

# Understanding the decision-making process of discrete choice modellers

Gabriel Nova\*<sup>1</sup>, Sander van Cranenburgh<sup>1</sup>, Stephane Hess<sup>2</sup>

<sup>1</sup>Transport and Logistics group, Delft University of Technology

<sup>2</sup>Choice Modelling Centre - Institute for Transport Studies, University of Leeds

# Relevance of modelling choices

People make choices every day and across dimensions.

- Products (e.g., smartphones, food), services (e.g., healthcare, transport), lifestyle choices (e.g., university, pensions).

# Choices

- create demand, influence markets and shape industries.
- impact on economy, environment and society.

## Choice modellers use choices

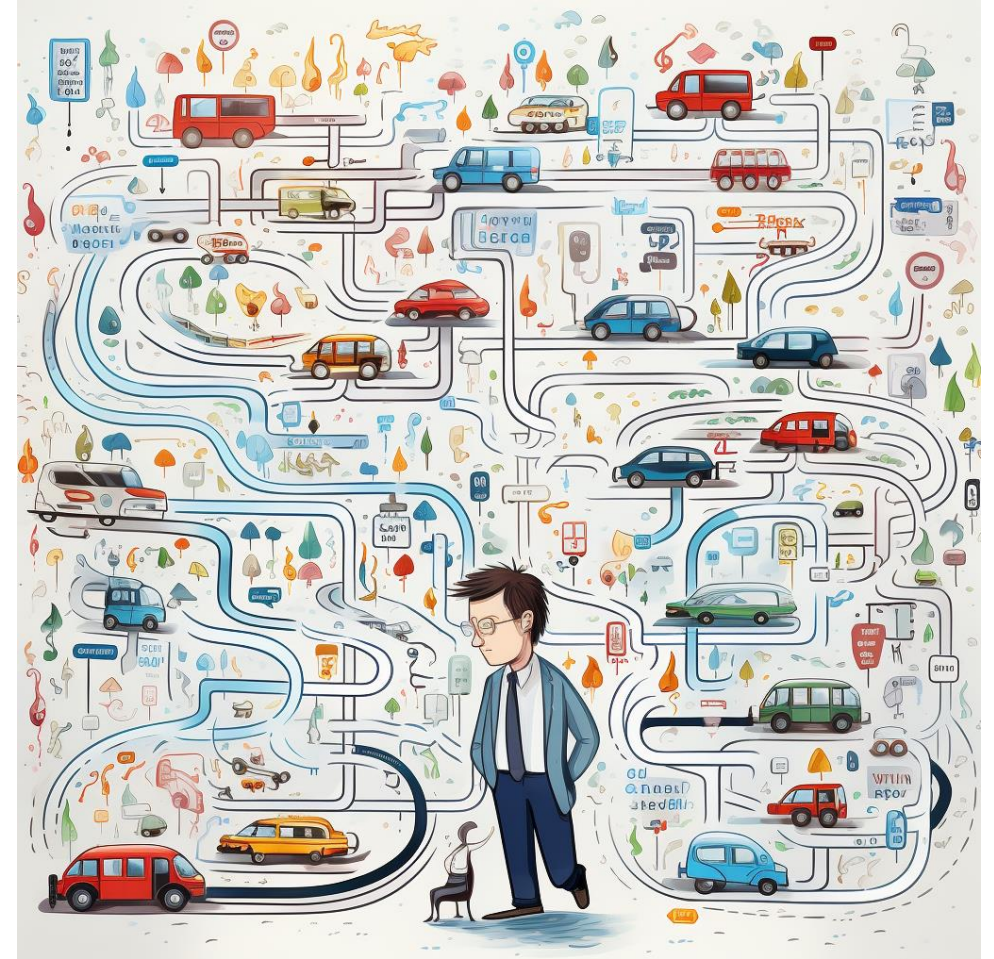
- to understand the factors that lead people to choose alternatives.
- to analyse policy and forecast demand.



# Background

Discrete choice modelling (DCM) involves mixing theoretical behavioral knowledge and statistical methods for understanding or predicting choice behaviour<sup>12</sup>.

- Researchers represent choices in a model and analyse their derived estimation results.
- DCM as an art: Multiple decisions across research phases
- Software packages have simplified several research processes
- Different workflows → variability in the modelling results



