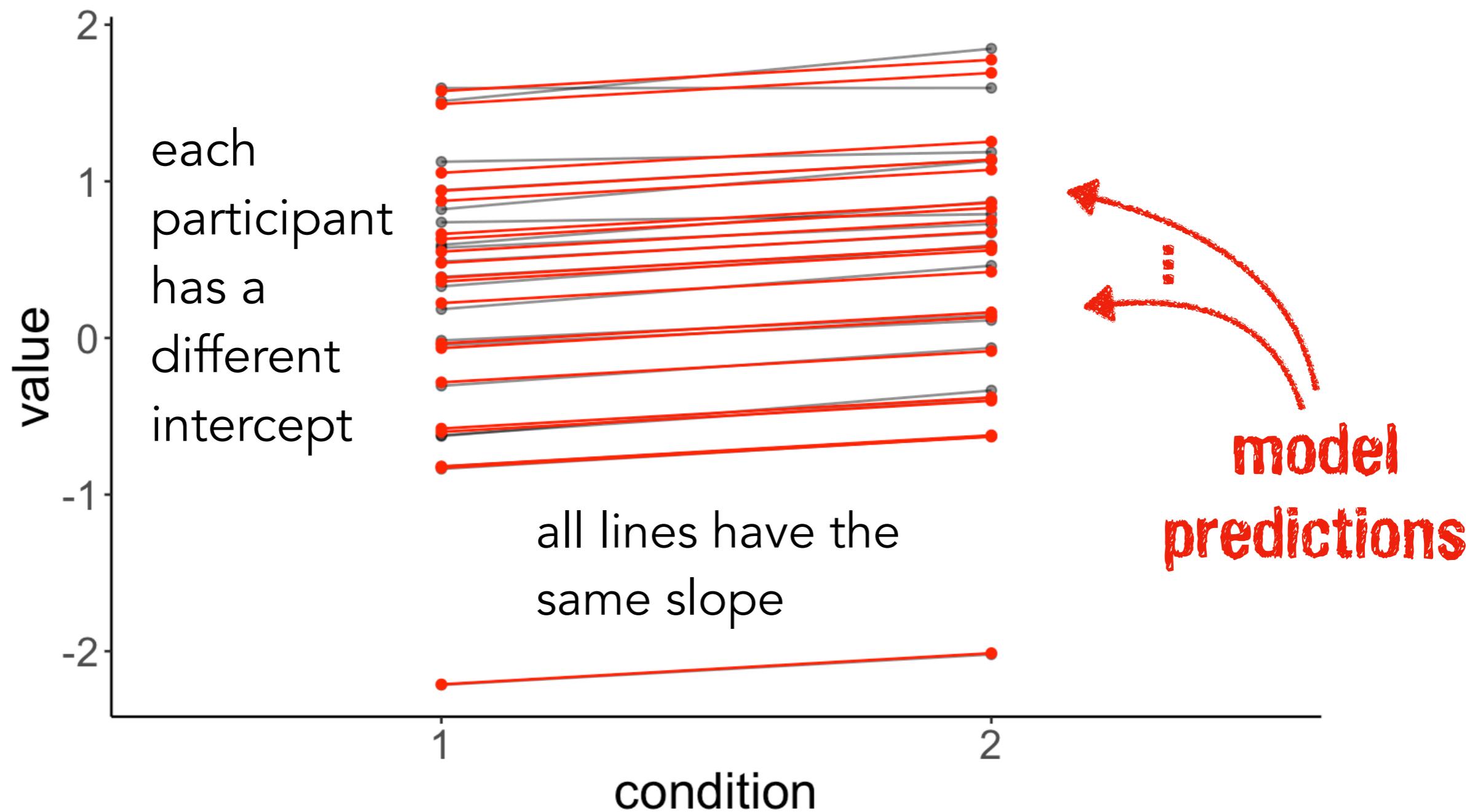


This is not much better than fitting a single line (point).

