Statistical Rethinking
Winter 2019

Lecture 09 / Week 5

Conditional Manatees

Stop testing, start thinking

- Off-the-shelf tools
  - *p*-values
  - information criteria
  - linear models
- Good decisions benefit from **bespoke** tools
  - bespoke risk analysis
  - bespoke models

Bespoke *Broke*

bespoke | bə'spōk |

adjective *chiefly British*

made for a particular customer or user: *a bespoke suit* | *bespoke kitchens* | *bespoke software systems* | *group tours and bespoke itineraries*.

- making or selling bespoke goods, especially clothing: *bespoke tailors*.

Leaders in New York and New Jersey Defend Shutdown for a Blizzard That Wasn't

By MATT FLEGENHEIMER JAN. 27, 2015

It was an unprecedented step for what became, in New York City, a common storm: For the first time in its 110-year history, the subway system was shut down because of snow.

Transit workers, caught off guard by the shutdown that Gov. Andrew M. Cuomo announced on Monday, scrambled to grind the network to a halt within hours.

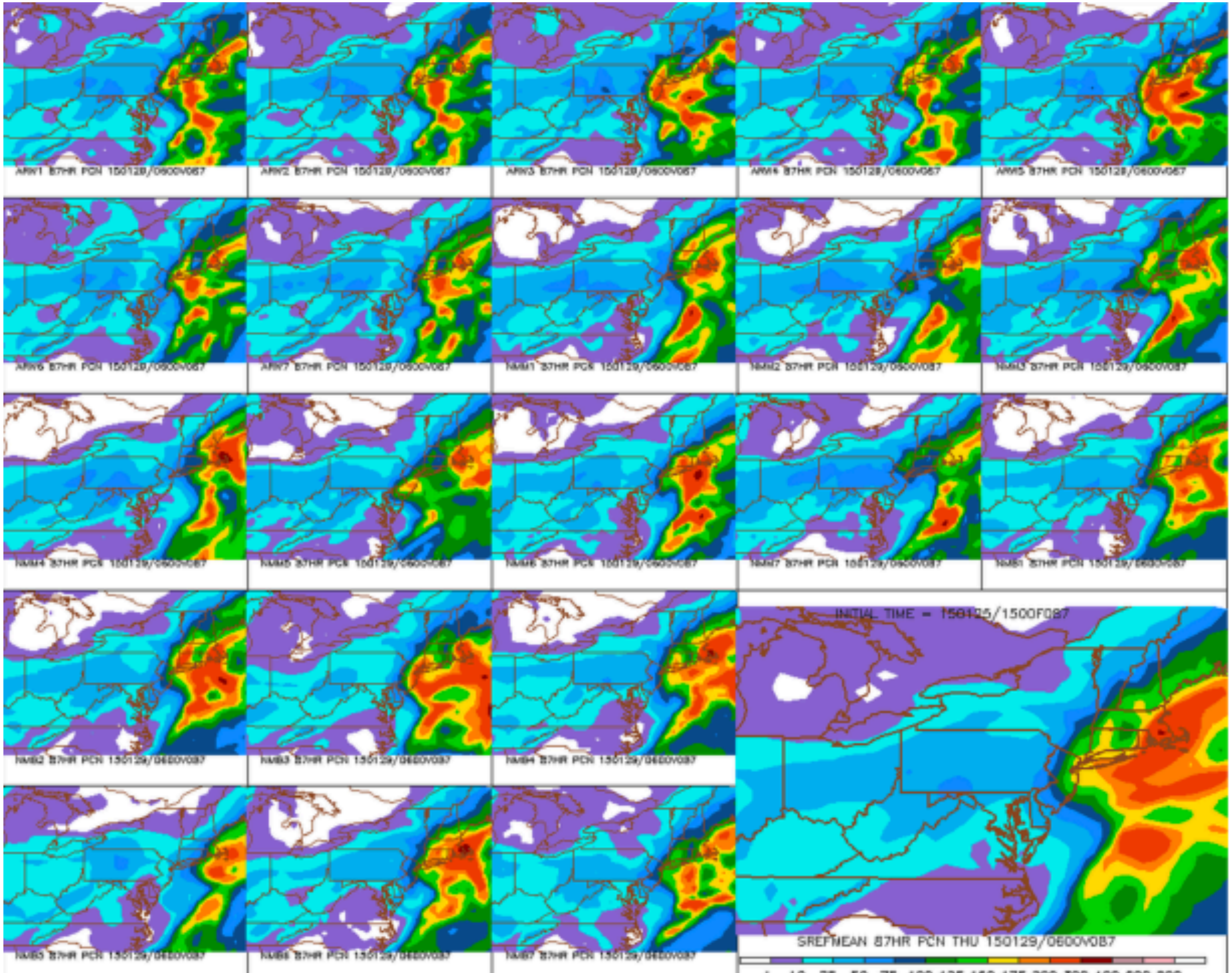
Residents moved quickly to find places

Mayor Bill de Blasio of New York City, with Sanitation Department workers in Manhattan on Monday, when he issued dire warnings about the storm. Yana Paskova for The New York Times

ECMWF

About Forecasts

World leader in global medium-range numerical weather prediction



Blizzard calibration

- Was it bad to predict NYC blizzard from ECMWF?
- Other models were more accurate
- But welfare enhanced by being prepared => use extreme forecasts
- **Accuracy always matters, but it isn't all that matters**

