The Importance of Random Slopes in Mixed Models for Bayesian Hypothesis Testing

Oberauer, K. (2022)

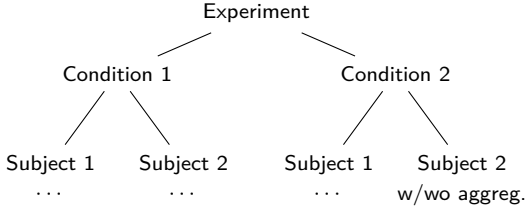
Zhengxiao Wei

October 24 & 31, 2023

Mixed Effects

A mixed-effects model considers the subject i at the lower level and the population at the higher level.

- Fixed effects are usually at the higher level that are assumed to be constant for every unit at the lower level.
- Random effects describe individual differences at the lower level.



Random Effects

Two-factor factorial design with blocking in *anovaBF*

$$Y_{ijk} = \mu + \sigma_\epsilon(s_i + \alpha_j + \beta_k + (\alpha\beta)_{jk}) + \epsilon_{ijk}, \quad \epsilon_{ijk} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\ \sigma_\epsilon^2) \tag{1}$$

Two-way repeated-measures design in `JASP`

$$Y_{ijk} = \mu + \sigma_\epsilon(s_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + (s\alpha)_{ij} + (s\beta)_{ik}) + \epsilon_{ijk} \tag{2}$$

1. Random intercepts: individual differences in the mean across all conditions.
2. Random slopes: differences between individuals in the direction and size of an effect.
3. Correlations between random effects are model parameters describing dependencies among random intercepts and random slopes of continuous predictors.

Mixed Model Example

Level 1: (within-subject growth and a time-varying predictor)

$$Y_{it} = \beta_{0i} + \beta_{1i} \cdot \text{Time}_{it} + \epsilon_{it}, \quad \epsilon_{it} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\ \sigma^2_\epsilon)$$

Level 2: (between-subjects differences in growth and a subject-level predictor)

$$\begin{aligned} \beta_{0i} &= \gamma_{00} + \gamma_{01} \cdot X_i + u_{0i} \\ \beta_{1i} &= \gamma_{10} + \gamma_{11} \cdot X_i + u_{1i}, \quad (u_{0i},\ u_{1i})^\top \sim \boldsymbol{\mathcal{N}}_2\left(\mathbf{0},\ \begin{pmatrix} \tau_0^2 & \rho\tau_0\tau_1 \\ \rho\tau_0\tau_1 & \tau_1^2 \end{pmatrix}\right) \\ &\qquad\qquad \succ 0 \end{aligned}$$

Conditional on X Growth Model:

$$Y_{it} = \underset{\substack{\text{fixed} \\ \text{intercept}}}{\gamma_{00}} + \underset{\substack{\text{random} \\ \text{intercepts}}}{u_{0i}} + \gamma_{01} \cdot X_i + (\underset{\substack{\text{fixed} \\ \text{slope}}}{\gamma_{10}} + \underset{\substack{\text{random} \\ \text{slopes}}}{u_{1i}}) \cdot \text{Time}_{it} + \gamma_{11} \cdot [\text{Time} \times X]_{it} + \epsilon_{it}$$

Modeling Fixed Effects

$\mathbf{Y}_i \equiv (Y_{i11}, \cdots, Y_{ip1}, \cdots, Y_{i1q}, \cdots, Y_{ipq})^\top, \qquad (\boldsymbol{\alpha\beta}) \equiv \left((\alpha\beta)_{11}, (\alpha\beta)_{p1}, \cdots, (\alpha\beta)_{1q}, (\alpha\beta)_{pq}\right)^\top$

$$\mathbf{Y} = \mu\mathbf{1} + \sigma_\epsilon(\mathbf{X}_s\boldsymbol{s} + \mathbf{X}^*_\alpha\boldsymbol{\alpha}^* + \mathbf{X}^*_\beta\boldsymbol{\beta}^* + \mathbf{X}^*_{(\alpha\beta)}(\boldsymbol{\alpha\beta})^*) + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \boldsymbol{\mathcal{N}}_{Npq}(\mathbf{0},\ \sigma^2_\epsilon\mathbf{I}), \qquad (1')$$

