Fully-crossed structures

Real data: Individual Differences

So simple, yet so effective

Ottavia M. Epifania University of Trento ottavia.epifania@unitn.it

Beyond Summer School

May, 30, 2025

How it started:

A random lesson by Professor Pastore on Generalized Linear Mixed Effects Models

A random lesson by Professor Pastore on Generalized Linear Mixed Effects Models

How it ended:

A publication on Psychological Methods



Psychological Methods

© 2024 American Psychological Association ISSN: 1082-989X

https://doi.org/10.1037/met0000708

A Guided Tutorial on Linear Mixed-Effects Models for the Analysis of Accuracies and Response Times in Experiments With Fully Crossed Design

> Ottavia M. Epifania, Pasquale Anselmi, and Egidio Robusto Department of Philosophy, Sociology, Education and Applied Psychology, University of Padova

> https://doi.org/10.1037/met0000708

Fully-crossed structures

An example: The SNARC effect

Fully-crossed structures

0000000



Small numbers: Perceived on the left Large numbers: Perceived on the right

An example: The SNARC effect

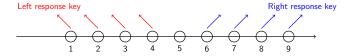
0000000

A sample of small numbers: 1, 2, 3, 4

A sample of large numbers: 6, 7, 8, 9

Two conditions:

The "natural" one (so-called compatible condition)



0000000

A sample of small numbers: A sample of large numbers:

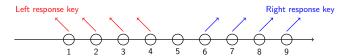
1, 2, 3, 4

6, 7, 8, 9

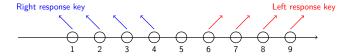
Real data: Individual Differences

Two conditions:

The "natural" one (so-called compatible condition)



The "innatural" one (so-called *incompatible* condition)



An example: The SNARC effect

$t = \{1, 2, \dots, T\}$: Number of trials (condition \times stimulus \times respondent)

		Small Numbers			Large Numbers				
	Condition	1	2	3	4	6	7	8	9
Jane	Compatible	y_{cj1}	y_{cj2}	y_{cj3}	y_{cj4}	y_{cj6}	y_{cj7}	y_{cj8}	$\sum_{t=1}^{T} y_{cj}/T$
	Incompatible	y_{ij1}	y_{ij2}	y_{ij3}	y_{ij4}	y_{ij6}	y_{ij7}	y_{ij8}	$\sum_{t=1}^{T} y_{ij}/T$
Mario	Compatible	y_{cm1}	y_{cm2}	y_{cm3}	y_{cm4}	y_{cm6}	y_{cm7}	y_{cm8}	$\sum_{t=1}^{T} y_{cm}/T$
	Incompatible	y_{im1}	y_{im2}	y_{im3}	y_{im4}	y_{im6}	y_{im7}	y_{im8}	$\sum_{t=1}^{T} y_{im}/T$

Scoring

Person-level scores

$$s = \frac{\bar{X}_{\rm comp} - \bar{X}_{\rm in}}{s d_{\rm pooled}}$$

Scoring

00000000

Person-level scores

$$s = \frac{\bar{X}_{\text{comp}} - \bar{X}_{\text{inc}}}{sd_{\text{pooled}}}$$



Advantages

Ease of computation Ease of interpretation 00000000

Person-level scores

Fully-crossed structures

$$s = \frac{\bar{X}_{\mathsf{comp}} - \bar{X}_{\mathsf{inc}}}{sd_{\mathsf{pooled}}}$$



Advantages

Ease of computation Ease of interpretation



(Implicit) Assumptions

- Being slow (less accurate) in one condition = being fast (or more accurate) in the opposite one: 0 means absence of bias
- 2 All stimuli have the same impact (fixed effects)

00000€00 The issue

A long tradition

i Respondents are random factors

Sampled from a larger population Need for acknowledging the sampling variability Results can be generalized to other respondents belonging to the same population

Real data: Individual Differences

Fully-crossed structures

i Respondents are random factors

Sampled from a larger population Need for acknowledging the sampling variability Results can be generalized to other respondents belonging to the same population

