An investigation of the context effect in the multidimensional forced-choice personality measurement

Kyosuke Bunji

Kobe University

Kensuke Okada

University of Tokyo

✉ bunji@bear.kobe-u.ac.jp

Outline

1

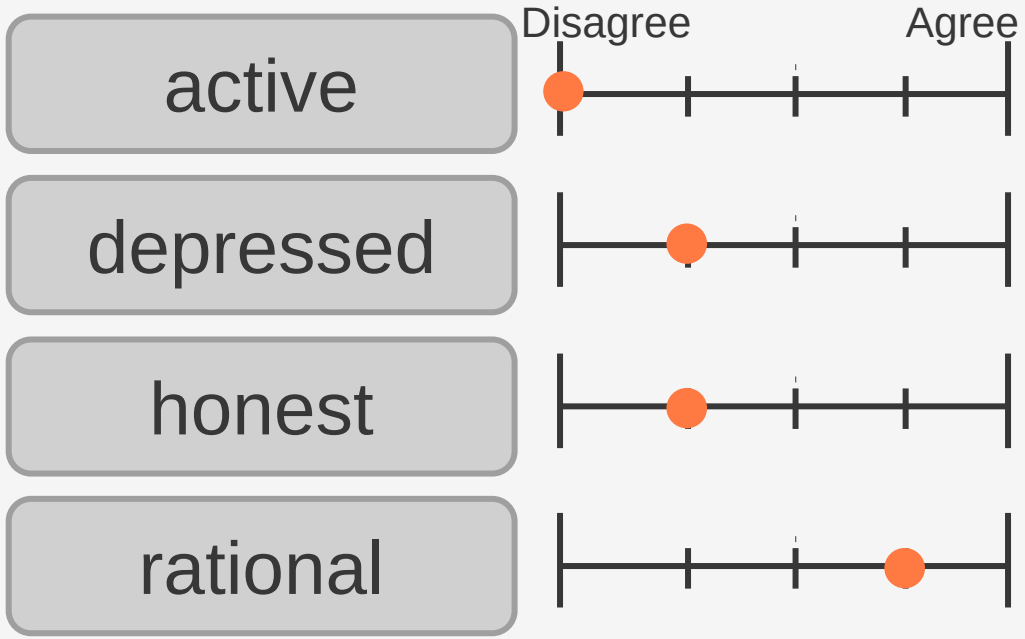
Introduction

Forced-Choice (FC) scale

Respondents are required to **rank the choice options**

Single-Stimulus (SS); Likert Scale

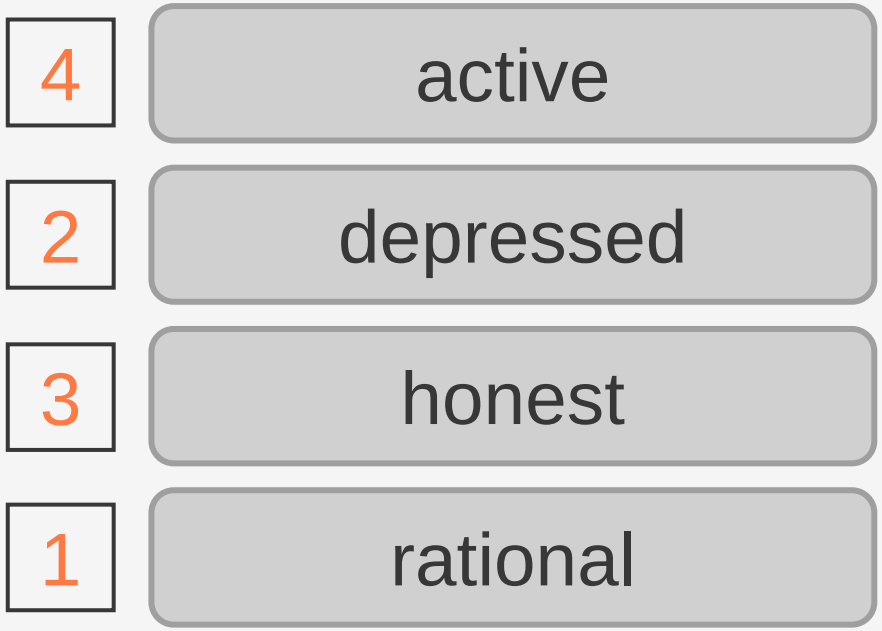
Q. To what extent do you agree with the following words about yourself?



... is frequently contaminated by systematic response biases

(4-Alternative) Forced-Choice

Q. Order the following words in the sequence that best describes you.



... is designed to reduce systematic response biases

Thurstonian IRT model (TIRT; Brown et al., 2011)

One of the most common models for the FC scale.

