Multiverse analysis in R

Exploratory and Inferential Approaches

Filippo Gambarota and Giulia Calignano

Senior Research Fellow
University of Padova
Department of Developmental and Social Psychology, DPSS

CSA Brixen, 13-15th February 2025

UNIVERSITÀ DEGLI STUDI DI PADOVA

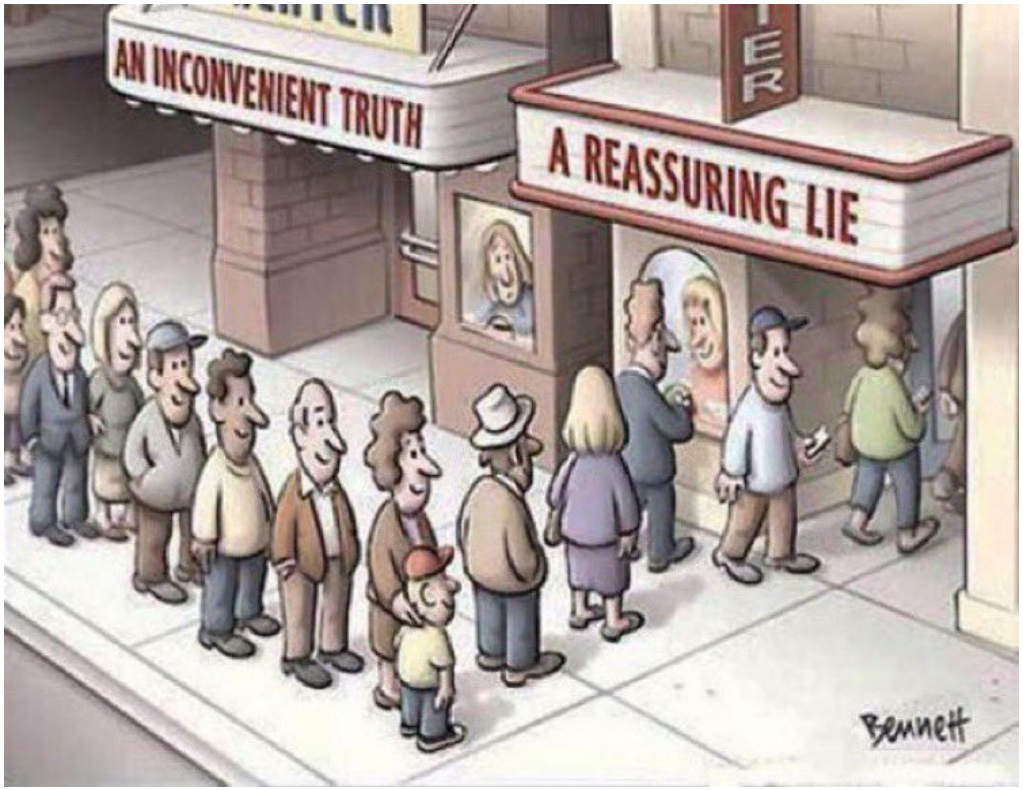
Contents

Theoretical Introduction to Multiverse Analysis

Origins of Multiverse Analysis

Origins of Multiverse Analysis: The concept of multiverse analysis emerged to systematically explore research decisions and their impact on scientific conclusions. It was formally introduced by Steegen et al. (2016) as a method to present a range of plausible research outcomes, enhancing transparency and reproducibility.

We rarely find data, we actively construct datasets



- A single data collection = a multiverse of possible datasets

(Steegen et al., 2016)

Any datum is not for granted

data

is the plural form of the Latin word *datum*,

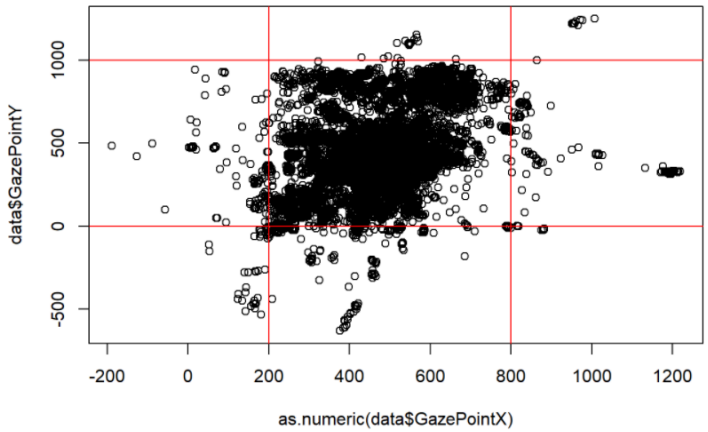
which means the 'thing given'

Degrees of Freedom (RDF) and Forking Paths

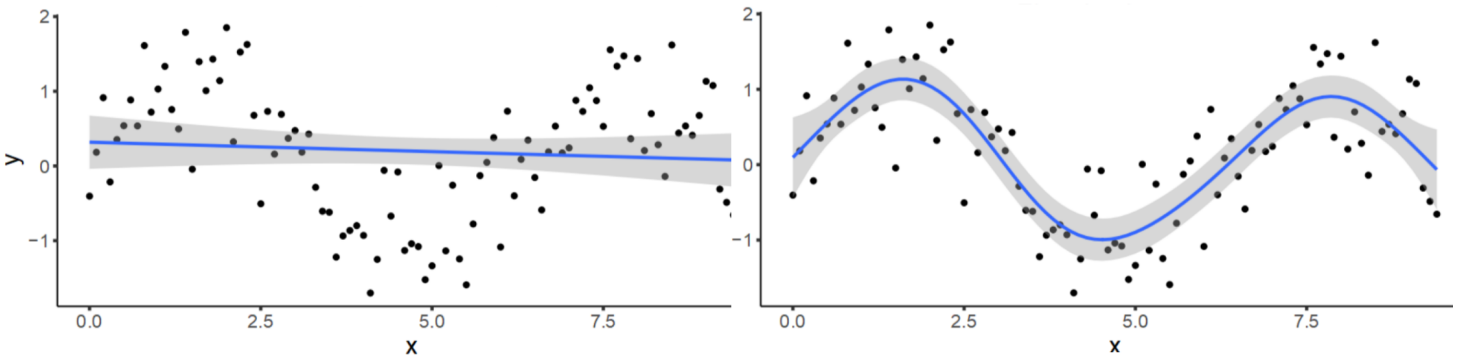
Researcher Degrees of Freedom (RDF): Multiverse analysis acknowledges that scientific research involves multiple reasonable choices, from data selection to analysis models. Each choice introduces 'researcher degrees of freedom' (RDF), which can drammatically influence findings (Gelman, 2014).

Degrees of freedom in pupil analysis

- Area of Interest (AOI) and implausible values e.g. outliers



- Baseline correction **!big issue in psychophysiology!**
- Statistical modeling **do not forget individual variability!**

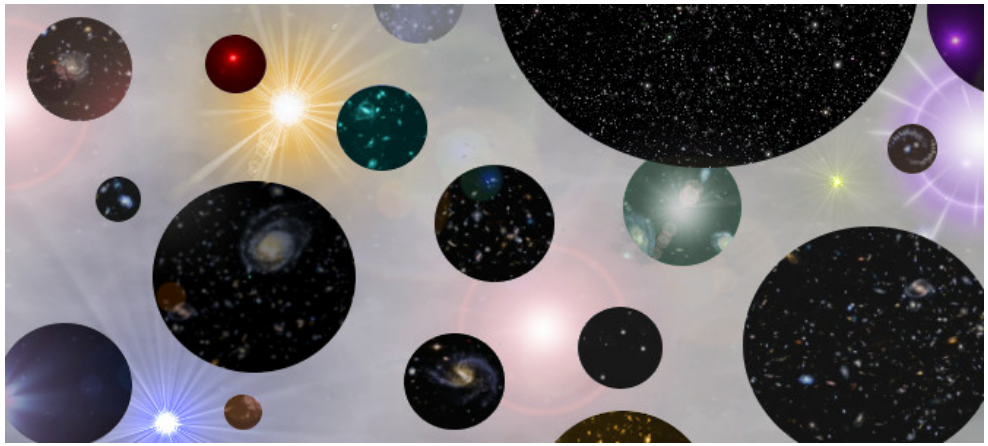


Multiverse: Robustness and Generalizability

Nosek's (2018) Approach:

		Data: Same	Data: Different
Analysis: Same		Reproducible e.g. FAIR	Replicable across cultural contexts
Analysis: Different		Robust e.g. multimethod - multiverse	Generalisable

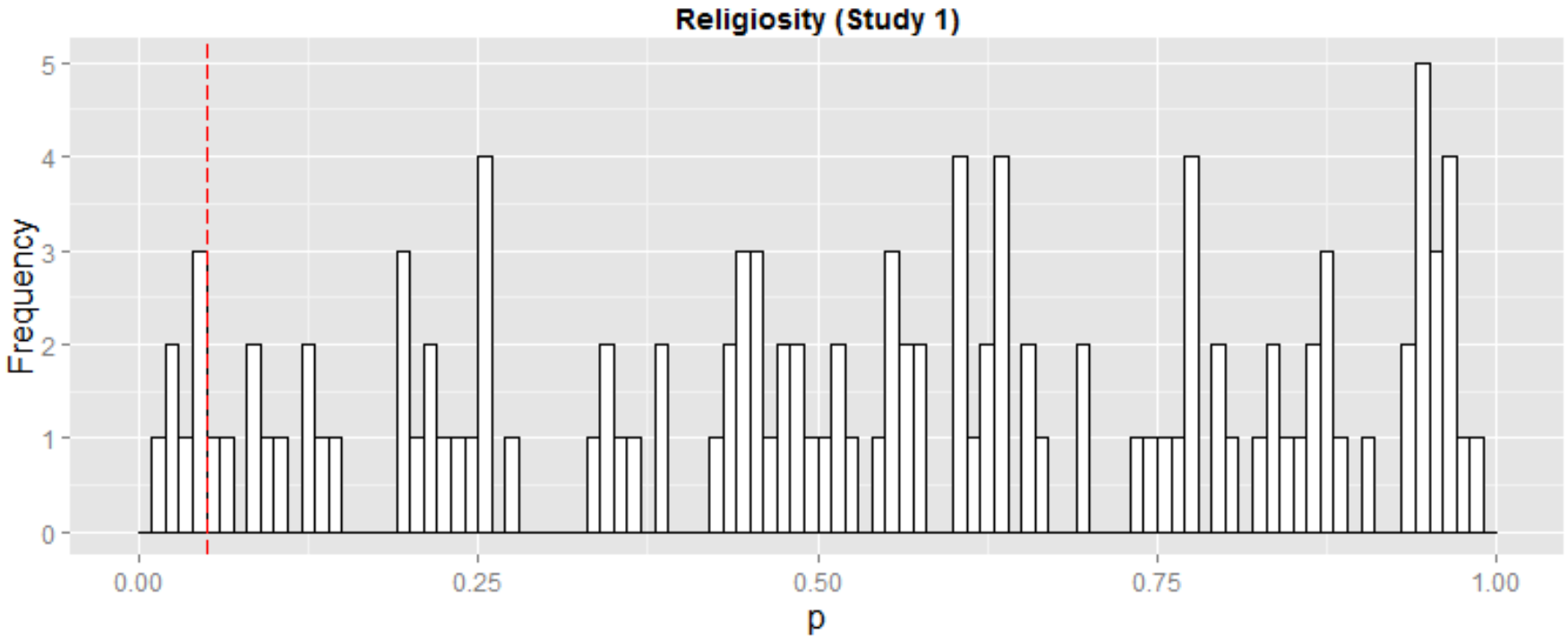
A Multiverse approach



- **a philosophy of statistical reporting** in the manuscript (not in the supllementary materials) the outcomes of many different statistical analyses showing how robust findings are (Dragicevic, et al.,2019)
- **robustness of a finding** across different options for all steps in data processing (Steegen et al., 2016).

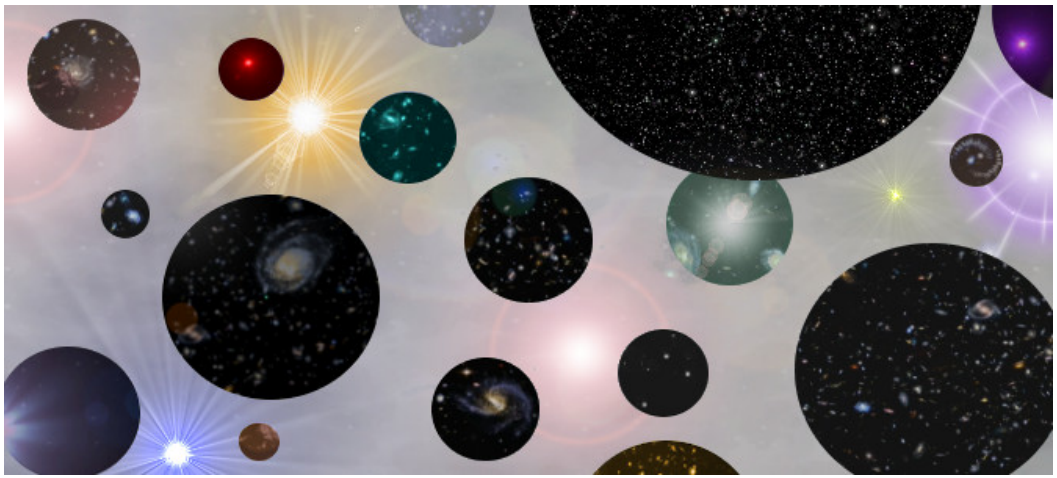
A Multiverse approach

- is the effect **robust** or is it driven by data processing choices?
- there is a multiverse of statistical results



(Steegen et al., 2016)

Building a reliable Psychophysiology and Cognitive Neuroscience



- **Importance of embracing (rather than be afraid of) the uncertainty in data**
- Data sharing and caring contribute to a full-multiverse approach