where $\mathbf{X}_s = \mathbf{I}_N \otimes \mathbf{1}_{pq}$, $\mathbf{X}_\alpha = \mathbf{1}_{Nq} \otimes \mathbf{I}_p$, $\mathbf{X}_\beta = \mathbf{1}_N \otimes \mathbf{I}_q \otimes \mathbf{1}_p$, $\mathbf{X}_{(\alpha\beta)} = \mathbf{X}_\beta \bullet \mathbf{X}_\alpha$,

Kronecker product

Face-splitting product

$$\mathbf{X}^*_\alpha = \mathbf{X}_\alpha \cdot \mathbf{Q}_p, \qquad \mathbf{X}^*_\beta = \mathbf{X}_\beta \cdot \mathbf{Q}_q, \qquad \text{and} \quad \mathbf{X}^*_{(\alpha\beta)} = \mathbf{X}^*_\beta \bullet \mathbf{X}^*_\alpha$$

Projecting a set of p effects into $p-1$ parameters with the property that the marginal g-prior on all p effects is identical. $\boldsymbol{\alpha}^{*\top} \equiv (\alpha^*_1, \cdots, \alpha^*_{p-1}) = (\alpha_1, \cdots, \alpha_p) \cdot \mathbf{Q}_p$

$\mathbf{I}_p - p^{-1}\mathbf{J}_p = \mathbf{Q}_p \cdot \mathbf{Q}_p^\top$, $\mathbf{Q}_p$ is a $p \times (p-1)$ matrix of the $p-1$ eigenvectors of unit length corresponding to the nonzero eigenvalues of the left side term.

For example, $\alpha^* = (\alpha_1 - \alpha_2)/\sqrt{2}$ when $p = 2$.
In the other direction (given $\alpha_1 + \alpha_2 = 0$), $\alpha_1 = \alpha^*/\sqrt{2}$ and $\alpha_2 = -\alpha^*/\sqrt{2}$.

Statement of Relevance

Remark 1: An inflation of type I errors

The omission of random slopes lead to an increase in false positives when the (fixed) effect of interest is actually zero.

Remark 2: A loss of statistical power

The inclusion of random slopes reduce the chance of obtaining compelling evidence for the (fixed) effect of interest if that effect is truly different from zero.

- in either frequentist or Bayesian hypothesis testing.

2×2 Repeated-Measures Analysis of Variance (RM-ANOVA)

For $i = 1, \cdots, N; \quad j, k \in \{1, 2\}; \quad r = 1, \cdots, n,$

$$\text{Level 1:} \quad Y_{ijkr} \sim \mathcal{N}(M_{ijk},\ 1) \tag{3}$$

$$M_{ijk} = \mu_i + b_{i1} \cdot C_1 + b_{i2} \cdot C_2 \tag{4}$$

$$\text{Level 2:} \quad \mu_i \sim \mathcal{N}(0,\ \sigma^2) \tag{5}$$

$$b_{i1} \sim \mathcal{N}(B_1,\ \sigma^2) \tag{6}$$

$$b_{i2} \sim \mathcal{N}(B_2,\ \sigma^2) \tag{7}$$

$\mu_i,\ b_{i1},\ b_{i2}$	subject-level effects
B_1, B_2	population-level effects
σ	between-subjects standard deviation of effects
C_1, C_2	contrast-coded independent variables and **no** interaction C_1C_2
$\overline{Y}_{ijk\cdot}$	aggregated data at the level of the design cell

Let $B_2 = .5, \quad C_1, C_2 = [-.5,\ .5], \quad N = 20, \quad$ and $n = 30$.
Vary $B_1 \in \{\underset{\text{null}}{0},\ .25,\ .5,\ .75\}$ and $\sigma \in \{.1,\ .25,\ .5,\ 1\}$ in simulations (50 runs each).