# Linear mixed effects model (accounting for dependence)

## Predictions by the linear mixed effects model

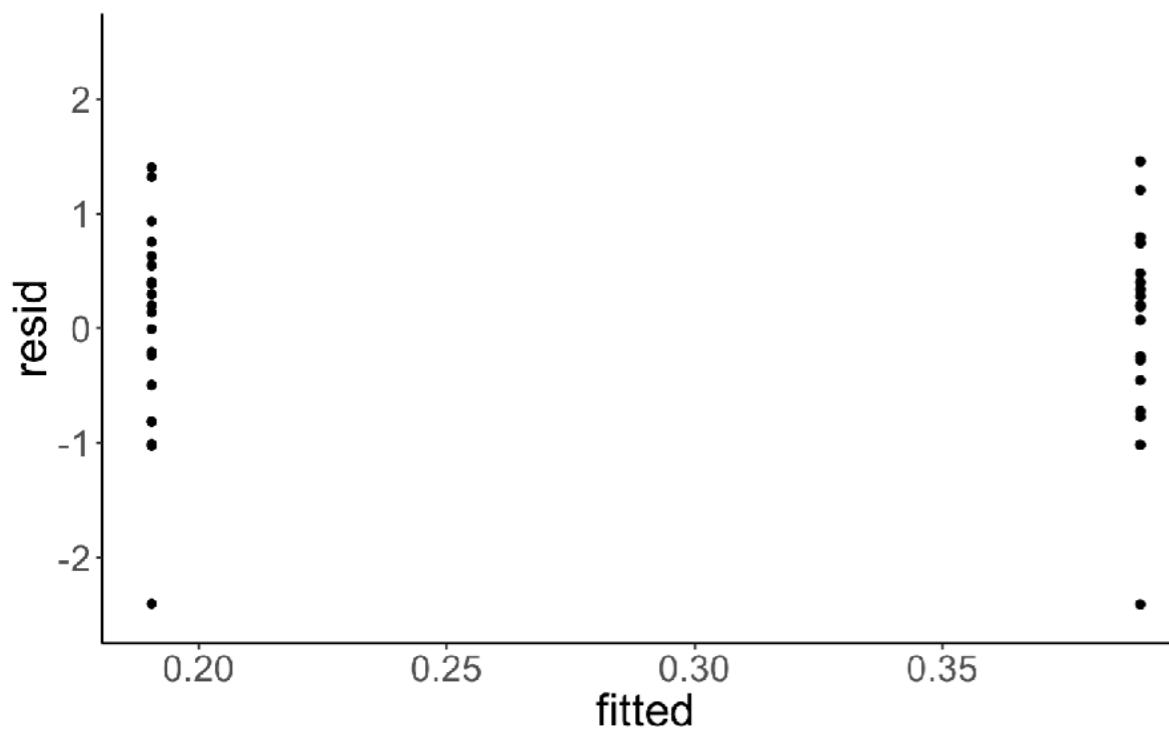
```
lmer (formula = value ~ condition + (1 | participant),  
      data = df.original)
```



