Chapter 7

Lin 205C
Santiago Barreda

Chapter Precap

- Introduce models that include factors with many levels, and multiple different factors.
- Discuss the concepts of within and between-subjects factors with respect to factorial designs, orthogonality, and interactions.
- Introduce posterior prediction, and the use of this for model checking.
- Interactions and interaction plots are discussed, as is the way that these can be used to understand main effects and simple main effects.
- A model with two factors and an interaction, and random effects for all predictors, is fit, and the model is discussed and interpreted.
- We then present Bayesian *R*2 as a simple measure of model fit.
- Finally, we discuss type S and type M errors, regions of practical equivalence, and the problem of how to know when effects are 'real'.

Comparing Many Groups

	A_1	A_2	A_3	A_4
Within-Subjects	S_1 S_2 S_3 S_4	S_1 S_2 S_3 S_4	S_1 S_2 S_3 S_4	S_1 S_2 S_3 S_4

	A_1	A_2	A_3	A_4
Between-Subjects	S_1 S_2 S_3 S_4	S_5 S_6 S_7 S_8	S_9 S_{10} S_{11} S_{12}	S_{13} S_{14} S_{15} S_{16}

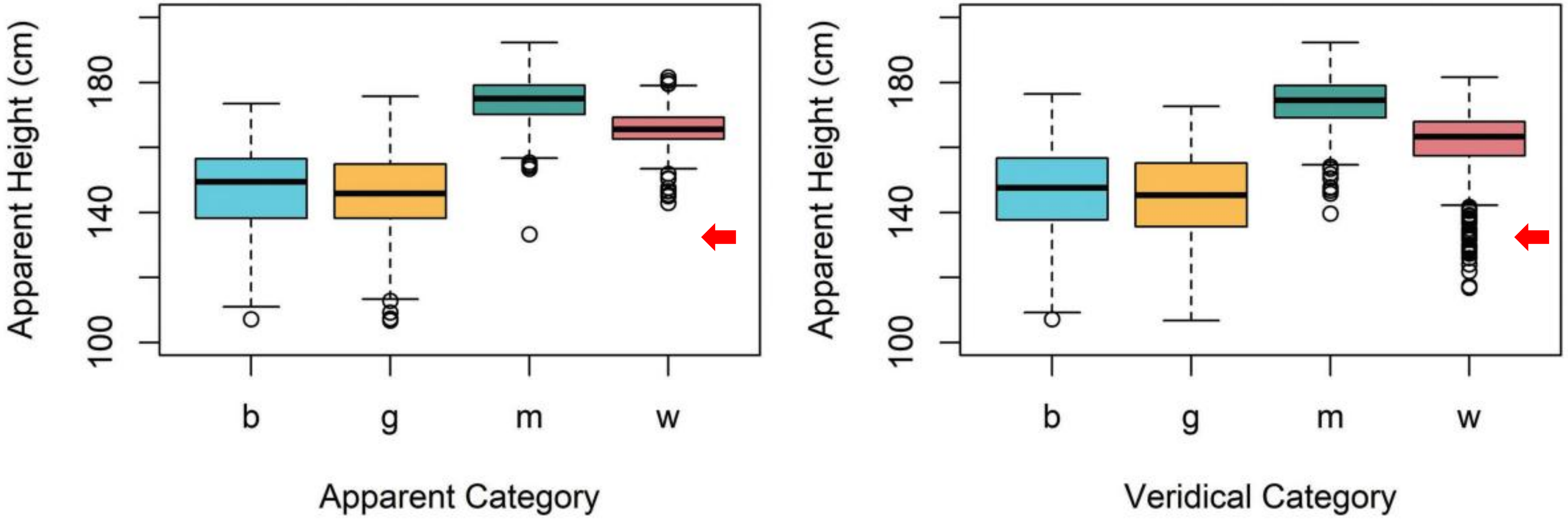
Data and Research Questions

```
library (brms)
library (bmmb)
options (contrasts = c('contr.sum','contr.sum'))
data (exp_data)
```

- `L`: A number from 1 to 15 indicating which *listener* responded to the trial.
- `C`: A letter representing the speaker *category* (`b` = boy, `g` = girl, `m` = man, `w` =woman) reported by the listener for each trial.
- `height`: A number representing the *height* (in centimeters) reported for the speaker on each trial.
- `S`: A number from 1 to 139 indicating which *speaker* produced the trial stimulus.

(Q1) Does apparent speaker height vary systematically across apparent speaker categories?

Apparent Speaker Category




```
# average correct category identification by category
tapply(exp_data$C == exp_data$C_v, exp_data$C_v, mean)
##      b      g      m      w
## 0.5778 0.6456 0.9274 0.7097
```

Comparing Many Groups

- When estimating a factor with J levels, you can only estimate (at most) J-1 unique values.

- The values of the J levels are <u>constrained to sum to zero</u> under sum coding.

- Therefore: The value of level J is equal to the *negative sum* of the J-1 levels.

$$A_1 + A_1 + A_3 + \cdots + A_{J-1} + A_J = 0$$
$$A_1 + A_1 + A_3 + \cdots + A_{J-1} = -A_J$$
$$-(A_1 + A_1 + A_3 + \cdots + A_{J-1}) = A_J$$

Description of the Model

```
height ~ C + (C|L) + (1|S)
```

Only 3 predicted category means

```
height ~ C1 + C2 + C3 + (C1 + C2 + C3|L) + (1|S)
```

```
height ~ x1*C1 + x2*C2 + x3*C3 + (x1*C1 + x2*C2 + x3*C3|L) + (1|S)
```

Not a real formula

Description of the Model

```
height ~ x1*C1 + x2*C2 + x3*C3 + (x1*C1 + x2*C2 + x3*C3|L) + (1|S)
```

```
contr.sum(1:4)
##   [,1] [,2] [,3]
## 1    1    0    0
## 2    0    1    0
## 3    0    0    1
## 4   -1   -1   -1
```

$$\mathrm{C}_{[1]} = 1, \mathrm{C}_{[2]} = 2, \mathrm{C}_{[3]} = 3, \mathrm{C}_{[4]} = 4 \ldots$$

$$\mu_{[i]} = x_1 \cdot C1 + x_2 \cdot C2 + x_3 \cdot C3$$

$$\mu_{[1]} = 1 \cdot C1 + 0 \cdot C2 + 0 \cdot C3 = C1$$

$$\mu_{[2]} = 0 \cdot C1 + 1 \cdot C2 + 0 \cdot C3 = C2$$

$$\mu_{[3]} = 0 \cdot C1 + 0 \cdot C2 + 1 \cdot C3 = C3$$

$$\mu_{[4]} = -1 \cdot C1 - 1 \cdot C2 - 1 \cdot C3 = -(C1 + C2 + C3) = C4$$

Description of the Model

$$\text{height}_{[i]} \sim \mathrm{t}\left(\nu, \mu_{[i]}, \sigma\right)$$

$$\mu_{[i]} = \text{Intercept} + C_{\left[\mathsf{C}_{[i]}\right]} + L_{\left[\mathsf{L}_{[i]}\right]} + C_{\left[\mathsf{C}_{[i]}\right]} : L_{\left[\mathsf{L}_{[i]}\right]} + S_{\left[\mathsf{S}_{[i]}\right]}$$

$$\text{Priors :}$$

$$S_{[\bullet]} \sim \mathrm{N}\left(0, \sigma_S\right)$$

$$\begin{bmatrix} L_{[\bullet]} \\ C_{[1]} : L_{[\bullet]} \\ C_{[2]} : L_{[\bullet]} \\ C_{[3]} : L_{[\bullet]} \end{bmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \Sigma \right)$$

$$\text{Intercept} \sim \mathrm{N}(156, 12)$$

$$C_{[\bullet]} \sim \mathrm{N}(0, 12)$$

$$\sigma_L, \sigma_{C[1]:L}, \sigma_{C[2]:L}, \sigma_{C[3]:L}, \sigma_S \sim \mathrm{N}(0, 12)$$

$$\nu \sim \text{gamma}(2, 0.1)$$

$$R \sim \text{LKJCorr}(2)$$

Comparing Models

$$\text{height}_{[i]} \sim \text{t}\left(\nu, \mu_{[i]}, \sigma\right)$$
$$\mu_{[i]} = \text{Intercept} + C_{[\mathsf{C}[i]]} + L_{[\mathsf{L}[i]]} + C_{[\mathsf{C}[i]]} : L_{[\mathsf{L}[i]]} + S_{[\mathsf{S}[i]]}$$

$$\text{Priors:}$$
$$S_{[\bullet]} \sim \text{N}(0, \sigma_S)$$
$$\begin{bmatrix} L_{[\bullet]} \\ C_{[1]} : L_{[\bullet]} \\ C_{[2]} : L_{[\bullet]} \\ C_{[3]} : L_{[\bullet]} \end{bmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \Sigma\right)$$

$$\text{Intercept} \sim \text{N}(156, 12)$$
$$C_{[\bullet]} \sim \text{N}(0, 12)$$
$$\sigma_L, \sigma_{C[1]:L}, \sigma_{C[2]:L}, \sigma_{C[3]:L}, \sigma_S \sim \text{N}(0, 12)$$
$$\nu \sim \text{gamma}(2, 0.1)$$
$$R \sim \text{LKJCorr}(2)$$

$$\text{height}_{[i]} \sim \text{t}\left(\nu, \mu_{[i]}, \sigma\right)$$
$$\mu_{[i]} = \text{Intercept} + A + L_{[\mathsf{L}[i]]} + A : L_{[\mathsf{L}[i]]} + S_{[\mathsf{S}[i]]}$$

$$\text{Priors:}$$
$$S_{[\bullet]} \sim \text{N}(0, \sigma_S)$$
$$\begin{bmatrix} L_{[\bullet]} \\ A : L_{[\bullet]} \end{bmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma\right)$$

$$\text{Intercept} \sim \text{t}(3, 156, 12)$$
$$A \sim \text{t}(3, 0, 12)$$
$$\sigma, \sigma_L, \sigma_{A:L}, \sigma_S \sim \text{t}(3, 0, 12)$$
$$\nu \sim \text{gamma}(2, 0.1)$$
$$R \sim \text{LKJCorr}(2)$$

$$\Sigma = \begin{bmatrix} \sigma_L & 0 \\ 0 & \sigma_{A:L} \end{bmatrix} \cdot R \cdot \begin{bmatrix} \sigma_L & 0 \\ 0 & \sigma_{A:L} \end{bmatrix}$$