Stimuli/items are fixed factors

Taken to be entire population
There is no sampling variability

There is no need to generalize the results because the stimuli are the population

00000●0
The issue

With long lasting consequences

- Generalization of the results is impaired
- Error variance everywhere, left free to bias everything
- The information at the stimulus level is lost

Fully-crossed structures

- Generalization of the results is impaired
- Error variance everywhere, left free to bias everything
- The information at the stimulus level is lost

$$\Xi$$

Linear Mixed Effects Models

Rasch model

The issue

With long lasting consequences

- Generalization of the results is impaired
- Error variance everywhere, left free to bias everything
- The information at the stimulus level is lost

 \sum

 ψ

Linear Mixed Effects Models

Rasch model

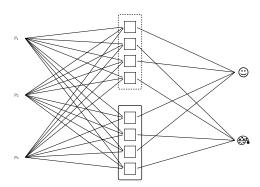


Rasch-like parametrization estimated with Linear Mixed Effects Models

When?

0000000

Fully-crossed structures



Sample-level differences:

Compatible and incompatible can be defined *a priori* (SNARC effect)

Individual differences:

Compatible and incompatible are defined within each respondent (Implicit Association Test)

A Classic of Psychometrics

The Rasch Model

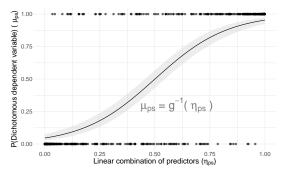
Fully-crossed structures

$$P(x_{ps}=1|\theta_p,b_s) = \frac{\exp(\theta_p-b_z)}{1+\exp(\theta_p-b_z)}$$

 θ_n : Latent trait of person p

 b_s : "challenging" power of stimulus s

A GLM for dichotmous responses



Logit link function g:

$$g(\eta_{ps}) = \log\left(\frac{\mu_{ps}}{1-\mu_{ps}}\right)$$

Inverse g^{-1}

$$g^{-1}=\frac{\exp(\eta_{ps})}{1+\exp(\eta_{ps})}$$

Rasch-like parametrization of response times

The log-normal model

$$E(t_{ps}|\tau_p,\delta_s) = \delta_s - \tau_p + \varepsilon$$

 au_p : the speed of person p

 $\delta_s \! \colon$ the time intensity of stimulus s

The log-normal model

$$E(t_{ps}|\tau_p,\delta_s) = \delta_s - \tau_p + \varepsilon$$

 τ_p : the speed of person p

 $\delta_s {:}\$ the time intensity of stimulus s

A linear model with an identity function!

Rasch

$$P(x_{ps}=1) = \frac{\exp(\theta_p - b_z)}{1 + \exp(\theta_p - b_z)}$$

i Log-normal

$$E(t_{ps}|\tau_p,\delta_s) = \delta_s - \tau_p + \varepsilon$$

GLM (inverse function)

 $P(x_{ps} = 1) = \frac{\exp(\theta_p + b_s)}{1 + \exp(\theta_n + b_s)}$

$$E(t_{ps}|\tau_p,\delta_s)=\delta_s+\tau_p+\varepsilon$$

Random Factors and Random Effects

In a LM:

$$\eta = \mathbf{X}\beta$$

X: Model Matrix

 β : Coefficients

Real data: Individual Differences

In a I M:

Fully-crossed structures

$$\eta = \mathbf{X}\beta$$

X: Model Matrix

 β : Coefficients

Needs to be extended:

$$\eta = \mathbf{X}\beta + \mathbf{Z}d$$

d: Random effects associated to the random factors in Z ... Not model parameters! Best Linear Unbiased Predictors

 Γ : Parameters estimated for the random factors in the model (variances and covariances)

Random structures

The maximal model

Address all the possible sources of random variability that can be expected

Random structures

The maximal model

Address all the possible sources of random variability that can be expected

The models that are useful for ones aim

Random structures

The maximal model

Address all the possible sources of random variability that can be expected

The models that are useful for ones aim

Common goal: Investigate the changes in the performance of the respondents between the associative conditions