- Consider a pair of statements (j, k) that reflect different factors (a, b)

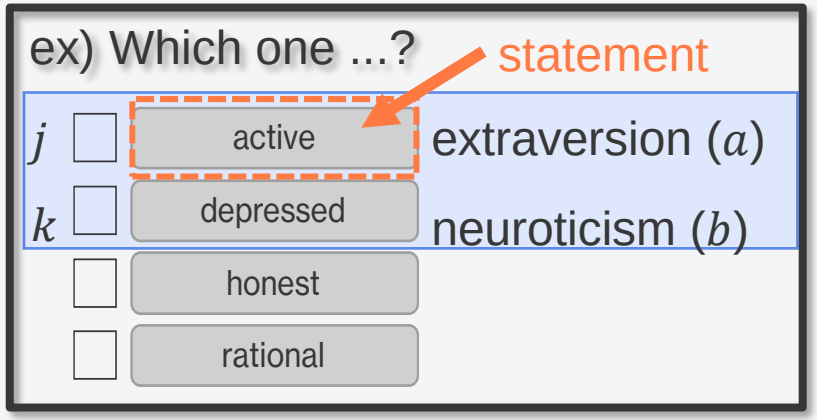
$$x_{jk} = u_j - u_k \qquad x_{jk} = 1 \;\text{ if }\; u_j > u_k$$

- The latent preference for one statement j is given as:

$$u_j = \mu_j + \beta_j \eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi_j^2\right)$$

- The probability $P\left(x_{jk} = 1 \middle| \boldsymbol{\eta}\right)$ is:

$$P\left(x_{jk} = 1 \middle| \boldsymbol{\eta}\right) = \Phi\left[\frac{\left(\mu_j - \mu_k\right) + \beta_j \eta_a - \beta_k \eta_b}{\sqrt{\Psi_j^2 + \Psi_k^2}}\right]$$



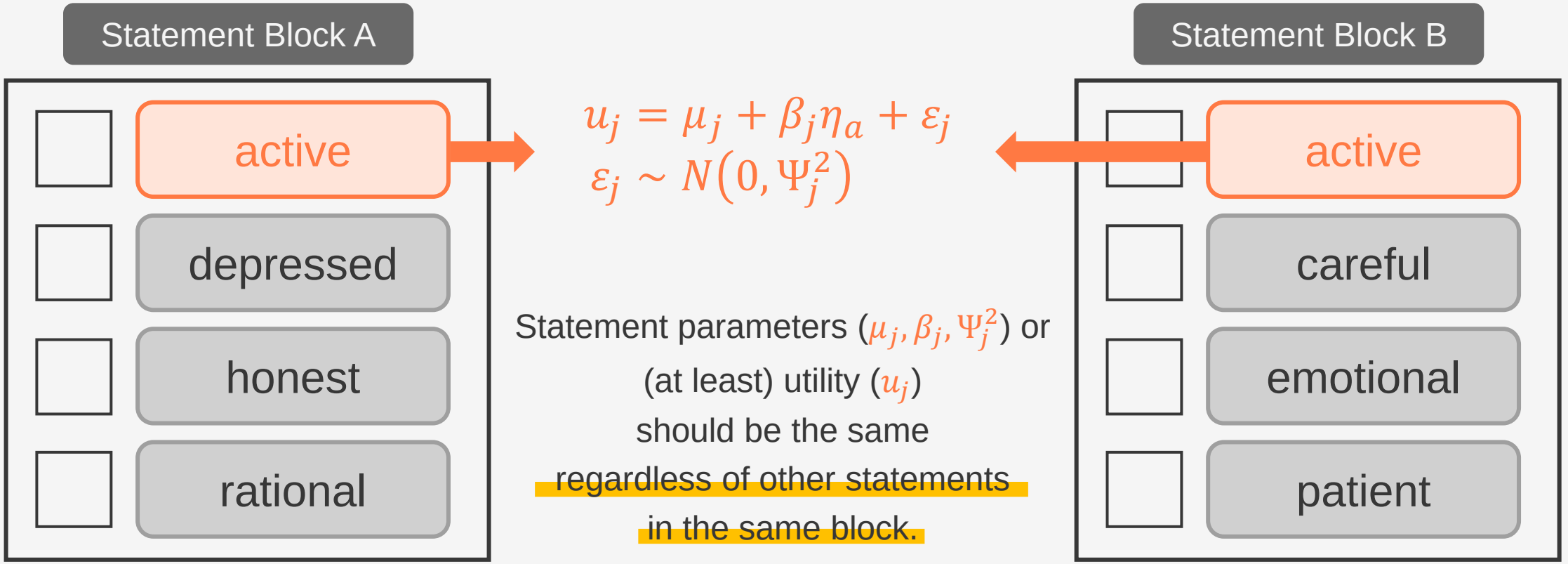
μ: mean utility of the statement

β : factor loading

η : factor score (trait)

Important assumption

- The utility of each choice option (u_j) should be invariant.