Fitting the Model

- No new priors!

```r
# Fit the model yourself
priors = c(brms::set_prior("student_t(3,156, 12)", class = "Intercept"),
           brms::set_prior("student_t(3,0, 12)", class = "b"),
           brms::set_prior("student_t(3,0, 12)", class = "sd"),
           brms::set_prior("lkj_corr_cholesky (2)", class = "cor"),
           brms::set_prior("gamma(2, 0.1)", class = "nu"),
           brms::set_prior("student_t(3,0, 12)", class = "sigma"))

model_four_groups =
  brms::brm (height ~ C + (C|L) + (1|S), data = exp_data, chains = 4,
             cores = 4, warmup = 1000, iter = 5000, thin = 4,
             prior = priors, family = "student")

# or download it from the GitHub page:
model_four_groups = bmmb::get_model ('7_model_four_groups.RDS')
```

The Model Fixed Effects

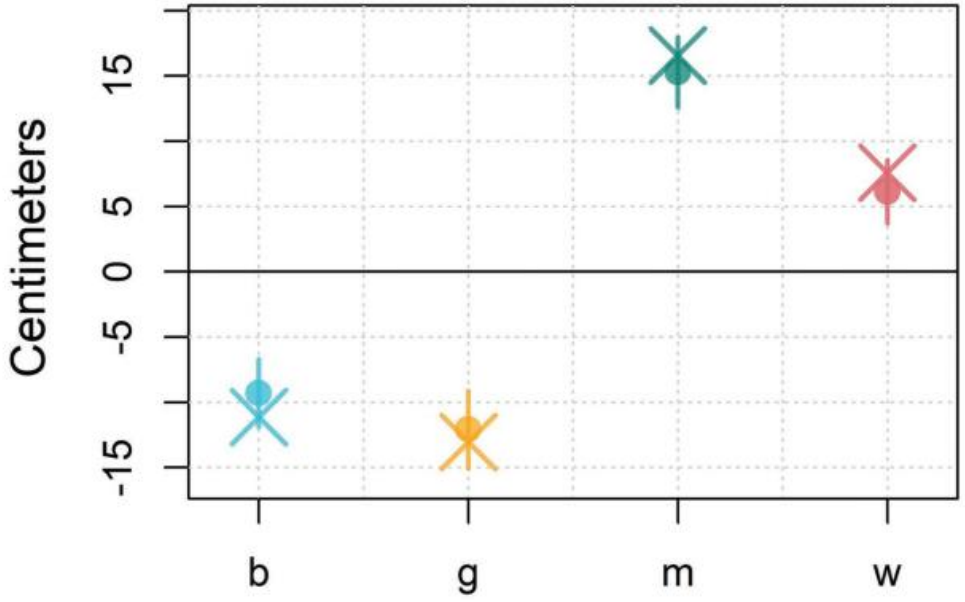
```
brms::fixef(model_four_groups)
##           Estimate Est.Error   Q2.5   Q97.5
## Intercept  158.581     1.083 156.41 160.685
## C1          -9.318     1.310 -11.87  -6.682
## C2         -12.082     1.494 -15.07  -9.110
## C3          15.286     1.327  12.59  17.950
```

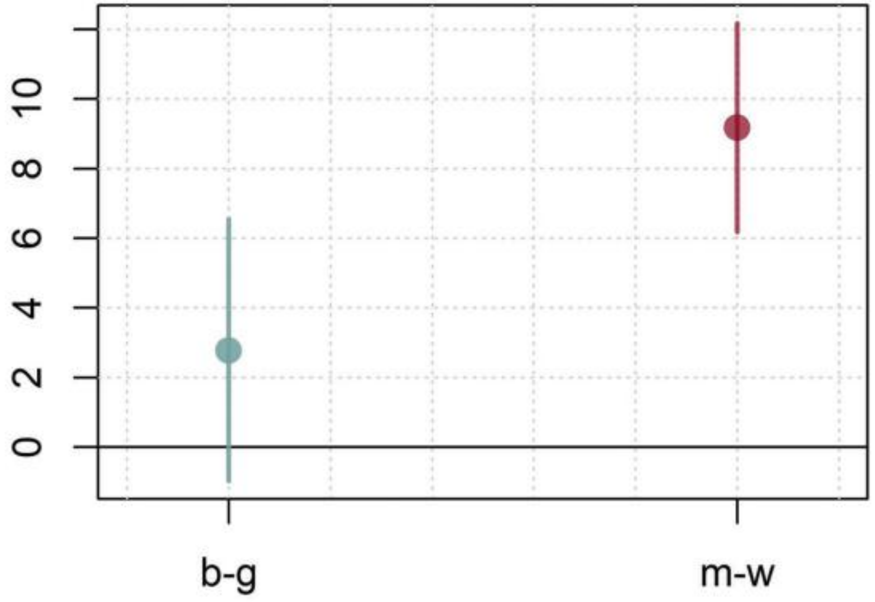
```
# group means
means = tapply (exp_data$height, exp_data$C, mean)
# Intercept = mean of means
mean (means)
## [1] 158

# Group effects = centered group means
means - mean (means)
##      b      g      m      w
## -11.12 -13.03  16.56   7.59
```

Working with the Fixed Effects

```
# missing group effect
bmmb::short_hypothesis (model_four_groups, c("-(C1+C2+C3) = 0"))
##     Estimate Est.Error  Q2.5 Q97.5        hypothesis
## H1     6.114     1.216 3.686 8.543 (-(C1+C2+C3)) = 0
```

```
# find differences between groups
comparisons = bmmb::short_hypothesis (model_four_groups,
                                      c("C1 = C2", "C3 = -(C1+C2+C3)"))
```





Multiple Factors at Once

Within-Subjects

	B_1	B_2
A_1	S_1 S_2 S_3 S_4	S_1 S_2 S_3 S_4
A_2	S_1 S_2 S_3 S_4	S_1 S_2 S_3 S_4

Between-Subjects

	B_1	B_2
A_1	S_1 S_2 S_3 S_4	S_5 S_6 S_7 S_8
A_2	S_9 S_{10} S_{11} S_{12}	S_{13} S_{14} S_{15} S_{16}

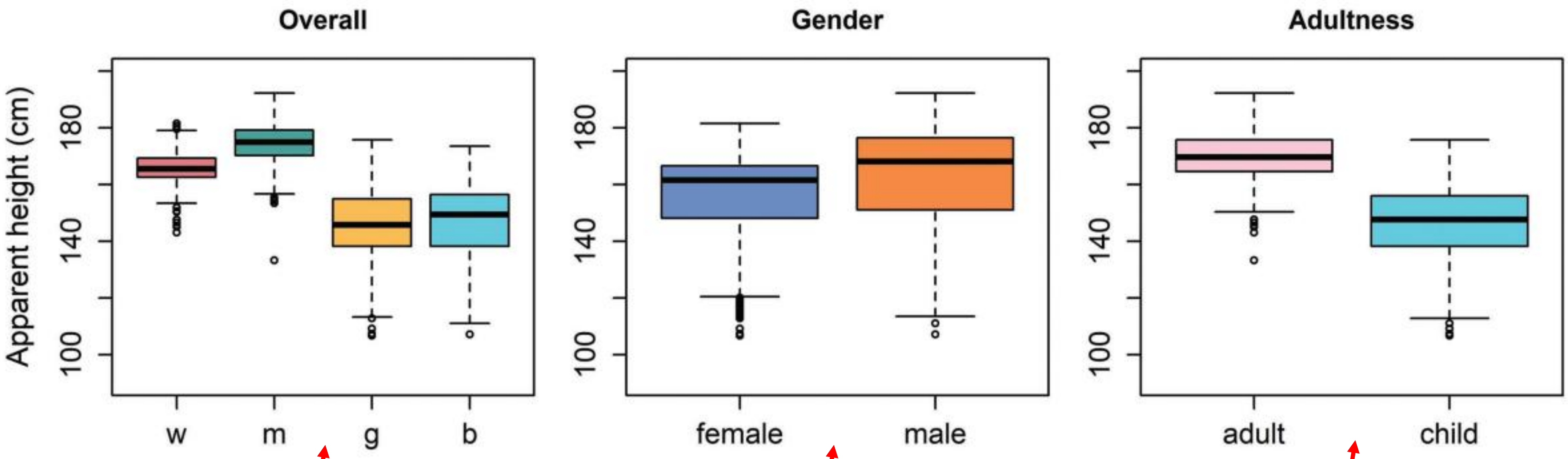
Mixed-design

	B_1	B_2
A_1	S_1 S_2 S_3 S_4	S_1 S_2 S_3 S_4
A_2	S_5 S_6 S_7 S_8	S_5 S_6 S_7 S_8

Factorial/orthogonal design

A ‘cell’ (i.e., a combination of factors)

Multiple ~~Factors~~ Experiments at Once



Are there any differences?

Are there differences because of apparent gender?

Are there differences because of apparent age?

Data and Research Questions

- `L`: A number from 1 to 15 indicating which *listener* responded to the trial.
- `height`: A number representing the *height* (in centimeters) reported for the speaker on each trial.
- `S`: A number from 1 to 139 indicating which *speaker* produced the trial stimulus.
- `G`: The *apparent gender* of the speaker indicated by the listener, `f` (female) or `m` (male).
- `A`: The *apparent age* of the speaker indicated by the listener, `a` (adult) or `c` (child).

(Q1) Does average apparent height differ across levels of apparent age?
(Q2) Does average apparent height differ across levels of apparent gender?