Two simple examples: Verbal memory RT, Accuracy and Pupillometry

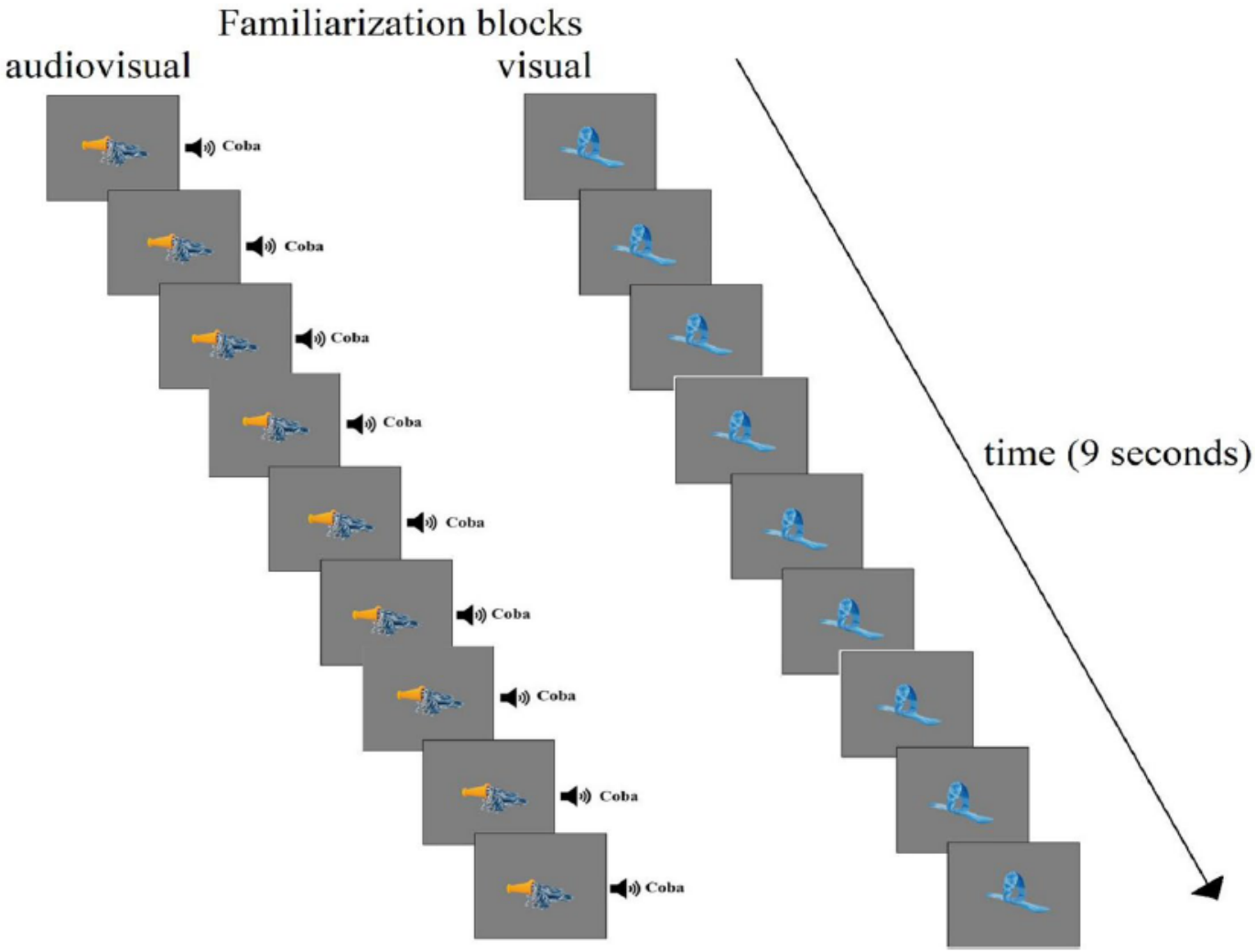
Concrete Example of RDF

Infancy research: pupil dilation and

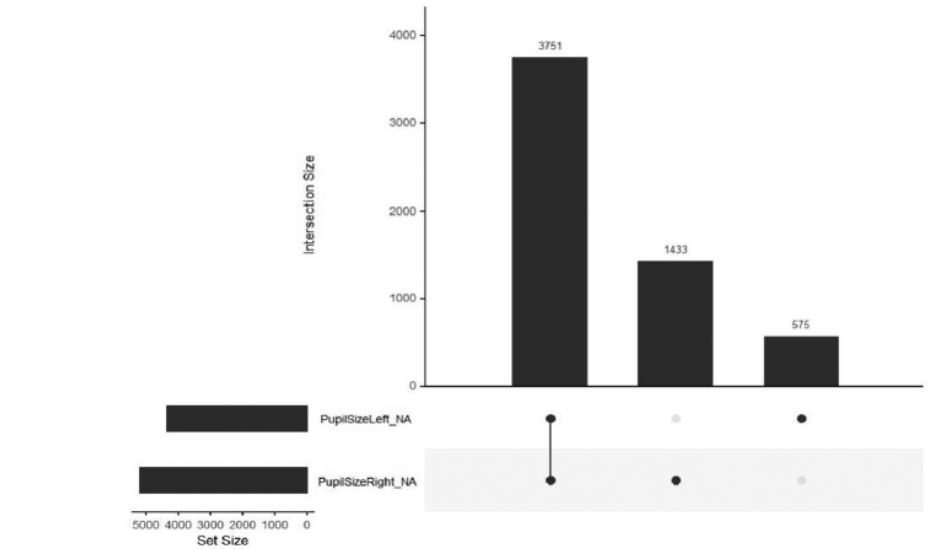
Case Study: Examining reaction times and pupil measurements within a multiverse framework involves systematic variations in pre-processing and statistical modeling.

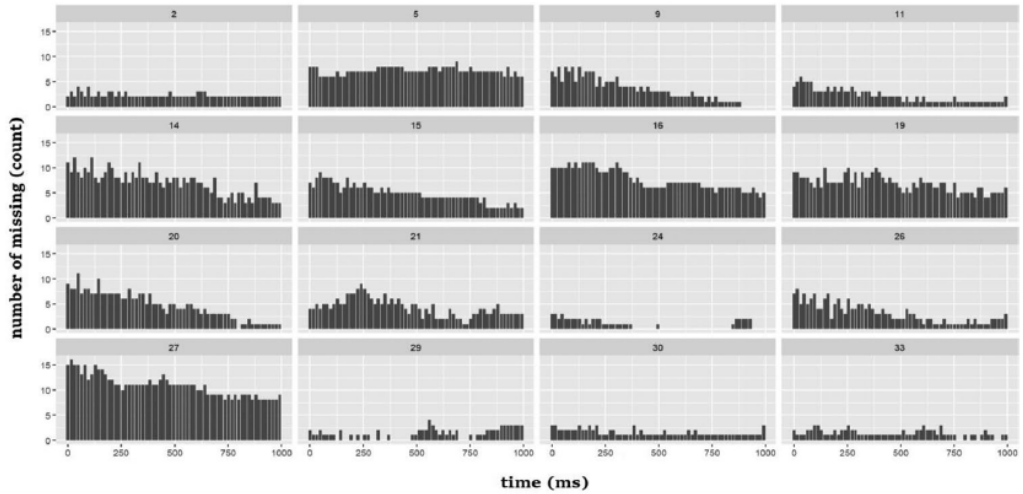
The garden of forking paths

An illustrative example



Are missing... data?





