Crossed random effects

Nora Wickelmaier

Last modified: January 31, 2025

Crossed random effects

- In many experiments in psychology the reaction of each subject ($j = 1, \ldots, N$) to a complete set of stimuli or items ($k = 1, \ldots, K$) is measured

$$y_{ijk} = \beta_0 + \beta_i x_i + \upsilon_{0j} + \eta_{0k} + \varepsilon_{ijk}$$

  with $\varepsilon_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$, $\upsilon_{0j} \overset{iid}{\sim} N(0, \sigma^2_\upsilon)$, and $\eta_{0k} \overset{iid}{\sim} N(0, \sigma^2_\eta)$
- Data are completely crossed: all subjects are presented with all items

		Subject				
		1	2	3	...	20
	1	1	1	1	...	1
	2	1	1	1	...	1
Item	3	1	1	1	...	1
	⋮	⋮	⋮	⋮	⋮	⋮
	10	1	1	1	...	1

Crossed random effects

```
> head(dat, 12)
   id  cond   item  av
1   1 cond1 item01 105
2   1 cond1 item02 116
3   1 cond1 item03 104
4   1 cond1 item04  81
5   1 cond1 item05  99
6   1 cond1 item06 109
7   1 cond1 item07 100
8   1 cond1 item08 103
9   1 cond1 item09  89
10  1 cond1 item10  94
11  2 cond1 item01 107
12  2 cond1 item02 100
```

```
> xtabs( ~ item + id, dat)
        id
item     1 2 3 4 5 6 7 8 9 ... 20
  item01 1 1 1 1 1 1 1 1 1 ...  1
  item02 1 1 1 1 1 1 1 1 1 ...  1
  item03 1 1 1 1 1 1 1 1 1 ...  1
  item04 1 1 1 1 1 1 1 1 1 ...  1
  item05 1 1 1 1 1 1 1 1 1 ...  1
  item06 1 1 1 1 1 1 1 1 1 ...  1
  item07 1 1 1 1 1 1 1 1 1 ...  1
  item08 1 1 1 1 1 1 1 1 1 ...  1
  item09 1 1 1 1 1 1 1 1 1 ...  1
  item10 1 1 1 1 1 1 1 1 1 ...  1
```

1. Lexical decision task

Lexical decision task (Baayen et al., 2008)

- This example will show how to include subjects and items as crossed, independent, random effects, as opposed to hierarchical or multilevel models in which random effects are assumed to be nested

Lexical decision task (Baayen et al., 2008)

- This example will show how to include subjects and items as crossed, independent, random effects, as opposed to hierarchical or multilevel models in which random effects are assumed to be nested
- Assume an example data set with three participants s1, s2 and s3 who each saw three items w1, w2, w3 in a priming lexical decision task under both short and long stimulus onset asynchrony (SOA) conditions

Lexical decision task (Baayen et al., 2008)

- This example will show how to include subjects and items as crossed, independent, random effects, as opposed to hierarchical or multilevel models in which random effects are assumed to be nested
- Assume an example data set with three participants s1, s2 and s3 who each saw three items w1, w2, w3 in a priming lexical decision task under both short and long stimulus onset asynchrony (SOA) conditions
- The data are generated by the following model with random intercepts for subject and item, and random slopes for subject

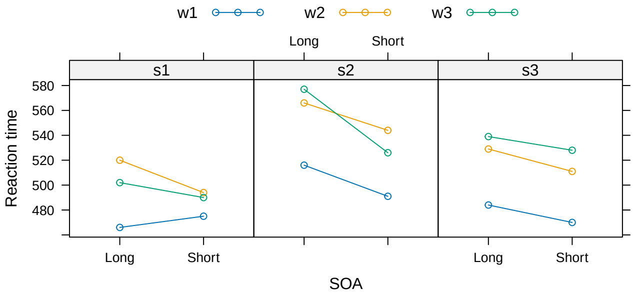
$$y_{ijk} = \beta_0 + \beta_1 SOA_k + \omega_{0j} + \upsilon_{0i} + \upsilon_{1i} SOA_k + \varepsilon_{ijk}$$

with $\boldsymbol{\upsilon} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} \end{pmatrix}\right)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ijk} \sim N(0, \sigma^2_\varepsilon)$, all i.i.d.