R Scripts: Simulation


```
N <- 20; n <- 30; C1 <- C2 <- c(-.5, .5); B1 <- 0; B2 <- .5; sd <- .5

set.seed(277)
s <- rnorm(N, 0, sd)   #random intercepts
b1 <- rnorm(N, B1, sd) #random slopes of factor A
b2 <- rnorm(N, B2, sd) #random slopes of factor B
M <- s + cbind(b1, b2) %*% rbind(rep(C1, 2), rep(C2, each=2)) #cell means
dat <- data.frame("DV" = rnorm(N*4*n, c(M), 1), #a1b1, a2b1, a1b2, a2b2
                  "ID" = factor(rep(paste0("s",1:N), 4*n)),
                  "A" = factor(rep(rep(c("a1","a2"), each=N), 2*n)),
                  "B" = factor(rep(rep(c("b1","b2"), each=N*2), n)),
                  "Trial" = factor(rep(paste0("r",1:n), each=N*4)))

library(BayesFactor) #v 0.9.12-4.5
set.seed(277)
BF_additive <- lmBF(DV ~ A + B + ID + A:ID + B:ID,
                    dat, whichRandom="ID", iterations=1e5, progress=F)
set.seed(277)
draws <- posterior(BF_additive, iterations=1e5, progress=F)
summary(draws)
```

Without aggregation (number of trials $n > 1$), `A:B:ID`, $(s\alpha\beta)_{ijk}$, is estimable.

R Output: Estimation

True values: $\mu = 0, \quad A = B_1 \cdot C_1 = 0, \quad B = B_2 \cdot C_2 = \pm.25, \quad \sigma = .5$, and $\sigma_\epsilon = 1$.

	Mean	SD	Naive SE	Time-series SE
mu	0.120172	0.13818	4.370e-04	4.370e-04
A-a1	-0.045653	0.07207	2.279e-04	2.279e-04
A-a2	0.045653	0.07207	2.279e-04	2.279e-04
B-b1	-0.256099	0.07067	2.235e-04	2.308e-04
B-b2	0.256099	0.07067	2.235e-04	2.308e-04
ID-s1	0.727577	0.16124	5.099e-04	5.099e-04
...	*#not show other subjects*			
A:ID-a1.&.s1	0.257360	0.10873	3.438e-04	3.438e-04
...				
A:ID-a2.&.s1	-0.257360	0.10873	3.438e-04	3.438e-04
...				
B:ID-b1.&.s1	-0.256027	0.10737	3.395e-04	3.491e-04
...				
B:ID-b2.&.s1	0.256027	0.10737	3.395e-04	3.491e-04
...				
sig2	0.938782	0.02734	8.645e-05	9.006e-05
g_A	1.420612	37.95095	1.200e-01	1.200e-01
g_B	2.457176	111.21710	3.517e-01	3.517e-01
g_ID	0.398090	0.14254	4.508e-04	4.940e-04
g_A:ID	0.224002	0.08337	2.636e-04	3.019e-04
g_B:ID	0.202324	0.07565	2.392e-04	2.750e-04

Bayes Factor (BF) Strategies

For testing the predictor of interest a, $BF_{01} \rightarrow$ constrained / additive models.

$$BF_{01}^{\text{No RS}} \rightarrow s + b \;/\; s + a + b \tag{8}$$

$$BF_{01}^{\text{RS 1}} \rightarrow s + b + sa \;/\; s + a + b + sa \tag{9}$$

$$BF_{01}^{\text{RS 2}} \rightarrow s + b + sb \;/\; s + a + b + sb \tag{10}$$

$$BF_{01}^{\text{RS}} \rightarrow s + b + sa + sb \;/\; s + a + b + sa + sb \tag{11}$$

$$BF_{01}^{\text{evidence}} \rightarrow \text{constrained / best-supported model by the data} \tag{12}$$

- The numerator of eq. 9 or 11 still contains sa because that model represents the null hypothesis that the population-level effect is zero, which does not necessarily mean that the effect is zero for each subject.

