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] Università Cattolica del Sacro Cuore, Milan (IT)

Respondents are random

Sampled from a lager population

Respondents are random

Sampled from a lager population

Need for acknowledging the sampling variability

Respondents are random

Sampled from a lager population

Need for acknowledging the sampling variability

Results can be generalized to other respondents belonging to the same population

Respondents are random

Sampled from a lager population

Need for acknowledging the sampling variability

Results can be generalized to other respondents belonging to the same population

Stimuli/item are fixed

Taken to be entire population

Respondents are random

Sampled from a lager population

Need for acknowledging the sampling variability

Results can be generalized to other respondents belonging to the same population

Stimuli/item are fixed

Taken to be entire population

There is no sampling variability

Respondents are random

Sampled from a lager population

Need for acknowledging the sampling variability

Results can be generalized to other respondents belonging to the same population

Stimuli/item are fixed

Taken to be entire population

There is no sampling variability

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

However...

The stimuli can also represent a sample of a larger universe

However...

The stimuli can also represent a sample of a larger universe

Processing speed of positive and negative attributes

However...

The stimuli can also represent a sample of a larger universe

Processing speed of positive and negative attributes

There is a universe of **positive attributes** as well as an universe of **negative attributes**

However...

The stimuli can also represent a sample of a larger universe

Processing speed of positive and negative attributes

There is a universe of **positive attributes** as well as an universe of **negative attributes**

Only samples of **positive attributes** (e.g., good, nice, ...) and **negative attributes** (e.g., bad, evil, ...) are administered

However...

The stimuli can also represent a sample of a larger universe

Processing speed of positive and negative attributes

There is a universe of **positive attributes** as well as an universe of **negative attributes**

Only samples of **positive attributes** (e.g., good, nice, ...) and **negative attributes** (e.g., bad, evil, ...) are administered

So... **there must be a sampling variability!**

Generalizability

Generalizability is bounded to the specific set of stimuli used in the experiment
Results can be generalized if and only if the exact same set of stimuli is used

Generalizability

Generalizability is bounded to the specific set of stimuli used in the experiment
Results can be generalized if and only if the exact same set of stimuli is used

Robustness of the results

Random variability at the stimulus level might inflate the probability of committing Type I errors
Averaging across stimuli to obtain person-level scores results in biased estimates due to the noise in the data

Generalizability

Generalizability is bounded to the specific set of stimuli used in the experiment
Results can be generalized if and only if the exact same set of stimuli is used

Robustness of the results

Random variability at the stimulus level might inflate the probability of committing Type I errors
Averaging across stimuli to obtain person-level scores results in biased estimates due to the noise in the data

Loss of information

Every stimulus is assumed to be equally informative
All the variability is not considered as well as all the information that can be obtained from it

This contribution

Focus on the loss of information...the other side of the coin

This contribution

Focus on the loss of information...the other side of the coin

The information at the stimulus level that can be retrieved from the accuracy responses (correct vs. incorrect) from a typical experiment where the response times are usually employed for scoring the data

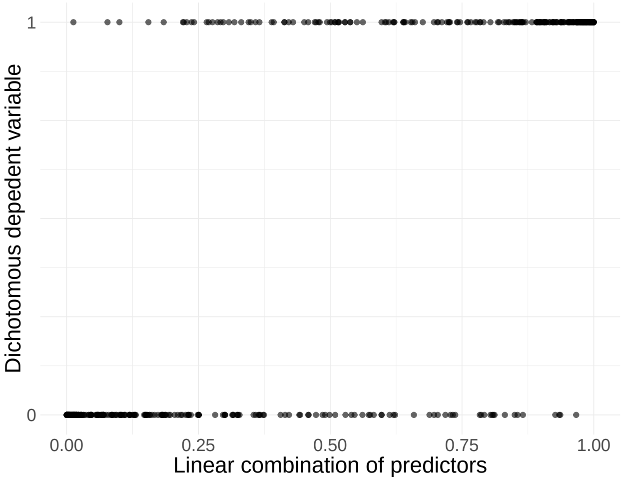
This contribution

Focus on the loss of information...the other side of the coin

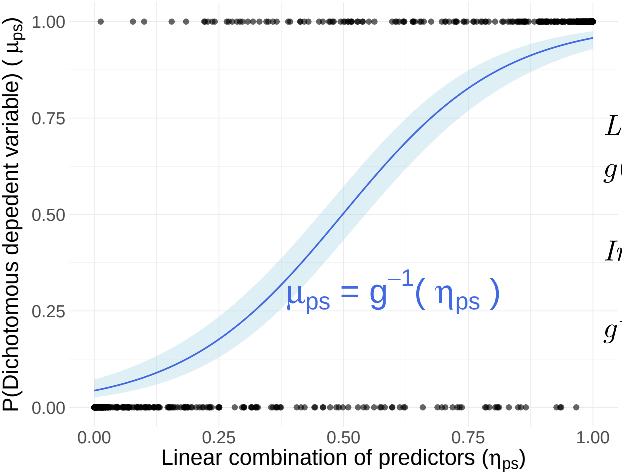
The information at the stimulus level that can be retrieved from the accuracy responses (correct vs. incorrect) from a typical experiment where the response times are usually employed for scoring the data