Description of the Model

~~`brm (height ~ A + (A|L) + (1|S)`~~

~~`brm (height ~ G + (G|L) + (1|S)`~~

```
height ~ A + G + (A + G|L) + (1|S)
```

$$\mu_{[i]} = \text{Intercept} + \left(C_{[\mathrm{C}[i]]} \right) \longrightarrow \mu_{[i]} = \text{Intercept} + \left(A_{[\mathrm{A}[i]]} + G_{[\mathrm{G}[i]]} \right)$$

Description of the Model

```
height ~ A + G + (A + G|L) + (1|S)
```

$$\text{height}_{[i]} \sim \mathrm{t}\left(\nu, \mu_{[i]}, \sigma\right)$$

$$\mu_{[i]} = \text{Intercept} + A + G + L_{\left[\mathsf{L}[i]\right]} + A : L_{\left[\mathsf{L}[i]\right]} + G : L_{\left[\mathsf{L}[i]\right]} + S_{\left[\mathsf{S}[i]\right]}$$

Priors:

$$S_{[\bullet]} \sim \mathrm{N}\left(0, \sigma_S\right)$$

$$\begin{bmatrix} L_{[\bullet]} \\ A : L_{[\bullet]} \\ G : L_{[\bullet]} \end{bmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \Sigma\right)$$

$$\text{Intercept} \sim \mathrm{N}(156, 12)$$

$$A, G \sim \mathrm{N}(0, 12)$$

$$\sigma_L, \sigma_{A:L}, \sigma_{G:L}, \sigma_S \sim \mathrm{N}(0, 12)$$

$$\sigma \sim \mathrm{N}(0, 12)$$

$$\nu \sim \text{gamma}(2, 0.1)$$

$$R \sim \text{LKJCorr}(2)$$

Fitting the Model

```
# Fit the model yourself
priors = c(brms::set_prior("student_t(3,156, 12)", class = "Intercept"),
           brms::set_prior("student_t(3,0, 12)", class = "b"),
           brms::set_prior("student_t(3,0, 12)", class = "sd"),
           brms::set_prior("lkj_corr_cholesky (2)", class = "cor"),
           brms::set_prior("gamma(2, 0.1)", class = "nu"),
           brms::set_prior("student_t(3,0, 12)", class = "sigma"))

model_both =
  brms::brm (height ~ A + G + (A + G|L) + (1|S), data = exp_data,
             chains = 4, cores = 4, warmup = 1000, iter = 5000,
             thin = 4, prior = priors, family = "student")

# Or download it from the GitHub page:
model_both = bmmb::get_model ('7_model_both.RDS')
```

Interpreting the Model

```
# inspect the fixed effects
brms::fixef (model_both)
##           Estimate Est.Error    Q2.5   Q97.5
## Intercept  158.654    1.1947 156.319 161.025
## A1           9.979    1.2107   7.562  12.282
## G1          -2.205    0.5524  -3.288  -1.124
```

```
# Intercept
mean (tapply (exp_data$height, exp_data$C, mean))
## [1] 158
```

```
# Age effect
diff (tapply (exp_data$height, exp_data$A, mean) ) / -2
##     c
## 12.16

# Gender effect
diff (tapply (exp_data$height, exp_data$G, mean) ) / -2
##      m
## -3.166
```

Predicting Group Means

```
means_pred = bmmb::short_hypothesis (model_both,
                                     c("Intercept + -A1 + -G1 = 0", # boys
                                       "Intercept + -A1 +  G1 = 0", # girls
                                       "Intercept +  A1 + -G1 = 0", # men
                                       "Intercept +  A1 +  G1 = 0")) # women

means_pred
##     Estimate Est.Error  Q2.5 Q97.5                 hypothesis
## H1     150.9    2.2410 146.6 155.3 (Intercept+-A1+-G1) = 0
## H2     146.5    2.3680 142.0 151.2  (Intercept+-A1+G1) = 0
## H3     170.8    1.2988 168.2 173.4  (Intercept+A1+-G1) = 0
## H4     166.4    0.6901 165.0 167.8   (Intercept+A1+G1) = 0
```

```
tapply (exp_data$height, exp_data$C, mean)
##     b     g     m     w
## 146.9 145.0 174.5 165.6
```

Posterior Prediction

- The posterior predictive distribution is the distribution of possible data given your parameter values and probability model.

model estimates

$$\tilde{y}_{[i]} \sim \mathrm{N}\left(\mu_{[i]}, \sigma\right)$$

- In a *posterior predictive check*, you compare your simulated data ($\tilde{y}$) to your actual data (y).

Using Models to Predict

$$y_{[i]} \sim \mathrm{N}\left(\mu_{[i]}, \sigma\right)$$

linear predictor/
expected value

$$\boxed{\mu_{[i]}} = \text{Intercept} + A_{\left[\mathsf{A}[i]\right]} + B_{\left[\mathsf{B}[i]\right]} + C_{\left[\mathsf{C}[i]\right]}$$

```
# linear predictor
y_lin_pred = fitted (model_both)

# posterior prediction
y_post_pred = predict (model_both)
```

The data your model predicts given its estimate of μ and σ for a given condition.

Posterior prediction includes the error!!

$$\boxed{\tilde{y}_{[i]}} \sim \mathrm{N}\left(\mu_{[i]}, \boxed{\sigma}\right)$$

Using Models to Predict

```
# linear predictions
head (y_lin_pred)
##        Estimate Est.Error  Q2.5 Q97.5
## [1,]      156.8     2.182 152.5 161.0
## [2,]      161.7     2.059 157.6 165.6
## [3,]      160.6     1.983 156.6 164.3
## [4,]      162.3     1.846 158.5 165.9
## [5,]      163.0     1.992 159.0 167.0
## [6,]      155.7     2.371 151.1 160.5
```

```
# posterior predictions
head (y_post_pred)
##        Estimate Est.Error  Q2.5 Q97.5
## [1,]      157.0     7.915 141.3 172.2
## [2,]      161.7     9.137 146.3 177.2
## [3,]      160.5     8.210 144.8 175.9
## [4,]      162.1     8.312 146.4 177.0
## [5,]      162.8     7.639 147.8 177.9
## [6,]      155.7     7.708 140.0 170.8
```

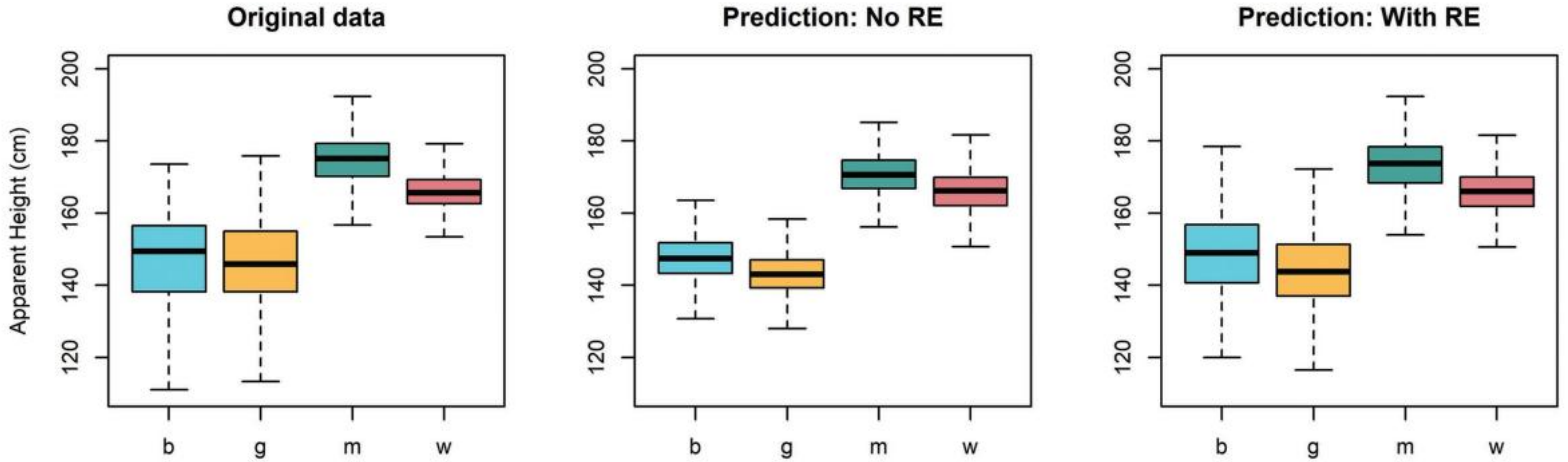
Fixed Effects Prediction

```
# posterior prediction
y_post_pred = predict (model_both)
```