# Model comparison

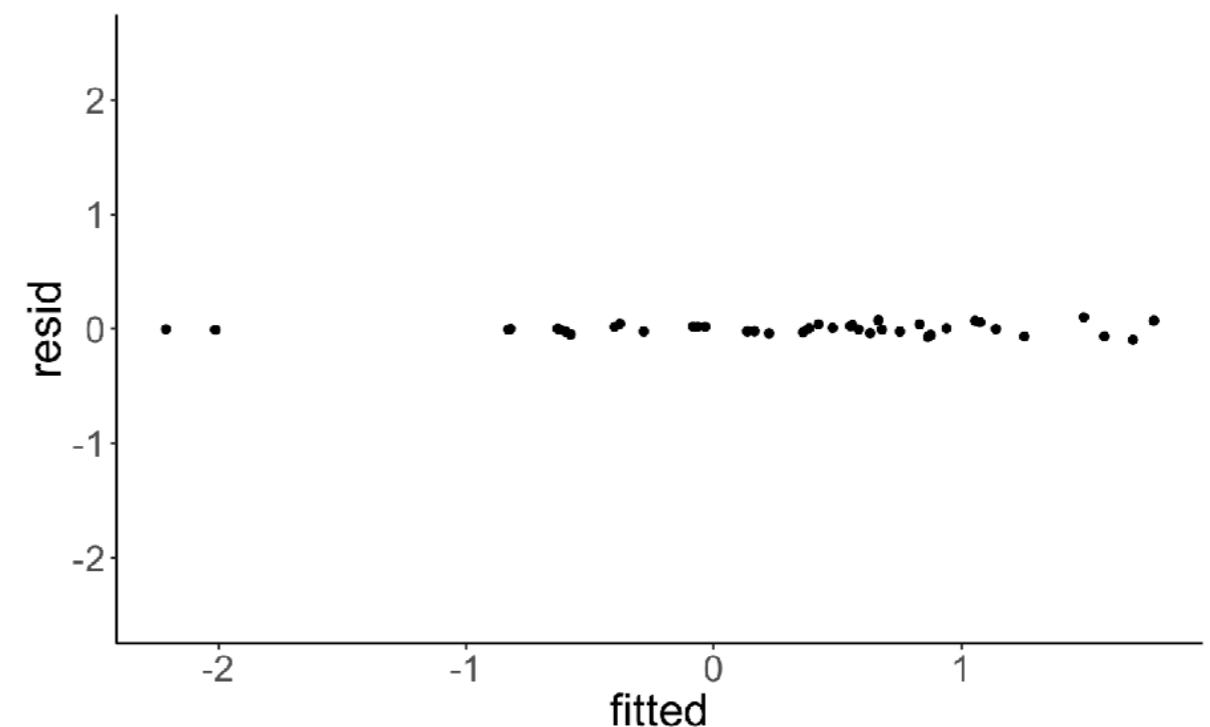
## Residual plots

```
lm(formula = value ~ 1 + condition,  
  data = df.original)
```



much variance left  
to be explained

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



almost all variance  
explained

# Model comparison

# Hypothesis test

Is taking into account individual differences worth it?

```
1 # fit models (without and with dependence)
2 fit.compact = lm(formula = value ~ 1 + condition,
3                   data = df.original)
4
5 fit.augmented = lmer(formula = value ~ 1 + condition + (1 | participant),
6                       data = df.original)
7
8 # compare models
9 # note: the lmer model has to be supplied first
10 anova(fit.augmented, fit.compact)
```

refitting model(s) with ML (instead of REML)

Data: df.original

Models:

fit.compact: value ~ 1 + condition

fit.augmented: value ~ 1 + condition + (1 | participant)

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
fit.compact	3	109.551	114.617	-51.775	103.551			
fit.augmented	4	17.849	24.605	-4.925	9.849	93.701	1	< 2.2e-16 ***
---								
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1								