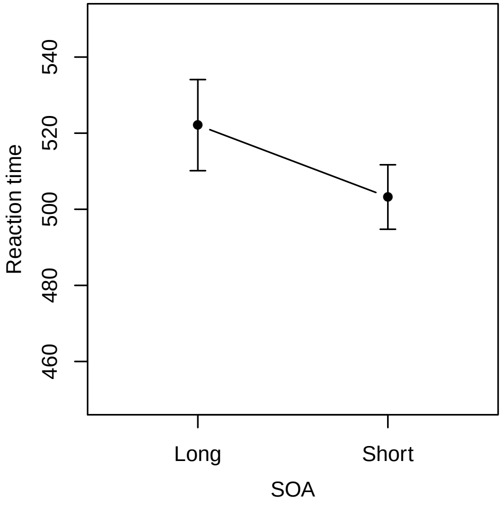
Structure of the data set

Subj	Item	SOA	RT
s1	w1	Long	466
s1	w2	Long	520
s1	w3	Long	502
s1	w1	Short	475
s1	w2	Short	494
s1	w3	Short	490
s2	w1	Long	516
s2	w2	Long	566
s2	w3	Long	577
s2	w1	Short	491
s2	w2	Short	544
s2	w3	Short	526
s3	w1	Long	484
s3	w2	Long	529
s3	w3	Long	539
s3	w1	Short	470
s3	w2	Short	511
s3	w3	Short	528

- When we collect data, we might get a data set like this
- We fit a model to the data to separate the structural and the stochastical parts



Aggregated data



Structure of the data set

Subj	Item	SOA	RT	Fixed		Random			Res
				Int	SOA	ItemInt	SubInt	SubSOA	
s1	w1	Long	466	522.2	0	−28.3	−26.2	0	−2.0
s1	w2	Long	520	522.2	0	14.2	−26.2	0	9.8
s1	w3	Long	502	522.2	0	14.1	−26.2	0	−8.2
s1	w1	Short	475	522.2	−19	−28.3	−26.2	11	15.4
s1	w2	Short	494	522.2	−19	14.2	−26.2	11	−8.4
s1	w3	Short	490	522.2	−19	14.1	−26.2	11	−11.9
s2	w1	Long	516	522.2	0	−28.3	29.7	0	−7.4
s2	w2	Long	566	522.2	0	14.2	29.7	0	0.1
s2	w3	Long	577	522.2	0	14.1	29.7	0	11.5
s2	w1	Short	491	522.2	−19	−28.3	29.7	−12.5	−1.5
s2	w2	Short	544	522.2	−19	14.2	29.7	−12.5	8.9
s2	w3	Short	526	522.2	−19	14.1	29.7	−12.5	−8.2
s3	w1	Long	484	522.2	0	−28.3	−3.5	0	−6.3
s3	w2	Long	529	522.2	0	14.2	−3.5	0	−3.5
s3	w3	Long	539	522.2	0	14.1	−3.5	0	6.0
s3	w1	Short	470	522.2	−19	−28.3	−3.5	1.5	−2.9
s3	w2	Short	511	522.2	−19	14.2	−3.5	1.5	−4.6
s3	w3	Short	528	522.2	−19	14.1	−3.5	1.5	13.2
						$\sigma^2_{\omega_0}$	$\sigma^2_{\upsilon_0}$	$\sigma^2_{\upsilon_1}$	σ^2_{ε}