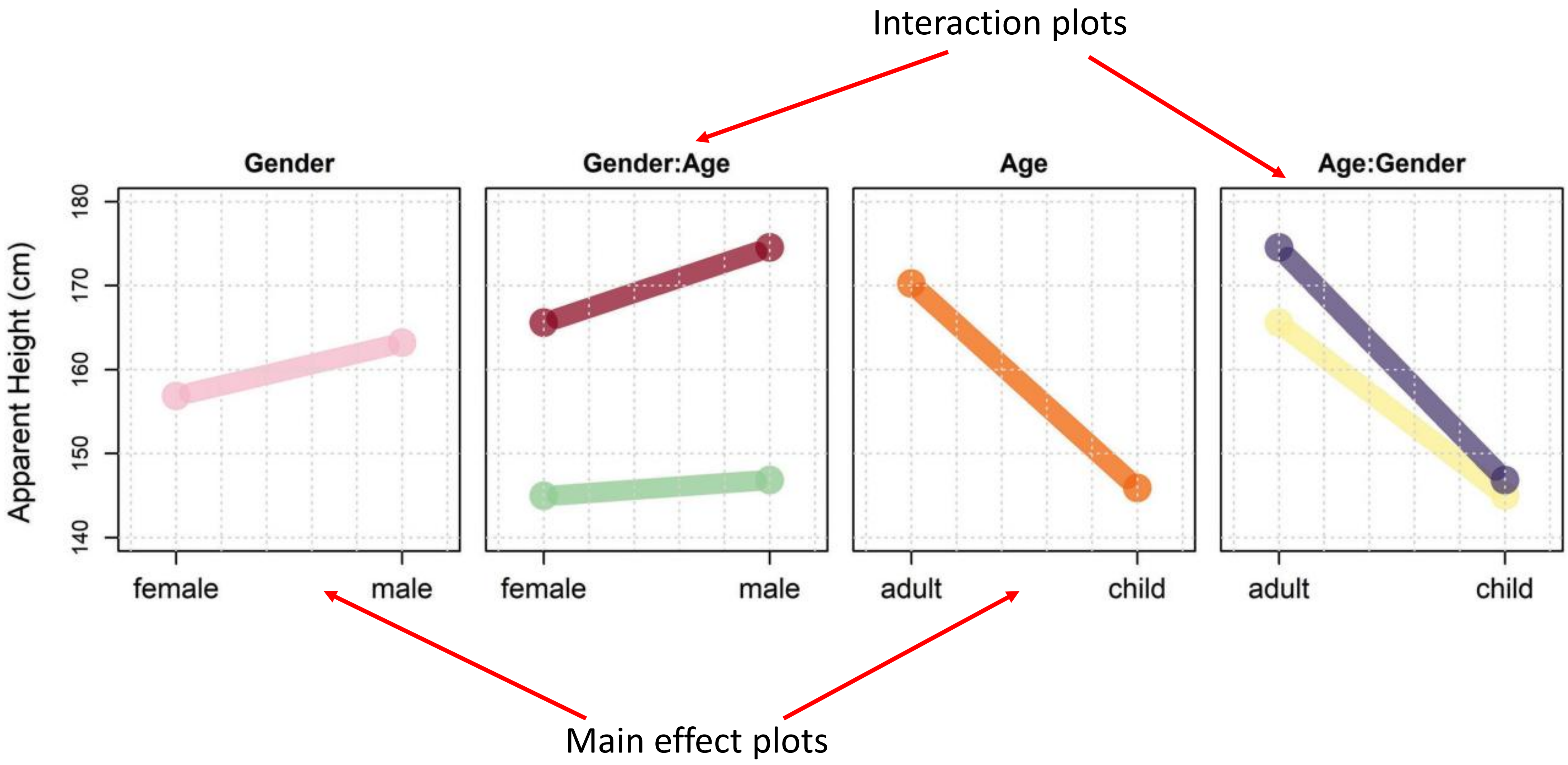
```
height ~ A + G + (A + G|S) + (1|L)
```

```
y_post_pred_no_re = predict (model_both, re_formula = NA)
```

```
height ~ A + G
```



Interaction Plots

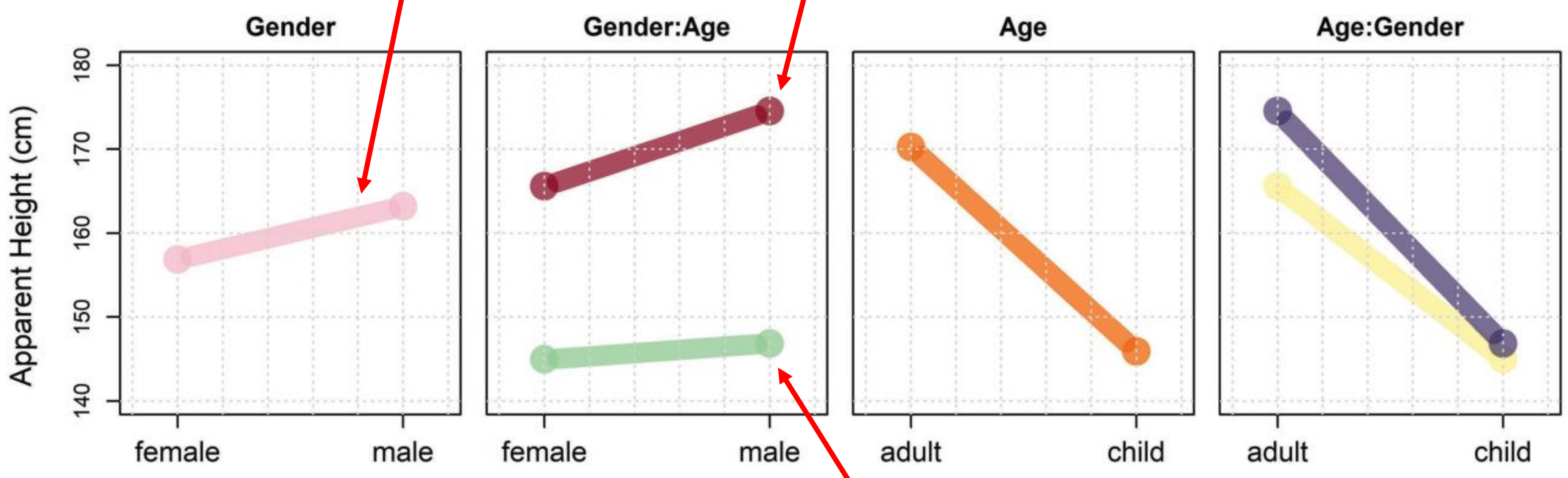


Simple (Main) Effects

- Simple (main) effects: the effect of one factor at <u>one</u> level of another factor.

Main effect for gender

Simple effect for gender
given apparent adult

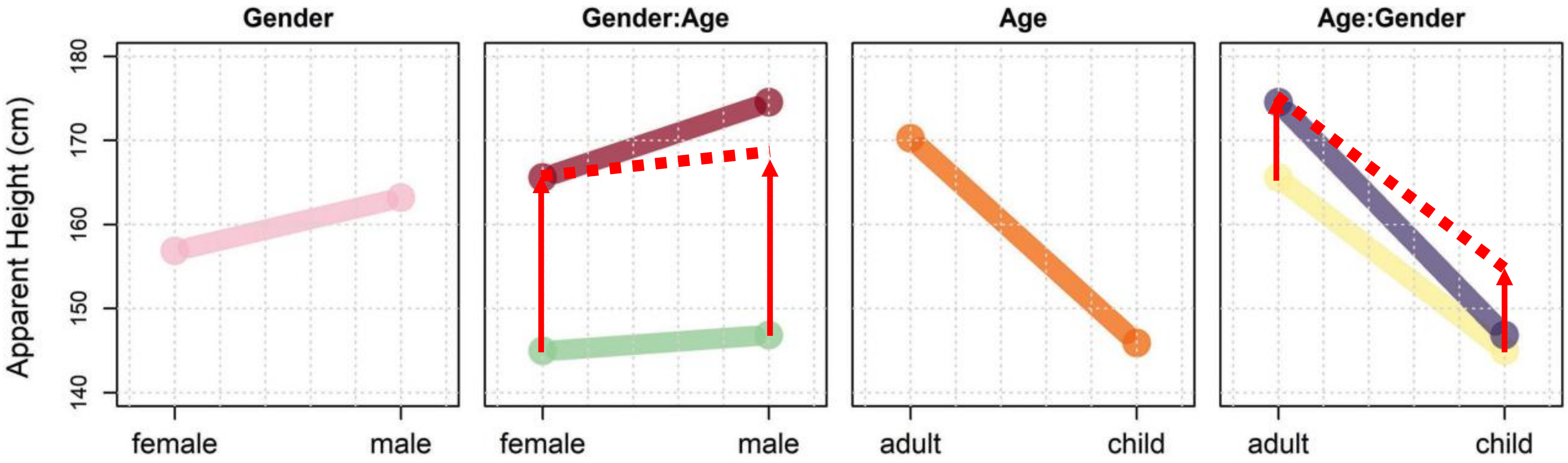


Simple effect for gender
given apparent child

Interaction Plots

- In the absence of an interaction, all points are shifted by the same amount (i.e., the main effect).

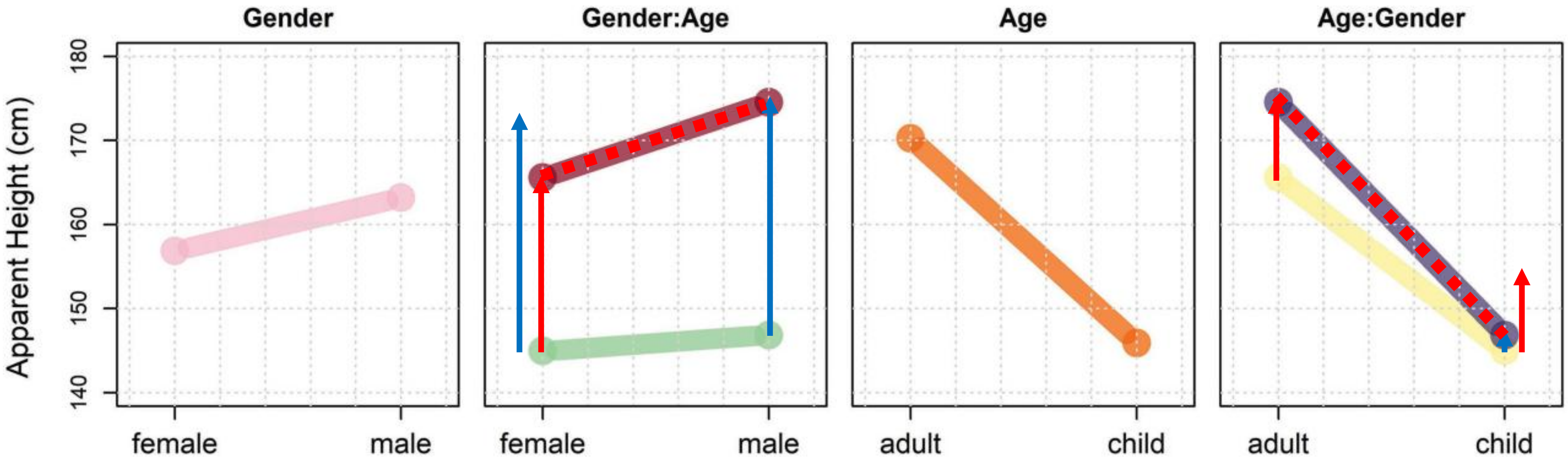
- The simple effects are parallel.