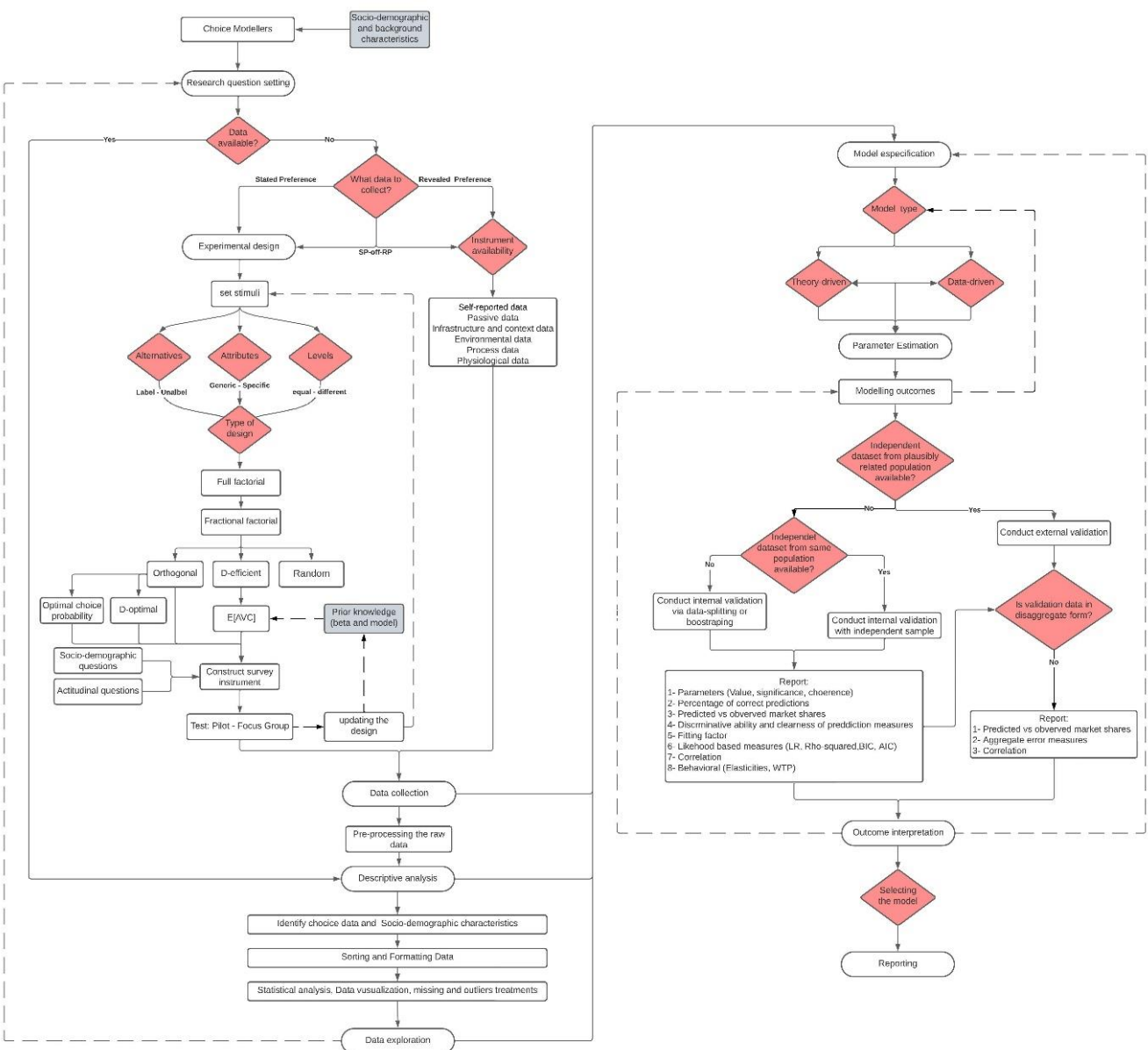
# Background

Discrete choice modelling (DCM) involves mixing theoretical behavioral knowledge and statistical methods for understanding or predicting choice behaviour<sup>12</sup>.

- Researchers represent choices in a model and analyse their derived estimation results.
- DCM as an art: Multiple decisions across research phases
- Software packages have simplified several research processes
- Different workflows → variability in the modelling results