Fly

Armstrong Whitworth Whitley Mk V

British Medium Bomber

Plastic model kit
Plastik-Modellbausatz
Plastikový model

1/72

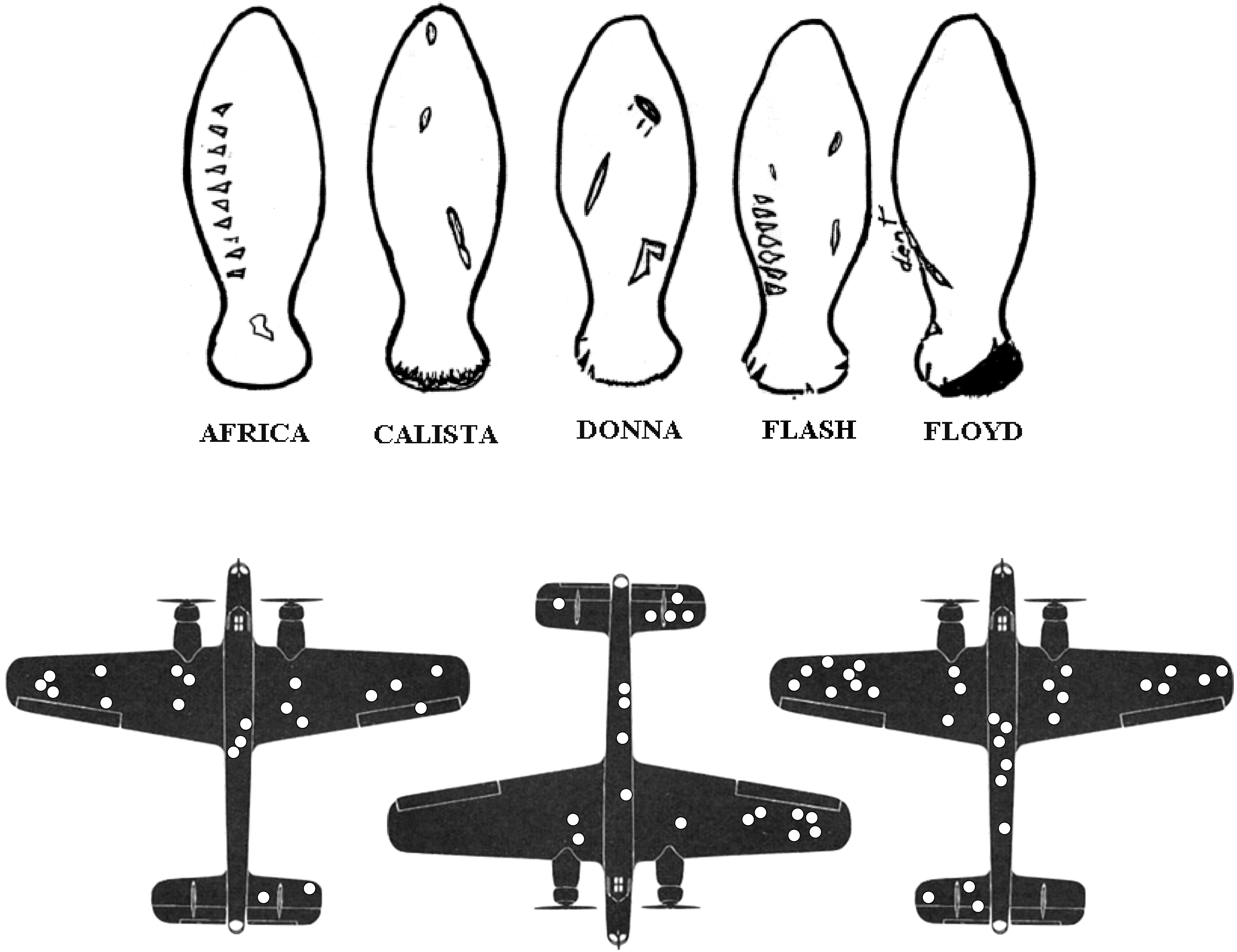
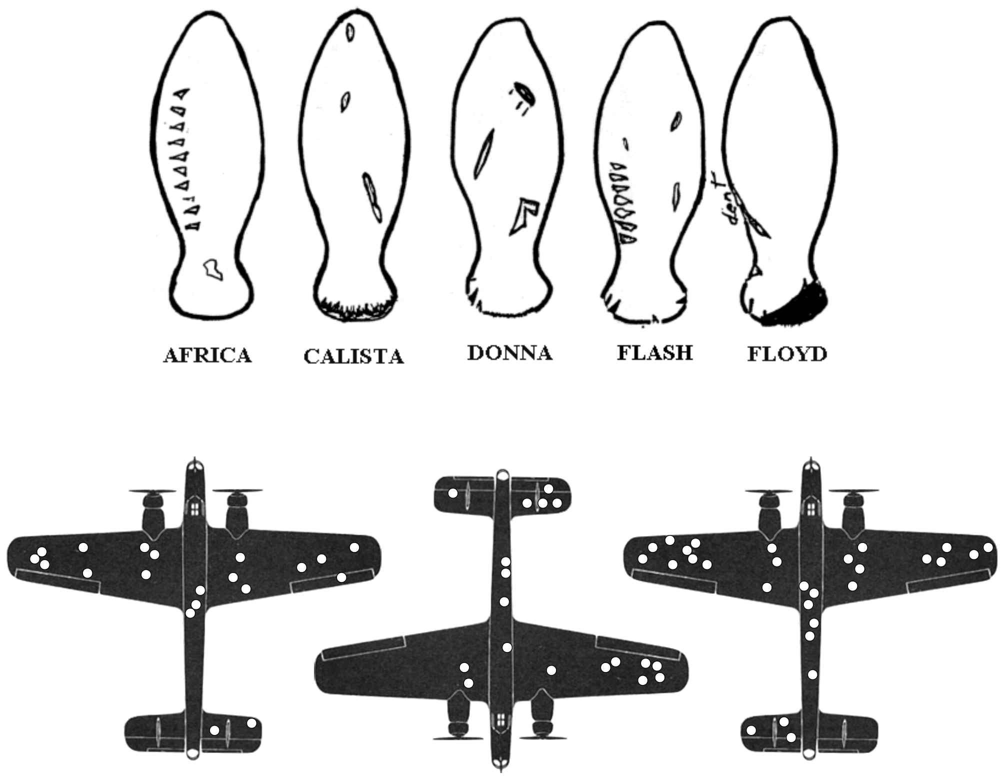


Figure 8.1



rotor/wing → survival ← keel/engine

Observe only:
undamaged
rotor/wing damaged

Figure 8.1

Manatees and bombers

- *Conditioning*: Dependence on state
- Everything is conditional
  - On data
  - On model
  - On information state
- Influence of variable conditional on other variable(s)

Interaction effects

- *Interactions*: Influence of predictor conditional on other predictor(s)
  - Influence of *sugar* in *coffee* depends on *stirring*
  - Influence of *gene* on *phenotype* depends on *environment*
  - Influence of *skin color* on *cancer* depends on *latitude*
- Generalized linear models (GLMs): All predictors interact to some degree
- Multilevel models: Massive interaction engines

Interaction effects in DAGs

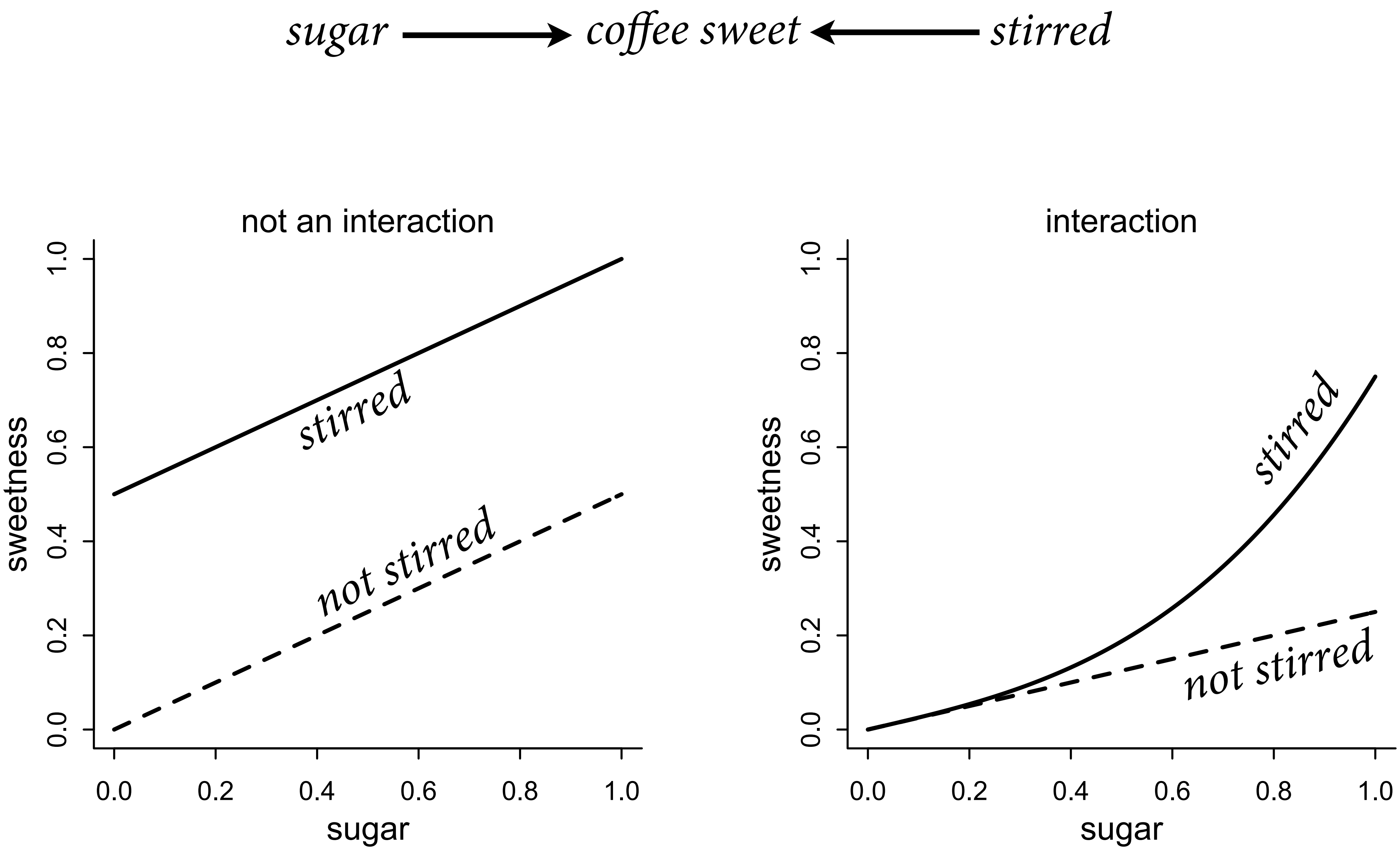
- In DAG, interaction doesn't look special:

$$\textit{sugar} \longrightarrow \textit{coffee sweet} \longleftarrow \textit{stirred}$$

- This just means:

$$\textit{coffee sweet} = f\,(\textit{sugar}, \textit{stirred})$$

- We have to figure out the function f.



sugar
coffee sweet
stirred
not an interaction
interaction
sweetness
sugar
stirred
not stirred
0.0
0.2
0.4
0.6
0.8
1.0


NETHERLANDS
BELGIUM
PORTUGAL
SWITZERLAND
FRANCE
SPAIN
GERMANY
ITALY
EASTERN EUROPE
UNITED STATES
INDIA
INDIA PART 2
CHINA
CHINA PART 2
UK
JAPAN

```
library(rethinking)
data(rugged)
d <- rugged
```