$\sigma_{\upsilon_0 \upsilon_1}$

True values

- We assume the following true parameters for a data simulation

Parameter	Model
β_0	522.11
β_1	-18.89
σ_ω	21.10
σ_{υ_0}	23.89
σ_{υ_1}	9.00
$\rho_{\upsilon_0\upsilon_1}$	-1.00
σ_ε	9.90

$$y_{ijk} = \beta_0 + \beta_1 SOA_k + \omega_{0j} + \upsilon_{0i} + \upsilon_{1i} SOA + \varepsilon_{ijk}$$

with $\boldsymbol{\upsilon} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} \end{pmatrix}\right)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ijk} \sim N(0, \sigma^2_\varepsilon)$

Fixed effects

```r
datsim <- expand.grid(subject = factor(c("s1", "s2", "s3")),
                      item = factor(c("w1", "w2", "w3")),
                      soa = factor(c("long", "short")))
datsim <- datsim |> sort_by( ~ subject)

# model matrix in dummy coding
model.matrix(~ soa, datsim)

beta0 <- 522.11
beta1 <- -18.89
b0 <- rep(beta0, 18)
b1 <- rep(rep(c(0, beta1), each = 3), 3)
cbind(b0, b1)
```

Random effects

```
sw  <- 21.1
sy0 <- 23.89; sy1 <- 9; ry <- -1
se  <- 9.9

w  <- rep(rnorm(3, mean = 0, sd = sw), 6)
e  <- rnorm(18, mean = 0, sd = se)
# draw from bivariate normal distribution
sig <- matrix(c(sy0^2, ry * sy0 * sy1, ry * sy0 * sy1, sy1^2), 2, 2)
y01 <- mvtnorm::rmvnorm(3, mean = c(0, 0), sigma = sig)
y0 <- rep(y01[,1], each = 6)
y1 <- rep(c(0, y01[1,2],
            0, y01[2,2],
            0, y01[3,2]), each = 3)
cbind(w, y0, y1, e)
```

Simulate data

```
datsim$rt <- b0 + b1 + w + y0 + y1 + e

# fit model
library(lme4)

lme1 <- lmer(rt ~ soa + (1 | item) + (soa | subject), datsim)
summary(lme1)
confint(lme1)

# btw
?pvalues
?convergence
```

Comparison of sample and model estimates

For this example, we are able to compare the "true" values to the parameter estimates

Parameter	Sample	Model
$\hat{\beta}_0$	522.2	522.11
$\hat{\beta}_1$	−19.00	−18.89
$\hat{\sigma}_\omega$	20.59	21.10
$\hat{\sigma}_{\upsilon_0}$	23.62	23.89
$\hat{\sigma}_{\upsilon_1}$	9.76	9.00
$\hat{\rho}_{\upsilon_0\upsilon_1}$	−0.71	−1.00
$\hat{\sigma}_\varepsilon$	8.55	9.90

$$y_{ijk} = \beta_0 + \beta_1 SOA_k + \omega_{0j} + \upsilon_{0i} + \upsilon_{1i} SOA_k + \varepsilon_{ijk}$$

with $\boldsymbol{\upsilon} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} \end{pmatrix}\right)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ijk} \sim N(0, \sigma^2_\varepsilon)$

Linear mixed-effects model

- The linear mixed-effects model has the general form

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \boldsymbol{\upsilon}_i + \boldsymbol{\varepsilon}_i$$

with fixed effects $\boldsymbol{\beta}$, random effects $\boldsymbol{\upsilon}_i$, and the design matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$ and the assumptions

$$\boldsymbol{\upsilon}_i \sim N(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon) \text{ i.i.d.}, \qquad \boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{n_i}) \text{ i.i.d.}$$

Linear mixed-effects model

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ 1 & x_{31} & x_{32} & \dots & x_{3p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & \dots & x_{Np} \end{pmatrix} \cdot \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} + \begin{pmatrix} z_{10} & z_{11} & \dots & z_{1q} & \dots \\ z_{20} & z_{21} & \dots & z_{2q} & \dots \\ z_{30} & z_{31} & \dots & z_{3q} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ z_{N0} & z_{N1} & \dots & z_{Nq} & \dots \end{pmatrix} \cdot \begin{pmatrix} \upsilon_{10} \\ \vdots \\ \upsilon_{1q} \\ \upsilon_{20} \\ \vdots \\ \upsilon_{Nq} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_N \end{pmatrix}$$