# Choice modelling research decision tree



---

# Research gap and objective

---

## Research gap

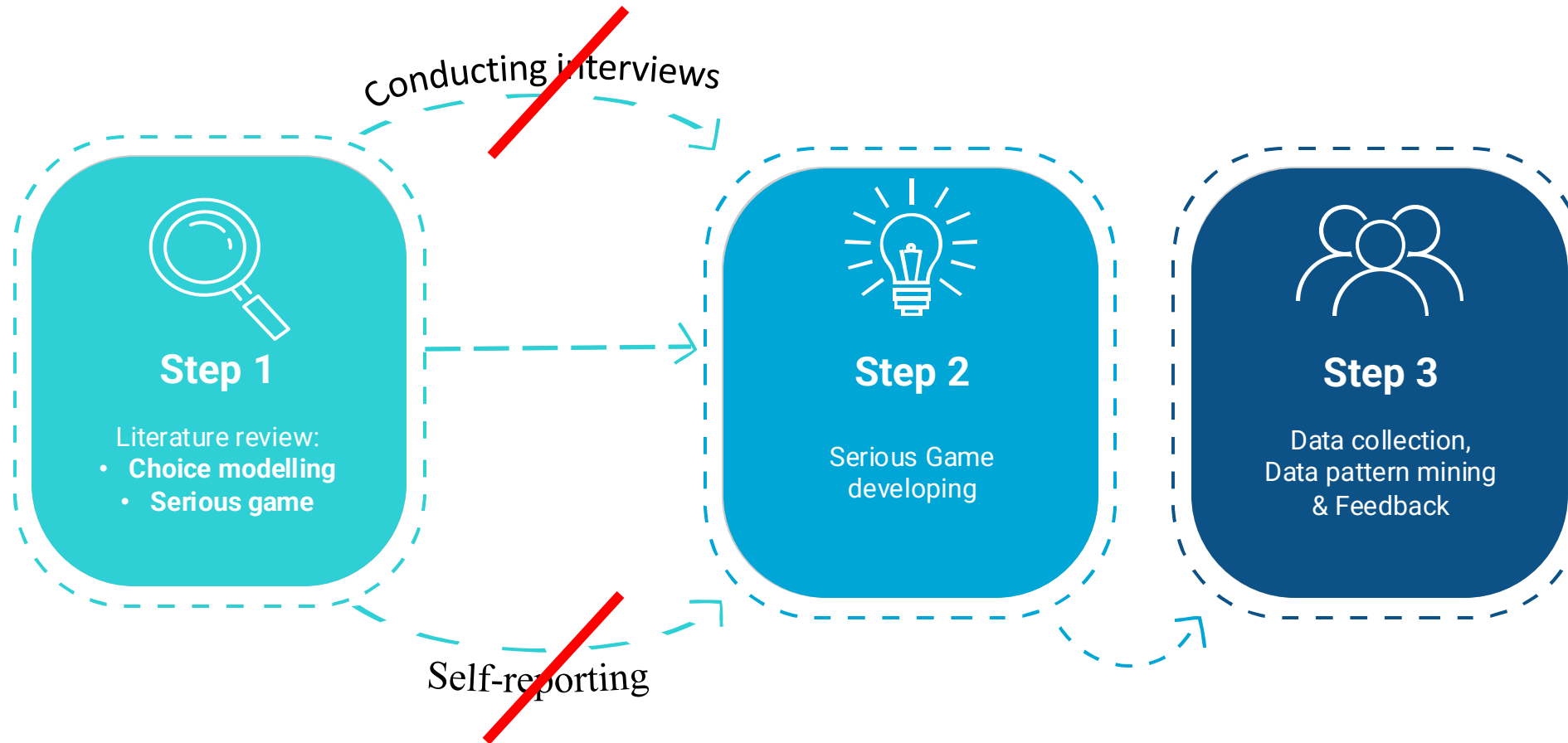
There is a lack of understanding of the decision-making process of the choice modelling process, the impact of the researcher's degrees of freedom<sup>1</sup> and the possible influence of forking paths<sup>2</sup> on the modelling process.

## Objective

Deepen understanding how the decisions taken within the choice modelling process co-determine subsequent ones and modelling results.

<sup>1</sup>Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011) <sup>2</sup>Gelman, A., & Loken, E. (2013)

# Methodology



# Serious Game

## Definition

- analogical or digital games beyond entertainment as their primary purpose
- focus on training, learning, behavioral change (... and in situ data collection)

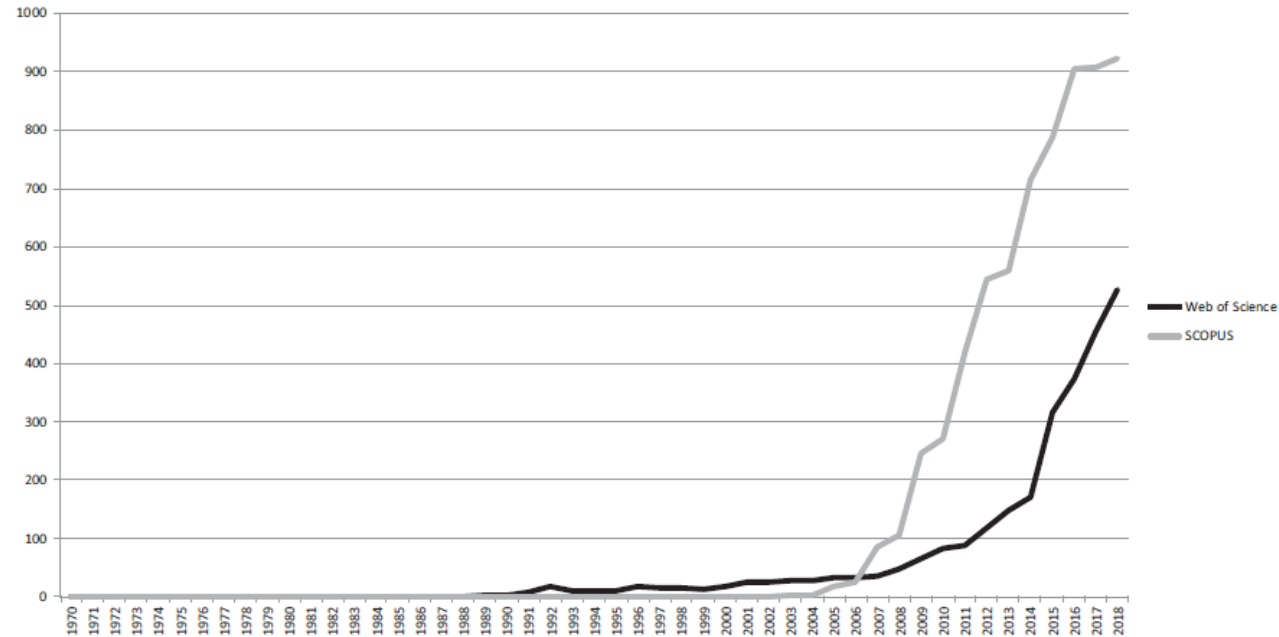


Figure1. Number of publications on serious games in Web of Science (WoS) and SCOPUS databases. (Olejniczak et al., 2020)