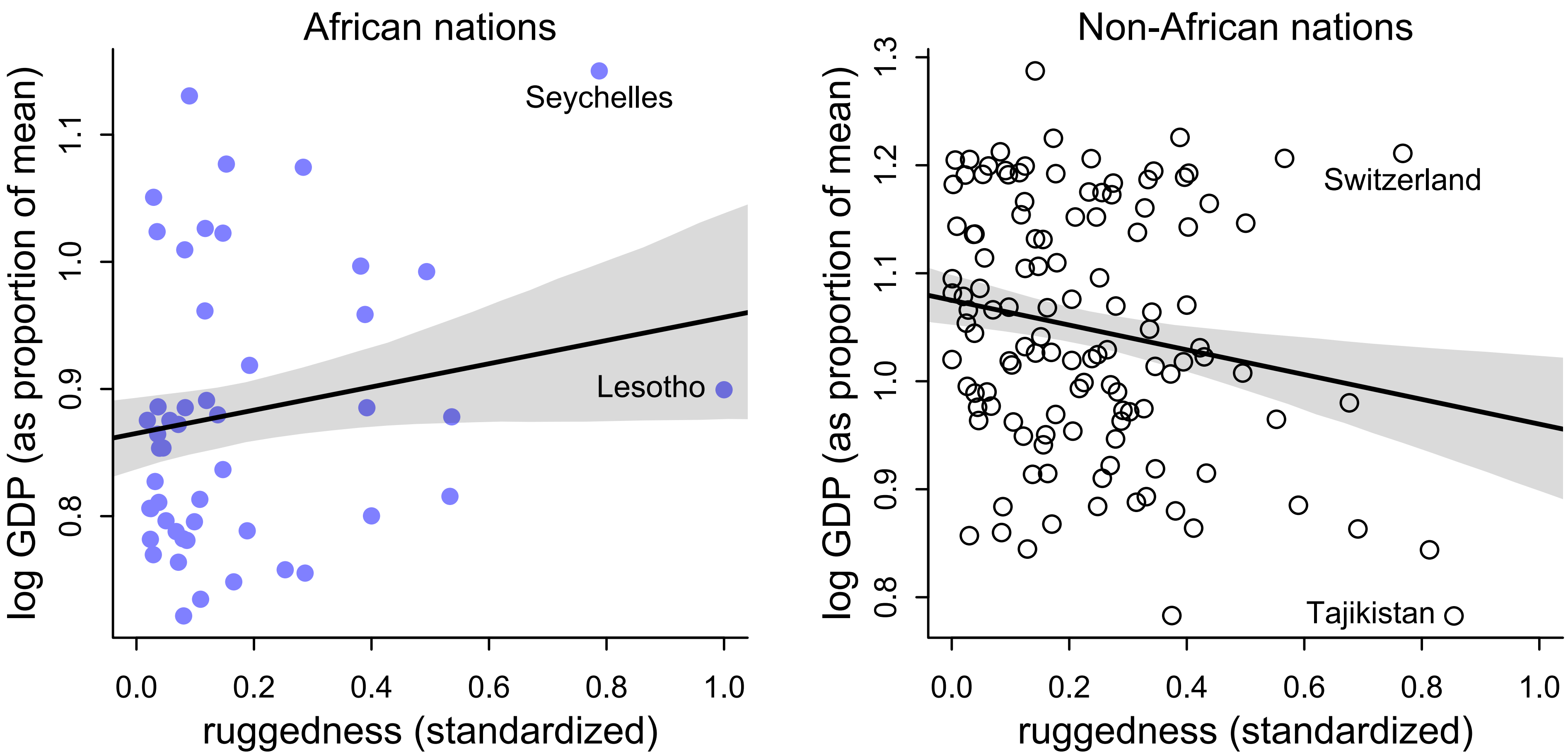


Figure 8.2

The sermon on priors

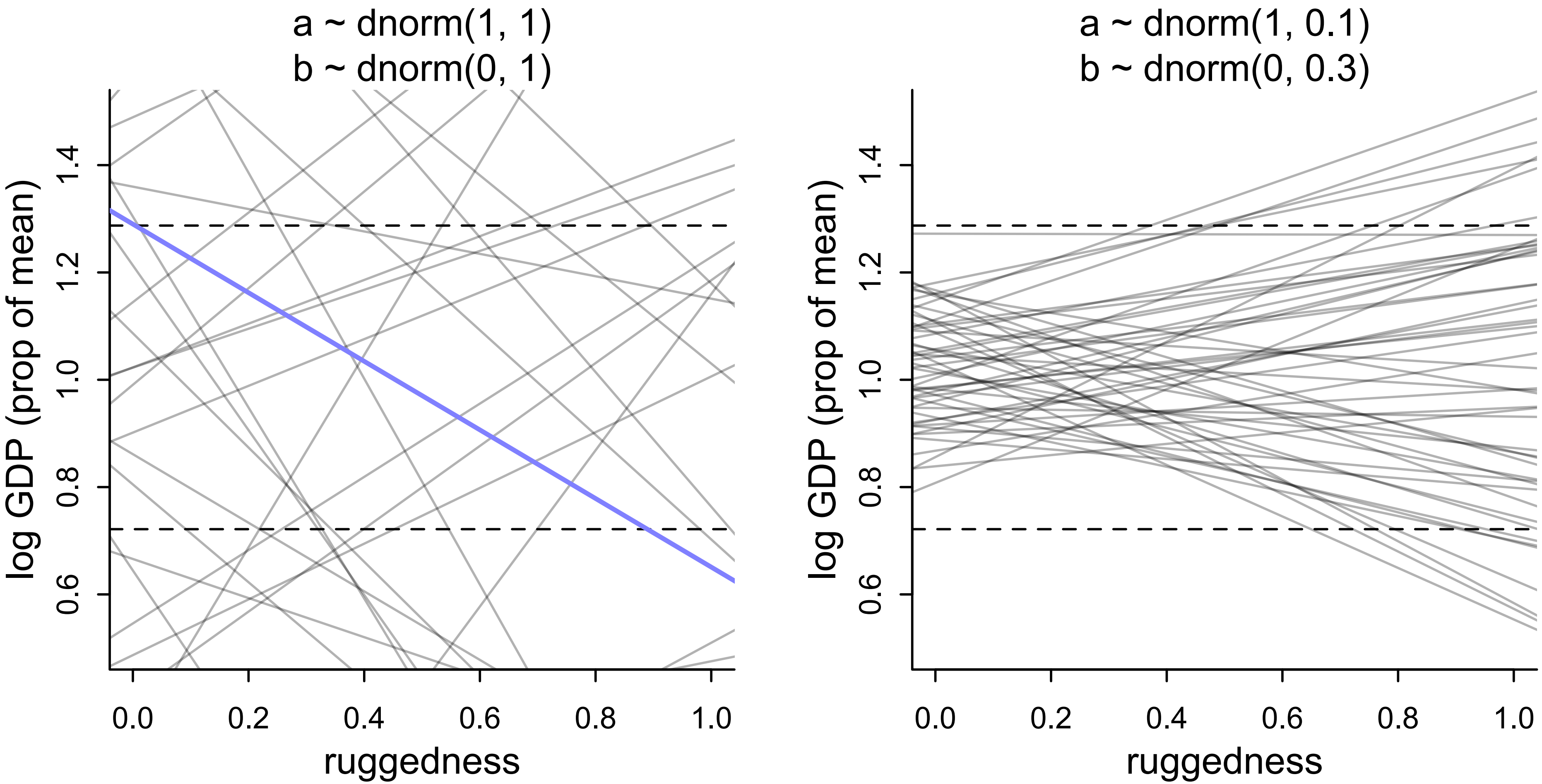
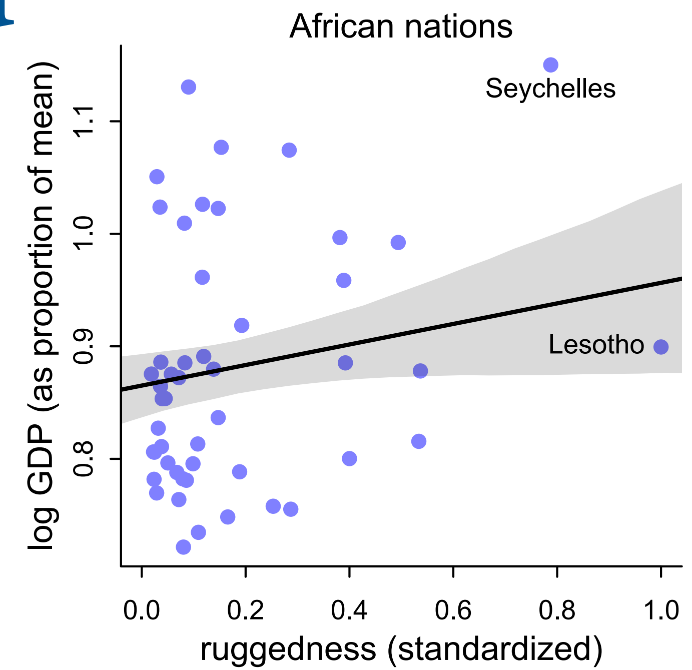