Degrees of freedom (1) Data filtering

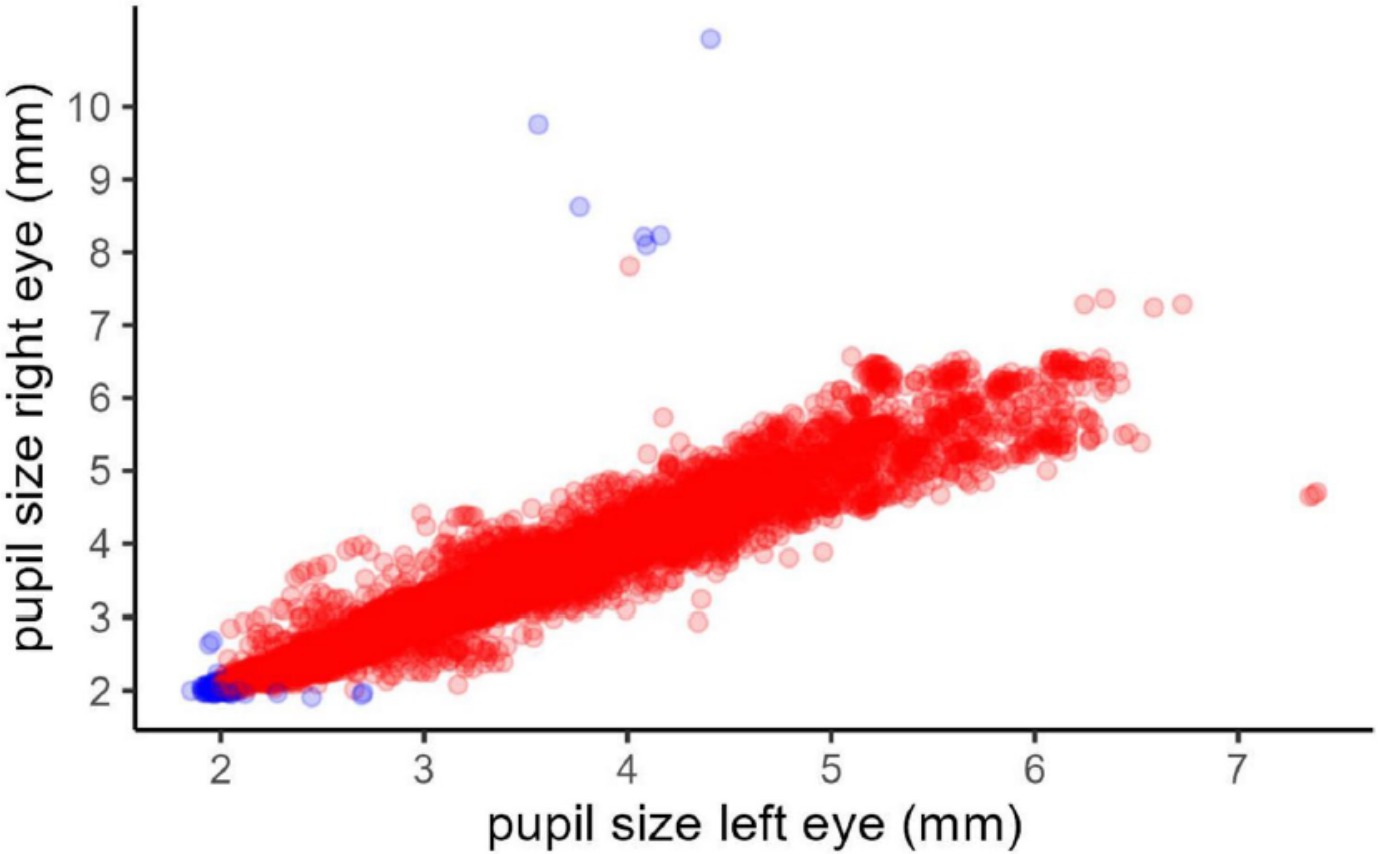
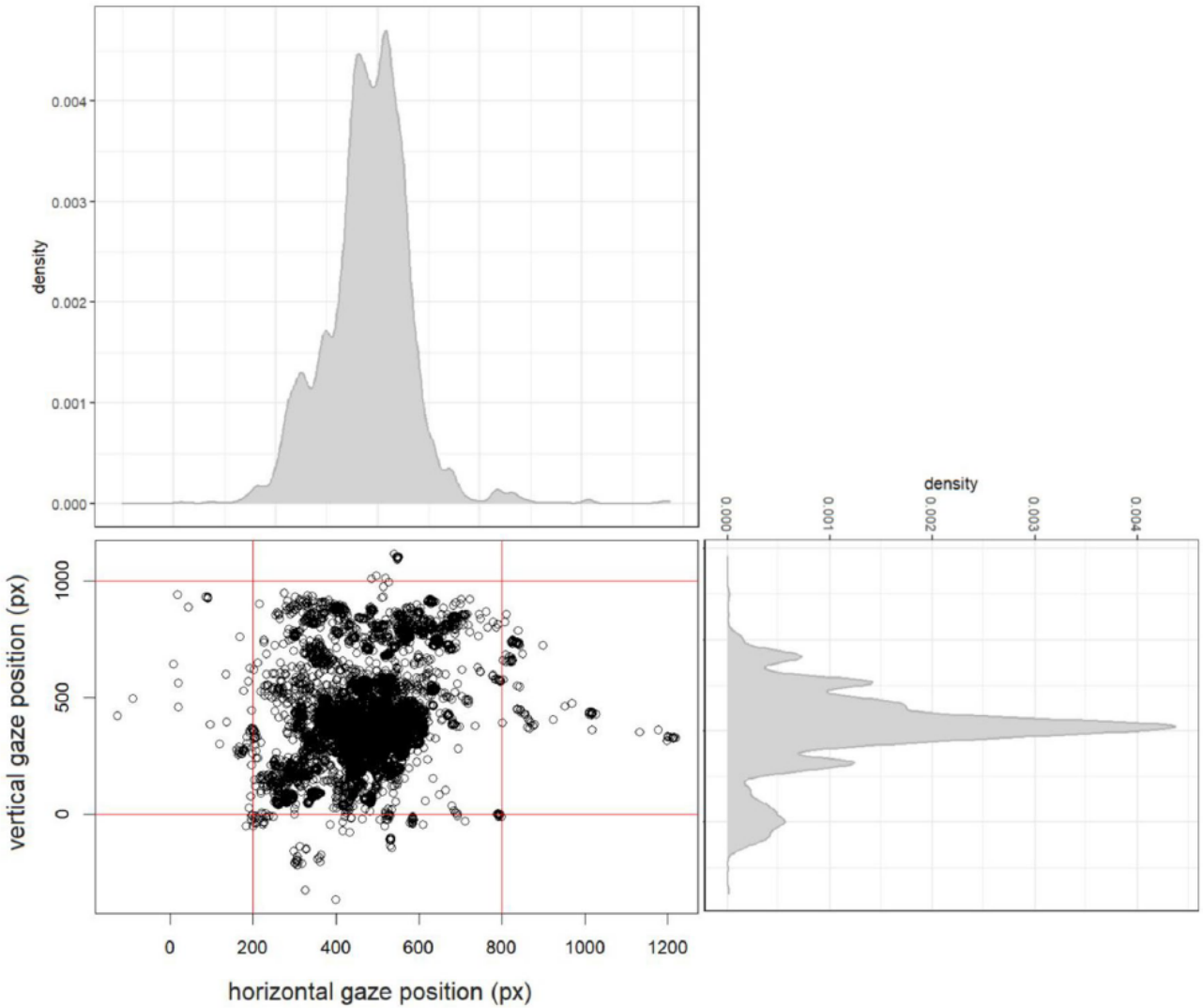


Fig. 3 Scatter plot correlating left and right eye's pupil size. *Blue points* indicate the values excluded in the second filtered dataset (trimmed dataset)

Degrees of freedom (2) Area of interest



Degrees of freedom (3) Dealing with blinks

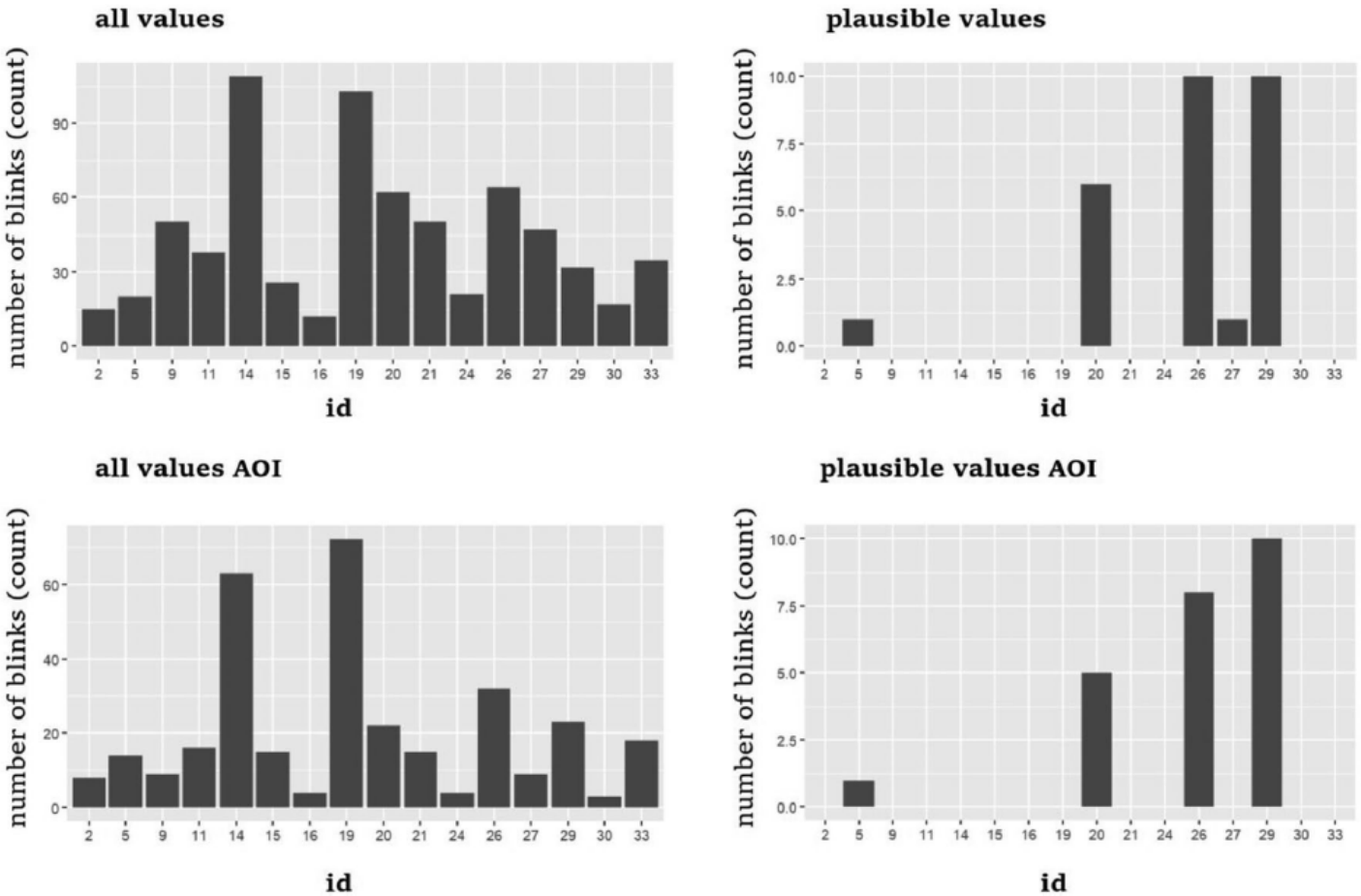


Fig. 5 Distribution of blink detected across the participants (id) in the four datasets, i.e., all values in the whole screen, plausible values in the whole screen, all values within the area of interest (AOI), plausible values within the AOI