# Serious Game

## Definition

- analogical or digital games beyond entertainment as their primary purpose
- focus on training, learning, and behavioral change

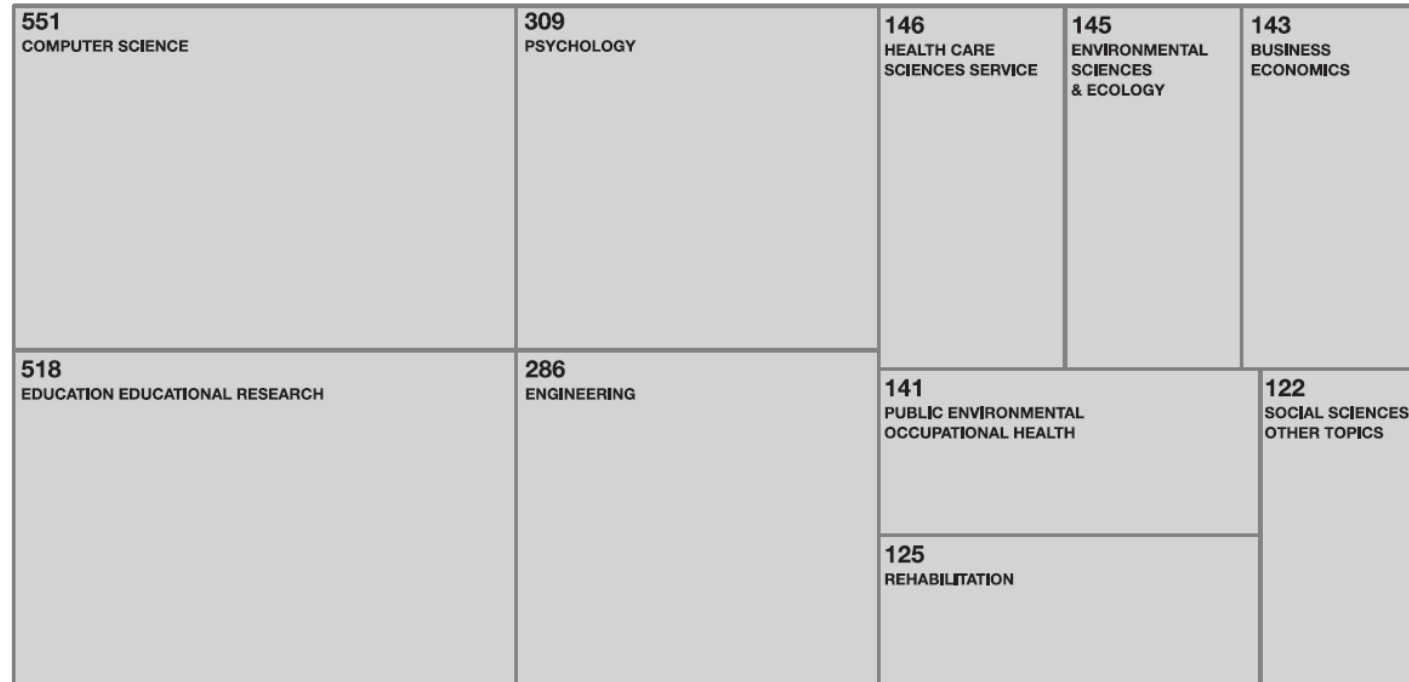


Figure1. Number of publications on serious games in Web of Science (WoS) and SCOPUS databases. (Olejniczak et al., 2020)