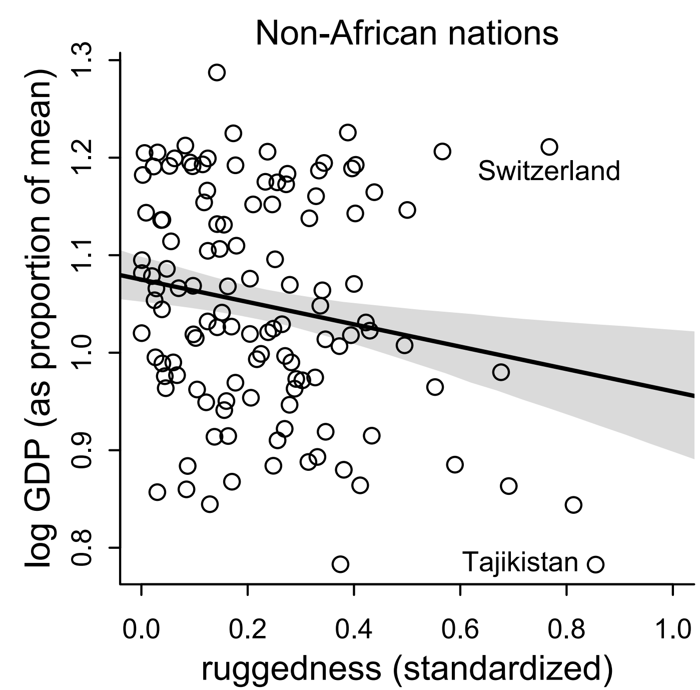
Figure 8.3

The value of being rugged

- Splitting the data is a bad idea:
  - No inference for how you split the data
  - Does not pool information

- How about adding a categorical variable for Africa?





Category doesn't work

- Index variable for continent:

$$\mu_i = \alpha_{\text{CID}[i]} + \beta(r_i - \bar{r})$$

```
m8.4 <- quap(
    alist(
        log_gdp_std ~ dnorm( mu , sigma ) ,
        mu <- a[cid] + b*( rugged_std - 0.215 ) ,
        a[cid] ~ dnorm( 1 , 0.1 ) ,
        b ~ dnorm( 0 , 0.3 ) ,
        sigma ~ dexp( 1 )
    ) ,
    data=dd )
```

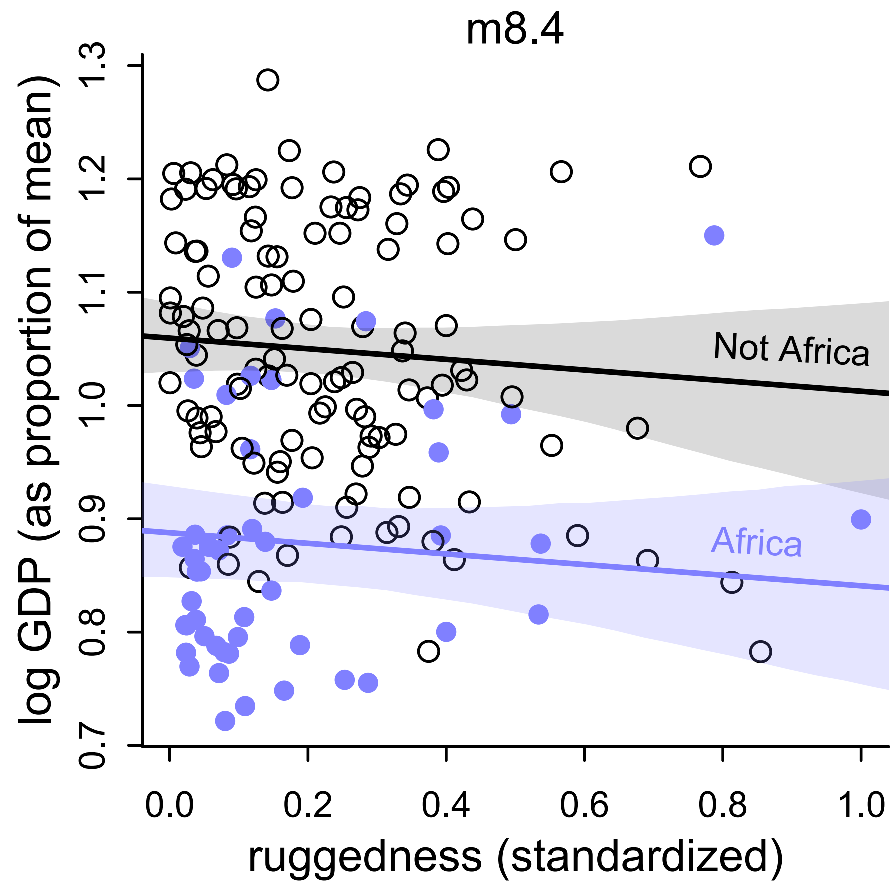


Figure 8.4

Interaction

- Need to allow effect of ruggedness to depend upon continent

$$\mu_i = \alpha_{\text{CID}[i]} + \beta_{\text{CID}[i]}(r_i - \bar{r})$$

Interaction

$$\mu_i = \alpha_{\text{CID}[i]} + \beta_{\text{CID}[i]}(r_i - \bar{r})$$

R code 8.13

```
m8.5 <- quap(
    alist(
        log_gdp_std ~ dnorm( mu , sigma ) ,
        mu <- a[cid] + b[cid]*( rugged_std - 0.215 ) ,
        a[cid] ~ dnorm( 1 , 0.1 ) ,
        b[cid] ~ dnorm( 0 , 0.3 ) ,
        sigma ~ dexp( 1 )
    ) ,
    data=dd )
```

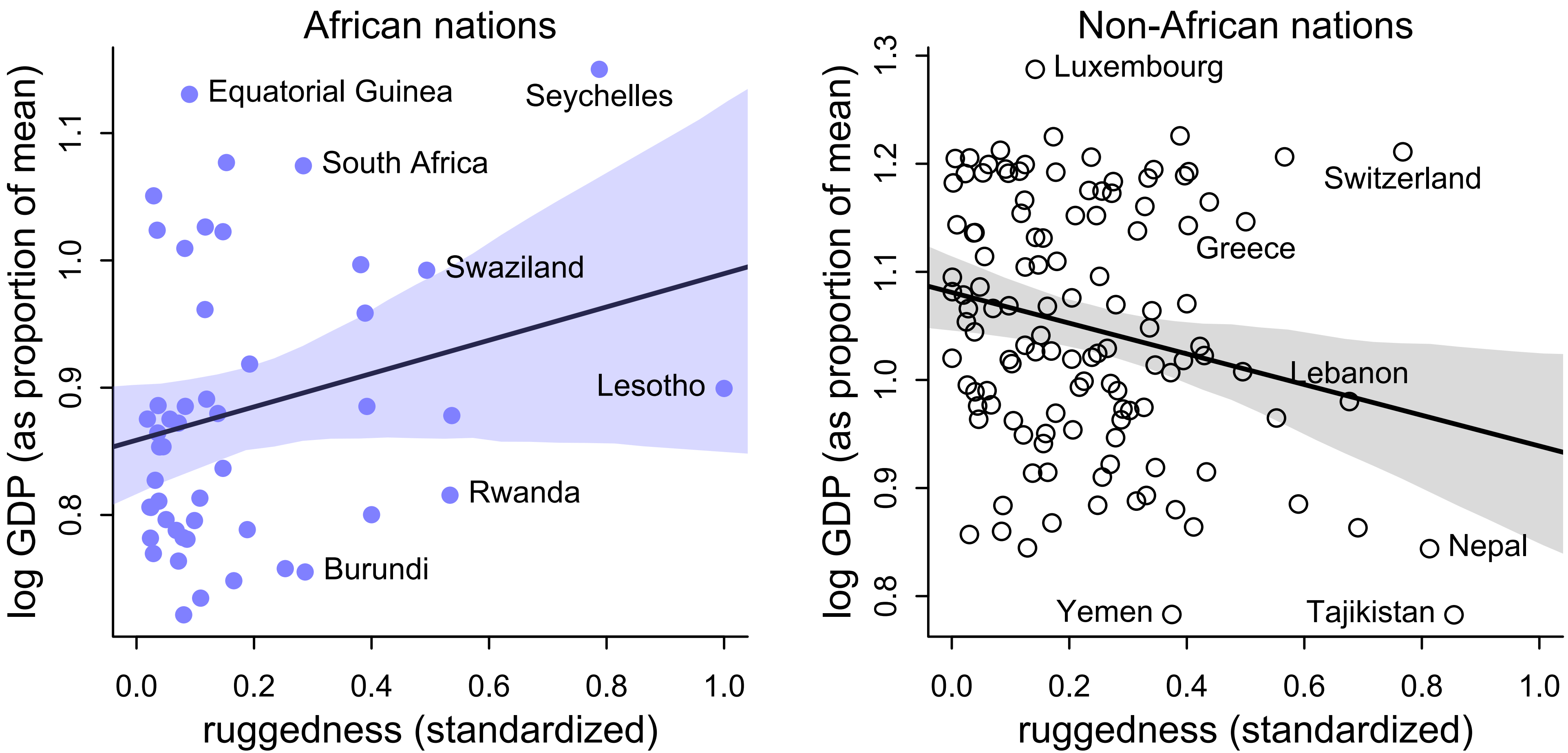
R code 8.13

```
m8.5 <- quap(
    alist(
        log_gdp_std ~ dnorm( mu , sigma ) ,
        mu <- a[cid] + b[cid]*( rugged_std - 0.215 ) ,
        a[cid] ~ dnorm( 1 , 0.1 ) ,
        b[cid] ~ dnorm( 0 , 0.3 ) ,
        sigma ~ dexp( 1 )
    ) ,
    data=dd )
```