Results of Remark 1

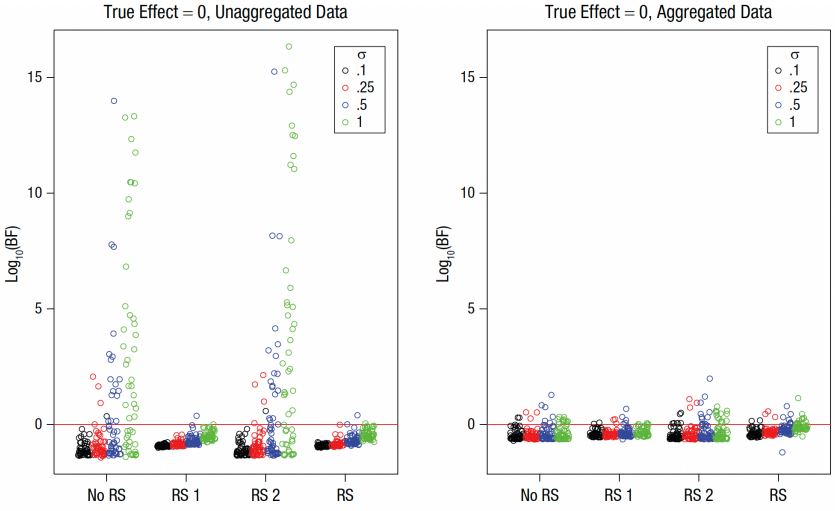


Fig. 1. Simulation 1: Bayes factors (BFs) for the effect of interest when the true effect is zero, as a function of model version and standard deviation of random effects, σ. Results are shown separately for unaggregated data (left) and aggregated data (right). The "no RS" model version included no random slopes, the "RS 1" version included random slopes for the effect of interest, the "RS 2" version included random slopes for the other effect, and the "RS" version included random slopes for both effects. Each point is a BF from one simulation run. BFs are plotted on a $\log_{10}$ scale for visibility. BFs above the red line indicate evidence for the effect; those below the red line indicate evidence against the effect.

Results of Remark 2

Not shown as a function of σ?

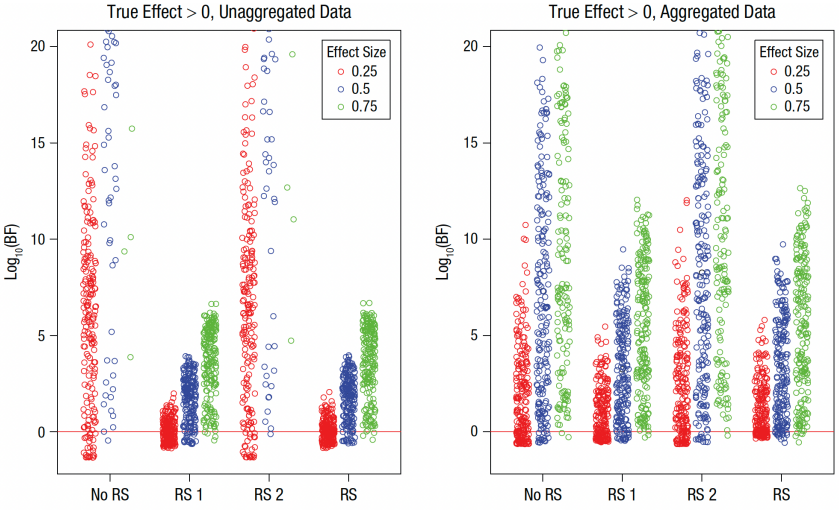


Fig. 3. Simulation 1: Bayes factors (BFs) from simulations in which the effect of interest was greater than 0, as a function of model version and true effect size. Results are shown separately for unaggregated data (left) and aggregated data (right). The "no RS" model version included no random slopes, the "RS 1" version included random slopes for the effect of interest, the "RS 2" version included random slopes for the other effect, and the "RS" version included random slopes for both effects. Each point is a BF from one simulation run. BFs are plotted on a $\log_{10}$ scale for visibility. BFs above the red line indicate evidence for the effect; those below the red line indicate evidence against the effect.