# Serious Game

## Data collection techniques

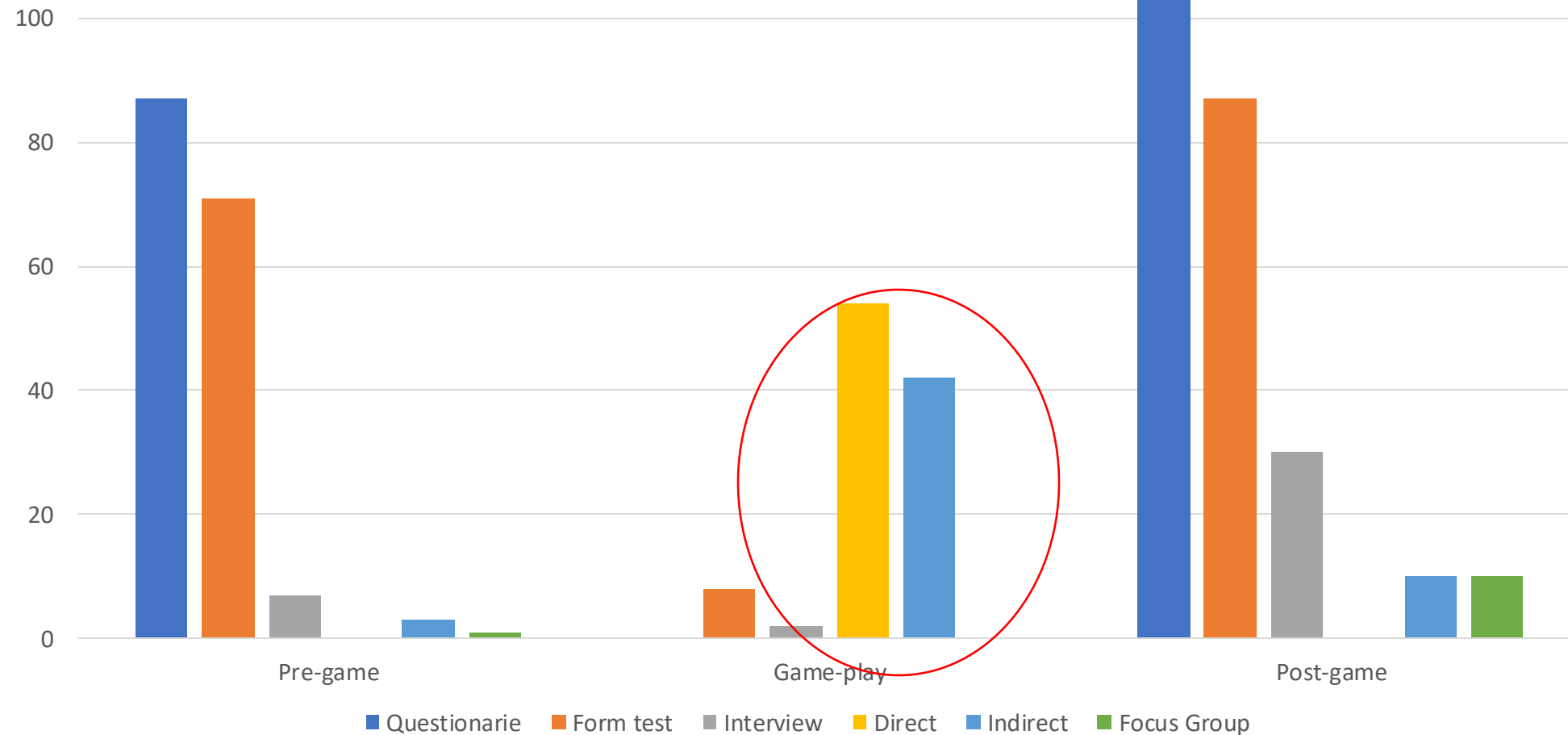


Figure 2. Number of specific data collection techniques used per phase of study. (Smith et al., 2015)

# Choice Modelling Serious Game

## Research game design

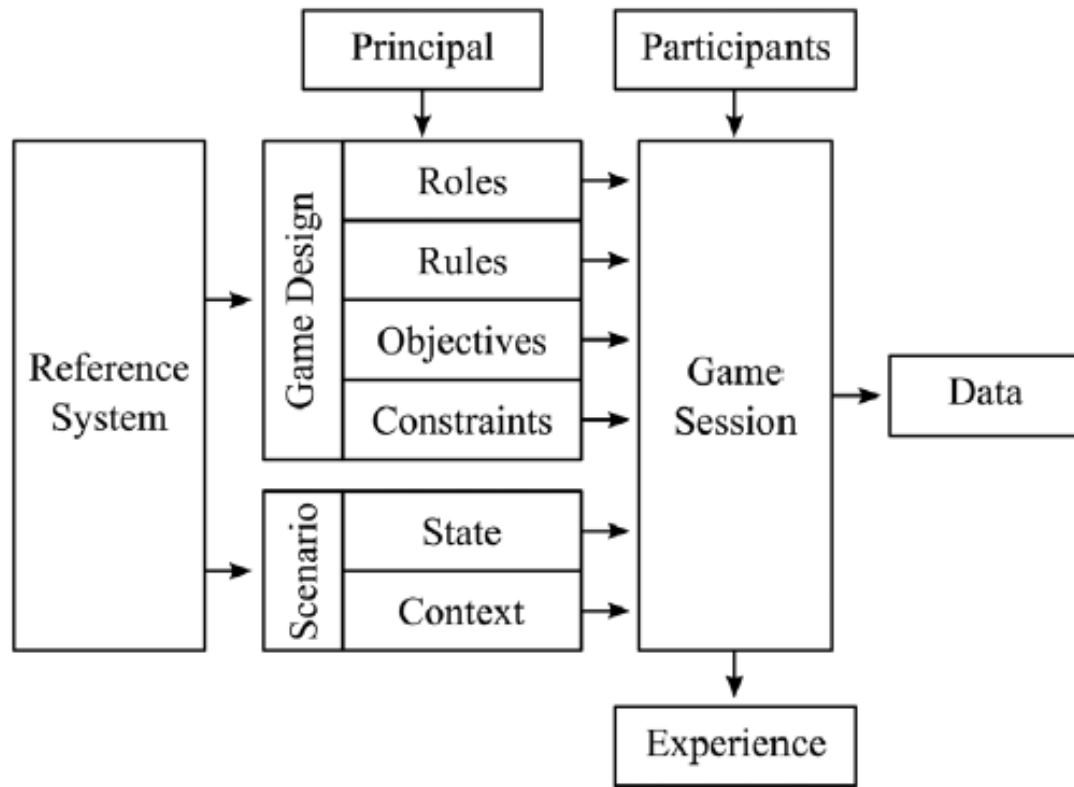


Figure 3. Element for a serious choice modelling game based on Meijer (2009)

System: Choice modelling

Roles: Play as a choice modeller

Rules: Conduct a research individually (35 min.)