## Linear model

```
lm(formula = value ~ 1 + condition,  
  data = df.original)
```

$$\text{value}_i = b_0 + b_1 \cdot \text{condition}_i + e_i$$

i = observation

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

3 parameters:  $b_0, b_1, s_{\text{error}}$

## Linear mixed effects model

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

$$\text{value}_{i,j} = b_0 + b_1 \cdot \text{condition}_{i,j} + U_i + e_i$$

i = participant,  
j = time point

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

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

$b_0, b_1$  = fixed effects

$U_i$  = random effect

 here: random intercept

4 parameters:  $b_0, b_1, s_{\text{error}}, s_U$

# Model coefficients

## Linear model

```
fit = lm(formula = value ~ 1 + condition,  
         data = df.original)  
coef(fit)
```

	(Intercept)	condition2
	0.1905239	0.1993528

- one intercept
- one slope for condition

## Linear mixed effects model

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

	participant	(Intercept)	condition2
1		-0.57839428	0.1993528
2		0.22299824	0.1993528
3		-0.82920677	0.1993528
4		1.49310938	0.1993528
5		0.36042775	0.1993528
6		-0.82060123	0.1993528
7		0.47929171	0.1993528
8		0.66401020	0.1993528
9		0.55135879	0.1993528
10		-0.28306703	0.1993528
11		1.57681676	0.1993528
12		0.38457642	0.1993528
13		-0.59969682	0.1993528
14		-2.21148391	0.1993528
15		1.05439374	0.1993528
16		-0.06476643	0.1993528
17		-0.03505690	0.1993528
18		0.93945348	0.1993528
19		0.87495531	0.1993528
20		0.63135911	0.1993528

```
attr(),"class")  
[1] "coef.mer"
```

- different intercept for each participant
- one slope for condition

# Understanding the `lmer()` summary

```
1 # fit a linear mixed effects model  
2 lmer(formula = value ~ condition + (1 | participant),  
3 data = df.original) %>%  
4 summary()
```

```
Linear mixed model fit by REML ['lmerMod']  
Formula: value ~ condition + (1 | participant)  
Data: df.original  
  
REML criterion at convergence: 17.3  
  
Scaled residuals:  
    Min     1Q   Median     3Q     Max  
-1.55996 -0.36399 -0.03341  0.34400  1.65823  
  
Random effects:  
Groups      Name        Variance Std.Dev.  
participant (Intercept) 0.816722 0.90373  
Residual             0.003796 0.06161  
Number of obs: 40, groups: participant, 20  
  
Fixed effects:  
            Estimate Std. Error t value  
(Intercept) 0.19052    0.20255   0.941  
condition2   0.19935    0.01948  10.231  
  
Correlation of Fixed Effects:  
          (Intr) condition2  
condition2 -0.048
```

**REML** = restricted maximum likelihood method for fitting models with **random effects**

# Understanding the **lmer()** summary

```
1 # fit a linear mixed effects model  
2 lmer(formula = value ~ condition + (1 | participant),  
3 data = df.original) %>%  
4 summary()
```

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

```
REML criterion at convergence: 17.3
```

```
Scaled residuals:
```

Min	1Q	Median	3Q	Max
-1.55996	-0.36399	-0.03341	0.34400	1.65823

```
Random effects:
```

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	0.816722	0.90373
Residual		0.003796	0.06161

```
Number of obs: 40, groups: participant, 20
```

```
Fixed effects:
```

	Estimate	Std. Error	t value
(Intercept)	0.19052	0.20255	0.941
condition2	0.19935	0.01948	10.231

```
Correlation of Fixed Effects:
```

	(Intr)
condition2	-0.048

fitting **lmer()** doesn't always work ...

**lmer()** complains when it didn't work

# Understanding the `lmer()` summary

```
1 # fit a linear mixed effects model  
2 lmer(formula = value ~ condition + (1 | participant),  
3 data = df.original) %>%  
4 summary()
```

```
Linear mixed model fit by REML ['lmerMod']  
Formula: value ~ condition + (1 | participant)  
Data: df.original  
  
REML criterion at convergence: 17.3
```

Scaled residuals:

Min	1Q	Median	3Q	Max
-1.55996	-0.36399	-0.03341	0.34400	1.65823

Random effects:

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	0.816722	0.90373
Residual		0.003796	0.06161

Number of obs: 40, groups: participant, 20

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	0.19052	0.20255	0.941
condition2	0.19935	0.01948	10.231

Correlation of Fixed Effects:

	(Intr)
condition2	-0.048

summary information  
about residuals

# Understanding the `lmer()` summary

```
1 # fit a linear mixed effects model
2 lmer(formula = value ~ condition + (1 | participant),
3           data = df.original) %>%
4 summary()
```

