Random effects for within-subject designs

Nora Wickelmaier

Last modified: April 1, 2025

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
- 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
- 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

1 Physical healing

Physical healing
(Aungle & Langer, 2023)

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

- How perceived time influences physical healing

Physical healing

(Aungle & Langer, 2023)

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

- 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

Physical healing

(Aungle & Langer, 2023)

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

- 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

Physical healing

(Aungle & Langer, 2023)

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

- 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

Physical healing
(Aungle & Langer, 2023)

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

- 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
- Pre and post pictures were presented to 25 raters (amount of healing with $0 =$ not at all healed, $5 =$ somewhat healed, $10 =$ completely healed)

Possible models

(Aungle & Langer, 2023)

- 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.

Possible models

(Aungle & Langer, 2023)

- 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.

Possible models

(Aungle & Langer, 2023)

- 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

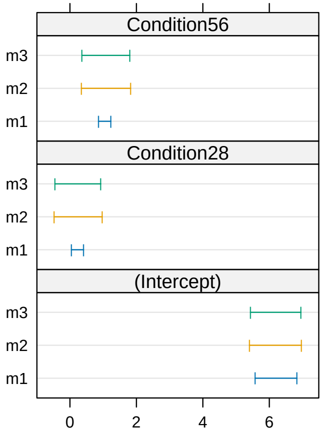
$$y_{ij} = \beta_0 + \beta_1 Condition_{28} + \beta_2 Condition_{56} + \omega_{0j} + \upsilon_{0i} + \upsilon_{1i} Condition_{28} + \upsilon_{2i} Condition_{56} + \varepsilon_{ij}$$

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

(Aungle & Langer, 2023)

```
library("lme4")
load("data/healing.RData")

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



Different random effects

Random intercept model

- What is the difference between these 3 models?

```
lapply(coef(m1), head, n = 3)
# $Subject
#         (Intercept) Condition28 Condition56
# 111191     5.759678   0.2272593    1.047163
# 117694     7.245319   0.2272593    1.047163
# 141451     4.276601   0.2272593    1.047163
#
# $ResponseId
#                    (Intercept) Condition28 Condition56
# R_1DZrj0mXFNlzerG     3.160095   0.2272593    1.047163
# R_1F99W1Qnk3uLGTg     5.440917   0.2272593    1.047163
# R_1I4p0OHhjngCBwT     3.914327   0.2272593    1.047163
```

Different random effects

Random slope model

- What is the difference between these 3 models?

```
lapply(coef(m2), head, n = 3)
# $Subject
#        (Intercept) Condition28 Condition56
# 111191    4.517425   0.9372819   4.0356794
# 117694    7.480483  -0.1545164   0.7669798
# 141451    3.717059   1.2029451   1.6695395
#
# $ResponseId
#                   (Intercept) Condition28 Condition56
# R_1DZrj0mXFNlzerG    3.115764   0.2462495    1.089368
# R_1F99W1Qnk3uLGTg    5.415579   0.2462495    1.089368
# R_1I4p0OHhjngCBwT    3.876277   0.2462495    1.089368
```

Different random effects

Random slope model without correlations

- What is the difference between these 3 models?

```
lapply(coef(m3), head, n = 3)
# $Subject
#         dummyCond28 dummyCond56 (Intercept) Condition28 Condition56
# 111191    0.5508955   2.8897054    4.596686   0.2390194    1.083531
# 117694   -0.3384845  -0.2573444    7.455698   0.2390194    1.083531
# 141451    0.8596727   0.4654271    3.768399   0.2390194    1.083531
#
# $ResponseId
#                   (Intercept) Condition28 Condition56
# R_1DZrj0mXFNlzerG    3.123060   0.2390194    1.083531
# R_1F99W1Qnk3uLGTg    5.422826   0.2390194    1.083531
# R_1I4p0OHhjngCBwT    3.883556   0.2390194    1.083531
```

Summary

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

Summary

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

Summary

- When we use mixed-effects models to fit data, we need to make an informed choice about the random effects we include in the model
- Complex random effect structures can lead to convergence problems and variance terms for 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 models to fit data, we need to make an informed choice about the random effects we include in the model
- Complex random effect structures can lead to convergence problems and variance terms for 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 models to fit data, we need to make an informed choice about the random effects we include in the model
- Complex random effect structures can lead to convergence problems and variance terms for 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/

2 Perceived risk of AI expert systems

Perceived risk of AI expert systems

Independent variables

- Participant ($N = 898$)
- Partner (AI vs. hu, within)
- Stakes (HS vs. LS, within)
- Context (edu vs. fin vs. law vs. med vs. psy, between)

Dependent variables

- Perceived risk (1 item, 7-point scale, from “None at all” to “Maximally”)
- Perceived trustworthiness (9 items, 7-point scale, averaged)