Interaction Plots

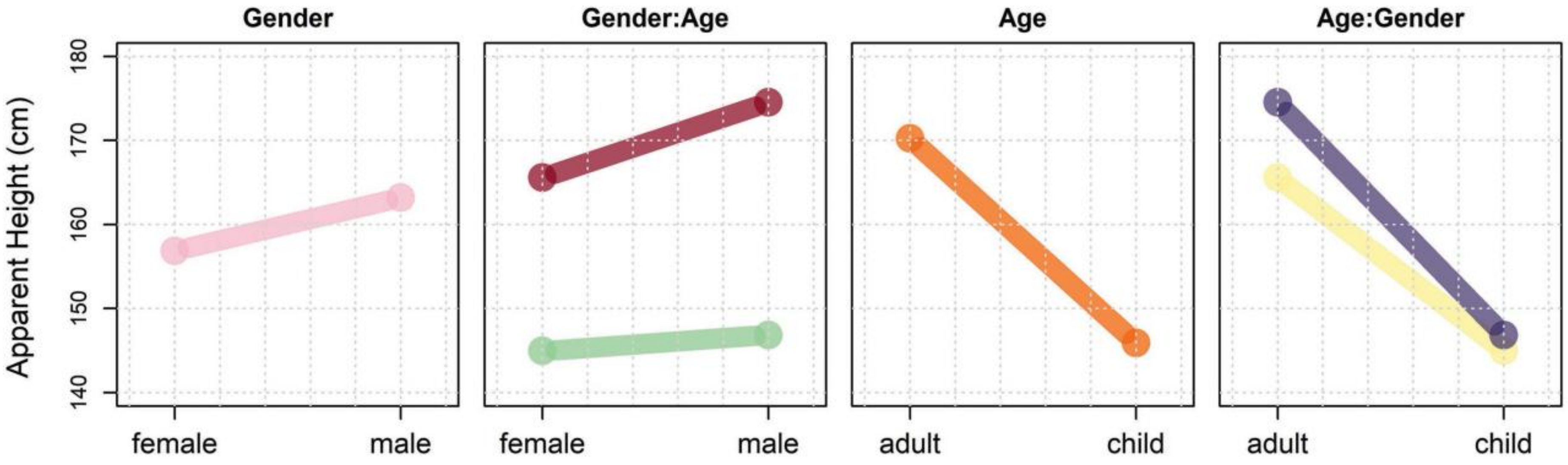
- If interactions exist, lines can be shifted by different amounts at different points (i.e., according to the conditional effects).

- Non-parallelism = Interaction!



Simple (Main) Effects

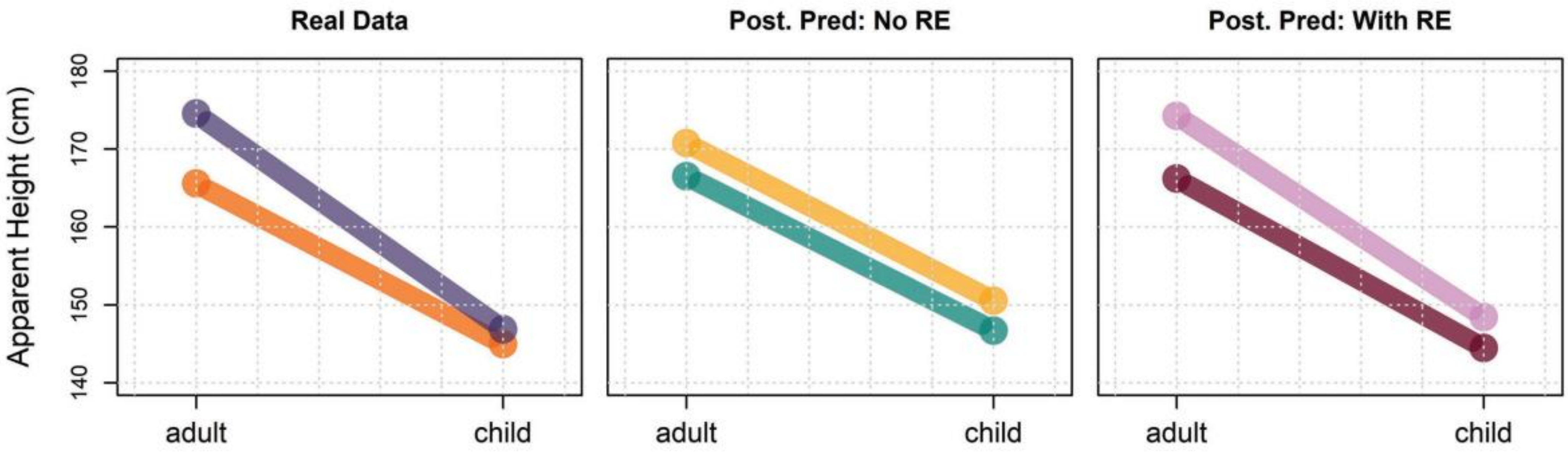
- When interactions are present, factors are interpreted in terms of their simple effects.
- E.g., below: We know the effect for gender varies as a function of apparent age. In what way does it vary?



No Interaction in Model = Bad

- Can’t even tell us if there are interactions: the simple effects are bound to equal.

- Our data suggests an interaction.



Description of the Model

```
height ~ A * G + (A * G|L) + (1|S)
```

```
height ~ A + G + A:G + (A + G + A:G|L) + (1|S)
```

$$\text{height}_{[i]} \sim \mathrm{t}\left(\nu, \mu_{[i]}, \sigma\right)$$

$$\mu_{[i]} = \text{Intercept} + A + G + A:G + L_{\left[\mathrm{L}[i]\right]} + A:L_{\left[\mathrm{L}[i]\right]} + G:L_{\left[\mathrm{L}[i]\right]} + A:G:L_{\left[\mathrm{L}[i]\right]} + S_{\left[\mathrm{S}[i]\right]}$$

Priors:

$$S_{[\bullet]} \sim \mathrm{t}(3, 0, \sigma_S)$$

$$\begin{bmatrix} L_{[\bullet]} \\ A:L_{[\bullet]} \\ G:L_{[\bullet]} \\ A:G:L_{[\bullet]} \end{bmatrix} \sim \text{MVNormal}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \Sigma\right)$$

$$\text{Intercept} \sim \mathrm{t}(3, 156, 12)$$

$$A, G, A:G \sim \mathrm{t}(3, 0, 12)$$

$$\sigma_L, \sigma_{A:L}, \sigma_{G:L}, \sigma_{A:G:L}, \sigma_S \sim \mathrm{t}(3, 0, 12)$$

$$\sigma \sim \mathrm{t}(3, 0, 12)$$

$$\nu \sim \text{gamma}(2, 0.1)$$

$$R \sim \text{LKJCorr}(2)$$

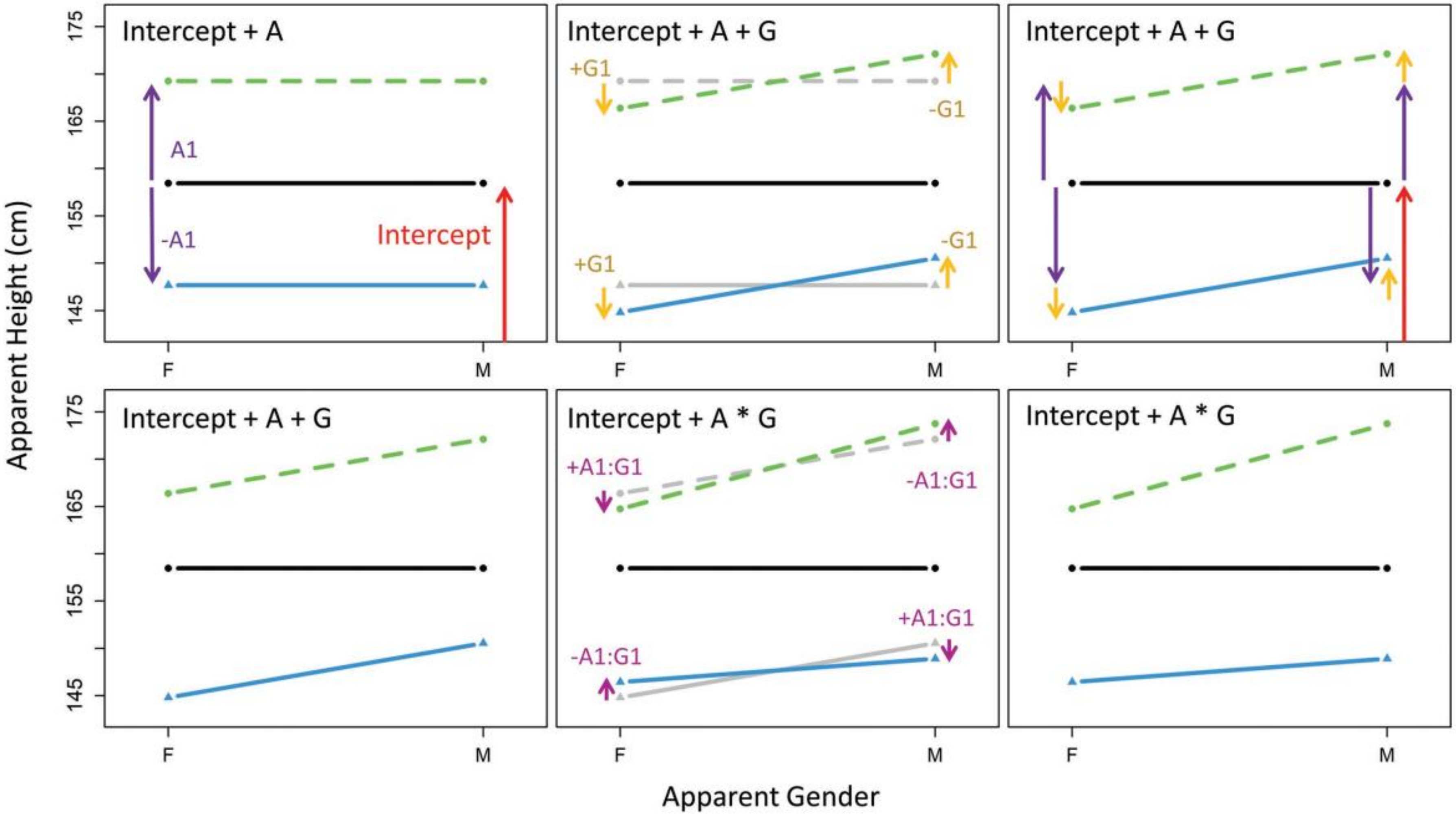
Fitting the Model

```
# Fit the model yourself
priors = c(brms::set_prior("student_t(3,156, 12)", class = "Intercept"),
           brms::set_prior("student_t(3,0, 12)", class = "b"),
           brms::set_prior("student_t(3,0, 12)", class = "sd"),
           brms::set_prior("lkj_corr_cholesky (2)", class = "cor"),
           brms::set_prior("gamma(2, 0.1)", class = "nu"),
           brms::set_prior("student_t(3,0, 12)", class = "sigma"))

model_interaction =
  brms::brm (height ~ A + G + A:G + (A + G + A:G|L) + (1|S),
             data = exp_data, chains = 4, cores = 4, warmup = 1000,
             iter = 5000, thin = 4, prior = priors, family = "student")

# Or download it from the GitHub page:
model_interaction = bmmb::get_model ('7_model_interaction.RDS')
```

```
bmmb::short_summary (model_interaction)
## Formula:  height ~ A + G + A:G + (A + G + A:G | L) + (1 | S)
##
## Group-Level Effects:
## ~L (Number of levels: 15)
##                        Estimate Est.Error l-95% CI u-95% CI
## sd(Intercept)              4.24      0.79     2.96     6.03
## sd(A1)                     4.47      0.85     3.13     6.42
## sd(G1)                     2.10      0.49     1.37     3.22
## sd(A1:G1)                  1.34      0.36     0.76     2.13
## cor(Intercept,A1)         -0.71      0.13    -0.90    -0.38
## cor(Intercept,G1)         -0.20      0.22    -0.59     0.24
## cor(A1,G1)                -0.24      0.21    -0.62     0.20
## cor(Intercept,A1:G1)       0.17      0.23    -0.29     0.58
## cor(A1,A1:G1)             -0.02      0.23    -0.46     0.43
## cor(G1,A1:G1)             -0.34      0.24    -0.74     0.18
##
## ~S (Number of levels: 139)
##                 Estimate Est.Error l-95% CI u-95% CI
## sd(Intercept)     2.36      0.31     1.79     2.99
##
## Population-Level Effects:
##             Estimate Est.Error l-95% CI u-95% CI
## Intercept     158.46      1.12   156.22   160.62
## A1             10.78      1.21     8.39    13.15
## G1             -2.87      0.60    -4.07    -1.72
## A1:G1          -1.64      0.41    -2.44    -0.81
##
## Family Specific Parameters:
##         Estimate Est.Error l-95% CI u-95% CI
## sigma     5.01      0.16     4.70     5.34
## nu        3.44      0.33     2.87     4.15
```