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

```
REML criterion at convergence: 17.3
```

```
Scaled residuals:
```

Min	1Q	Median	3Q	Max
-1.55996	-0.36399	-0.03341	0.34400	1.65823

```
Random effects:
```

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	0.816722	0.90373
Residual		0.003796	0.06161

```
Number of obs: 40, groups: participant, 20
```

```
Fixed effects:
```

	Estimate	Std. Error	t value
(Intercept)	0.19052	0.20255	0.941
condition2	0.19935	0.01948	10.231

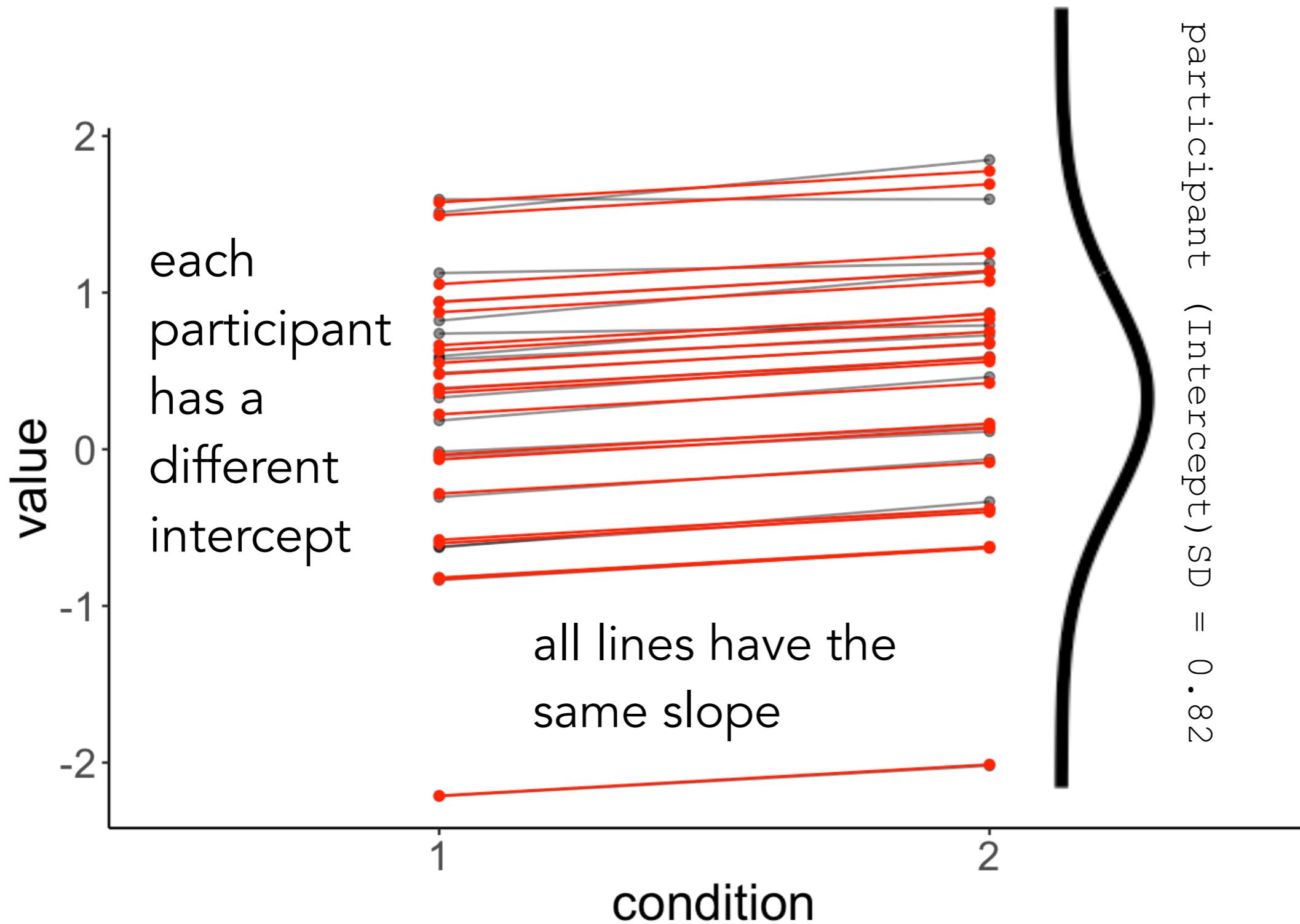
```
Correlation of Fixed Effects:
```

	(Intr)
condition2	-0.048

one parameter to capture the variance between participants (gives us a sense for whether there are interindividual differences)

one parameter to capture the residual variance (just like sigma in an `lm()`)

# Understanding the `lmer()` summary



# Understanding the **lmer()** summary

```
1 # fit a linear mixed effects model  
2 lmer(formula = value ~ condition + (1 | participant),  
3 data = df.original) %>%  
4 summary()
```

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

```
REML criterion at convergence: 17.3
```

```
Scaled residuals:
```

Min	1Q	Median	3Q	Max
-1.55996	-0.36399	-0.03341	0.34400	1.65823

```
Random effects:
```

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	0.816722	0.90373
Residual		0.003796	0.06161

```
Number of obs: 40, groups: participant, 20
```

```
Fixed effects:
```

	Estimate	Std. Error	t value
(Intercept)	0.19052	0.20255	0.941
condition2	0.19935	0.01948	10.231

```
Correlation of Fixed Effects:
```

	(Intr)
condition2	-0.048