```
dat <- read.table("data/data-nico.csv", sep = ",", header = TRUE,
                  stringsAsFactors = TRUE)
```

Choose a mixed-effects model

- Let us look at perceived risk

Choose a mixed-effects model

- Let us look at perceived risk
- Hypothesis: In certain contexts, people will perceive the risk to consult an AI expert sytem as much higher compared to a human expert when stakes are high

Choose a mixed-effects model

- Let us look at perceived risk
- Hypothesis: In certain contexts, people will perceive the risk to consult an AI expert sytem as much higher compared to a human expert when stakes are high
- Draw a hypothesis plot

Choose a mixed-effects model

- Let us look at perceived risk
- Hypothesis: In certain contexts, people will perceive the risk to consult an AI expert sytem as much higher compared to a human expert when stakes are high
- Draw a hypothesis plot
- What mixed-effects model is suited to test this hypothesis?

Choose a mixed-effects model

- Let us look at perceived risk
- Hypothesis: In certain contexts, people will perceive the risk to consult an AI expert sytem as much higher compared to a human expert when stakes are high
- Draw a hypothesis plot
- What mixed-effects model is suited to test this hypothesis?
- How many parameters does this model have?

Choose a mixed-effects model

- Let us look at perceived risk
- Hypothesis: In certain contexts, people will perceive the risk to consult an AI expert sytem as much higher compared to a human expert when stakes are high
- Draw a hypothesis plot
- What mixed-effects model is suited to test this hypothesis?
- How many parameters does this model have?
- Which parameter do we need to look at in order to test the hypothesis?

Choose a mixed-effects model

- Let us look at perceived risk
- Hypothesis: In certain contexts, people will perceive the risk to consult an AI expert sytem as much higher compared to a human expert when stakes are high
- Draw a hypothesis plot
- What mixed-effects model is suited to test this hypothesis?
- How many parameters does this model have?
- Which parameter do we need to look at in order to test the hypothesis?
- Which random effects are needed to represent the experimental design?

Testing three-way interaction

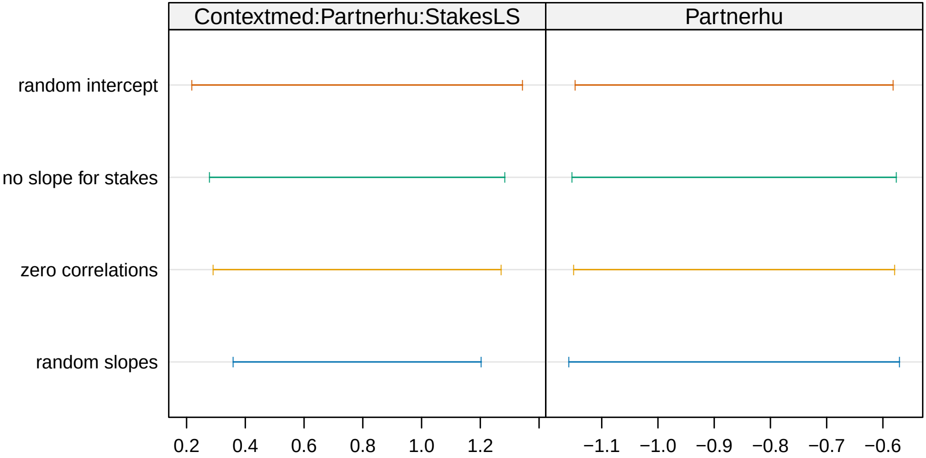
```
m0 <- lmer(Risk ~ (Context + Partner + Stakes)^2 +
           (1 + Partner + Stakes | Participant), data = dat)

m1 <- lmer(Risk ~ Context * Partner * Stakes +
           (1 + Partner + Stakes | Participant), data = dat)