Interpreting the Fixed Effects

```
## Population-Level Ef
##                Estimate
## Intercept        158.46
## A1                10.78
## G1                -2.87
## A1:G1             -1.64
##
```



Recreating the Group Means

```
# intercept, boys, girls, men, women
means_pred_interaction = bmmb::short_hypothesis (
  model_interaction,
  c("Intercept + -A1 + -G1 +  A1:G1 = 0",  # boys
    "Intercept + -A1 +  G1 + -A1:G1 = 0",  # girl
    "Intercept +  A1 + -G1 + -A1:G1 = 0",  # men
    "Intercept +  A1 +  G1 +  A1:G1 = 0")) # women
```

```
# actual data means
tapply (exp_data$height, exp_data$C, mean)
##     b     g     m     w
## 146.9 145.0 174.5 165.6

# predictions with no interaction term
means_pred[,1]
## [1] 150.9 146.5 170.8 166.4

# predictions with interaction term
means_pred_interaction[,1]
## [1] 148.9 146.4 173.8 164.7
```

Calculating Simple Effects

- Simple effects can be calculated by adding main effects to interactions.

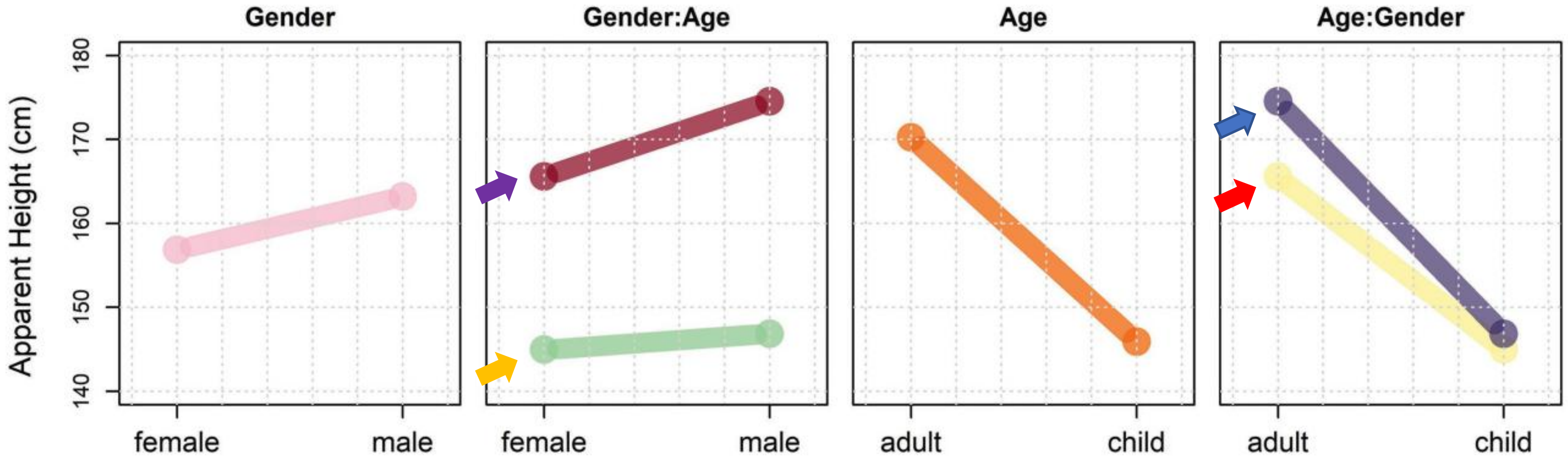
- For example:
  - A1 = the effect for apparent adultness
  - G1 = the effect for apparent femaleness
  - -G1 = the effect for apparent maleness (a.k.a. G2)
  - A1:G1 = interaction between apparent adultness and femaleness.
  - -A1:G1 = interaction between apparent adultness and maleness (a.k.a. A1:G2).

A1 + A1:G1 = The effect for adultness given apparent femaleness

A1 - A1:G1 = The effect for adultness given apparent maleness

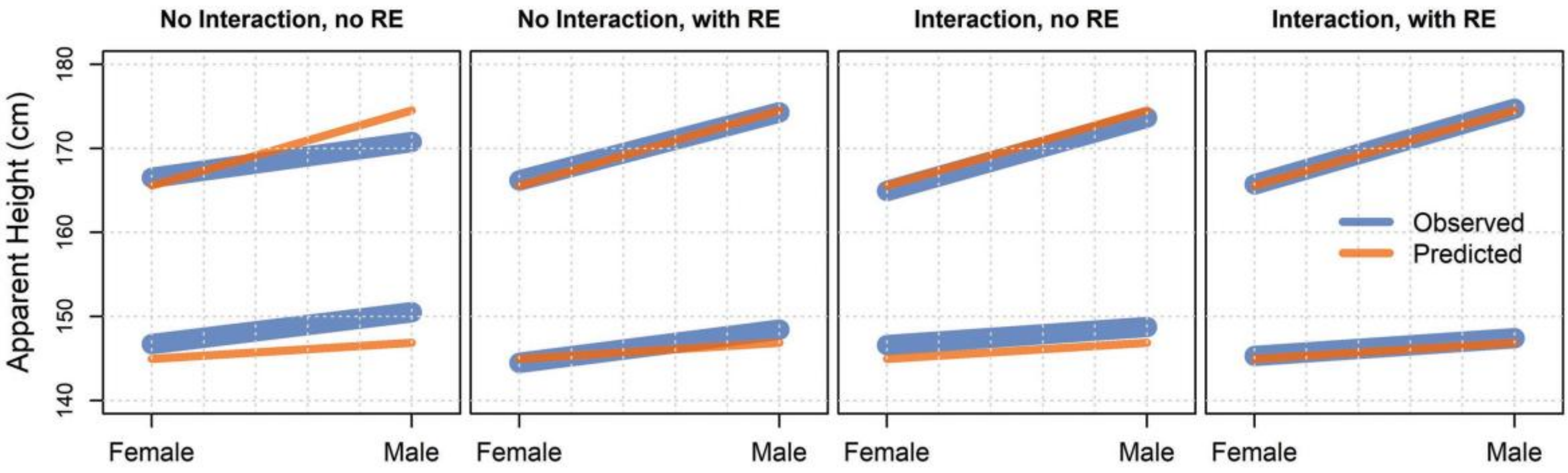
Calculating Simple Effects

```
# intercept, boys, girls, men, women
simple_effects = bmmb::short_hypothesis (
  model_interaction,
  c("A1 + A1:G1 = 0",   # effect for apparent age for females (G1)
    "A1 - A1:G1 = 0",   # effect for apparent age for males     (-G1)
    "G1 + A1:G1 = 0",   # effect for apparent gender for adults (A1)
    "G1 - A1:G1 = 0"))  # effect for apparent gender for children (-A1)

# predictions with interaction term
simple_effects
##     Estimate Est.Error   Q2.5   Q97.5      hypothesis
## H1     9.142    1.2792  6.639 11.5977 (A1+A1:G1) = 0
## H2    12.424    1.2707  9.938 14.9913 (A1-A1:G1) = 0
## H3    -4.514    0.6174 -5.731
## H4    -1.232    0.8192 -2.860
```