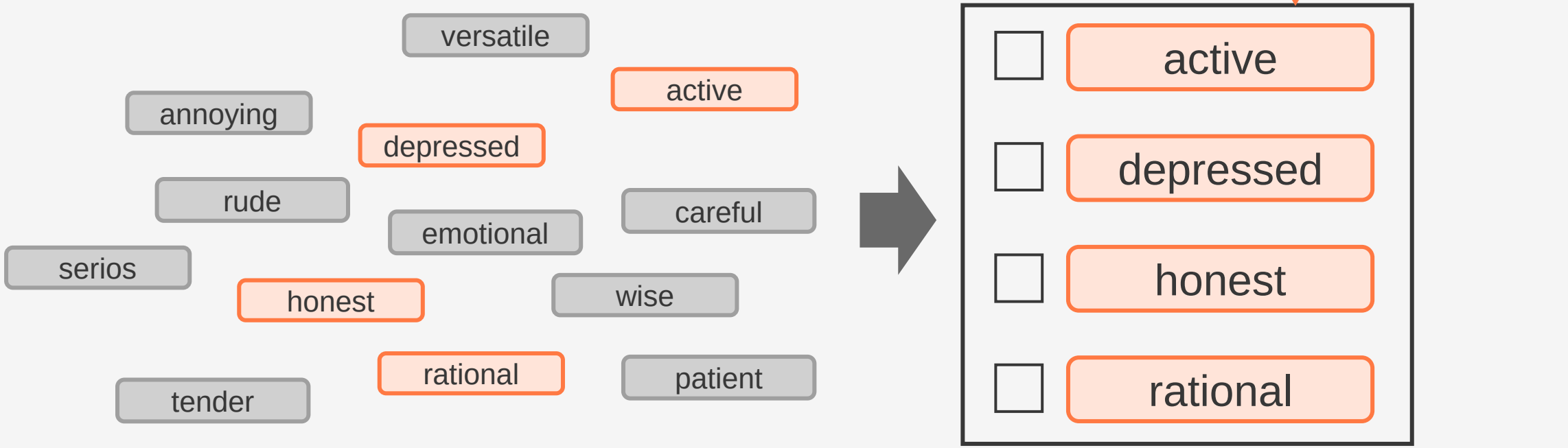


We want to assume there is **no** context effect.

Objective of the study

We examined the existence of the context effect.

- If context effect does not exist…
  - We can reduce the number of parameters by using the same statement.
    → stabilize the result and enable efficient calculation
  - We can make adaptive measurements in the most adaptive way.