Simulate data using model matrices

```
X <- model.matrix( ~ soa, datsim)
Z <- model.matrix( ~ 0 + item + subject + subject:soa, datsim,
  contrasts.arg =
    list(subject = contrasts(datsim$subject, contrasts = FALSE)))

# fixed effects
beta  <- c(beta0, beta1)

# random effects
u <- c(w = unique(w),
       y0 = y01[,1],
       y1 = y01[,2])

datsim$rt2 <- X %*% beta + Z %*% u + e
```

Exercise

- Change the data simulation from the previous slides for $N = 30$ subjects instead of only 3
- You can choose if you want to use model matrices or create the vectors "manually"

Summary

- Mixed-effects model with crossed random effects allow to include random effects from different sources (e. g., subjects and items)
- These models have more power than models on aggregated data like ANOVAs or a paired t test in this example (see e. g., Jaeger, 2008)
- But more importantly (IMHO), they allow us to assume a much more flexible data generating process that seems to be closer to reality

2 Physical healing as a function of perceived time

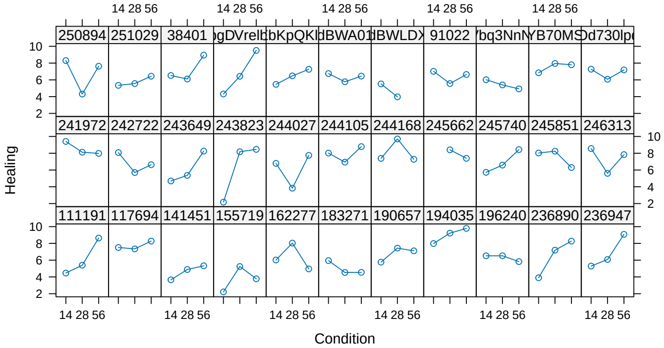
Physical healing (Aungle & Langer, 2023)

- Aungle and Langer (2023) investigate how perceived time influences physical healing
- They used cupping to induce bruises on 33 subjects, then took a picture, waited for 28 min and took another picture
- Subjects participated in all three conditions over a two week period
- Subjective time was manipulated to feel like 14, 28, or 56 min
- The pre and post pictures were presented to 25 raters who rated the amount of healing on a 10-point-scale with 0 = not at all healed, 5 = somewhat healed, 10 = completely healed

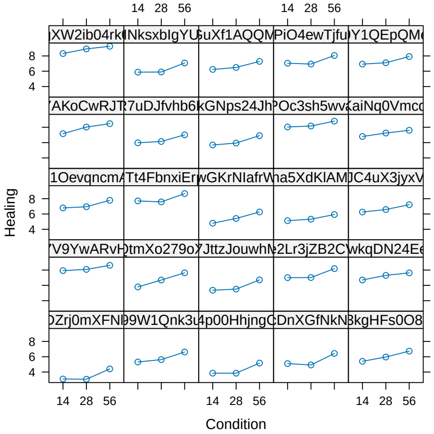
Condition	Mean	SD	$N_{subjects}$	$N_{ratings}$
14-min	6.17	2.59	32	800
28-min	6.43	2.54	33	825
56-min	7.30	2.25	32	800