Other Simulations

2. $2 \times q$ RM-ANOVA With Many Design Cells

$j \in \{1, 2\}$, $k = 1, \cdots, q = 2,\ 4,\ 6,$ or 10, $\sigma = .5$, $B_1 \in \{0,\ .5\}$, $B_2 = .5$, $N = 20$, and $n = 30$.

3. Two-Level Regression With Correlated Random Slopes (*brms* and *bridgesampling*)

$p \overset{\text{ordinal}}{=} q = 5$, $\rho \in \{0,\ .2,\ .5,\ .7\}$, $\sigma \in \{.1,\ .25,\ .5,\ 1\}$, $B_1 \in \{0,\ .25\}$, $B_2 = .25$, $N = 20$, and $n = 30$.

```
brm(DV ~ A + B + (1 + A + B | ID), ...)
```

- single | leaves correlations free; double || sets correlations to 0; see also *lme4::lmer*

4. Effects of N and n in the 2×2 RM-ANOVA

$\sigma = .5$, $B_1 \in \{0,\ .25\}$, $B_2 = .5$, $N \in \{20,\ 30,\ 40,\ 50\}$, and $n \in \{20,\ 30,\ 40,\ 60,\ 90\}$.

5. *BF* Calibration in the 2×2 RM-ANOVA

The model parameters were sampled from the JZS priors in Rouder et al. (2012).
$B_1, B_2 \overset{\text{i.i.d.}}{\sim}$ Cauchy$(0,\ .5)$, $\sigma \sim$ Gamma$(2,$ rate$=10)$, $C_1, C_2 = [-1/\sqrt{2},\ 1/\sqrt{2}]$, $N = 20$, and $n = 30$.

BF Calibration*

*disambiguation

$$p(\mathcal{M}_\xi) = \iint p(\mathcal{M}_\xi \mid \boldsymbol{y}) \cdot p(\boldsymbol{y} \mid \mathcal{M}) \cdot p(\mathcal{M}) \, \mathrm{d}\mathcal{M} \, \mathrm{d}\boldsymbol{y} \qquad \text{for } \xi \in \{0, 1\} \qquad (13)$$

Choose a model from the model prior, $p(\mathcal{M})$, e.g., $p(\mathcal{M}_0) = p(\mathcal{M}_1) = .5$.

Simulate data based on this model, $p(\boldsymbol{y} \mid \mathcal{M}) = \int p(\boldsymbol{y} \mid \boldsymbol{\theta}, \mathcal{M}) \cdot p(\boldsymbol{\theta} \mid \mathcal{M}) \, \mathrm{d}\boldsymbol{\theta}$.

Infer a model from the simulated data, $p(\mathcal{M}_\xi \mid \boldsymbol{y})$. Recall $p(\mathcal{M}_1 \mid \boldsymbol{y}) = \frac{BF_{10}}{BF_{10} + p(\mathcal{M}_0)/p(\mathcal{M}_1)}$

Repeat.

If the computation of Bayes factors and posterior model probabilities is performed correctly, then the average posterior probability for a model over simulation runs should approximate its prior probability in eq. 13 (Schad et al., 2022, p. 6-7).