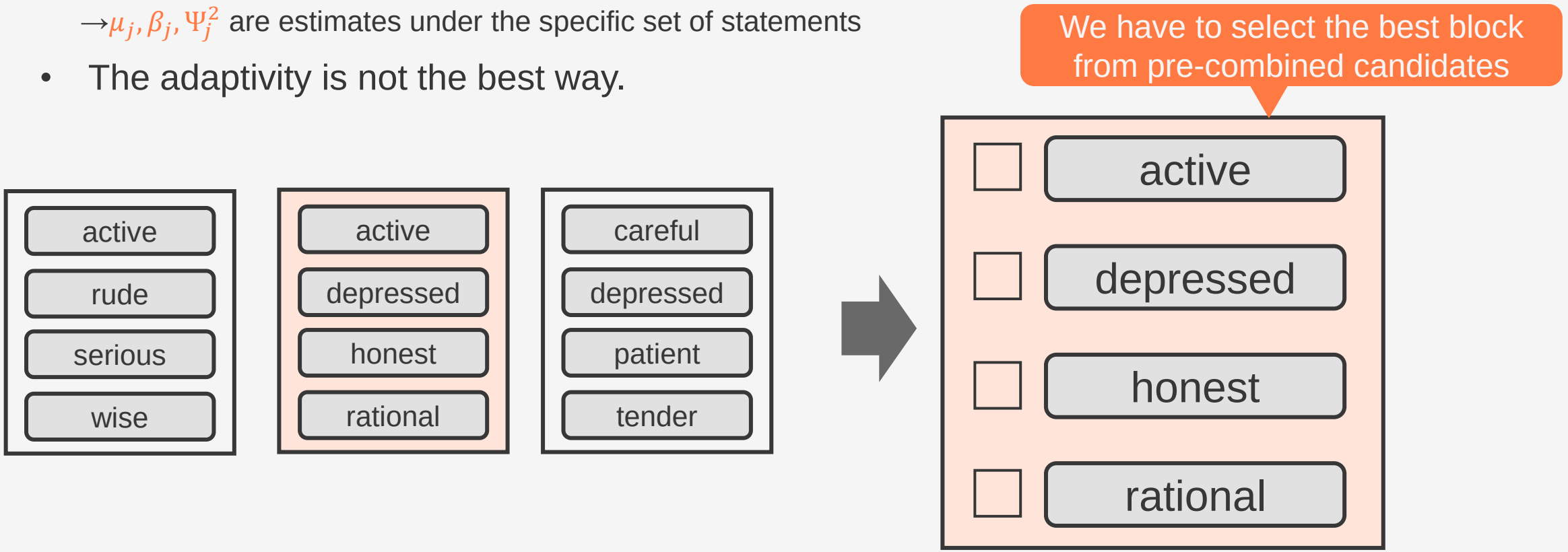


Objective of the study

We examined the existence of the context effect.

- If context effect cannot be ignored…
  - The interpretation of statement parameters becomes difficult.

    →μ_j, β_j, Ψ_j^2 are estimates under the specific set of statements
  - The adaptivity is not the best way.



2

Methods

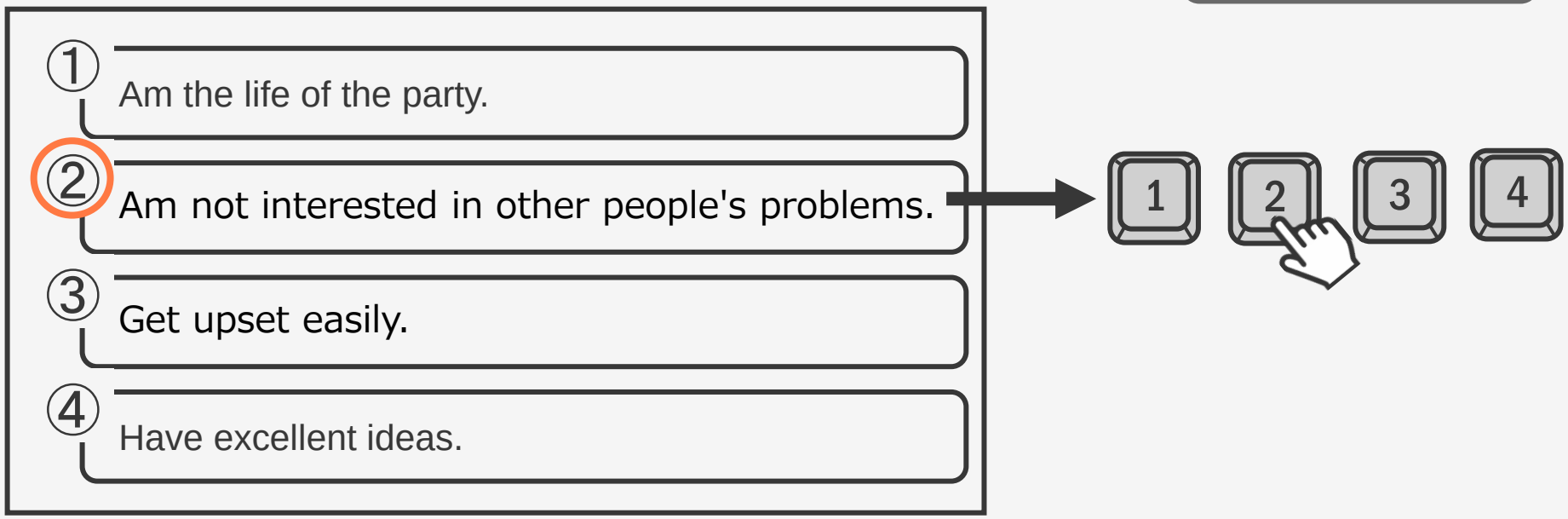
Data

Participants 484 Japanese (184 Males; 296 Females; 4 Unanswered)
→ Collected via crowdsourcing platform → answered online

Statements Mini-IPIP Scale (Donnellan, 2006) Big-Five factors; 20 sentences
BIDR-J (Tani, 2008) Social desirability; 4 statements

6 factors,
24 statements

(example)