Fully-crossed structures

The maximal mode

Address all the possible sources of random variability that can be expected

The models that are useful for ones aim

Common goal: Investigate the changes in the performance of the respondents between the associative conditions

Less common: Investigate the changes in the functioning of the stimuli between the associative conditions

Preliminarities

Index	Meaning	Variable
	Respondent Stimulus Associative condition Trial	respondents stimuli condition

Accuracy: GLMM
$$y = [0, 1]$$

$$\begin{array}{c} \text{Log-time response} \\ \text{LMM} \\ y = [0, +\infty) \text{ (log-transformed)} \end{array}$$

$$\varepsilon \sim \mathcal{N}(0, \sigma^2)$$

Fully-crossed structures

Model 1

i Mathematical Notation

$$y = \beta_c X_c + \alpha_p[i] + \alpha_s[i]$$

1me4 notation

Asch-like parametrization

	GLMM	LMM
respondents stimuli	$egin{pmatrix} heta_p \ b_s \ \end{matrix}$	$\tau_p \\ \delta_s$

Fully-crossed structures

Model 2

i Mathematical Notation

$$y = \beta_c X_c + \alpha_p[i] + \beta_s[i]c_i$$

1me4 notation

Rasch-like parametrization

	GLMM	LMM
respondents stimuli	$\begin{matrix}\theta_p\\b_{sc}\end{matrix}$	$\begin{matrix}\tau_p\\\delta_{sc}\end{matrix}$

Model 3

i Mathematical Notation

$$y = \beta_c X_c + \beta_p[i] c_i + \alpha_s[i]$$

1me4 notation

y ~ 0 + condition + (1|stimuli) + (0+condition|respondents)

Assch-like parametrization

	GLMM	LMM
respondents stimuli	$\begin{matrix}\theta_{pc}\\b_s\end{matrix}$	$\begin{matrix} \tau_{pc} \\ \delta_s \end{matrix}$

Fully-crossed structures

All models are wrong...

Find the useful model via model comparison: AIC and BIC

The lower the value, the better the model

AIC, BIC, and model complexity:

Total number of parameters: β and Γ NOT the levels in d.

Model 2 and Model 3: Same complexity, different focus

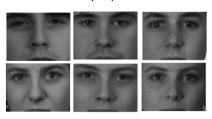
The chosen model is the least wrong model given the models considered: Relativity applies everywhere

Real data: Individual Differences

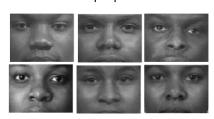
The Implicit Association Test

12 Object stimuli

White people faces



Black people faces



16 Attribute stimuli

Positive attributes Good, laughter, pleasure, glory, peace, happy, joy, love

Negative attributes

Evil, bad, horrible, terrible, nasty, pain, failure, hate

Get set

library(lme4) # Fitting LMMs
library(ggplot2) # Plots

Real data: Individual Differences

Fully-crossed structures

```
respondent condition
                         stimuli accuracy latency
1
             Whitegood
                            hate
                                               1224
                                               5160
           1 Whitegood
                            bf14
                                               1214
             Whitegood laughter
             Whitegood
                            bf56
                                               1143
5
             Whitegood
                            evil
                                                827
             Whitegood
                             wf3
                                               1859
```

Number of trials \times condition \times respondent:

table(data\$respondent, data\$condition)

	Whitebad	Whitegood
1	60	60
2	60	60
3	60	60
4	60	60
5	60	60
6	60	60

Model 2:
$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{p[i]} + \beta_{s[i]} c_i)$$
 accuracy = glmer(accuracy ~ 0 + condition + (0 + condition|stimuli) + (1|respondent), data = data, family = "binomial")

Model 3:
$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{s[i]} + \beta_{p[i]}c_i)$$
 accuracy3 = glmer(accuracy ~ 0 + condition + (1|stimuli) +
$$(0 + \text{condition}|\text{respondent}),$$
 data = data,
$$\text{family} = \text{"binomial"})$$