# Test interaction with Likelihood-ratio Test
anova(m0, m1)
```

Testing three-way interaction

```
m0 <- lmer(Risk ~ (Context + Partner + Stakes)^2 +
           (1 + Partner + Stakes | Participant), data = dat)

m1 <- lmer(Risk ~ Context * Partner * Stakes +
           (1 + Partner + Stakes | Participant), data = dat)

# Test interaction with Likelihood-ratio Test
anova(m0, m1)
```

- Calculate the confidence intervals for Model `m1` and compare them to a model with only random intercepts for `Participant`
- What would you expect based on the results we looked at for Aungle and Langer (2023)?

Confidence intervals for three way interaction



3 Simulations

Selection of random effects

- What is the difference between the two designs we looked at?

Selection of random effects

- What is the difference between the two designs we looked at?
- Maybe it's just one of the data sets behaving weird?

Selection of random effects

- What is the difference between the two designs we looked at?
- Maybe it's just one of the data sets behaving weird?
- Let's check and simulate some data ...

2×2 within design

$$y = \beta_0 + \beta_1 a_2 + \beta_2 b_2 + \beta_3 a_2 b_2 + \upsilon_0 + \upsilon_1 a_2 + \upsilon_2 b_2 + \varepsilon$$

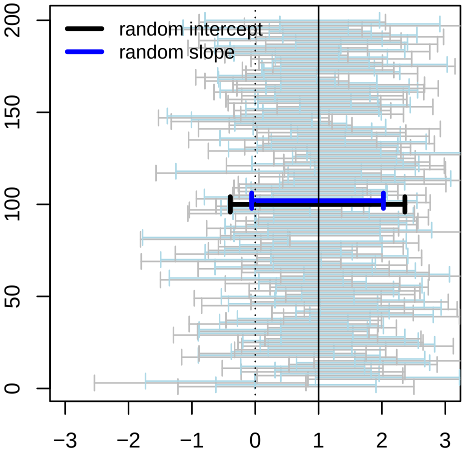
with $\boldsymbol{\upsilon} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} & \sigma_{\upsilon_0\upsilon_2} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} & \sigma_{\upsilon_1\upsilon_2} \\ \sigma_{\upsilon_0\upsilon_2}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

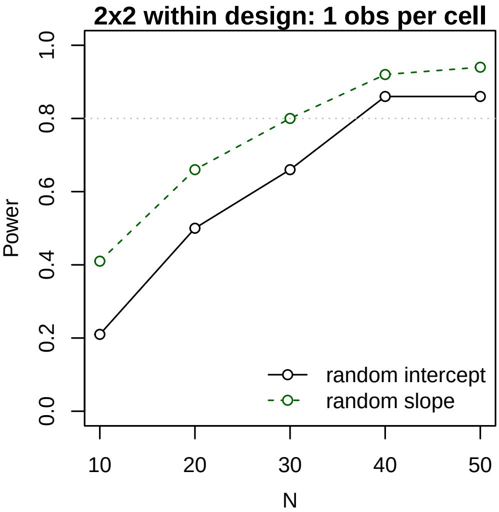
```
dat <- expand.grid(A = factor(c("a1", "a2")), B = factor(c("b1", "b2")),
                   id = factor(1:10))

beta  <- c(3, .5, .5, 1)
sp    <- c(1, .8, .6)
r     <- -.5
S     <- r * sp %o% sp; diag(S) <- sp^2
se    <- 1
```

2 × 2 within design

```
Lt    <- chol(S) / se
theta <- t(Lt)[lower.tri(Lt, diag = TRUE)]