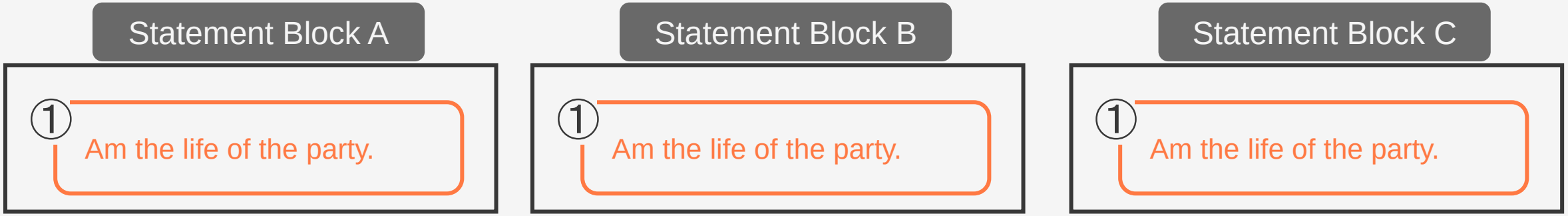
Specific manipulation in the Data

Each of the 24 statements is **placed in three blocks.**

Statement Block A	Statement Block B	Statement Block C
① Am the life of the party.	① Am the life of the party.	① Am the life of the party.
② Am not interested in other people's problems.	② Feel others' emotions.	② Am not really interested in others.
③ Get upset easily.	③ Have frequent mood swings.	③ Get chores done right away.
④ Have excellent ideas.	④ Never swear.	④ Do not have a good imagination.

1st strategy to investigate the Research Question

We compared the results obtained from the following two models.



[**Variable** model] statement parameters are different in different blocks (i.e., assumes context effect)

$$u_{j(A)} = \mu_{j(A)} + \beta_{j(A)}\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_{j(A)}\right)$$

$$u_{j(B)} = \mu_{j(B)} + \beta_{j(B)}\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_{j(B)}\right)$$

$$u_{j(C)} = \mu_{j(C)} + \beta_{j(C)}\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_{j(C)}\right)$$

[**Invariable** model] statement parameters are the same (i.e., assumes no context effect)

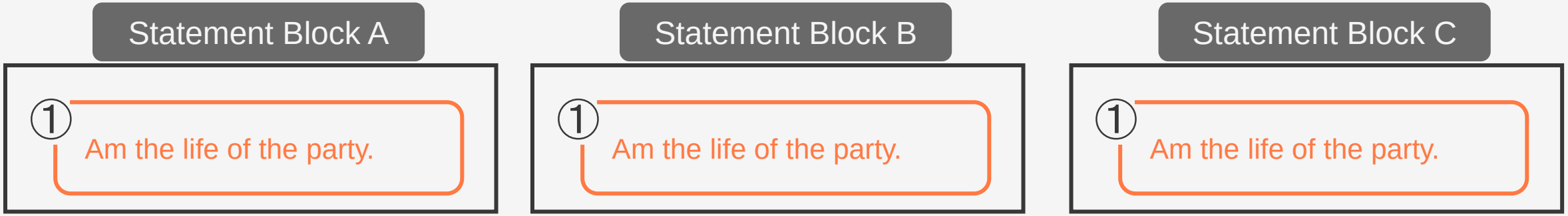
$$u_j = \mu_j + \beta_j\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_j\right)$$

$$u_j = \mu_j + \beta_j\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_j\right)$$

$$u_j = \mu_j + \beta_j\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_j\right)$$

2nd strategy to investigate the Research Question

- We directly compared the parameter estimates.



[**Variable** model] statement parameters are different in different blocks (i.e., assumes context effect)

$$u_{j(A)} = \mu_{j(A)} + \beta_{j(A)}\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_{j(A)}\right)$$

$$u_{j(B)} = \mu_{j(B)} + \beta_{j(B)}\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_{j(B)}\right)$$

$$u_{j(C)} = \mu_{j(C)} + \beta_{j(C)}\eta_a + \varepsilon_j \qquad \varepsilon_j \sim N\left(0, \Psi^2_{j(C)}\right)$$

[If context effect exists…]

$$\mu_{j(A)} \neq \mu_{j(B)} \neq \mu_{j(C)}$$

$$u_{j(A)} \neq u_{j(B)} \neq u_{j(C)}$$

[If context effect **does not** exist…]

$$\mu_{j(A)} = \mu_{j(B)} = \mu_{j(C)}$$

$$u_{j(A)} = u_{j(B)} = u_{j(C)}$$

A problem

Statement parameters are relative in the block.

Statement Block A

① Am the life of the party.

② Am not interested in other people's problems.

③ Get upset easily.