R code 8.14

```
precis( m8.5 , depth=2 )
```

```
       mean   sd  5.5% 94.5%
a[1]   0.89 0.02  0.86  0.91
a[2]   1.05 0.01  1.03  1.07
b[1]   0.13 0.07  0.01  0.25
b[2]  -0.14 0.05 -0.23 -0.06
sigma  0.11 0.01  0.10  0.12
```



African nations
Non-African nations
log GDP (as proportion of mean)
ruggedness (standardized)
0.8
0.9
1.0
1.1
1.2
1.3
0.0
0.2
0.4
0.6
0.8
1.0
Equatorial Guinea
Seychelles
South Africa
Swaziland
Lesotho
Rwanda
Burundi
Luxembourg
Switzerland
Greece
Lebanon
Nepal
Yemen
Tajikistan

Figure 8.5

Interpreting interactions

- Is hard
  - Add interaction => other parameters change meaning
  - Influence of predictor depends upon multiple parameters and their covariation

R code 8.14

```
precis( m8.5 , depth=2 )
```

```
       mean   sd  5.5% 94.5%
a[1]   0.89 0.02  0.86  0.91
a[2]   1.05 0.01  1.03  1.07
b[1]   0.13 0.07  0.01  0.25
b[2]  -0.14 0.05 -0.23 -0.06
sigma  0.11 0.01  0.10  0.12
```

Interactions are symmetric

- Effect of ruggedness depends upon continent:

$$\mu_i = \alpha_{\textsc{cid}[i]} + \beta_{\textsc{cid}[i]}(r_i - \bar{r})$$

- Effect of continent depends upon ruggedness:

$$\mu_i = \underbrace{(2 - \textsc{cid}_i)(\alpha_1 + \beta_1(r_i - \bar{r}))}_{\textsc{cid}[i]=1} + \underbrace{(\textsc{cid}_i - 1)(\alpha_2 + \beta_2(r_i - \bar{r}))}_{\textsc{cid}[i]=2}$$

Interactions are symmetric

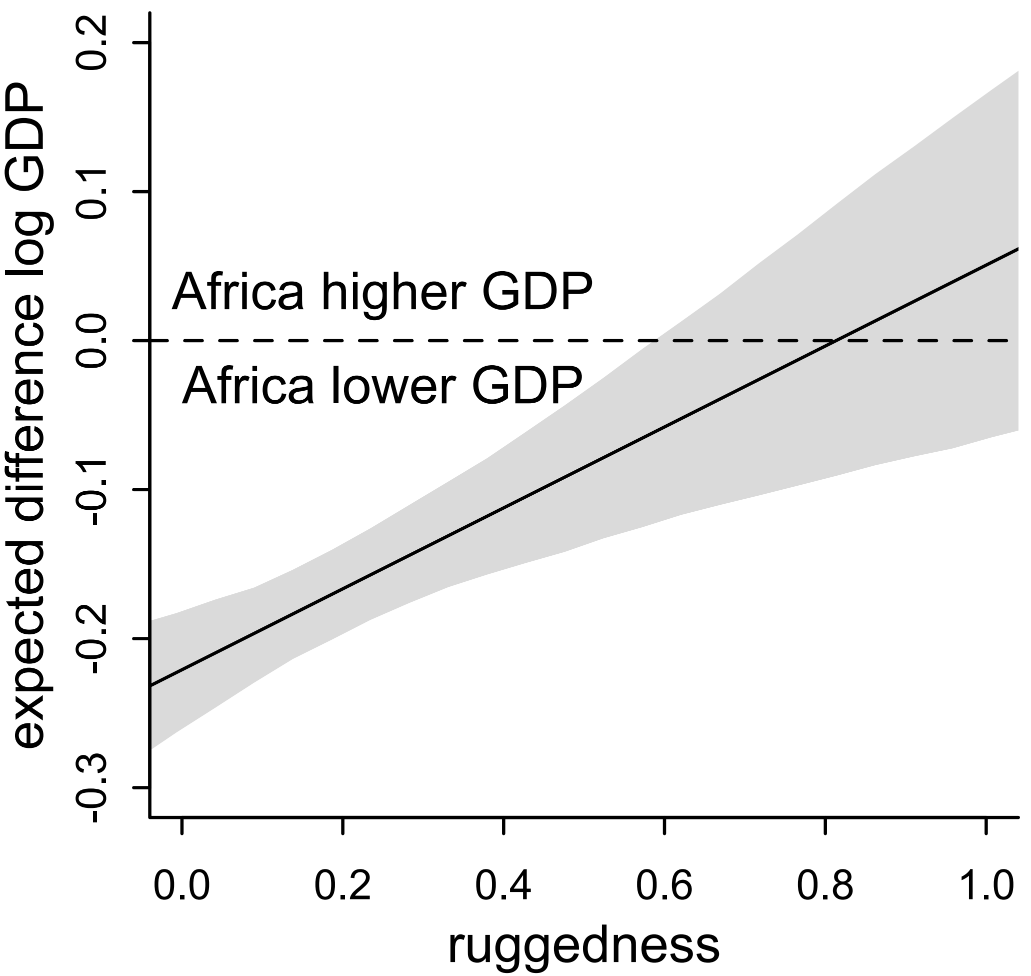
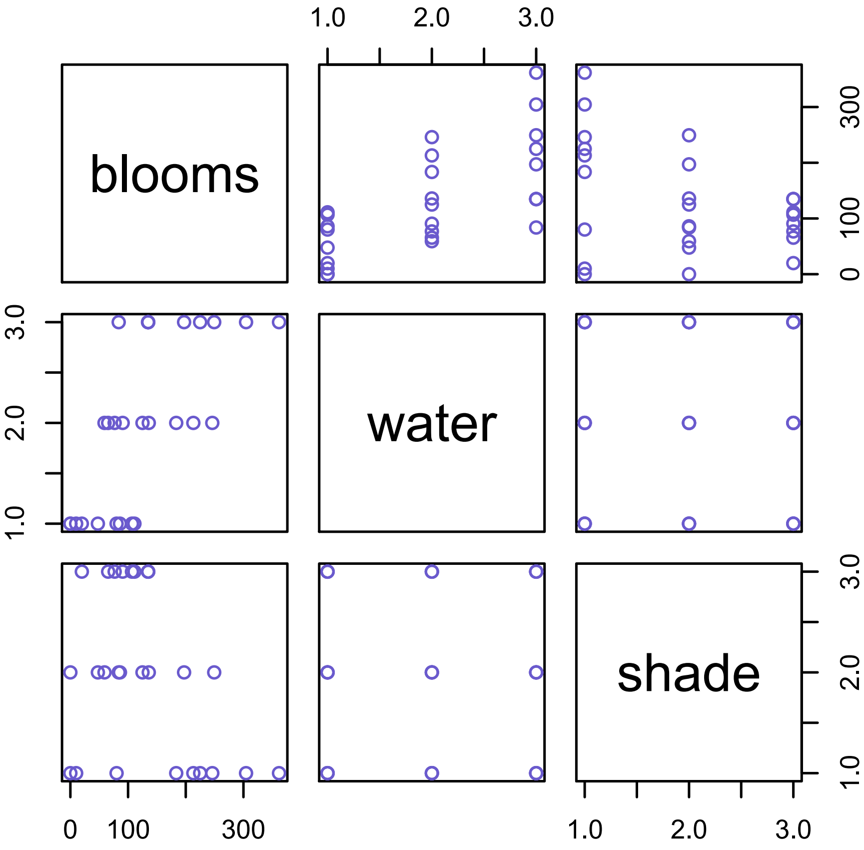