Objective: Develop a choice model to infer willingness to pay for reducing noise pollution

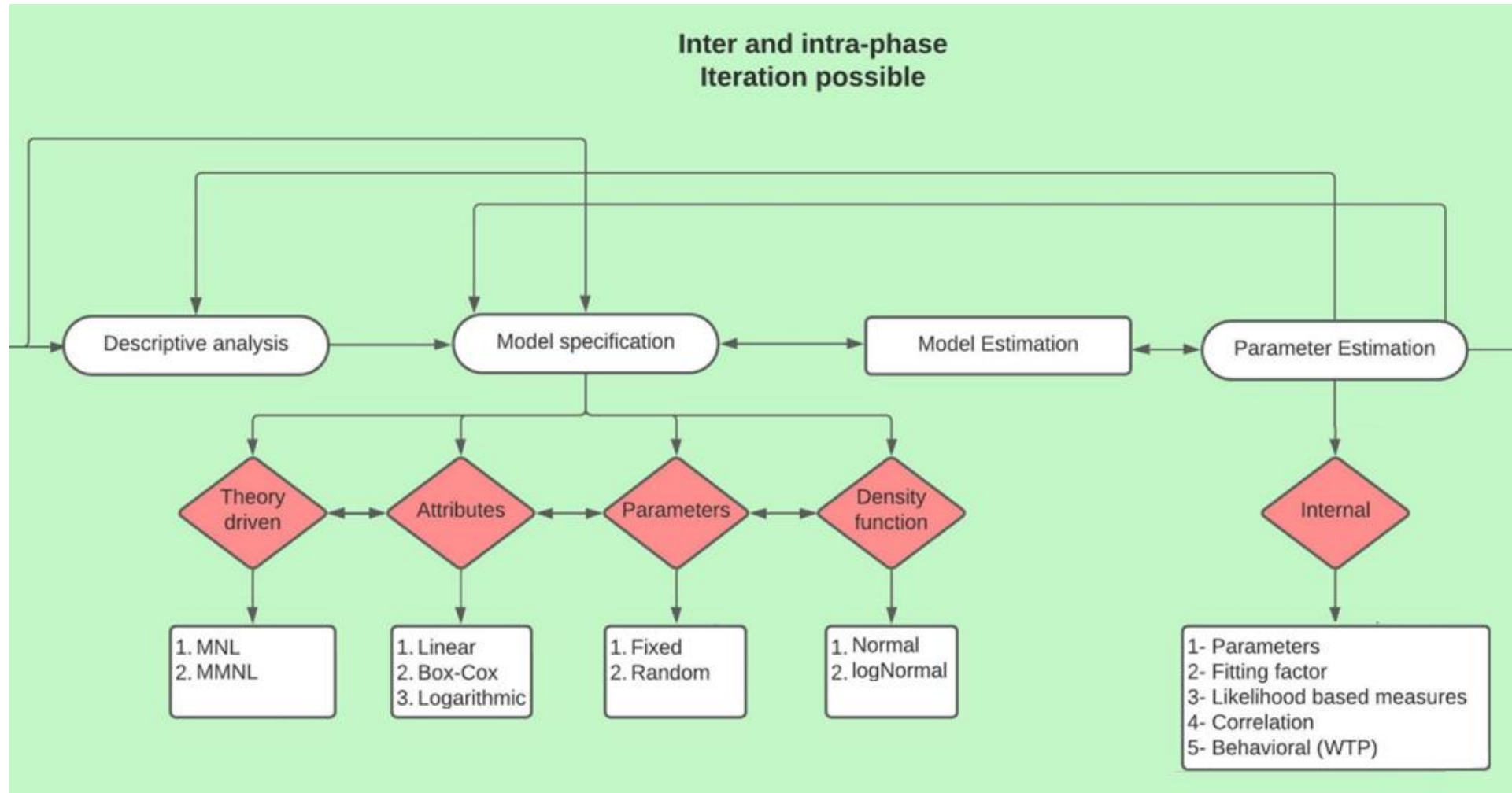
Constraints: Not all pre-estimated Mixed Logit models.

State: Descriptive analysis (16), specifications (34) and interpretation (15) variable decisions.