④ Have excellent ideas.

$$u_1 = \mu_1 + \beta_1\eta_a + \varepsilon_1$$
$$\varepsilon_1 \sim N(0, \Psi_1^2)$$

$$u_2 = \mu_2 + \beta_2\eta_a + \varepsilon_2$$
$$\varepsilon_2 \sim N(0, \Psi_2^2)$$

$$u_3 = \mu_3 + \beta_3\eta_a + \varepsilon_3$$
$$\varepsilon_3 \sim N(0, \Psi_3^2)$$

$$u_4 = \mu_4 + \beta_4\eta_a + \varepsilon_4$$
$$\varepsilon_4 \sim N(0, \Psi_4^2)$$

The following constraints are usually imposed in the TIRT model:

- The sum of μ_j is set to 0.
  (or directly estimate $\gamma_{jk} = \mu_j - \mu_k$)
- One of Ψ_j^2s is set to 1.

Parameter estimates obtained from different blocks **cannot be compared.**

A solution to the problem

- Parameters of only one statement are assumed to be variable.

Statement Block A

① Am the life of the party.

② Am not interested in other people's problems.

③ Get upset easily.

④ Have excellent ideas.

[**Partially Variable** model]

[Variable]

$$u_{1(A)} = \mu_{1(A)} + \beta_{1(A)}\eta_a + \varepsilon_1$$
$$\varepsilon_1 \sim N\left(0, \Psi^2_{1(A)}\right)$$

[Invariable]

$$u_2 = \mu_2 + \beta_2\eta_a + \varepsilon_2$$
$$\varepsilon_2 \sim N(0, \Psi^2_2)$$

$$u_3 = \mu_3 + \beta_3\eta_a + \varepsilon_3$$
$$\varepsilon_3 \sim N(0, \Psi^2_3)$$

$$u_4 = \mu_4 + \beta_4\eta_a + \varepsilon_4$$
$$\varepsilon_4 \sim N(0, \Psi^2_4)$$

Other statements (assumed to be invariable) act as the anchors between blocks.

Parameters of [Variable] statement **can be compared.**

Several measures

* Parameters were estimated via MCMC (cmdstanr).

We checked the following measures.

[Strategy 1: Between variable and invariable models]

1. Correlation of trait scores ($\boldsymbol{\eta}$)
2. Bayes factor
3. An information criterion (WAIC)

[Strategy 2: Directly compare the parameter estimates of the same statement]

4. 95% credible interval (HDI) of the difference of the estimates (e.g., $\mu_{j(A)} - \mu_{j(B)}$)
5. Overlap ratio of posterior distributions of utilities ($P(u_{j(\cdot)}|\boldsymbol{X})$)

3

Results

Result 1: Correlation of trait scores (η)

- The correlations between the two models are at least 0.997.

	Ext	Agr	Con	Emo	Int	Soc
Ext	**1.000**	0.390	0.146	0.299	0.433	0.234
Agr	0.383	**0.999**	0.288	0.175	0.476	0.078
Con	0.156	0.289	**0.999**	0.234	0.229	0.237
Emo	0.297	0.162	0.251	**0.999**	0.280	0.682
Int	0.422	0.463	0.224	0.274	**0.999**	0.180
Soc	0.257	0.085	0.265	0.706	0.194	**0.997**

Intertrait correlations within the variable model

Intertrait correlations within the invariable model

* The maximum absolute difference in the intertrait correlations was 0.027

Result 2, 3: Bayes factor and WAIC

- [WAIC] The variable model performs *slightly* better than the invariable model.

the difference was about 1/10 of SE

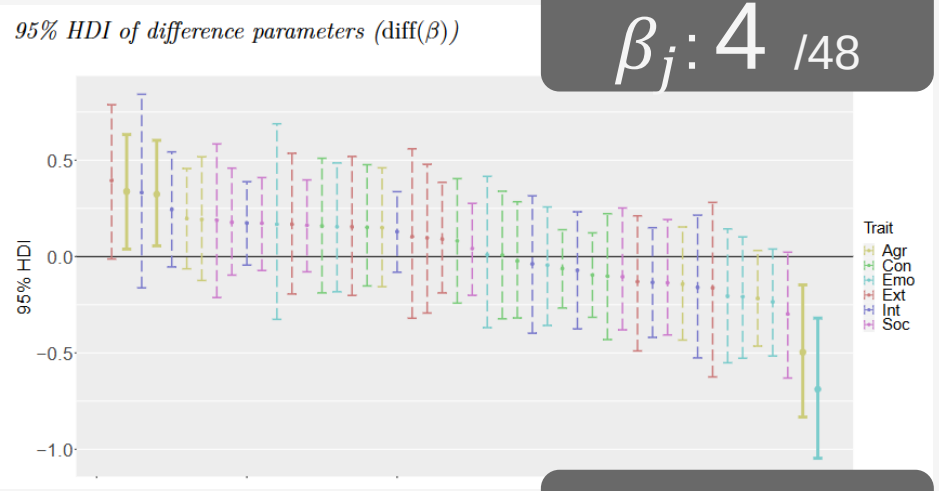
- [Bayes factor] The invariable model was strongly supported than the variable model.