Posterior Prediction

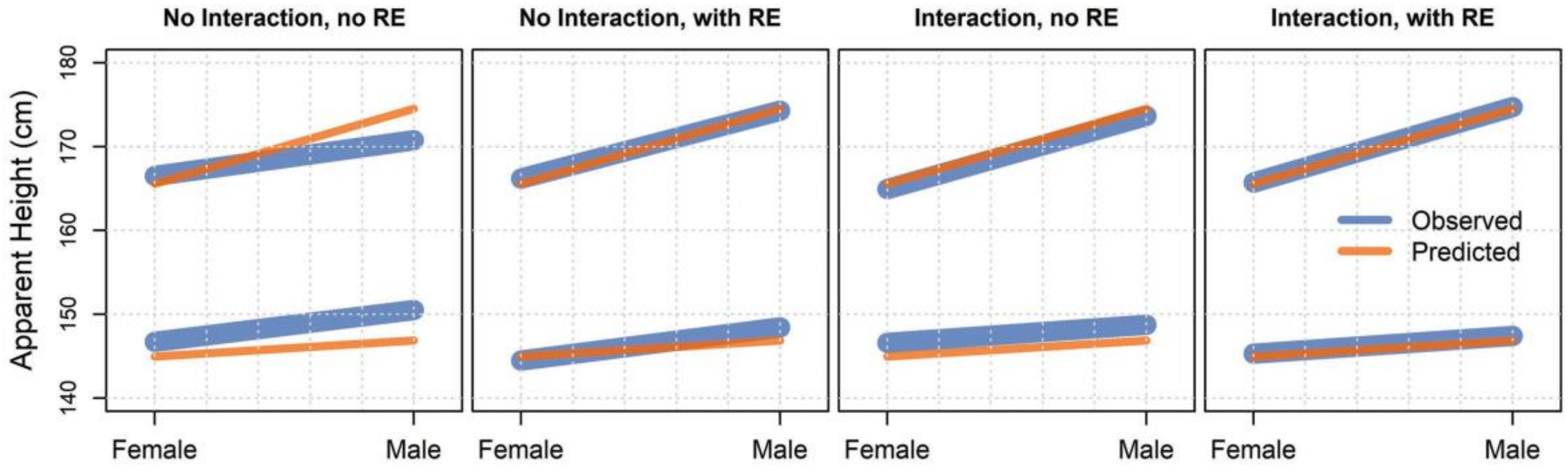
```
y_post_pred_int = predict (model_interaction)
y_post_pred_no_re_int = predict (model_interaction, re_formula = NA)
```



Model Comparison

```
model_both = brms::add_criterion (model_both, criterion="loo")
model_interaction = brms::add_criterion (model_interaction, criterion="loo")
```

```
brms::loo_compare (model_both, model_interaction)
##                    elpd_diff se_diff
## model_interaction    0.0       0.0
## model_both         -29.1      11.2
```



R^2 ('R-squared')

Variance decomposition → $$\sigma^2_{\text{total}} = \sigma^2_{\text{explained}} + \sigma^2_{\text{error}}$$

Ratio of explained variance to total → $$R^2 = \frac{\sigma^2_{\text{explained}}}{\sigma^2_{\text{total}}} = \frac{\sigma^2_{\text{explained}}}{\sigma^2_{\text{explained}} + \sigma^2_{\text{error}}}$$

Bayesian equivalent, for posterior sample S. → $$R_s^2 = \frac{V_{i=1}^{n}\ \hat{y}_i^s}{V_{i=1}^{n}\ \hat{y}_i^s + V_{i=1}^{n}\ \hat{e}_i^s}$$

$$\hat{e}_i^s = \hat{y}_i^s - y_i$$

residuals

predicted data

observed data

R^2 Comparison

```r
r2_both = r2_bayes(model_both)
r2_interaction = r2_bayes(model_interaction)
```

```r
r2_both
##        Estimate Est.Error   Q2.5  Q97.5
## [1,]     0.7755  0.005566 0.7641 0.7858

r2_interaction
##        Estimate Est.Error   Q2.5  Q97.5
## [1,]     0.7802  0.005347 0.7694 0.7901
```

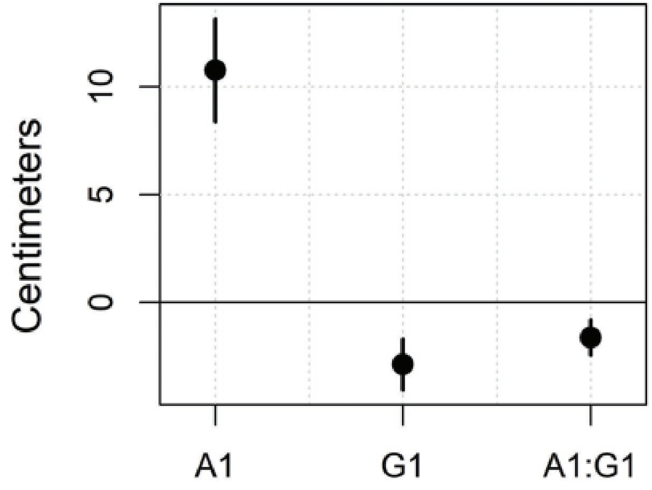
```r
r2_both_no_re = r2_bayes(model_both, re_formula = NA)
r2_interaction_no_re = r2_bayes(model_interaction, re_formula = NA)
```

```r
r2_both_no_re
##        Estimate Est.Error   Q2.5  Q97.5
## [1,]      0.503   0.07258 0.3411 0.6216
```

```r
r2_interaction_no_re
##        Estimate Est.Error   Q2.5  Q97.5
## [1,]     0.5732   0.05909 0.4389 0.6678
```

Answering Our Research Questions

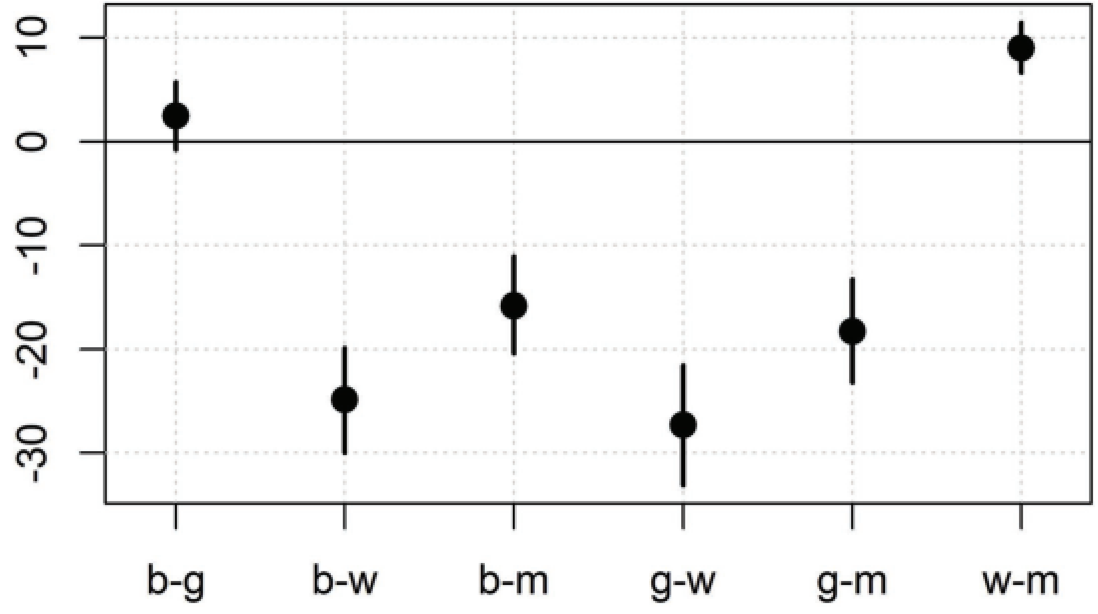
Results indicate that the average apparent height across all speaker groups (i.e. the intercept) was 158.5 (s.d. = 1.12, 95% C.I = [156.22, 160.62]). We also found an average effect of 10.8 cm for apparent speaker age (s.d. = 1.21, 95% C.I = [8.39, 13.15]) and −2.9 cm for apparent speaker gender (s.d. = 0.6, 95% C.I = [−4.07, −1.72]). In addition, we found an interaction between the effects of apparent age and apparent gender on apparent heights (mean = −1.64, s.d. = 0.41, 95% C.I = [−2.44, -0.81]). The result of these effects is that apparent adults were perceived as taller than apparent children and apparent males were perceived as taller than apparent females. However, the difference in apparent height due to apparent gender was larger for adults than for children (and the effect for apparent age was larger for males than for females). Figure 7.11 presents the model's fixed effects other than the intercept (whose value is too large to plot in this range).



Pairwise Group Differences

```
# get group mean predictions
C = attributes(means_pred_interaction)$samples

# find pairwise differences
pairwise_diffs = cbind("b-g"=C[,1]-C[,2], "b-w"=C[,1]-C[,3],
                       "b-m"=C[,1]-C[,4], "g-w"=C[,2]-C[,3],
                       "g-m"=C[,2]-C[,4], "w-m"=C[,3]-C[,4])

# summarize these
pairwise_diffs_summary = posterior_summary(pairwise_diffs)
```

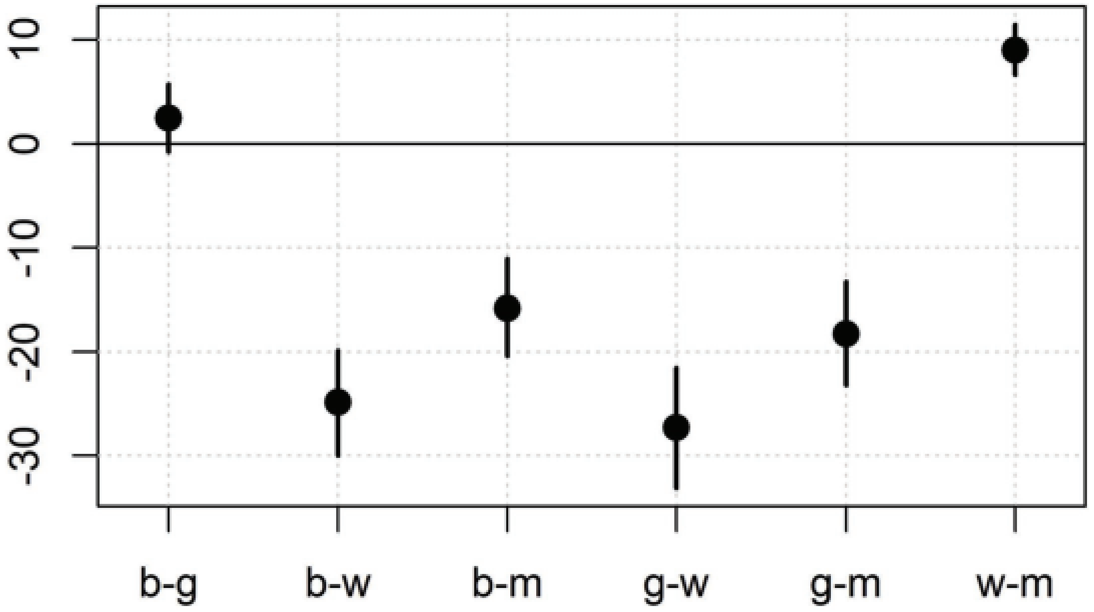
Which are ‘real’?