Results

2. Aggregating does not reduce the false positive rate in complex designs.

3. Omitting the correlations between random slopes from the model when they are in the data is **innocuous**.

4. Increasing the sample size and the number of trials increases the power.

5. Always including random slopes, as well as using the **parsimonious** model-comparison approach (i.e., including random slopes only when there was evidence for them), yielded well-calibrated *BF*s at least for one common design.

$$BF_{\text{RS evidence}} \rightarrow s + a + b + sa + sb \;/\; s + a + b \tag{14}$$

R Scripts: Parsimonious Model Comparison

Data Source

```
df <- read.table(file.choose(), header=T, stringsAsFactors=T) #A: resp;  B: align
# set.seed(277)
# anovaBF(RT ~ resp * align + subj,
#         df, whichRandom="subj", whichModels="top", iterations=1e5, progress=F)

set.seed(277)
full_RS <- lmBF(RT ~ resp * align + subj + resp:subj + align:subj,
                df, whichRandom="subj", iterations=1e5, progress=F)

set.seed(277)
full_noRS <- lmBF(RT ~ resp * align + subj,
                  df, whichRandom="subj", iterations=1e5, progress=F)

full_RS / full_noRS #1.83207e+26 ±5.75%
```

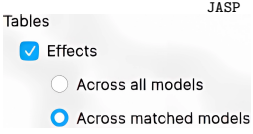
The Paradox: Factor B is statistically significant in RM-ANOVA but not supported in the *anovaBF* function (v 0.9.12-4.5 or lower). One can break the data into two parts, separated by the factor A, and conduct separate Bayesian tests for the factor of interest B in each condition of A. Under these circumstances, a very clear effect B emerges in both analyses. This discrepancy is not the product of a contrast between a liberal and a conservative test (Bub, Masson, & van Noordenne, 2021, p. 80).

Model-Averaged *BF*

For testing the predictor of interest a,
the inclusion *BF*s of models containing the effect relative to models stripping of it
(always including random slopes and enforcing the principle of marginality for fixed effects) are

$$BF_{\text{incl}} \rightarrow \frac{[s + a + b + ab + sa + sb] + [s + a + b + sa + sb] + [s + a + sa + sb]}{[s + b + sa + sb] + [s + sa + sb]} \quad (15)$$

$$BF_{\text{incl}}^{\text{match}} \rightarrow \frac{[s + a + b + sa + sb] + [s + a + sa + sb]}{[s + b + sa + sb] + [s + sa + sb]} \quad (16)$$



$BF^{RS}_{\xi\cdot}$ to BF_{incl}

Prior Model Probabilities, s.t. $\sum_{\xi} p(\mathcal{M}_\xi) = 1$; Posterior Model Probabilities, s.t. $\sum_{\xi} p(\mathcal{M}_\xi \mid \boldsymbol{y}) = 1$

Prior Inclusion Probability, $p(\mathcal{M}_{incl}) = \sum_{\eta} p(\mathcal{M}_\eta)$

Posterior Inclusion Probability,

$$p(\mathcal{M}_{incl} \mid \boldsymbol{y}) = \sum_{\eta} p(\mathcal{M}_\eta \mid \boldsymbol{y}) = \frac{\sum_\eta p(\boldsymbol{y} \mid \mathcal{M}_\eta) \cdot p(\mathcal{M}_\eta)}{\sum_\xi p(\boldsymbol{y} \mid \mathcal{M}_\xi) \cdot p(\mathcal{M}_\xi)} = \frac{\sum_\eta BF^{RS}_{\eta\cdot} \cdot p(\mathcal{M}_\eta)}{\sum_\xi BF^{RS}_{\xi\cdot} \cdot p(\mathcal{M}_\xi)}$$

$$BF_{incl} = \frac{p(\mathcal{M}_{incl} \mid \boldsymbol{y}) \ / \ (1 - p(\mathcal{M}_{incl} \mid \boldsymbol{y}))}{p(\mathcal{M}_{incl}) \ / \ (1 - p(\mathcal{M}_{incl}))} \tag{17}$$

$$BF^{match}_{incl} = \frac{p(\mathcal{M}^{match}_{incl} \mid \boldsymbol{y}) \ / \ p(\mathcal{M}^{match}_{excl} \mid \boldsymbol{y})}{p(\mathcal{M}^{match}_{incl}) \ / \ p(\mathcal{M}^{match}_{excl})} \tag{18}$$

Note that $p(\mathcal{M}^{match}_{excl}) \neq 1 - p(\mathcal{M}^{match}_{incl})$ and $p(\mathcal{M}^{match}_{excl} \mid \boldsymbol{y}) \neq 1 - p(\mathcal{M}^{match}_{incl} \mid \boldsymbol{y})$.