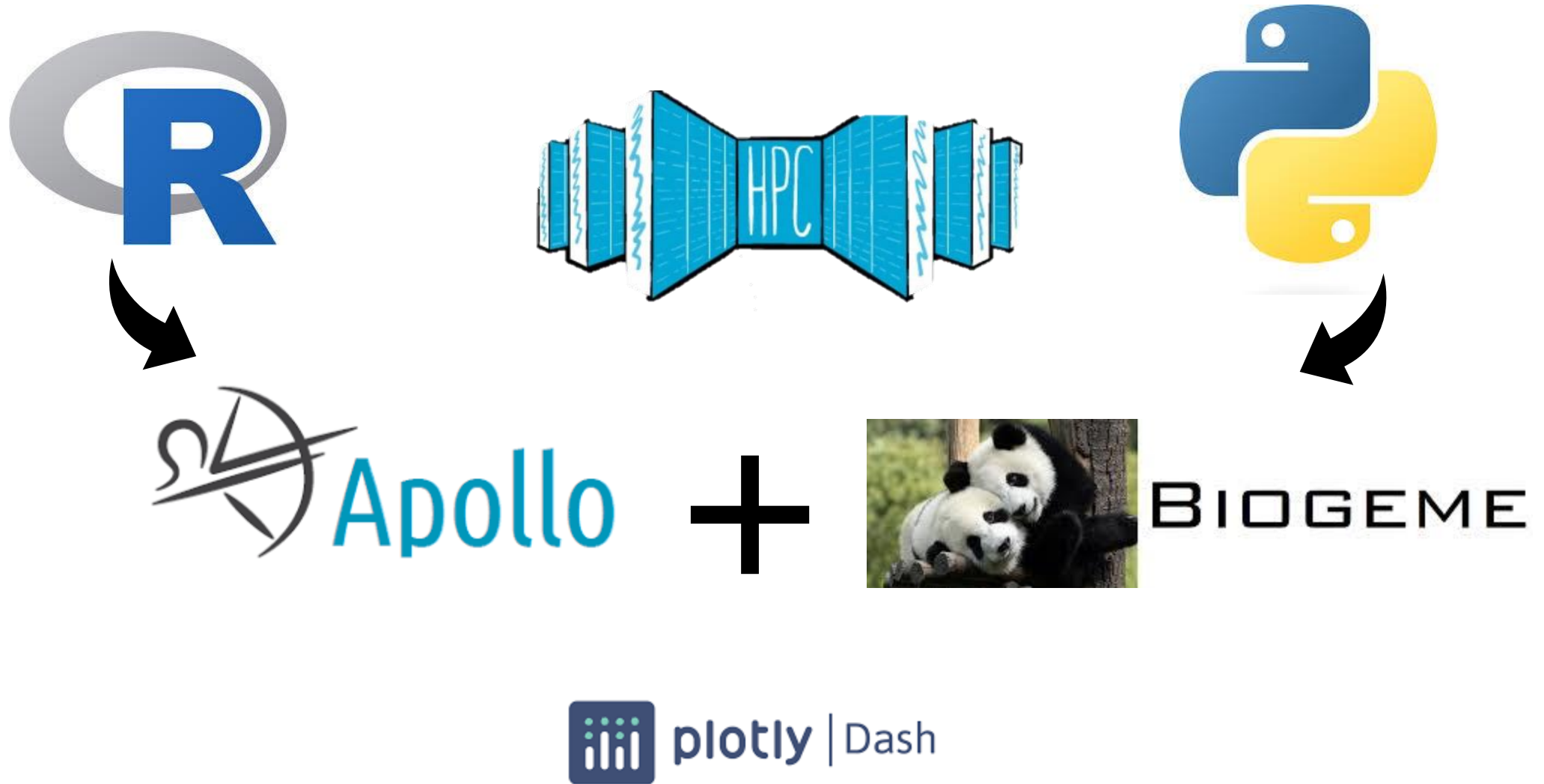
Context: Help with analysing a SC data set on residential location choices

# Choice Modelling Serious Game

## Game session



# Serious Choice Modelling Game Development



# Serious Game Context

## Game context:

- Imagine a colleague has asked for your help in analyzing a stated choice dataset on residential location preferences.
- This dataset has been designed to determine the willingness to pay (WTP) for reducing noise pollution.
- Respondents were faced with 3 unlabeled neighbourhoods (A,B,C) and asked to select the one they preferred to live in.
- The data were collected across four different cities and are representative of the target population.

## Game objective:

- **Develop a choice model to estimate the Willingness-to-Pay (WTP) for noise pollution reduction.**
- Your WTP estimate will be used by policymakers to make informed urban planning decisions.

## Game rules:

- Navigate freely between game phases: descriptive analysis, model specification, and outcome interpretation.
- Use the provided navigation buttons to move between phases as much as you like.
- The game takes around 45 minutes.
- Please avoid sharing information with co-workers during the game (if applicable).

What would you like to do next?

Instructions

Descriptive  
analysis

Model  
specification

Estimate  
and save

Reporting  
results

End  
game

# Choice Modelling Serious Game

## Descriptive analysis

What would you like to do in this Descriptive analysis?

Show	Missing value handling	Plot	Sort dataset by
Data dictionary		Boxplot	
First 5 rows	Display missing values	Histogram	Respondent
Descriptive variables	Delete missing values	Correlation	
Frequency of choices		Scatter plot	Choice task
Example of a choice task	Replace missing values	Pie chart	
		Bar chart	

# Choice Modelling Serious Game

## Model Specification

What is the model name?

What type of model do you want to specify?

Multinomial Logit Model × ▾

Which attributes would you like to consider?