Degrees of freedom (4) Baseline correction

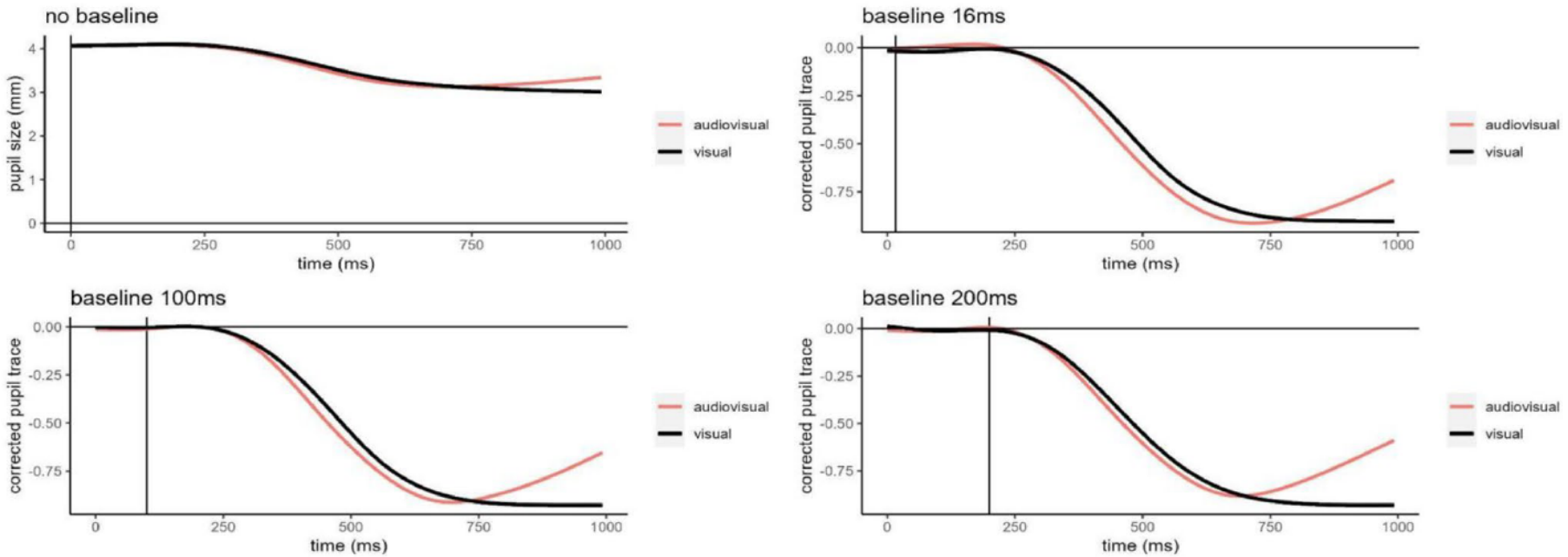


Fig. 6 Average pupil size variation (no baseline) and pupil changes relative to baseline (16, 100, and 200 ms) smoothed across time, we used the dataset with trimmed values filtered by the AoI and interpolated blinks for illustrative purposes. The *red* and *black lines* represent the audiovisual and visual familiarization, respectively. The *vertical line* indicates the end of the baseline (when present)

Degrees of freedom (5) Participants inclusion

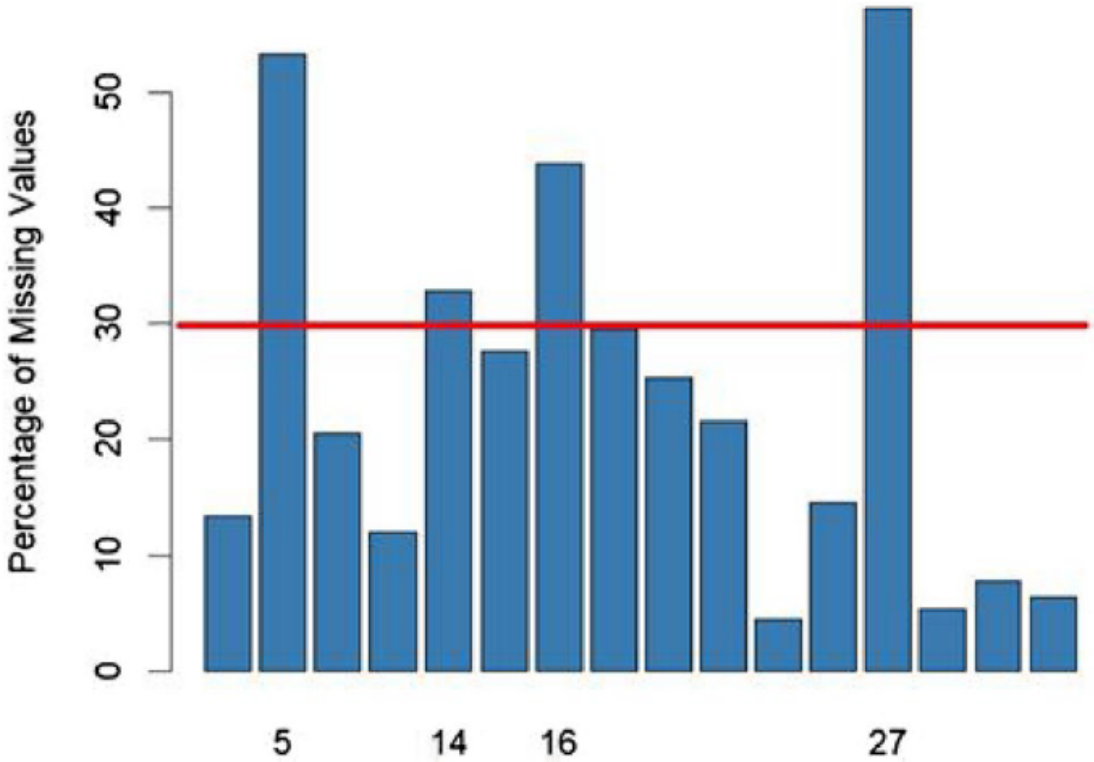
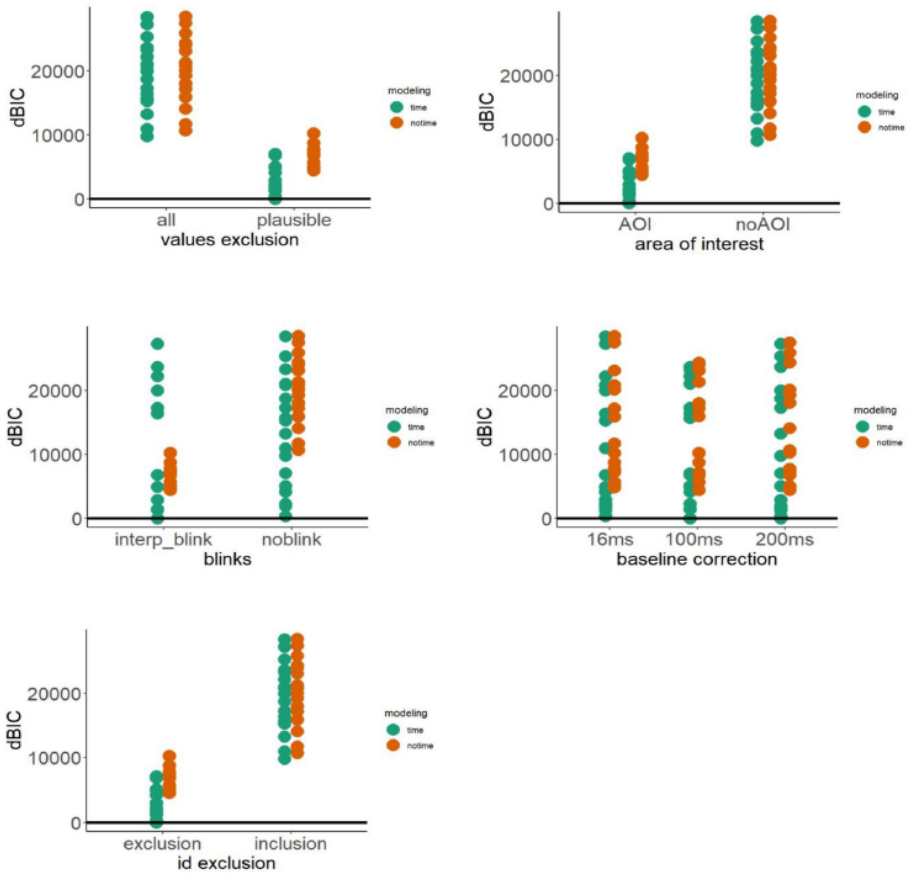