one parameter for the global intercept (value for the baseline condition)

one parameter for the condition effect (difference between the two conditions)

interpretation the same as for **lm()**, also: we can use contrasts!

# Understanding the `lmer()` summary

```
1 # fit a linear mixed effects model
2 lmer(formula = value ~ condition + (1 | participant),
3           data = df.original) %>%
4 summary()
```

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

```
REML criterion at convergence: 17.3
```

```
Scaled residuals:
```

Min	1Q	Median	3Q	Max
-1.55996	-0.36399	-0.03341	0.34400	1.65823

```
Random effects:
```

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	0.816722	0.90373
Residual		0.003796	0.06161

```
Number of obs: 40, groups: participant, 20
```

```
Fixed effects:
```

	Estimate	Std. Error	t value
(Intercept)	0.19052	0.20255	0.941
condition2	0.19935	0.01948	10.231

```
Correlation of Fixed Effects:
  (Intr) condition2
condition2 -0.048
```

correlation between intercept and condition2

The "correlation of fixed effects" output doesn't have the intuitive meaning that most would ascribe to it. Specifically, is not about the correlation of the variables. It is in fact about the expected correlation of the regression coefficients.

# general points about `lmer()`

- **fixed effects:**
  - parameters are estimated
  - often: factors that we manipulate experimentally
- **random effects:**
  - variation we want to control for
  - often: differences between participants (or items) in our experiment
  - sampling viewpoint: we explicitly model the variation in participants