Figure 8.6

Continuous interactions

- data(tulips): 27 replicate blooms across three levels of both water and shade



Tulip blooms

No interaction:
water and shade have independent effects

$$b_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_w (w_i - \bar{w}) + \beta_s (s_i - \bar{s})$$

Interaction:
water and shade have interdependent effects

$$b_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_w (w_i - \bar{w}) + \beta_s (s_i - \bar{s}) + \beta_{ws} (w_i - \bar{w})(s_i - \bar{s})$$

How is interaction formed?

$$b_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_w(w_i - \bar{w}) + \beta_s(s_i - \bar{s}) + \beta_{ws}(w_i - \bar{w})(s_i - \bar{s})$$

How is interaction formed?

$$b_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_w(w_i - \bar{w}) + \beta_s(s_i - \bar{s}) + \beta_{ws}(w_i - \bar{w})(s_i - \bar{s})$$

$$\mu_i = \alpha + \gamma_{w,i} W_i + \beta_s S_i$$
$$\gamma_{w,i} = \beta_w + \beta_{ws} S_i$$

How is interaction formed?

$$
\begin{aligned}
b_i &\sim \text{Normal}(\mu_i, \sigma) \\
\mu_i &= \alpha + \beta_w(w_i - \bar{w}) + \beta_s(s_i - \bar{s}) + \beta_{ws}(w_i - \bar{w})(s_i - \bar{s})
\end{aligned}
$$

$$
\begin{aligned}
\mu_i &= \alpha + \gamma_{w,i} W_i + \beta_s S_i \\
\gamma_{w,i} &= \beta_w + \beta_{ws} S_i
\end{aligned}
$$

$$
\mu_i = \alpha + \underbrace{(\beta_w + \beta_{ws} S_i)}_{\gamma_{w,i}} W_i + \beta_s S_i = \alpha + \beta_w W_i + \beta_s S_i + \beta_{ws} S_i W_i
$$

Tulip model – no interaction

```
m8.6 <- quap(
    alist(
        blooms_std ~ dnorm( mu , sigma ) ,
        mu <- a + bw*water_cent + bs*shade_cent ,
        a ~ dnorm( 0.5 , 0.25 ) ,
        bw ~ dnorm( 0 , 0.25 ) ,
        bs ~ dnorm( 0 , 0.25 ) ,
        sigma ~ dexp( 1 )
    ) ,
    data=d )
```

R code 8.23

Plotting interaction

- Slope changes with values of other predictor, so use more than one plot
- Here, need three plots, *triptych*

trip•tych | ˈtriptik |

noun

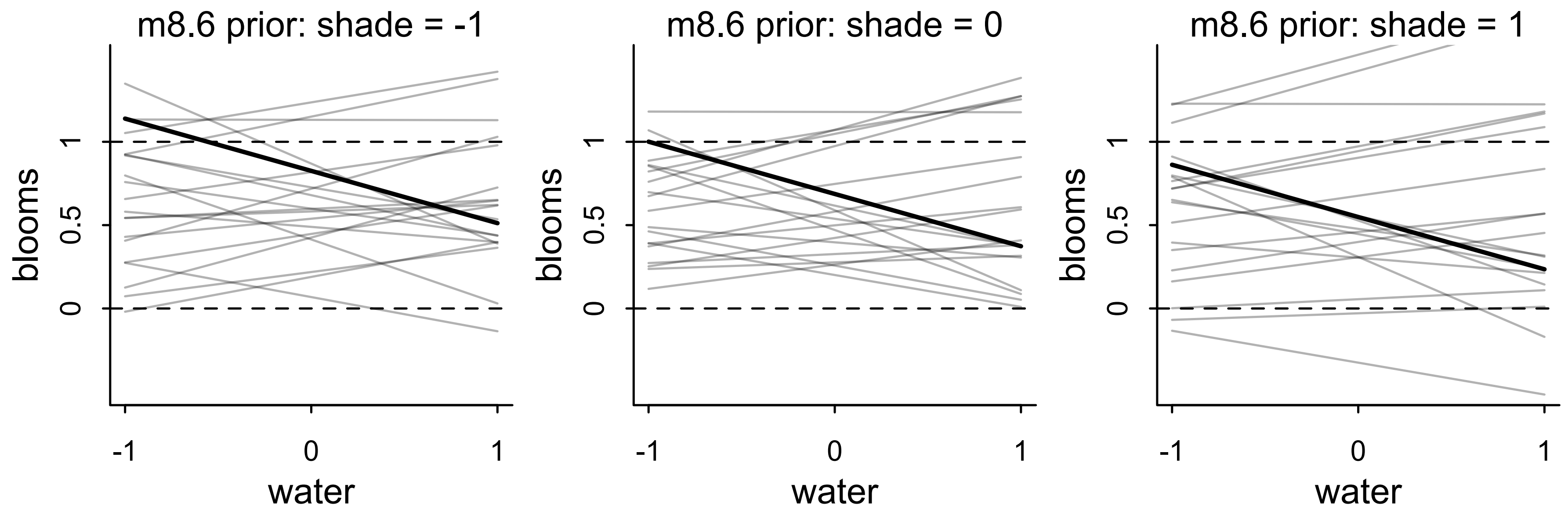
a picture or relief carving on three panels, typically hinged together side by side and used as an altarpiece.
• a set of three associated artistic, literary, or musical works intended to be appreciated together.

ORIGIN mid 18th cent. (denoting a set of three writing tablets hinged or tied together): from TRI- ***'three,'*** on the pattern of *diptych* .

Lewis Powell (1844–1865), before his hanging for conspiracy to assassinate Abraham Lincoln.

Prior predictions

No interaction



F

Posterior predictions

No interaction

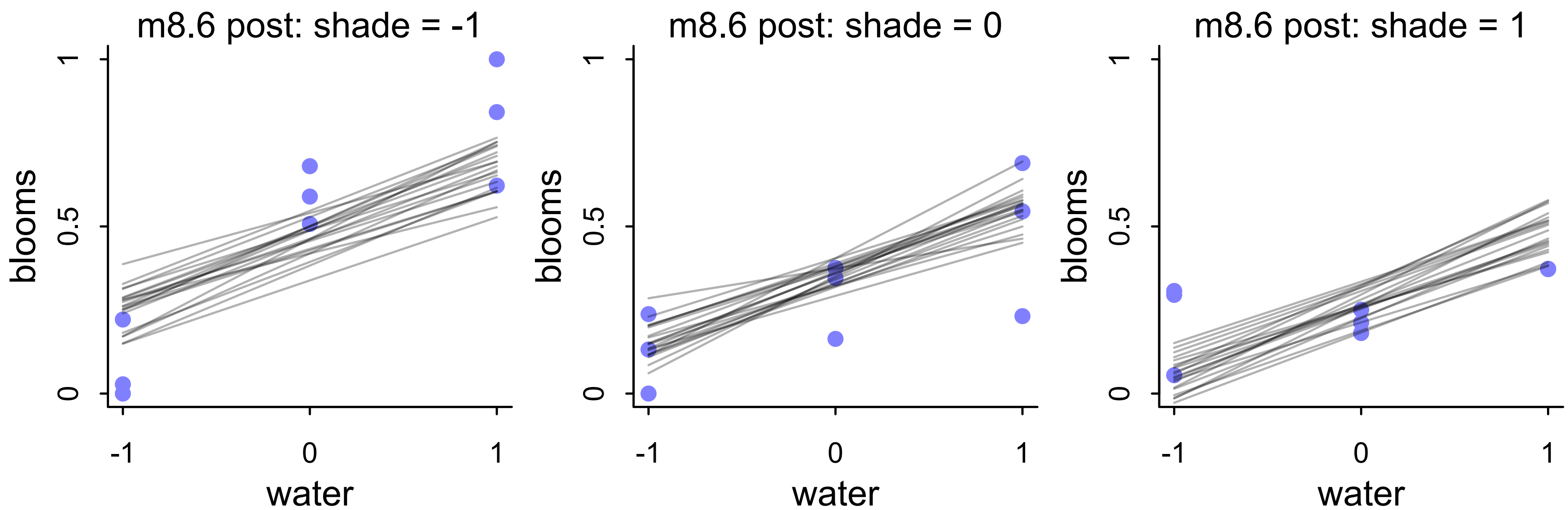


Figure 8.7

Tulip model – interaction

R code 8.24

```
m8.7 <- quap(
    alist(
        blooms_std ~ dnorm( mu , sigma ) ,
        mu <- a + bw*water_cent + bs*shade_cent + bws*water_cent*shade_cent ,
        a ~ dnorm( 0.5 , 0.25 ) ,
        bw ~ dnorm( 0 , 0.25 ) ,
        bs ~ dnorm( 0 , 0.25 ) ,
        bws ~ dnorm( 0 , 0.25 ) ,
        sigma ~ dexp( 1 )
    ) ,
    data=d )
```

Interpreting parameters very hard! Plot.