Visualization of data

Subjects

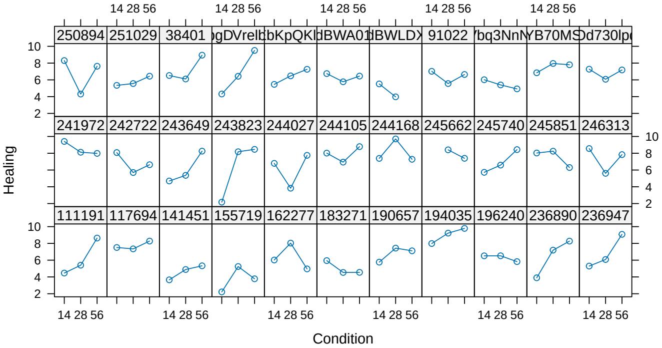


Raters

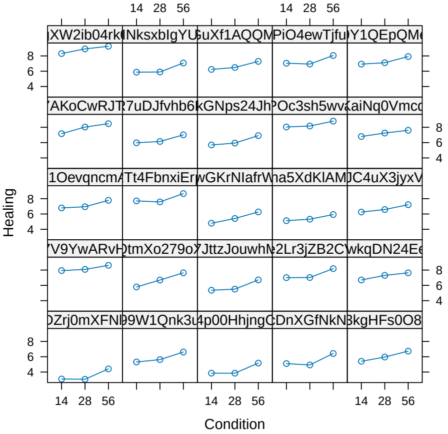


Visualization of data

Subjects

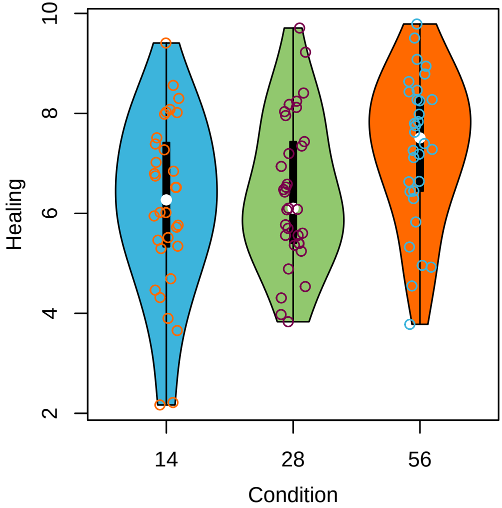


Raters



What model would you choose?

Aggregated data



Fitted model

In the paper, a mixed-effects model with random intercepts for subjects and raters was fitted to the data

$$y_{ijk} = \beta_0 + \beta_{1k} Condition_k + \omega_{0j} + \upsilon_{0i} + \varepsilon_{ijk}$$

with $\upsilon_{0i} \sim N(0, \sigma^2_{\upsilon})$, $\omega_{0j} \sim N(0, \sigma^2_{\omega})$, $\varepsilon_{ijk} \sim N(0, \sigma^2_{\varepsilon})$, all i.i.d.

	Est.	Std.	t value
(Intercept)	6.20	0.32	19.56
Condition28	0.23	0.09	2.44
Condition56	1.05	0.09	11.10
Random effects			
Residual	1.873		
Subj (Intercept)	1.079		
Rater (Intercept)	1.233		

```r
load("data/healing.RData")
dat <- DFmodel

m1 <- lmer(Healing ~ Condition + (1 | Subject) + (1 | ResponseId), dat)
```

Selection of random effects

- How to choose an appropriate random effects structure for a mixed-effects model is widely discussed in the literature (e. g., Barr et al., 2013; Bates et al., 2018; Gelman & Brown, 2024)

Selection of random effects

- How to choose an appropriate random effects structure for a mixed-effects model is widely discussed in the literature (e. g., Barr et al., 2013; Bates et al., 2018; Gelman & Brown, 2024)
- A prominent view is, that the random effects structure needs to represent the experimental design

Selection of random effects

- How to choose an appropriate random effects structure for a mixed-effects model is widely discussed in the literature (e. g., Barr et al., 2013; Bates et al., 2018; Gelman & Brown, 2024)
- A prominent view is, that the random effects structure needs to represent the experimental design
- In our example here, leaving out random slopes for the subjects implies that there are no individual effects for subjects

Selection of random effects

- How to choose an appropriate random effects structure for a mixed-effects model is widely discussed in the literature (e. g., Barr et al., 2013; Bates et al., 2018; Gelman & Brown, 2024)
- A prominent view is, that the random effects structure needs to represent the experimental design
- In our example here, leaving out random slopes for the subjects implies that there are no individual effects for subjects
- Hence, it is assumed that the experimental manipulation has the same effect for every subject

Selection of random effects

- How to choose an appropriate random effects structure for a mixed-effects model is widely discussed in the literature (e. g., Barr et al., 2013; Bates et al., 2018; Gelman & Brown, 2024)
- A prominent view is, that the random effects structure needs to represent the experimental design
- In our example here, leaving out random slopes for the subjects implies that there are no individual effects for subjects
- Hence, it is assumed that the experimental manipulation has the same effect for every subject
- What about the raters in this example? Could they be influenced by the conditions?

