When randomness opens new possibilities: Acknowledging the stimulus sampling variability in Experimental Psychology

Ottavia M. Epifania[1,2,3], Pasquale Anselmi[1], Egidio Robusto[1]

[1] University of Padova (IT)

[2] Psicostat Group, Univeristy of Padova (IT)

[3] Catholic Univerisity of the Sacred Heart, Milan (IT)

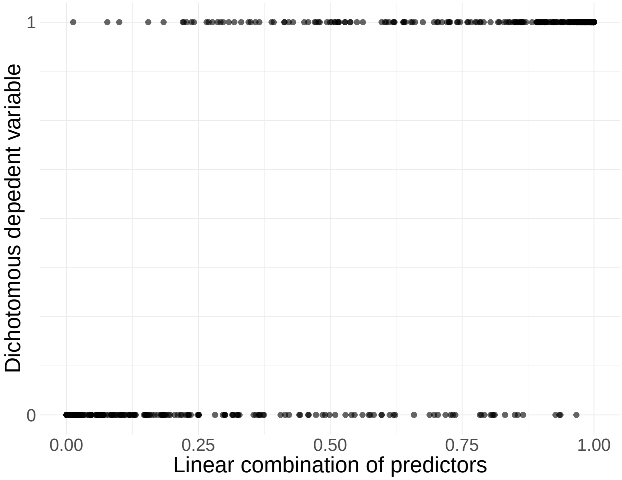
Introduction

Stimuli are fixed, respondents are random

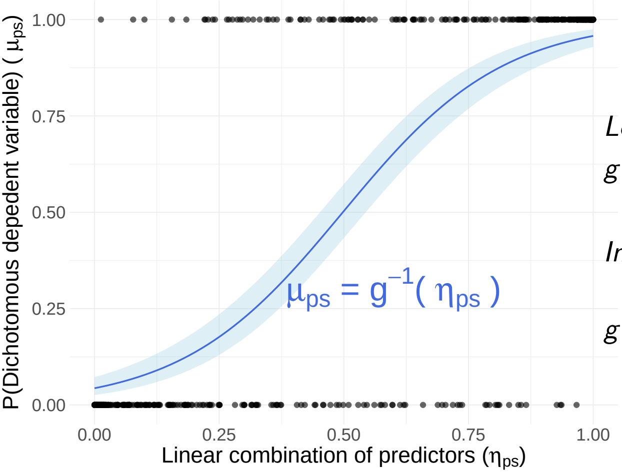
Introduction

What if

Generalized linear model for dichotomous responses



Generalized linear model for dichotomous responses



Logit link function g

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

Inverse g^{-1}

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

Random effects and random factors

Linear component in a (G)LM:

$$\eta = \beta X, \tag{1}$$

where β indicates the coefficients of the fixed intercept and slope(s), and X is the model-matrix.
Linear components in a (Generalized) Linear Mixed-Effects Model (GLMM):

$$\eta = \beta X \, Zd, \tag{2}$$

where Z is the matrix and d is the vector of the random effects (not parameters!)

Random effects and random factors

Linear component in a (G)LM:

$$\eta = \beta X, \tag{1}$$

where β indicates the coefficients of the fixed intercept and slope(s), and X is the model-matrix.
Linear components in a (Generalized) Linear Mixed-Effects Model (GLMM):

$$\eta = \beta X \, Zd, \tag{2}$$

where Z is the matrix and d is the vector of the random effects (not parameters!)

Best Linear Unbiased Predictors

The Rasch model

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

where:

θ_p: ability of respondent p (i.e., latent trait level of respondent p)

b_s: difficulty of stimulus s (i.e., "challenging" power of stimulus s)

The Rasch model

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

where:
θ_p: ability of respondent p (i.e., latent trait level of respondent p)
b_s: difficulty of stimulus s (i.e., "challenging" power of stimulus s)

Standard

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

GLM

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

The Rasch model

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

where:

θ_p: ability of respondent p (i.e., latent trait level of respondent p)

b_s: difficulty of stimulus s (i.e., "challenging" power of stimulus s)

Standard

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

GLM

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

Random stimuli in Experimental Psychology

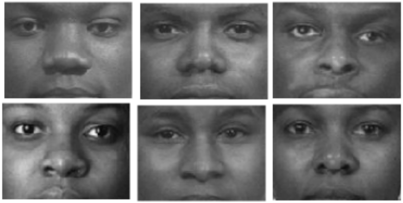
Experiment

The stimuli

12 Object stimuli

White people faces

Black people faces



16 Attribute stimuli

Positive attributes

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

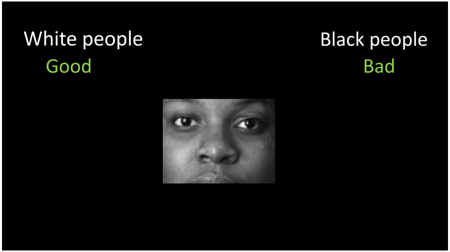
Negative attributes

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

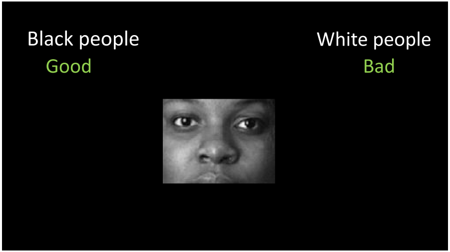
The task

Two experimental conditions

White-Good/Black-Bad (WGBB): 60 trials



Black-Good/White-Bad (BGWB): 60 trials



Random stimuli in Experimental Psychology

Models

The expected response y for the observation $i = 1, \ldots, I$ for respondent $p = 1, \ldots, P$ on stimulus $s = 1, \ldots, S$ in condition $c = 1, \ldots, C$:

Model 1:

$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{p[i]} + \alpha_{s[i]} + \varepsilon_i)$$
$$\alpha_p \sim \mathcal{N}(0, \sigma_p^2),$$
$$\alpha_s \sim \mathcal{N}(0, \sigma_s^2).$$

Model 2:

$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{p[i]} + \beta_{s[i]} c_i + \varepsilon_i)$$
$$\alpha_p \sim \mathcal{N}(0, \sigma_p^2),$$
$$\beta_s \sim \mathcal{MVN}(0, \Sigma_{sc}).$$

Model 3:

$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{s[i]} + \beta_{p[i]} c_i + \varepsilon_i)$$
$$\alpha_s \sim \mathcal{N}(0, \sigma_s^2),$$
$$\beta_p \sim \mathcal{MVN}(0, \Sigma_{pc}).$$

Accuracy: $\epsilon \sim Logistic(0, \sigma^2)$

Models

The expected response y for the observation $i = 1, \ldots, I$ for respondent $p = 1, \ldots, P$ on stimulus $s = 1, \ldots, S$ in condition $c = 1, \ldots, C$:

Model 1:

$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{p[i]} + \alpha_{s[i]} + \varepsilon_i)$$
$$\alpha_p \sim \mathcal{N}(0, \sigma_p^2),$$
$$\alpha_s \sim \mathcal{N}(0, \sigma_s^2).$$

Model 2:

$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{p[i]} + \beta_{s[i]} c_i + \varepsilon_i)$$
$$\alpha_p \sim \mathcal{N}(0, \sigma_p^2),$$
$$\beta_s \sim \mathcal{MVN}(0, \Sigma_{sc}).$$

Model 3:

$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{s[i]} + \beta_{p[i]} c_i + \varepsilon_i)$$
$$\alpha_s \sim \mathcal{N}(0, \sigma_s^2),$$
$$\beta_p \sim \mathcal{MVN}(0, \Sigma_{pc}).$$

Accuracy: $\epsilon \sim Logistic(0, \sigma^2)$

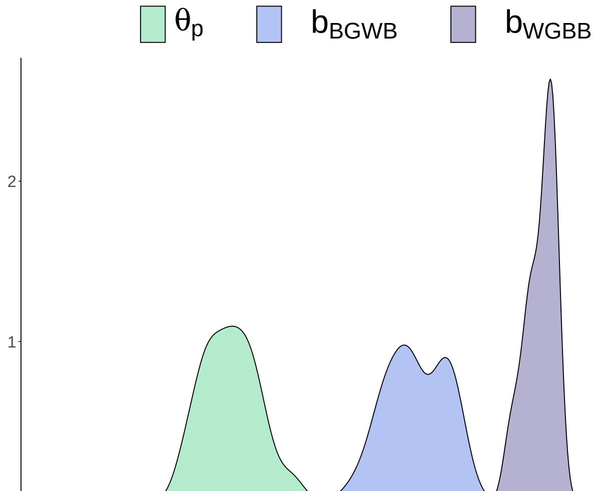
Fixed Effects Random structure

Random stimuli in Experimental Psychology

Results

Model 2 is the least wrong model

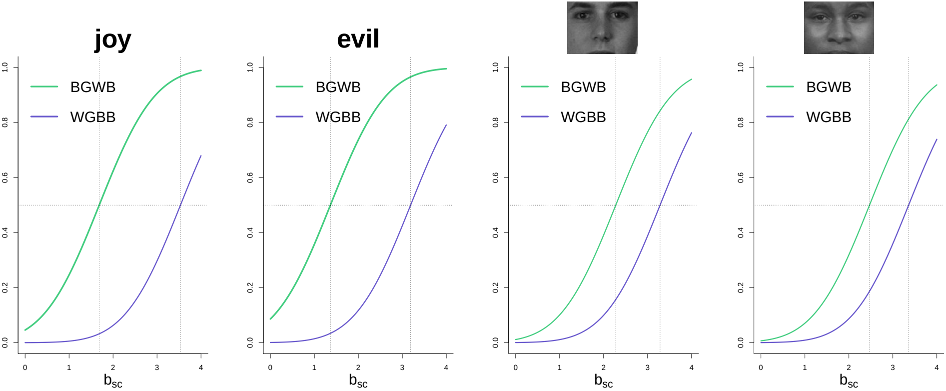
Rasch model:
Model 2



Condition–specific easiness

Highly contributing stimuli

Lowly contributing stimuli



- Improve generalizability of the results to other sets of stimuli
- Control for random variance in the data
- Allow for obtaining a Rasch-like parametrization of the data
- Possibility of extending the (linear) model to other dependent variables (e.g., response times)