Prior predictions

No interaction

Interaction

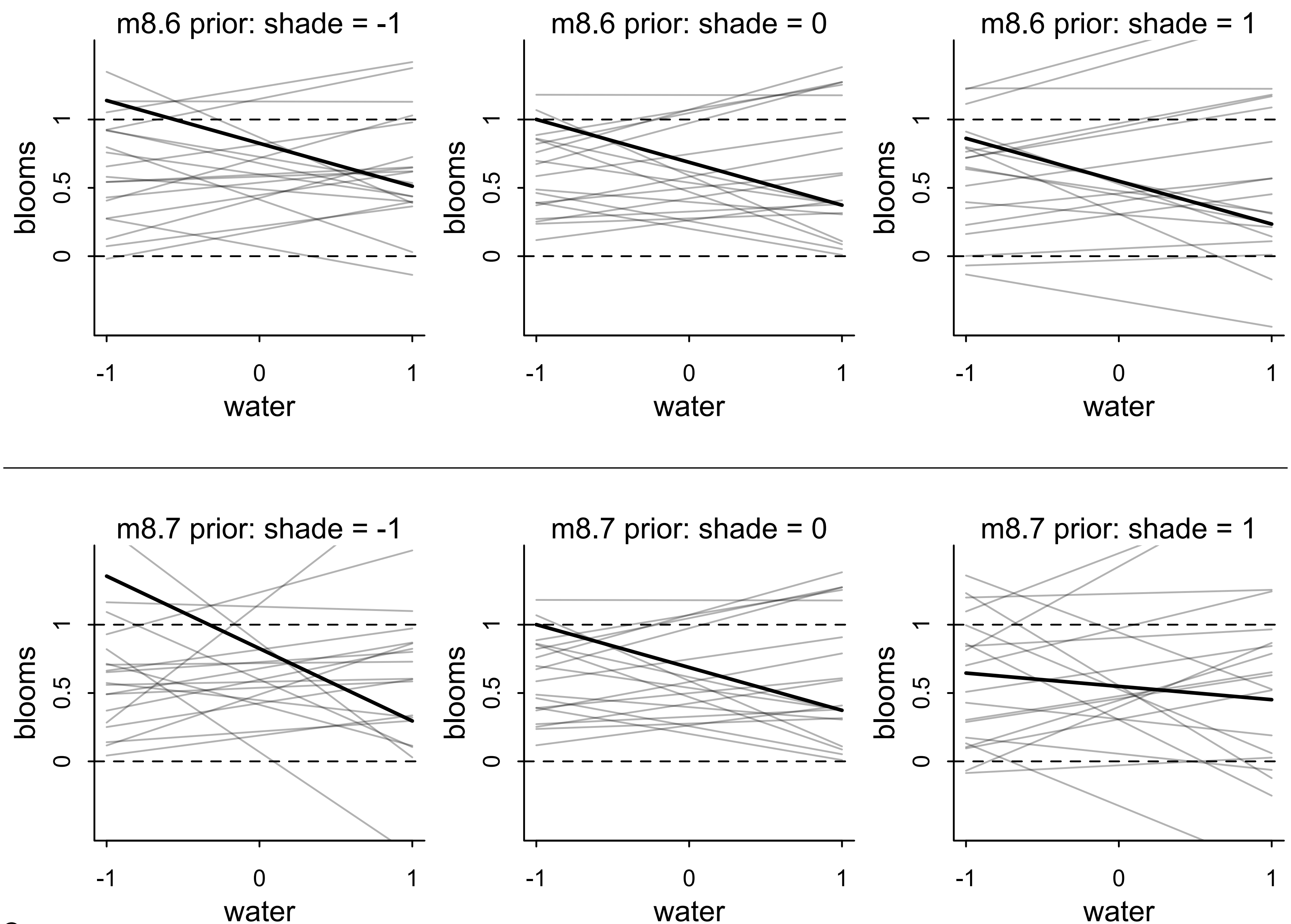


Figure 8.8

Posterior predictions

No interaction

Interaction

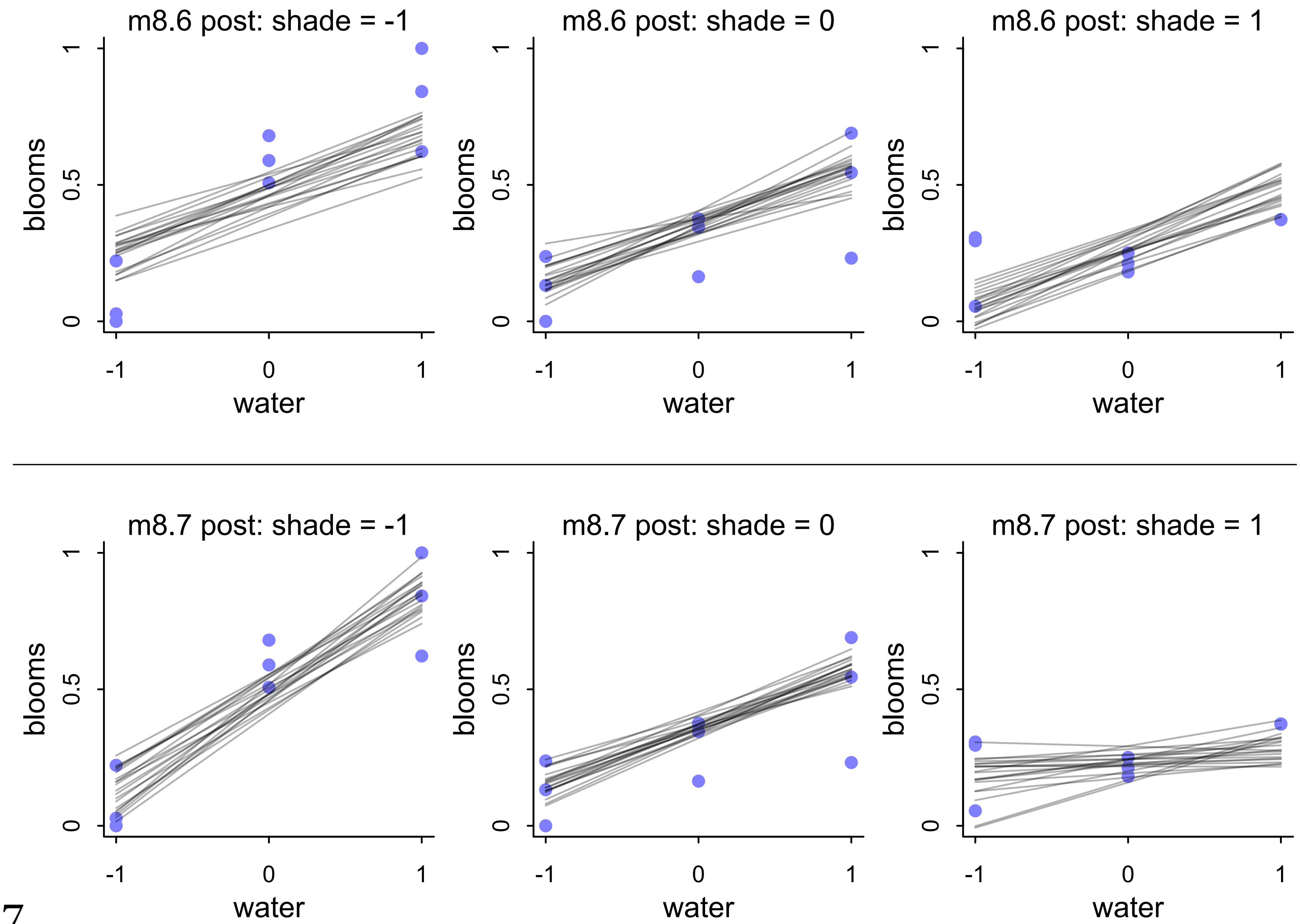


Figure 8.7

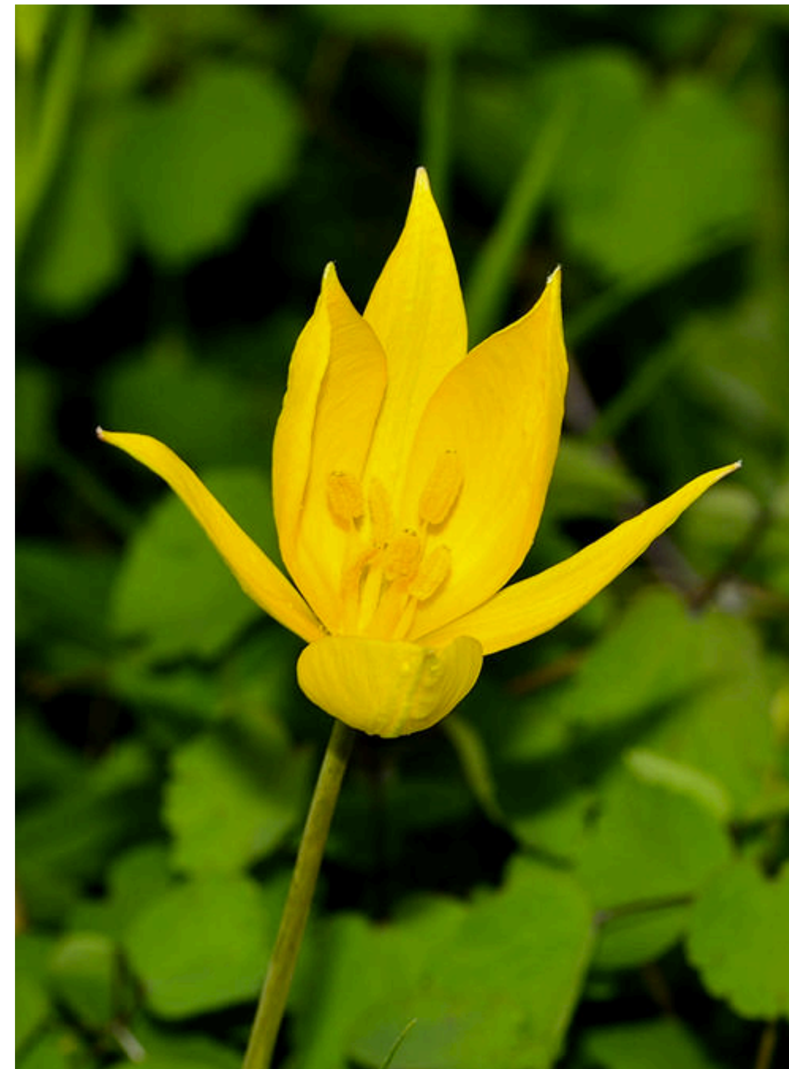
Causal thinking

- Tulip experiment:

$$\begin{tikzcd} S \arrow[dr] & & W \arrow[dl] \\ & B & \end{tikzcd}$$

- Tulip reality:

$$\begin{tikzcd} S \arrow[rr] \arrow[dr] & & W \arrow[dl] \\ & B & \end{tikzcd}$$



Interactions not always linear

- Suppose all tulip data collected under "cool" temperatures
- Under "hot" temperature, tulips do not bloom
- Interaction, but not a linear one
    - blooms goes to zero at threshold

Higher order interactions

- Just keep multiplying:

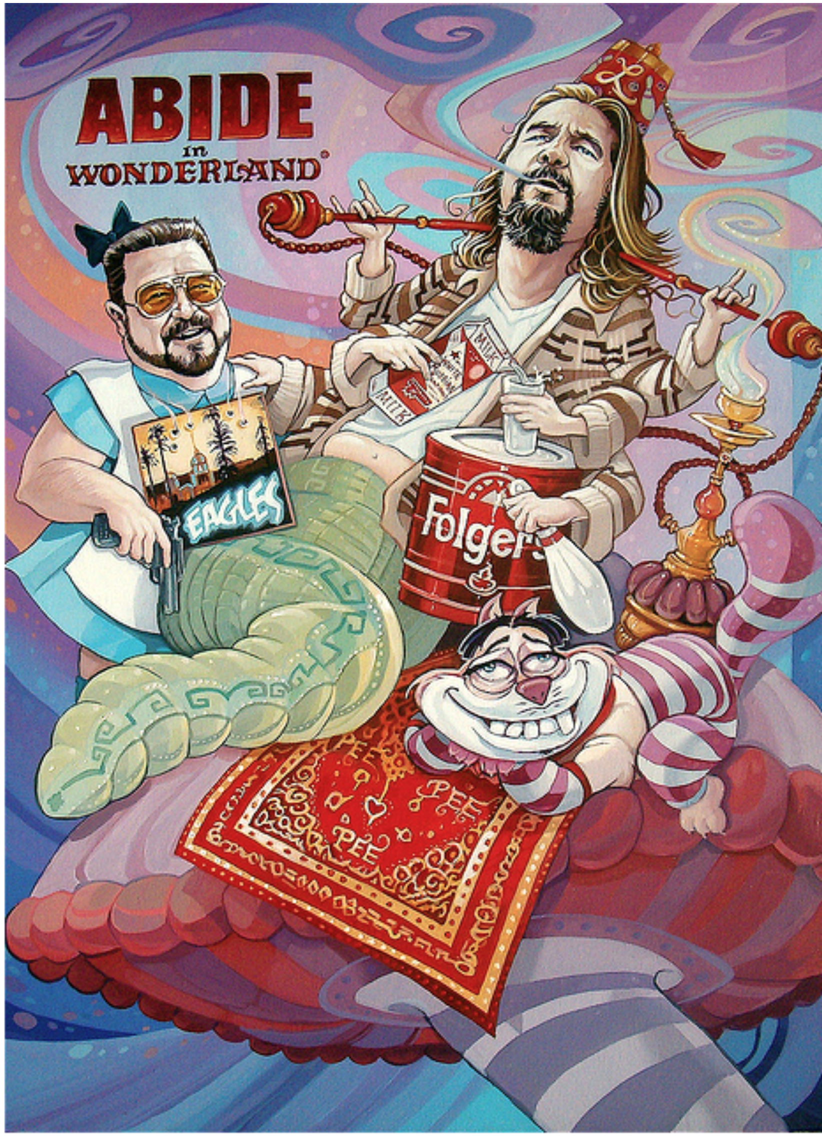
main effects

$$
\begin{aligned}
y_i &\sim \text{Normal}(\mu_i, \sigma), \\
\mu_i &= \alpha + \boxed{\beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i}} \\
&\quad + \boxed{\beta_{12} x_{1i} x_{2i} + \beta_{13} x_{1i} x_{3i} + \beta_{23} x_{2i} x_{3i}} \quad \textit{2-way interactions} \\
&\quad + \boxed{\beta_{123} x_{1i} x_{2i} x_{3i}.}
\end{aligned}
$$

3-way interaction

Higher order interactions

- Dangers of high-order interactions
  - Hard to interpret: "The extent to which the effect of x_1 depends upon the value of x_2 depends upon the value of x_3, dude."
  - Hard to estimate: need lots of data, must regularize
  - But you might really need them, because conditionality runs deep



The Dude abides high-order interactions

Higher order interactions

- `data(Wines2012)`
- Judgment of Princeton, 2012
  - New Jersey wines vs fine French wines
- Outcome variable: score
- Predictors:
  - region (NJ/FR)
  - nationality of judge (USA/FR-BE)
  - flight (red/white)

Higher order interactions

- Predictors: region, nationality of judge, flight
- Consider interactions:
    - Interaction of **region** and **judge** is bias.
      Bias depends upon **flight**.
    - Interaction of **judge** and **flight** is preference.
      Preference depends upon **region**.
    - Interaction of **region** and **flight** is comparative advantage.
      Advantage depends upon **judge**.

Interaction everywhere

- Interaction, regularization, responsibility
- Next time: Markov Chain Monte Carlo