Fig. 7 Percentage of missing values by subject (ID). The *red line* indicates the cut-off value. Note that all that only ID participants above the cut-off are shown in the *x*-axis

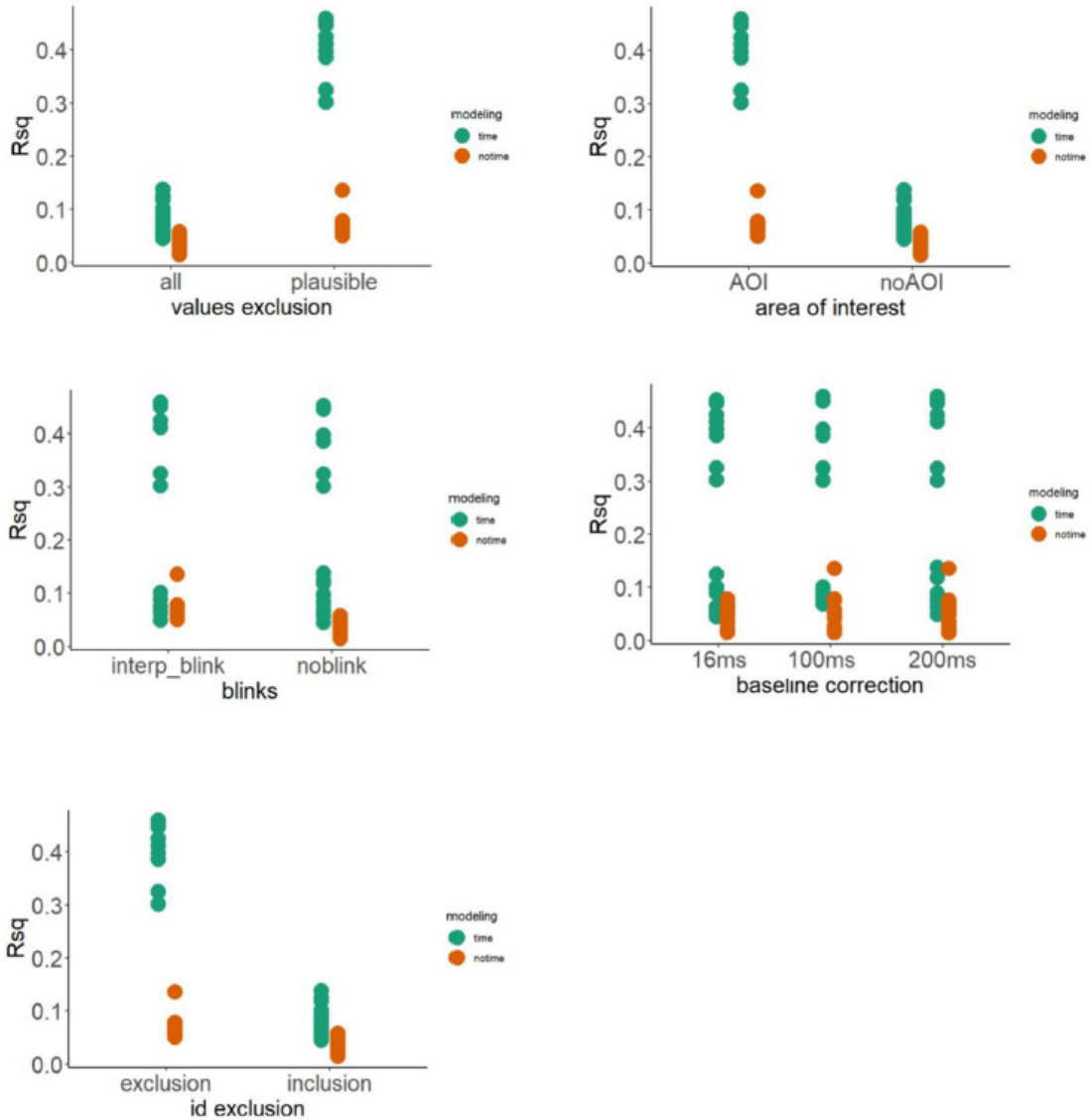
Degrees of freedom (6) A multiverse of models

$$Y = \alpha + \beta X + \mathrm{g1}(X, \mathrm{id}) + \varepsilon$$

$$Y = \alpha + \beta X + \mathrm{g1}(t, X) + \mathrm{g2}(t, \mathrm{id}) + \varepsilon$$