# general points about `lmer()`

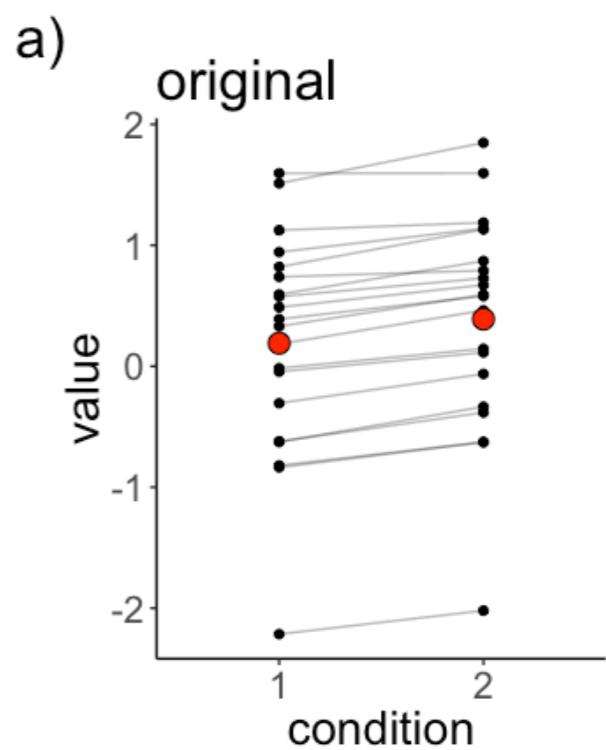
- Why don't we just run individual regressions?
  - overfitting ...
  - inflating type 1 error
  - larger uncertainty in parameter estimates because only few data points are used
- Why don't we just run a regression on the means?
  - we throw away a lot of information
- Mixed effects models:
  - make use of all available information
  - more next time ...

# we just performed a paired t-test ...

```
1 t.test(df.original$value[df.original$condition == "1"],  
2         df.original$value[df.original$condition == "2"],  
3         alternative = "two.sided",  
4         paired = T)
```

```
Paired t-test  
  
data: df.original$value[df.original$condition == "1"] and  
df.original$value[df.original$condition == "2"]  
t = -10.231, df = 19, p-value = 3.636e-09  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
 -0.2401340 -0.1585717  
sample estimates:  
mean of the differences  
 -0.1993528
```

If we take the differences for each participant between condition 1 and condition 2, are these difference scores significantly different from 0?



# we just performed a paired t-test ...

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

- explicitly models the interindividual variation
- much more flexible ...

```
1 t.test(df.original$value[df.original$condition == "1"],  
2         df.original$value[df.original$condition == "2"],  
3         alternative = "two.sided",  
4         paired = T)
```

Paired t-test

```
data: df.original$value[df.original$condition == "1"] and  
df.original$value[df.original$condition == "2"]  
t = -10.231, df = 19, p-value = 3.636e-09  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
-0.2401340 -0.1585717  
sample estimates:  
mean of the differences  
-0.1993528
```

```
Linear mixed model fit by REML ['lmerMod']  
Formula: value ~ condition + (1 | participant)  
Data: df.original  
  
REML criterion at convergence: 17.3  
  
Scaled residuals:  
    Min     1Q Median     3Q    Max  
-1.55996 -0.36399 -0.03341  0.34400  1.65823  
  
Random effects:  
Groups   Name        Variance Std.Dev.  
participant (Intercept) 0.816722 0.90373  
Residual           0.003796 0.06161  
Number of obs: 40, groups: participant, 20  
  
Fixed effects:  
            Estimate Std. Error t value  
(Intercept) 0.19052   0.20255  0.941  
condition2  0.19935   0.01948 10.231  
  
Correlation of Fixed Effects:  
          (Intr)  
condition2 -0.048
```

# Summary

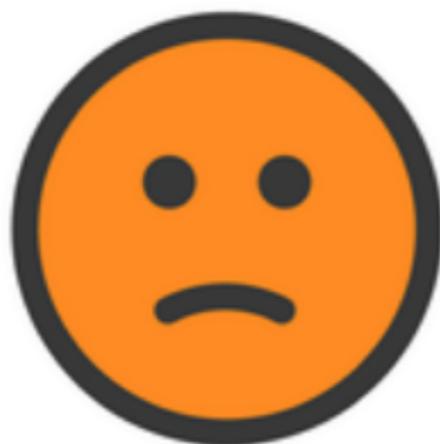
- Model comparison
  - Cross-validation
  - AIC and BIC
- Mediation
- Moderation
- Linear mixed effects model
  - modeling dependence in data

# **Feedback**

# How was the pace of today's class?

much    a little    just    a little    much  
too        too        right      too        too  
slow      slow                                    fast      fast

# How happy were you with today's class overall?



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

Thank you!