It can actually help in disentangling what is known to be a shortcoming of the score usually employed for analyzing the data of this experiment

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

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

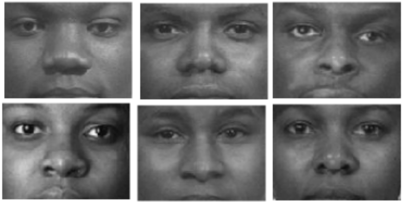
GLM

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

12 Object stimuli

White people faces

Black people faces



16 Attribute stimuli

Positive attributes

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

Negative attributes

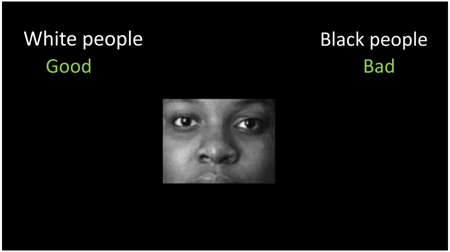
Evil, bad, horrible, terrible, nasty, pain, failure, hate

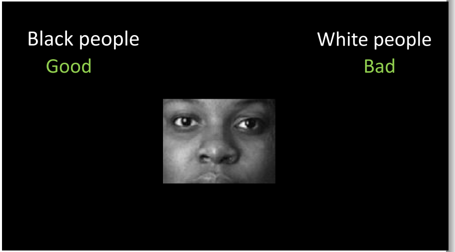
Participants: 62 (F $= 48.39\%$, Age $= 24.92 \pm 2.11$ years)

Two experimental conditions

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

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





$$D = \frac{M_{\text{BGWB}} - M_{\text{WGBB}}}{s_{\text{BGWB, WGBB}}}$$

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

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:

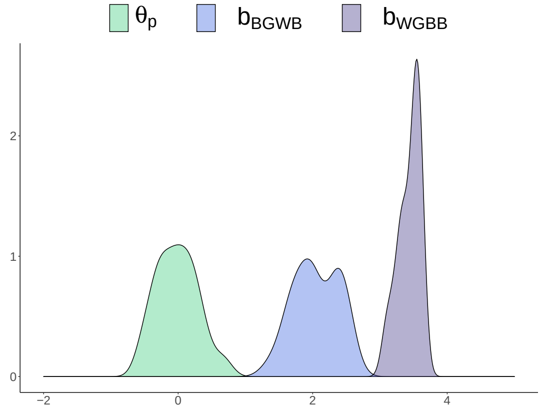
$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{s[i]} + \beta_{p[i]} c_i + \varepsilon_i)$$
$$\alpha_s \sim \mathcal{N}(0, \sigma_s^2),$$
$$\beta_p \sim \mathcal{MVN}(0, \Sigma_{pc}).$$

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

Fixed Effects Random structure

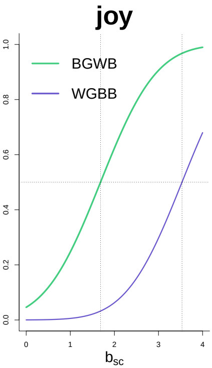
Model 2 is the least wrong model

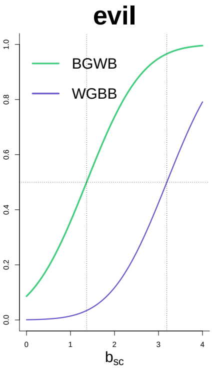
$$y_i = logit^{-1}(\alpha + \beta_c X_c + \alpha_{p[i]} + \beta_{s[i]} c_i + \varepsilon_i)$$



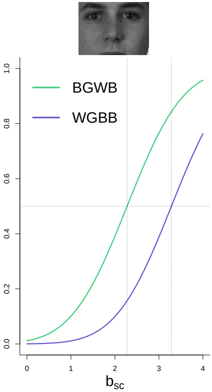
Condition–specific easiness

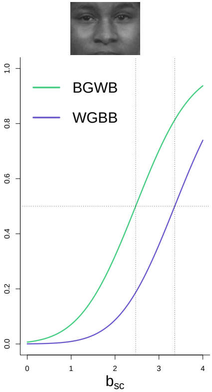
Highly contributing stimuli





Lowly contributing stimuli





- Acknowledge and gather the information at the stimulus level
- 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)

Thank you!

ottavia.epifania@unipd.it

Slides: https://ottaviae.github.io/presentations/Berlino2023/CMS-epifania.pdf