Random slope model

- Let us add a random slope to the model

```
m2 <- lmer(Healing ~ Condition + (1 + Condition | Subject) +
           (1 | ResponseId), dat)
```

Random slope model

- Let us add a random slope to the model

```
m2 <- lmer(Healing ~ Condition + (1 + Condition | Subject) +
           (1 | ResponseId), dat)
```

- What will change?

Random slope model

- Let us add a random slope to the model

```
m2 <- lmer(Healing ~ Condition + (1 + Condition | Subject) +
           (1 | ResponseId), dat)
```

- What will change?
- What could be the problem with this model?

(Some) Possible models

- Model with random intercepts for subjects and random intercepts for raters

$$y_{ij} = \beta_0 + \beta_1 Condition_{28} + \beta_2 Condition_{56} + \omega_{0j} + \upsilon_{0i} + \varepsilon_{ij}$$

with $\upsilon_{0i} \sim N(0, \sigma^2_\upsilon)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$, all i.i.d.

(Some) Possible models

- Model with random intercepts for subjects and random intercepts for raters

$$y_{ij} = \beta_0 + \beta_1 Condition_{28} + \beta_2 Condition_{56} + \omega_{0j} + \upsilon_{0i} + \varepsilon_{ij}$$

with $\upsilon_{0i} \sim N(0, \sigma^2_\upsilon)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$, all i.i.d.

- Model with random slopes for subjects and random intercepts for raters

$$y_{ij} = \beta_0 + \beta_1 Condition_{28} + \beta_2 Condition_{56} + \omega_{0j} + \upsilon_{0i} + \upsilon_{1i} Condition_{28} + \upsilon_{2i} Condition_{56} + \varepsilon_{ij}$$

with $\boldsymbol{\upsilon} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} & \sigma_{\upsilon_0\upsilon_2} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} & \sigma_{\upsilon_1\upsilon_2} \\ \sigma_{\upsilon_0\upsilon_2} & \sigma_{\upsilon_1\upsilon_2} & \sigma^2_{\upsilon_2} \end{pmatrix}\right)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$, all i.i.d.

(Some) Possible models

- Model with random intercepts for subjects and random intercepts for raters

$$y_{ij} = \beta_0 + \beta_1 Condition_{28} + \beta_2 Condition_{56} + \omega_{0j} + \upsilon_{0i} + \varepsilon_{ij}$$

with $\upsilon_{0i} \sim N(0, \sigma^2_\upsilon)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$, all i.i.d.

- Model with random slopes for subjects and random intercepts for raters

$$y_{ij} = \beta_0 + \beta_1 Condition_{28} + \beta_2 Condition_{56} + \omega_{0j} + \upsilon_{0i} + \upsilon_{1i} Condition_{28} + \upsilon_{2i} Condition_{56} + \varepsilon_{ij}$$

with $\boldsymbol{\upsilon} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} & \sigma_{\upsilon_0\upsilon_2} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} & \sigma_{\upsilon_1\upsilon_2} \\ \sigma_{\upsilon_0\upsilon_2} & \sigma_{\upsilon_1\upsilon_2} & \sigma^2_{\upsilon_2} \end{pmatrix}\right)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$, all i.i.d.