Is an Effect 'Real'?

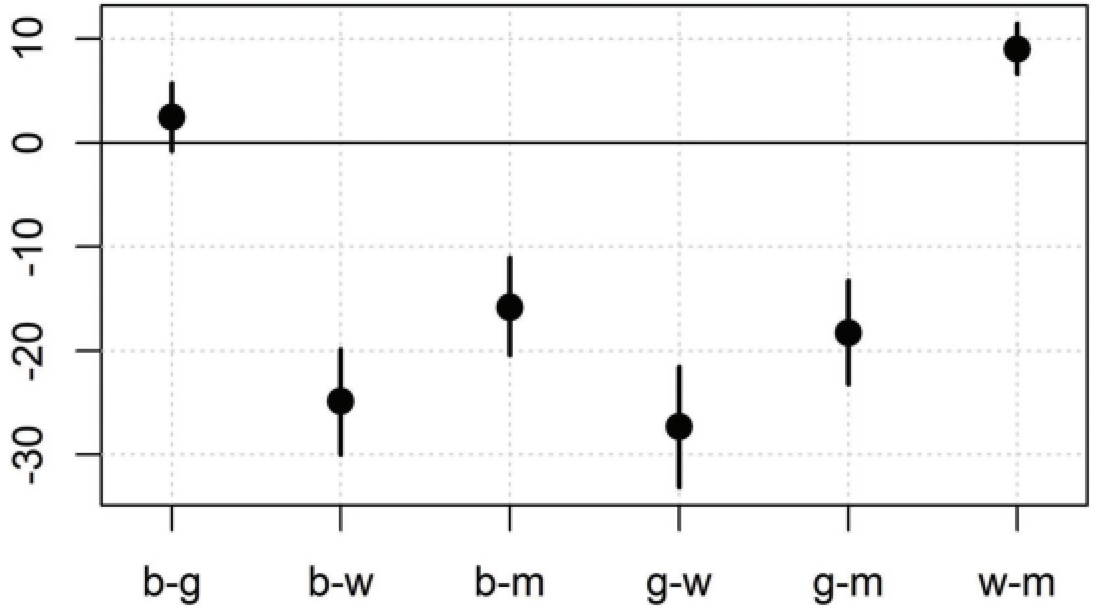
- Most 95% Credible intervals don't cross 0. Does this make the effects 'real'?

- The b-g credible intervals crosses 0. Does this mean the effect equals 0?
  - Even if an interval does cross zero, its most likely value is still not usually zero.



Type S and M Errors

- Better way to think of it: Avoid type S and M errors.
- Type S (sign) errors: Is the value negative when you think its positive (or vice versa)?
- Type M (magnitude) errors: Is the value small when you think its large (or vice versa)?

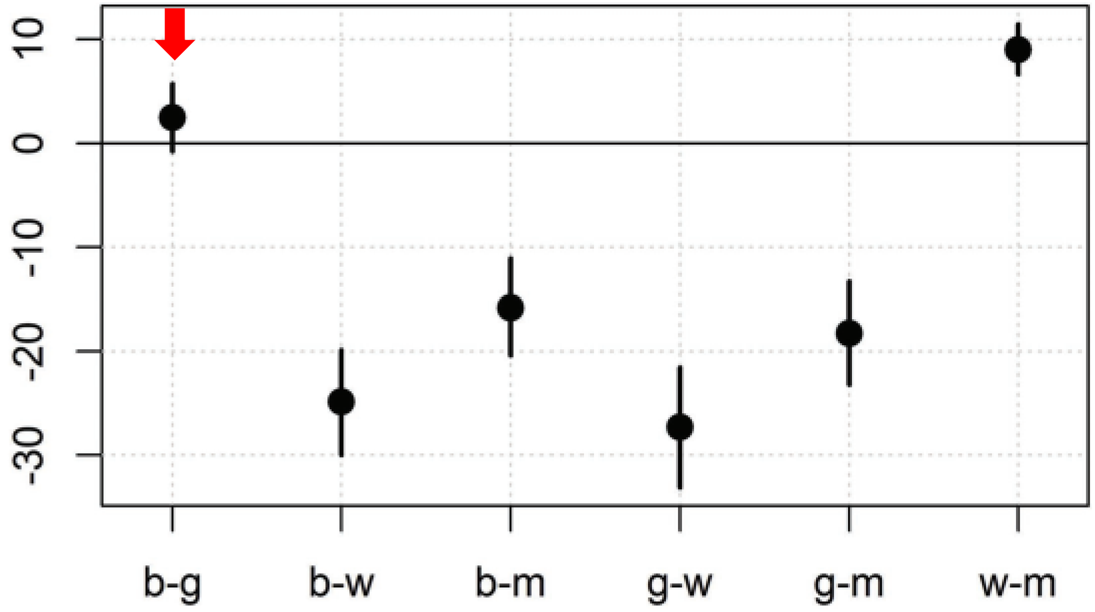


Region of Practical Equivalence

- Region of Practical Equivalence (ROPE): How small does an effect have to be before an effect *might as well* be zero?

- Example: For human height, differences under 0.5 cm have little practical meaning. A ‘significant’ difference of 0.01 cm might as well be zero.

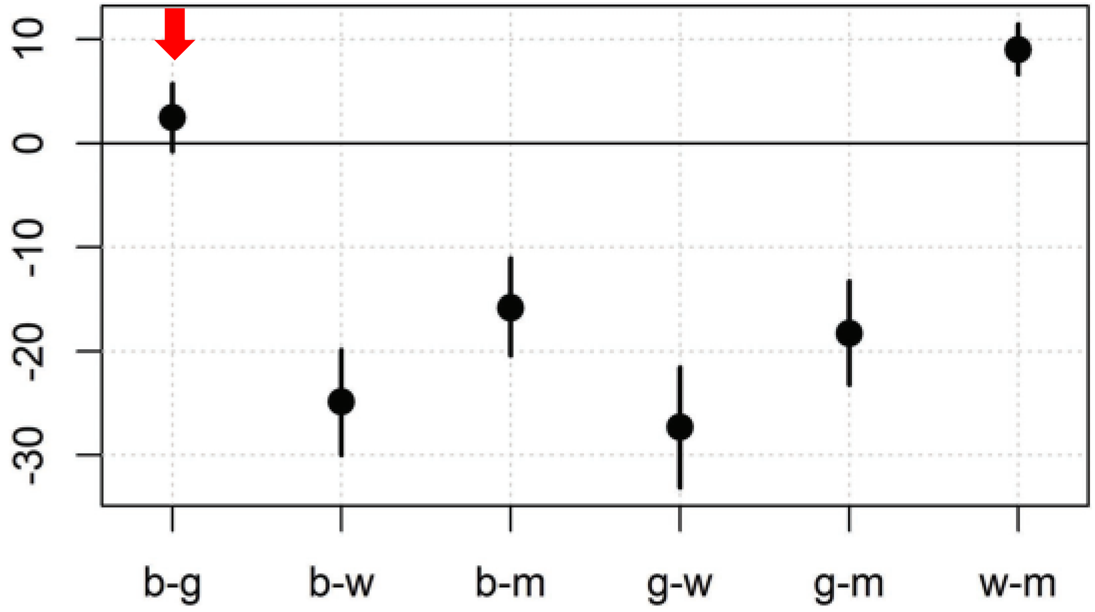
```
# summary of difference between boy and girl means
pairwise_diffs_summary[1,]
##   Estimate Est.Error      Q2.5     Q97.5
##     2.4643    1.6384   -0.7837    5.7199
```



So, is the b-g Difference ‘Real’?

- This is an epistemological or ontological claim.
- Statistics can help you make these claims, but <u>you</u> are responsible for them.
- Final word: The difference may be real and positive, but it is likely small. Further investigation is necessary.

```r
# summary of difference between boy and girl means
pairwise_diffs_summary[1,]
##   Estimate Est.Error      Q2.5     Q97.5
##     2.4643    1.6384   -0.7837    5.7199
```



Factors with more Levels

$$\begin{bmatrix} A1 & A2 & A3 & (A4) \end{bmatrix}$$

$$\begin{bmatrix} B1 & B2 & B3 & (B4) \end{bmatrix}$$

$$\begin{bmatrix} A1 & A2 & A3 & -(A1+A2+A3) \end{bmatrix}$$

$$\begin{bmatrix} B1 & B2 & B3 & -(B1+B2+B3) \end{bmatrix}$$

Factors with more Levels

$$\begin{bmatrix} A1:B1 & A2:B1 & A3:B1 & (A4:B1) \\ A1:B2 & A2:B2 & A3:B2 & (A4:B2) \\ A1:B3 & A2:B3 & A3:B3 & (A4:B3) \\ (A1:B4) & (A2:B4) & (A3:B4) & (A4:B4) \end{bmatrix}$$

$$\begin{bmatrix} A1:B1 & A2:B1 & A3:B1 & -(A1:B1+A2:B1+A3:B1) \\ A1:B2 & A2:B2 & A3:B2 & (A4:B2) \\ A1:B3 & A2:B3 & A3:B3 & (A4:B3) \\ -(A1:B1+A1:B2+A1:B3) & (A2:B4) & (A3:B4) & (A4:B4) \end{bmatrix}$$

Exercises

Use the data in 'exp_ex' to do one of the following. You may also use your own data to answer a related question. In any case, describe the model, present and explain the results, and include at least two figures.

1. Medium: Fit and interpret the pre-fit model (or a similar model that you fit). Report the results of the model. Recreate the predicted group means and report and interpret at least one simple effect.

2. Hard: Fit and interpret some other model that includes at least two factors with at least two levels each, and their interactions. Report the results of the model. Recreate the predicted group means and report and interpret at least one simple effect.