Attributes (✔):

ASC? Yes: ☐ Stores  
☐ Transport  
☐ No ☐ City  
☒ Yes ☐ Noise  
☐ Green  
☐ Cost

Generic (✔) / Specific (□):

☐ Stores  
☐ Transport  
☐ City  
☐ Noise  
☐ Green  
☐ Cost

Which transformations would you like to apply for each attribute?

No, linear-additive × ▾

Which interactions would you like to consider?

No, linear-additive × ▾

Save before estimate

# Choice Modelling Serious Game

## Modelling outcomes

What would you like to analyse from the estimated model results?

Number of individuals

Estimated parameters

$\rho^2$  vs equal shares

Time taken

Number of rows in database

LL(start)

Adj.  $\rho^2$  vs equal shares

Number of modelled outcomes

LL at equal shares, LL(0)

AIC

WTP

Number of cores used

LL(final)

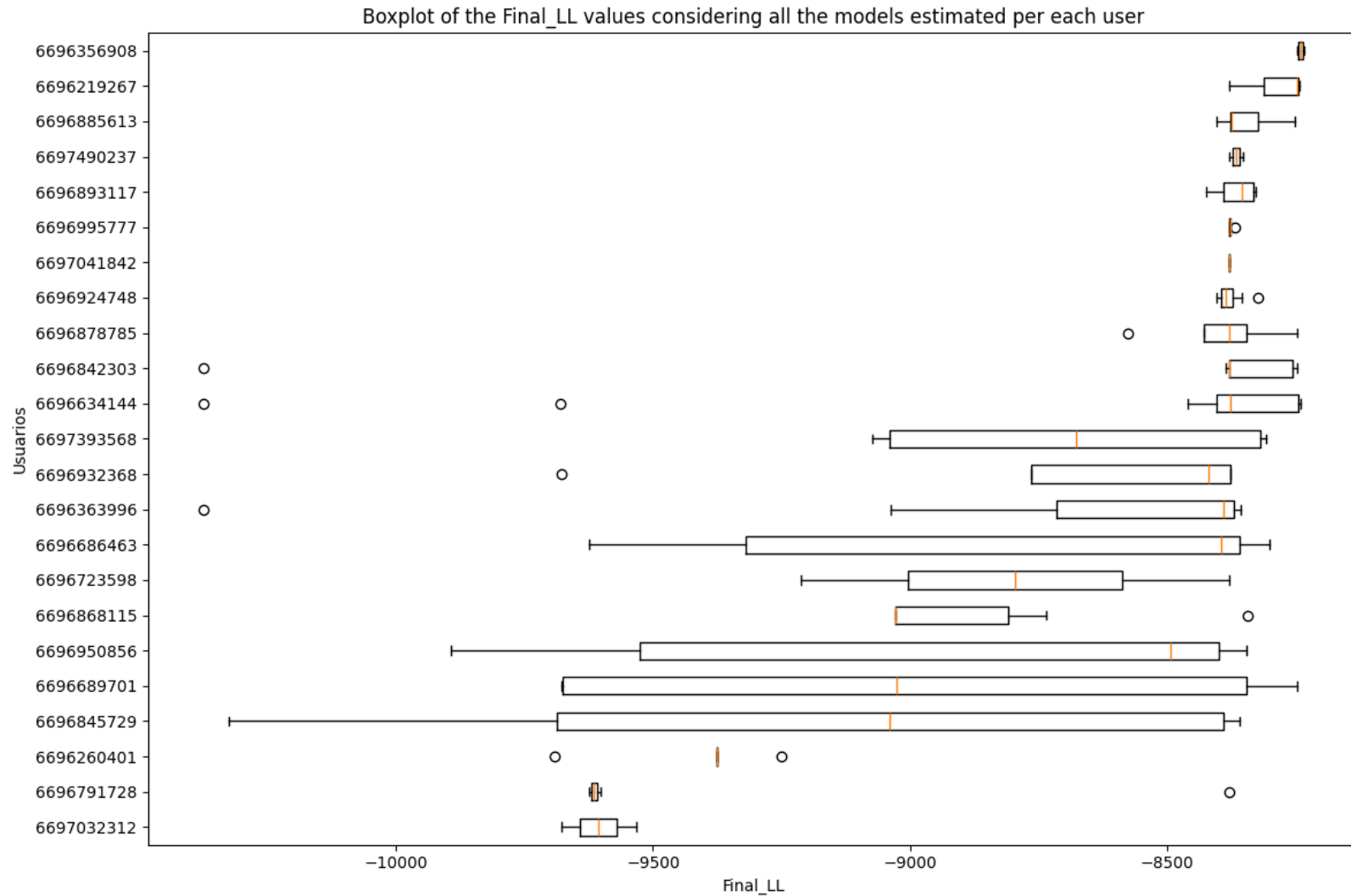
BIC

Parameters format

# Game session



# Preliminary results



# Understanding the decision-making process of discrete choice modellers

Gabriel Nova\*<sup>1</sup>, Sander van Cranenburgh<sup>1</sup>, Stephane Hess<sup>2</sup>

<sup>1</sup>Transport and Logistics group, Delft University of Technology

<sup>2</sup>Choice Modelling Centre - Institute for Transport Studies, University of Leeds