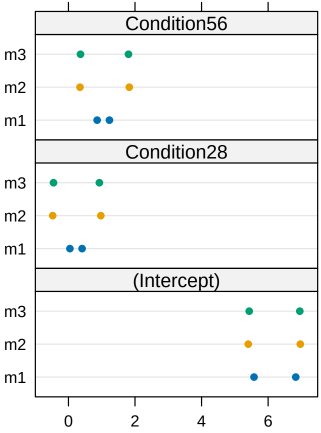
- Model with random slope for subjects and random intercepts for raters, zero correlations

$$y_{ij} = \beta_0 + \beta_1 Condition_{28} + \beta_2 Condition_{56} + \omega_{0j} + \upsilon_{0i} + \upsilon_{1i} Condition_{28} + \upsilon_{2i} Condition_{56} + \varepsilon_{ij}$$

with $\boldsymbol{\upsilon} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & 0 & 0 \\ 0 & \sigma^2_{\upsilon_1} & 0 \\ 0 & 0 & \sigma^2_{\upsilon_2} \end{pmatrix}\right)$, $\omega_{0j} \sim N(0, \sigma^2_\omega)$, $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$, all i.i.d.

Model comparisons

```
m1 <- lmer(Healing ~ Condition +
  (1 | Subject) + (1 | ResponseId),
  data = dat)
m2 <- lmer(Healing ~ Condition +
  (Condition | Subject) + (1 | ResponseId),
  data = dat)
m3 <- lmer(Healing ~ Condition +
  (1 | Subject) +
  (0 + dummy(Condition, "28") | Subject) +
  (0 + dummy(Condition, "56") | Subject) +
  (1 | ResponseId),
  data = dat)
```



Exercise

- Fit the models from the previous slide
- Profile the models with `profile(<model>)`
- Use the functions `xyplot()`, `densityplot()`, `splom()` from the lattice package to take a closer look at the estimated random parameters
- Compare the three models with likelihood ratio tests
- What is the best model in your opinion?

Summary

- When we use mixed-effects model to fit data, we need to make an informed choice about the random effects we include into the model

Summary

- When we use mixed-effects model to fit data, we need to make an informed choice about the random effects we include into the model
- Complex random effect structures can lead to convergence problems and random slopes are not always easy to estimate

Summary

- When we use mixed-effects model to fit data, we need to make an informed choice about the random effects we include into the model
- Complex random effect structures can lead to convergence problems and random slopes are not always easy to estimate
- The random effects structure strongly influences the confidence intervals for the fixed effects which we are often interested in

Summary

- When we use mixed-effects model to fit data, we need to make an informed choice about the random effects we include into the model
- Complex random effect structures can lead to convergence problems and random slopes are not always easy to estimate
- The random effects structure strongly influences the confidence intervals for the fixed effects which we are often interested in
- This is especially relevant in a confirmatory setting

Summary

- When we use mixed-effects model to fit data, we need to make an informed choice about the random effects we include into the model
- Complex random effect structures can lead to convergence problems and random slopes are not always easy to estimate
- The random effects structure strongly influences the confidence intervals for the fixed effects which we are often interested in
- This is especially relevant in a confirmatory setting
- For some critical discussion of the healing paper and their choice of random effects see Gelman and Brown (2024) and Gelman's blog post and discussion here: https://statmodeling.stat.columbia.edu/2025/01/23/slopes/

References I

Aungle, P., & Langer, E. (2023).Physical healing as a function of perceived time. *Scientific Reports*, *13*(1), 22432. https://doi.org/10.1038/s41598-023-50009-3

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008).Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013).Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models. https://doi.org/10.48550/arXiv.1506.04967

Gelman, A., & Brown, N. J. L. (2024).How statistical challenges and misreadings of the literature combine to produce unreplicable science: An example from psychology. *Advances in Methods and Practices in Psychological Science*, *7*(4), 25152459241276398. https://doi.org/10.1177/25152459241276398

References II

Jaeger, T. F. (2008).Categorical data analysis: Away from anovas (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*(4), 434–446. https://doi.org/10.1016/j.jml.2007.11.007

Wickelmaier, F. (2022).Simulating the power of statistical tests: A collection of R examples. *ArXiv*. https://arxiv.org/abs/2110.09836