cis <- replicate(100, {
  y <- simulate(~ A * B + (A + B | id),
                newdata = dat,
                newparams = list(beta = beta, theta = theta, sigma = se))$sim_1
  m0 <- lmer(y ~ A * B + (1 | id), dat)
  m1 <- lmer(y ~ A * B + (A + B | id), dat)
  matrix(c(confint(m0, par = "Aa2:Bb2", method = "Wald") |> as.numeric(),
           confint(m1, par = "Aa2:Bb2", method = "Wald") |> as.numeric()),
         nrow = 2, byrow = TRUE)
  }, simplify = FALSE
)
dat_ci <- as.data.frame(do.call(rbind, cis))
names(dat_ci) <- c("lb", "ub")
dat_ci$model <- factor(c("random intercept", "random slope"))
```

2 × 2 within design





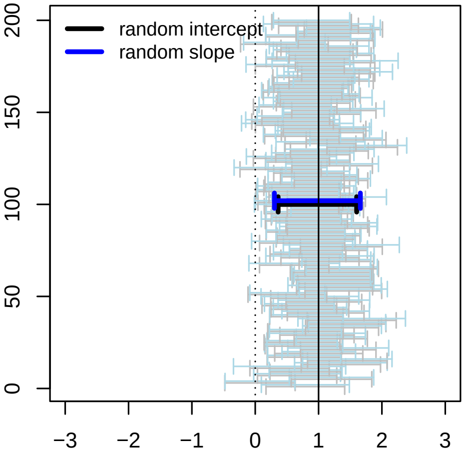
2×2 within design with several items

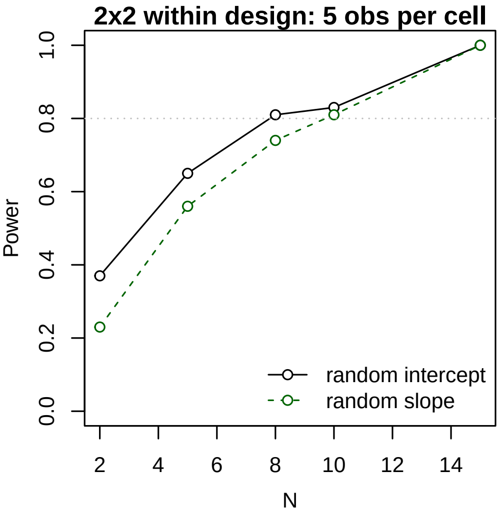
$$y = \beta_0 + \beta_1 a_2 + \beta_2 b_2 + \beta_3 a_2 b_2 + \upsilon_0 + \upsilon_1 a_2 + \upsilon_2 b_2 + \upsilon_3 a_2 b_2 + \omega_0 + \varepsilon$$

$$\text{with } \boldsymbol{\upsilon} \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_\upsilon = \begin{pmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} & \sigma_{\upsilon_0\upsilon_2} & \sigma_{\upsilon_0\upsilon_3} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} & \sigma_{\upsilon_1\upsilon_2} & \sigma_{\upsilon_1\upsilon_3} \\ \sigma_{\upsilon_0\upsilon_2} & \sigma_{\upsilon_1\upsilon_2} & \sigma^2_{\upsilon_2} & \sigma_{\upsilon_2\upsilon_3} \\ \sigma_{\upsilon_0\upsilon_3} & \sigma_{\upsilon_1\upsilon_3} & \sigma_{\upsilon_2\upsilon_3} & \sigma^2_{\upsilon_3} \end{pmatrix}\right), \ \omega_0 \sim N(0, \sigma^2_\omega), \ \varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$$

```
dat <- expand.grid(A = factor(c("a1", "a2")), B = factor(c("b1", "b2")),
                   item = factor(1:5), id = factor(1:10))
beta   <- c(3, .5, .5, 1)
sp     <- c(1, .8, .6)
r      <- -.5
se     <- 1
S      <- r * sp %o% sp; diag(S) <- sp^2
sw     <- 1
```

2×2 within design with several items





Summary

- The discussion "Keep it maximal" (Barr et al., 2013) vs. "Parsimonious mixed models" (Bates et al., 2018) is even more complicated than it looks (as always)

Summary

- The discussion “Keep it maximal” (Barr et al., 2013) vs. “Parsimonious mixed models” (Bates et al., 2018) is even more complicated than it looks (as always)
- I still think that the random effects structure should represent the experimental design, since this aligns with thinking about the structure of the random effects as the data generating process

Summary

- The discussion "Keep it maximal" (Barr et al., 2013) vs. "Parsimonious mixed models" (Bates et al., 2018) is even more complicated than it looks (as always)
- I still think that the random effects structure should represent the experimental design, since this aligns with thinking about the structure of the random effects as the data generating process
- If you want to interpret the random effects or even test hypotheses about them, you need to be careful if the estimates are good enough, however

Summary

- The discussion "Keep it maximal" (Barr et al., 2013) vs. "Parsimonious mixed models" (Bates et al., 2018) is even more complicated than it looks (as always)
- I still think that the random effects structure should represent the experimental design, since this aligns with thinking about the structure of the random effects as the data generating process
- If you want to interpret the random effects or even test hypotheses about them, you need to be careful if the estimates are good enough, however
- And always: Do not go through the motions, try to understand the model that you fit and what its structure implies

References

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

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