R Scripts: BF_{incl} and BF_{incl}^{match}

```
inclusionBF <- function(data, match_models=F, prior_prob=rep(.2, 5), ...) {
  #' data:          two-way repeated-measures data whose column names contain "DV", "ID", "A", and "B"
  #' match_models: whether effects are across matched models (default is FALSE)
  #' prior_prob:   prior probabilities of [1]full, [2]additive, [3]A, [4]B, and [5]null models (default is being equal)
  #' Output -      a list of the JASP-replicating (v 0.16.3 or later) Bayes factors from the "BayesFactor" dependency
  if(round(sum(prior_prob), 3) != 1 || length(prior_prob) != 5) stop("Prior model probabilities must sum to 1!")
  BF_full <- BayesFactor::lmBF(DV ~ A * B + ID + A:ID + B:ID, data, whichRandom="ID", ...)
  BF_additive <- BayesFactor::lmBF(DV ~ A + B + ID + A:ID + B:ID, data, whichRandom="ID", ...)
  BF_A <- BayesFactor::lmBF(DV ~ A + ID + A:ID + B:ID, data, whichRandom="ID", ...)
  BF_B <- BayesFactor::lmBF(DV ~ B + ID + A:ID + B:ID, data, whichRandom="ID", ...)
  BF_0 <- BayesFactor::lmBF(DV ~ ID + A:ID + B:ID, data, whichRandom="ID", ...)
  BF <- data.frame(c(BF_full, BF_additive, BF_A, BF_B, BF_0) / BF_0) #compare to null
  BF_weighted <- BF$bf * prior_prob; BF_avg <- sum(BF_weighted)
  if(match_models) {
    indices_incl <- list(c(2, 3), c(2, 4), 1)       #indices of inclusion matched models
    indices_excl <- list(c(4, 5), c(3, 5), 2)       #indices of exclusion matched models
  } else {
    indices_incl <- list(c(1, 2, 3), c(1, 2, 4), 1) #indices of inclusion models
    indices_excl <- list(c(4, 5), c(3, 5), 2:5)     #indices of exclusion models
  }
  out <- data.frame("Effects" = c("A", "B", "AB"),
                    "Prior.Incl" = sapply(indices_incl, function(x) sum(prior_prob[x])), #prior inclusion probabilities
                    "Prior.Excl" = sapply(indices_excl, function(x) sum(prior_prob[x])), #prior exclusion probabilities
                    "Post.Incl" = sapply(indices_incl, function(x) sum(BF_weighted[x])) / BF_avg,
                    "Post.Excl" = sapply(indices_excl, function(x) sum(BF_weighted[x])) / BF_avg)
  out$BF.Incl <- (out$Post.Incl * out$Prior.Excl) / (out$Post.Excl * out$Prior.Incl)     #inclusion Bayes factors
  OG <- data.frame("Models" = row.names(BF), "Prior.M" = prior_prob, "Post.M" = BF_weighted / BF_avg,
                   "BF10" = BF$bf, "error.per" = BF$error * 100)             #JASP model comparison
  OG$BF.M <- (OG$Post.M * (1 - prior_prob)) / ((1 - OG$Post.M) * prior_prob) #model Bayes factors
  OG <- OG[order(OG$BF10, decreasing=T), c(1:3,6,4,5)]; row.names(OG) <- NULL
  list("MRE" = OG, "BMA" = out) #MRE: maximal set of random effects; BMA: Bayesian model averaging
}
```

Model Misspecification

The three elimination strategies — not testing nuisance effects, eliminating models with interactions and no corresponding main effects, and **eliminating subject-by-treatment interactions** — reduces the number of models considerably. For example, in the four-factor case with factors of *distance*, *presentation*, *age*, and *subjects*, the strategy reduces the number of models from 32,769 to 19 (Rouder et al., 2017, p. 314).