An information criterion (WAIC and its standard error) and $\log_{10}\mathrm{BF}_{10}$ *on each model*

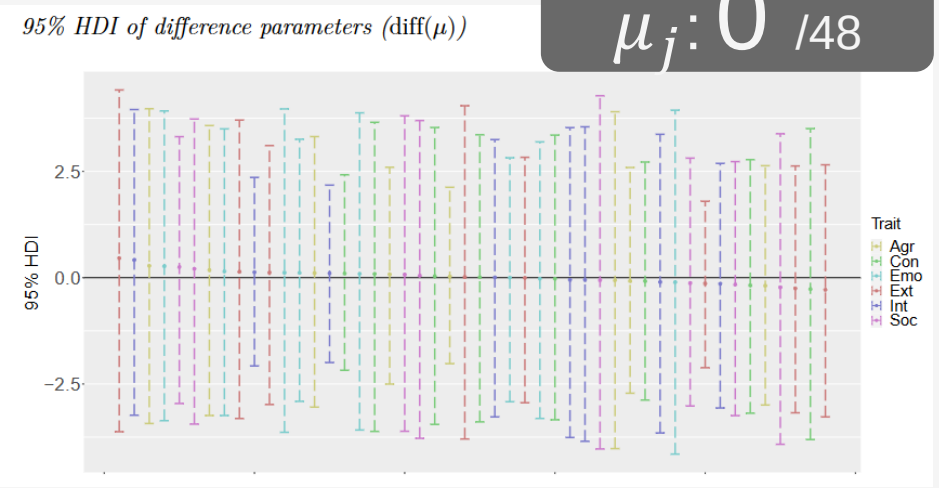
model	WAIC	SE(WAIC)	$\log_{10}\mathrm{BF}_{10}$
invariable	48420.92	523.40	
variable	48370.42	522.76	-146.30

Result 4: 95% HDI of the difference

- Most of the difference parameters contained 0 in the 95% HDI.



β_j: 4 /48

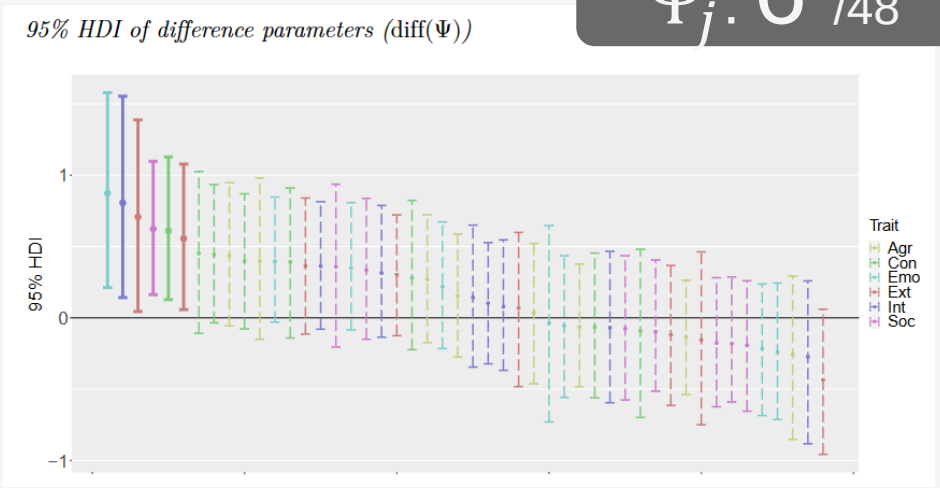


μ_j: 0 /48

* We obtained two difference parameters on each statement.

$$\mathrm{diff}(\mu_{j(B)}) = \mu_{j(B)} - \mu_{j(A)}$$
$$\mathrm{diff}(\mu_{j(C)}) = \mu_{j(C)} - \mu_{j(A)}$$