Degrees of freedom (6) A multiverse of models



Degrees of freedom (6) A multiverse of models

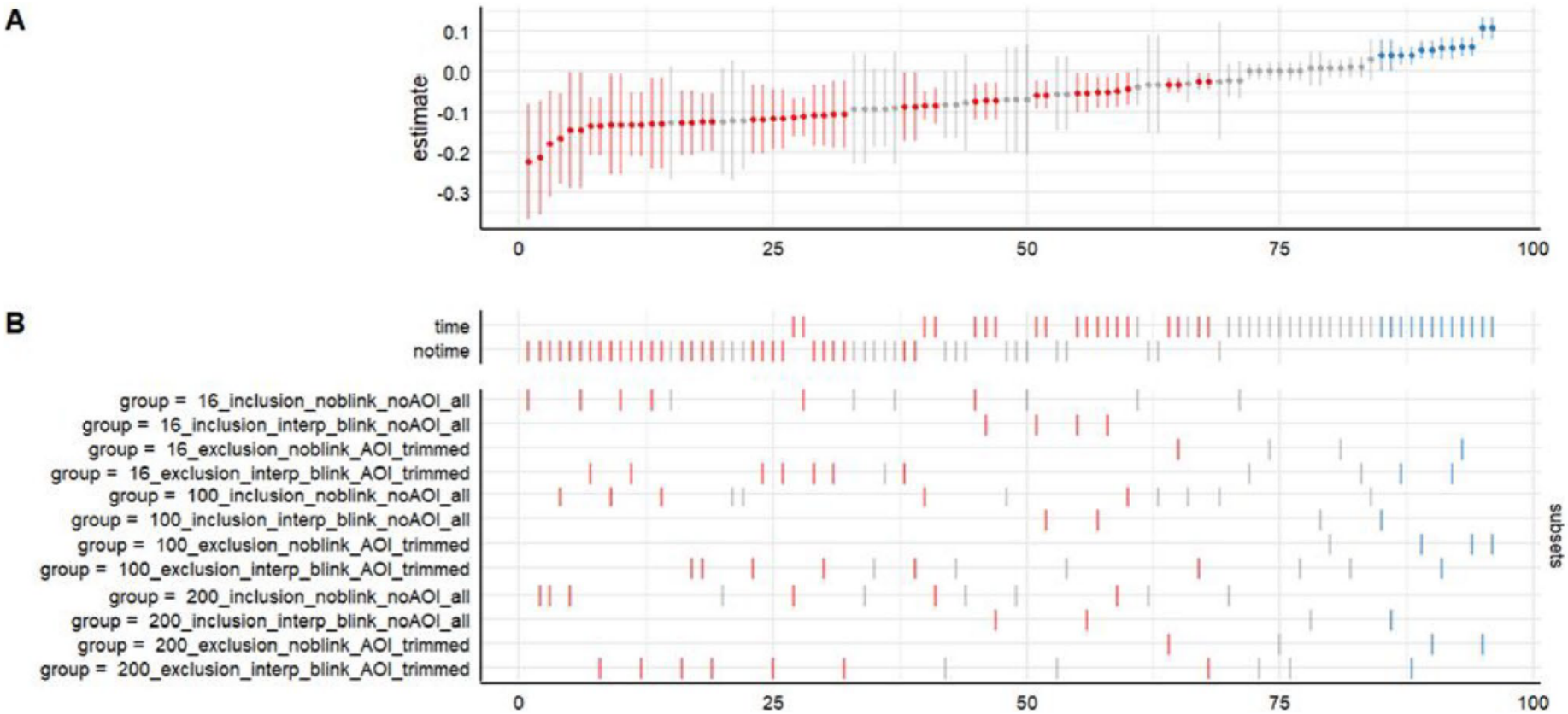
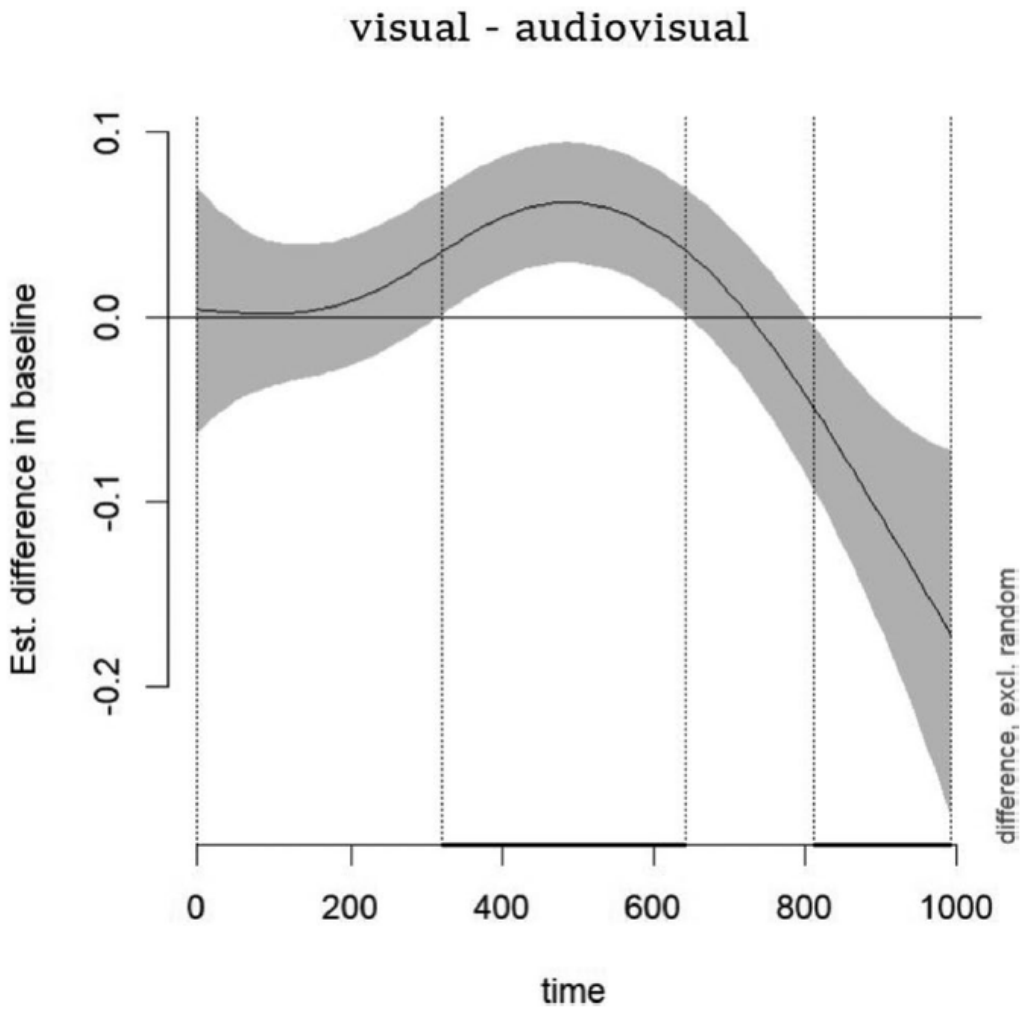
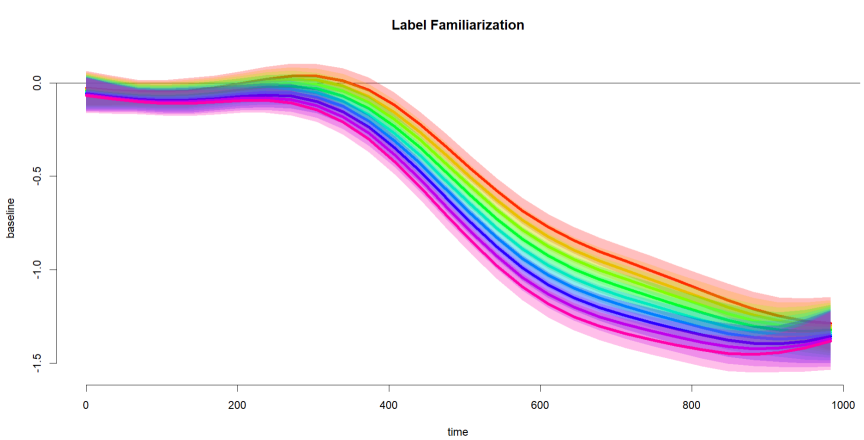


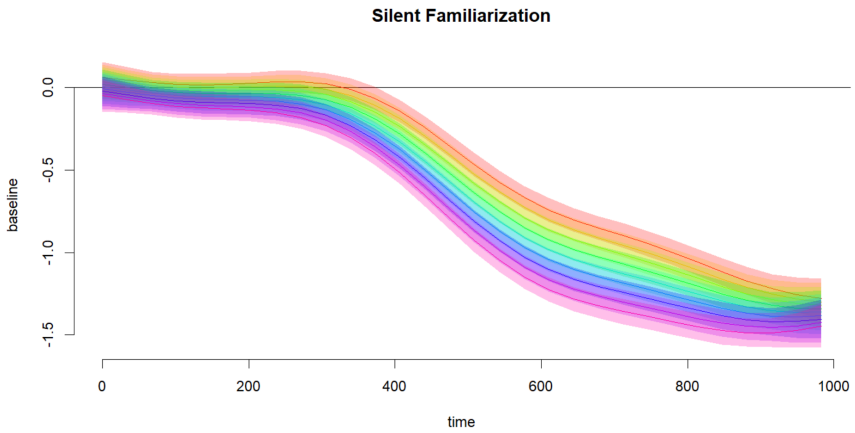
Fig. 9 **A** The 96 coefficient's estimates and relative 95% CI related to the Visual vs. Audiovisual regressor. **B** Relative combinations by the six degrees of freedom of the multiverse analysis. The direction of the significant results are highlighted (negative = *red*, positive = *blue*, *gray* = non-significant). Note that positive estimates (*in blue*) indicate higher pupil dilation for the Audiovisual condition and negative estimates (*in red*) indicate higher pupil dilation for the Visual condition. The *x*-axis represents the model number, while the *y*-axis represents the estimated coefficient