As currently implemented, the proposed Bayesian ANOVA assumes that there are **no individual differences** in the magnitude of effects. We suspect that this assumption is neither obvious to nor desired by most analysts because it is untenable in most applications (van den Bergh et al., 2023, p. 1).

Analysts should ask whether it is plausible that there are no individual differences if an effect is present (van den Bergh et al., 2023, p. 6).

RM-ANOVA Assumptions

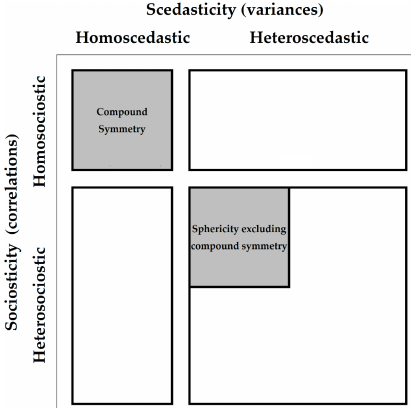
Sphericity (Circularity)

The variances of the differences between any pair of within-subject conditions are equal.

Sphericity does not apply if there are only two levels of a within-subject factor.

Compound Symmetry $\epsilon_{ijk} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\, \sigma^2_\epsilon)$

All conditions have equal variance and all pairs of conditions have equal covariance.



(Cousineau, 2019, p. 232)

Between-subjects designs:
subjects are nested within factors

Within-subject designs:
subjects are crossed with the levels of the factors

Recommendations

When analyzing unaggregated data, one should always start with a model including the random slopes corresponding to all fixed effects. Researchers could test whether all random slopes are warranted by the data by comparing the full model with a model that has a reduced random-effects structure. If the *BF* unambiguously supports the reduced model, it should be fairly safe to continue testing fixed effects in models using the reduced random-effects structure.

When one analyzes aggregated data, doing so without random slopes should be reasonably safe with simple designs. For more complex within-subjects designs that have many design cells, Oberauer (2022) recommended starting with a model that includes the random slopes and excluding them only if the evidence speaks unambiguously against them.

References

Bub, D. N., Masson, M. E. J., & van Noordenne, M. (2021). Motor representations evoked by objects under varying action intentions. *Journal of Experimental Psychology: Human Perception and Performance*, *47*, 53–80. https://doi.org/10.1037/xhp0000876

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1-28. https://doi.org/10.18637/jss.v080.i01

Cousineau, D. (2019). Correlation-adjusted standard errors and confidence intervals for within-subject designs: A simple multiplicative approach. *The Quantitative Methods for Psychology*, *15*, 226–241. https://doi.org/10.20982/tqmp.15.3.p226

Morey, R. D., & Rouder, J. N. (2023). *BayesFactor: Computation of Bayes factors for common designs*. R package version 0.9.12-4.5. https://cran.r-project.org/package=BayesFactor

Oberauer, K. (2022). The importance of random slopes in mixed models for Bayesian hypothesis testing. *Psychological Science*, *33*, 648-665. https://doi.org/10.1177/09567976211046884

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, *22*, 304-321. https://doi.org/10.1037/met0000057

Schad, D. J., Nicenboim, B., Bürkner, P. C., Betancourt, M., & Vasishth, S. (2022). Workflow techniques for the robust use of Bayes factors. *Psychological Methods*, 1-23. https://doi.org/10.1037/met0000472

van den Bergh, D., Wagenmakers, E.-J., & Aust, F. (2023). Bayesian repeated-measures analysis of variance: An updated methodology implemented in JASP. *Advances in Methods and Practices in Psychological Science*, *6*, 1-11. https://doi.org/10.1177/25152459231168024

Thank you! The Alt. Cov. Str. Project.

Contact: zhengxiao@uvic.ca