Random Factors and Random Effects

Real data: Individual Differences

Model 1:
$$y_i = \alpha + \beta_c X_c + \alpha_{p[i]} + \alpha_{s[i]} + \varepsilon_i$$
 logtime1 = lmer(log(latency) ~ 0 + condition + (1|stimuli) + (1|respondent), data = data, REML = FALSE)

Model 2:
$$y_i = \alpha + \beta_c X_c + \alpha_{p[i]} + \beta_{s[i]} c_i + \varepsilon_i$$
 logtime2 = lmer(log(latency) ~ 0 + condition + (0 + condition|stimuli) + (1|respondent), data = data, REML = FALSE)

Model 3:
$$y_i = \alpha + \beta_c X_c + \alpha_{s[i]} + \beta_{p[i]} c_i + \varepsilon_i$$
 logtime3 = lmer(log(latency) ~ 0 + condition + (1|stimuli) + (0 + condition|respondent), data = data, REML = FALSE)

A Classic of Psychometrics

Fully-crossed structures

GLMM



The use of the anova() function is just for the convenience of having all the information on the same page!

anova(accuracy1, accuracy2, accuracy3)

LMMs



The use of the anova() function is just for the convenience of having all the information on the same page!

anova(logtime1, logtime2, logtime3)

Fully-crossed structures

GLMMs Model~2 θ_p b_{WGRB} and b_{RGWB}

The IAT effect is mostly due to variations in the *stimuli functioning* between conditions, while the performance of the respondents seems unaltered

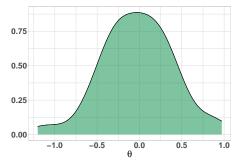
Results should be interpreted together!

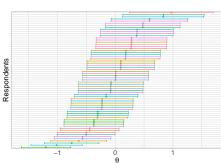
Real data: Individual Differences

The IAT effect is mostly due to variations in the *performance of the respondents* between conditions, while the functioning of the stimuli appears not affected

Rasch-like estimates

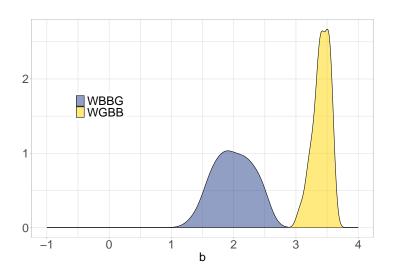
 θ



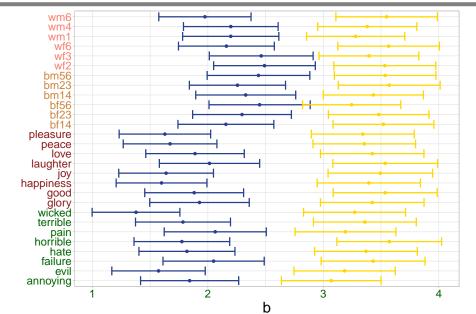


Rasch-like estimates

b_{WGBB} and b_{WGBB}

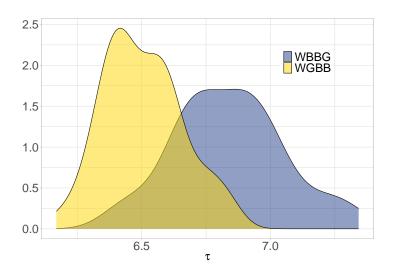


Rasch-like estimates

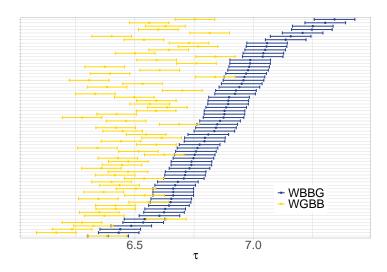


Log-normal estimates

$\tau_{\rm WGBB}$ and $\tau_{\rm BGWB}$



Log-normal estimates



Log-normal estimates

δ