➡ $2 \times 24 = 48$ parameters



Ψ_j: 6 /48

Result 5: Overlap rate of the utility $P(u_{j(\cdot)}|\boldsymbol{Y})$

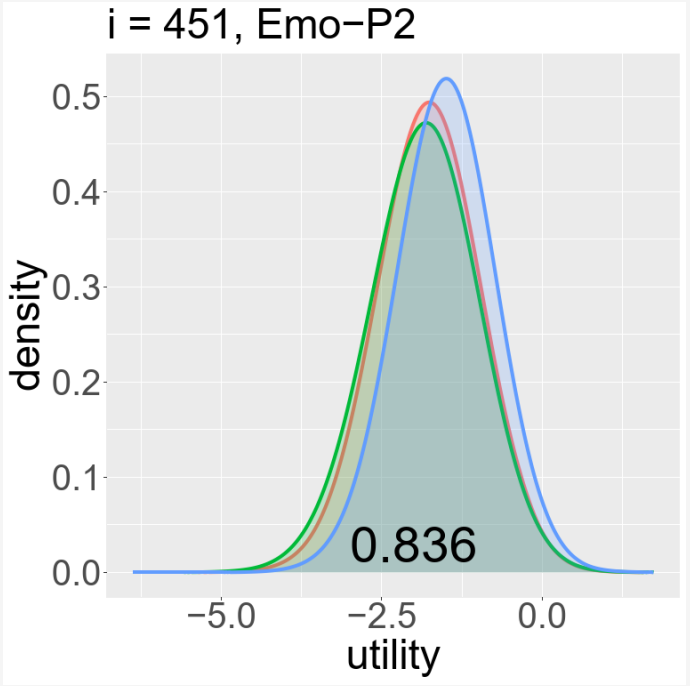
- The overall mean OR was 0.836.

  Only 0.413% showed smaller than 0.5.

OR is calculated on every combination of respondent (484) x statement (24)

→ Total number of OR was 11,616.

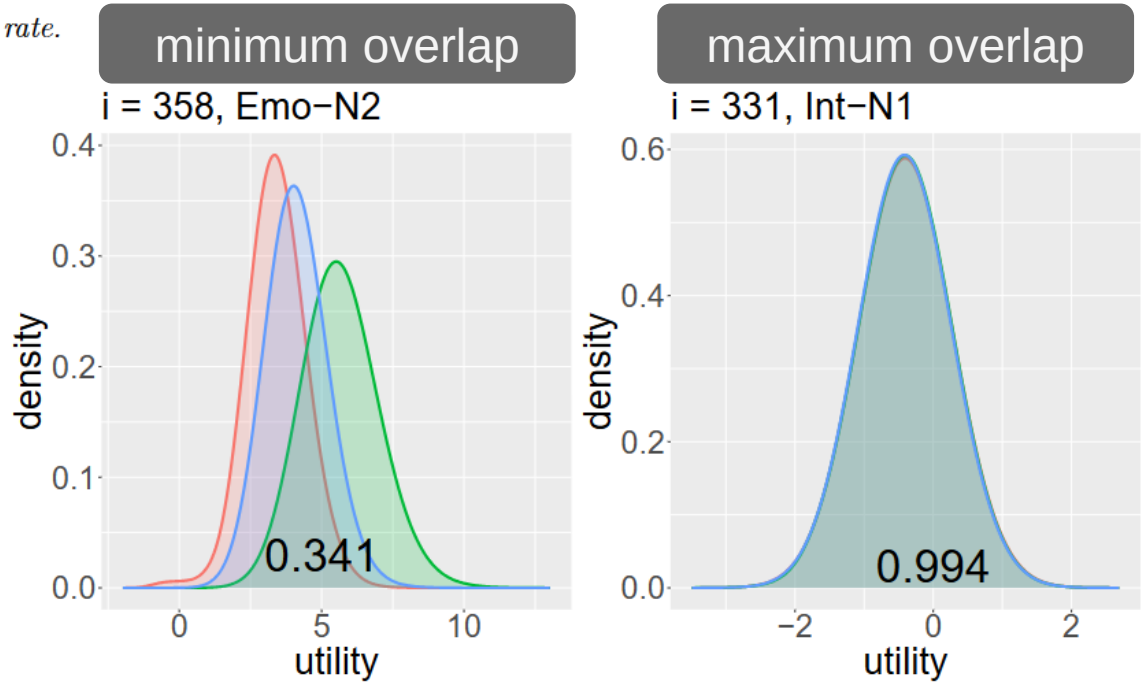
Closest to the overall mean OR



Density plot of utilities with minimum (left panel) and maximum (right panel) overlap rate.

minimum overlap

maximum overlap



4

Summary and Discussion

Results and future work

- We can conclude the context effect was **negligibly small**.

	Measure	Result at a glance
1	Trait correlation	Both models can obtain the almost same scores.
2	Bayes factor	The variable model was favored but the difference was small.
3	WAIC	The invariable model overwhelmed the variable model.
4	95% HDI of difference	There was little evidence of the existence of the effect.
5	Overlap rate of utility	The bulk of the utility distribution overlapped on average.

- It is also true there does exist context effect to some extent.

[Future work]

Examine the context effect with different traits

Examine the other effects (e.g., order, situation, response format…)

Thank you for your attention!

An investigation the context effect
in the multidimensional forced-choice
personality measurement

Kyosuke Bunji

Kobe University

Kensuke Okada

University of Tokyo

✉ bunji@bear.kobe-u.ac.jp