Degrees of freedom (6) A multiverse of models



Degrees of freedom (6) A multiverse of models





ManyBabies 2 : Theory of Mind in Infancy

MB2 - ToM

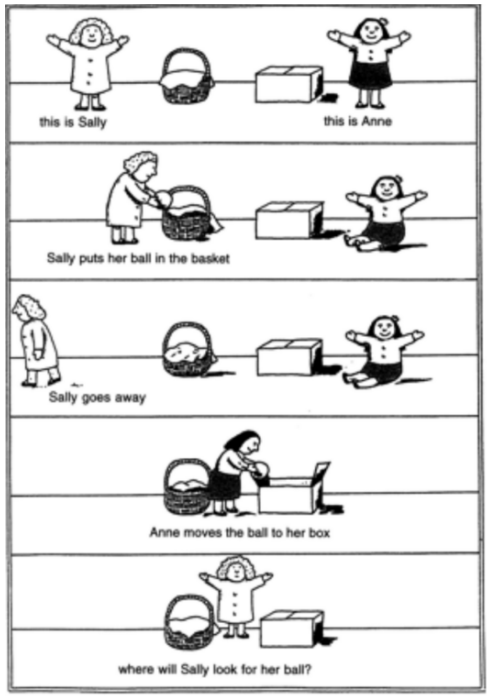
MB2

Theory of Mind in Infancy

Stage 1 Registered Report

*Schuwerk, T.**, *Kampis, D.**, Baillargeon, R., Biro, S., Bohn, M., Byers-Heinlein, K., Dörrenberg, S., Fisher, C., Franchin, L., Fulcher, T., Garbisch, I., Geraci, A., Grosse Wiesmann, C., Hamlin, K., Haun, D. B. M., Hepach, R., Hunnius, S., Hyde, D. C., Karman, P., ... Rakoczy, H. (accepted pending data collection). **Action anticipation based on an agent's epistemic state in toddlers and adults**. *Child Development*. PsyArXiv. https://doi.org/10.31234/osf.io/x4jbm (**co-first authors*)

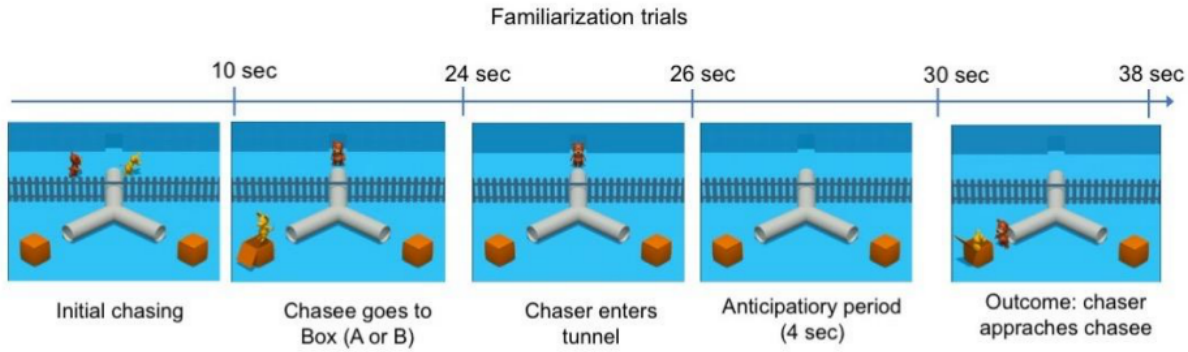
MB2 - Do toddlers and adults engage in spontaneous Theory of Mind (ToM)?



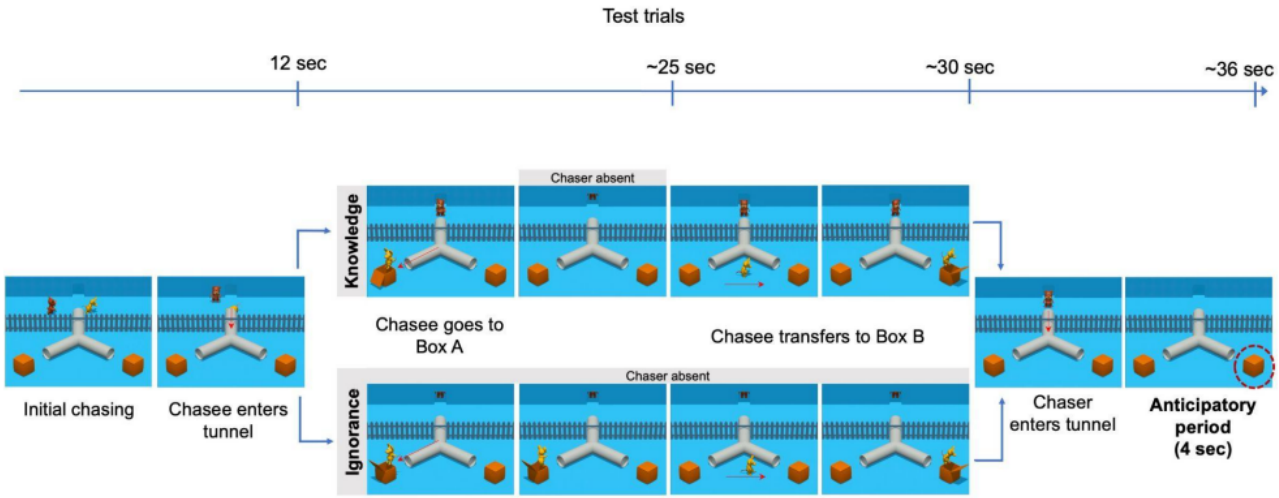
- Evidence from **anticipatory looking (AL)** studies
- But a growing body of **failed replication studies** raised **the need to test the robustness**
- Question: 18- to 27-month olds' and adults' anticipatory looks between **two forms of epistemic states**: knowledge and ignorance.

MB2 - ToM - Paradigm

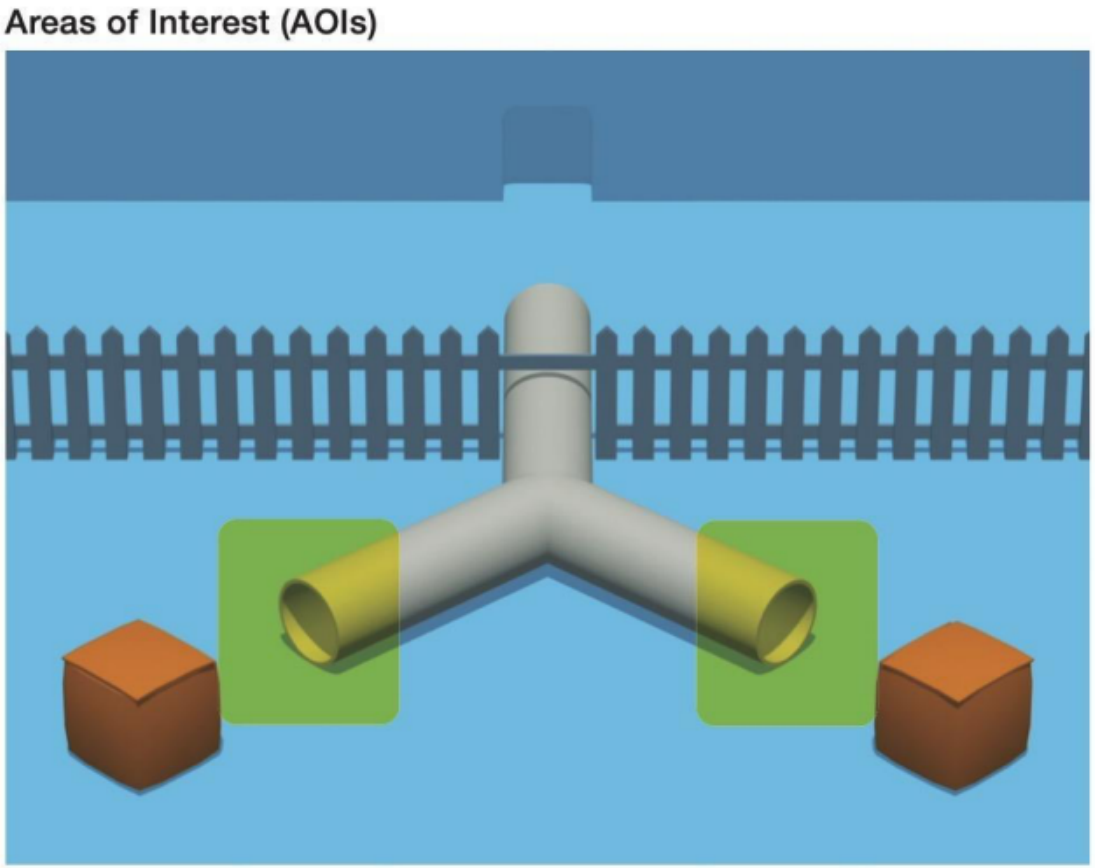
Timeline of the familiarization trials.



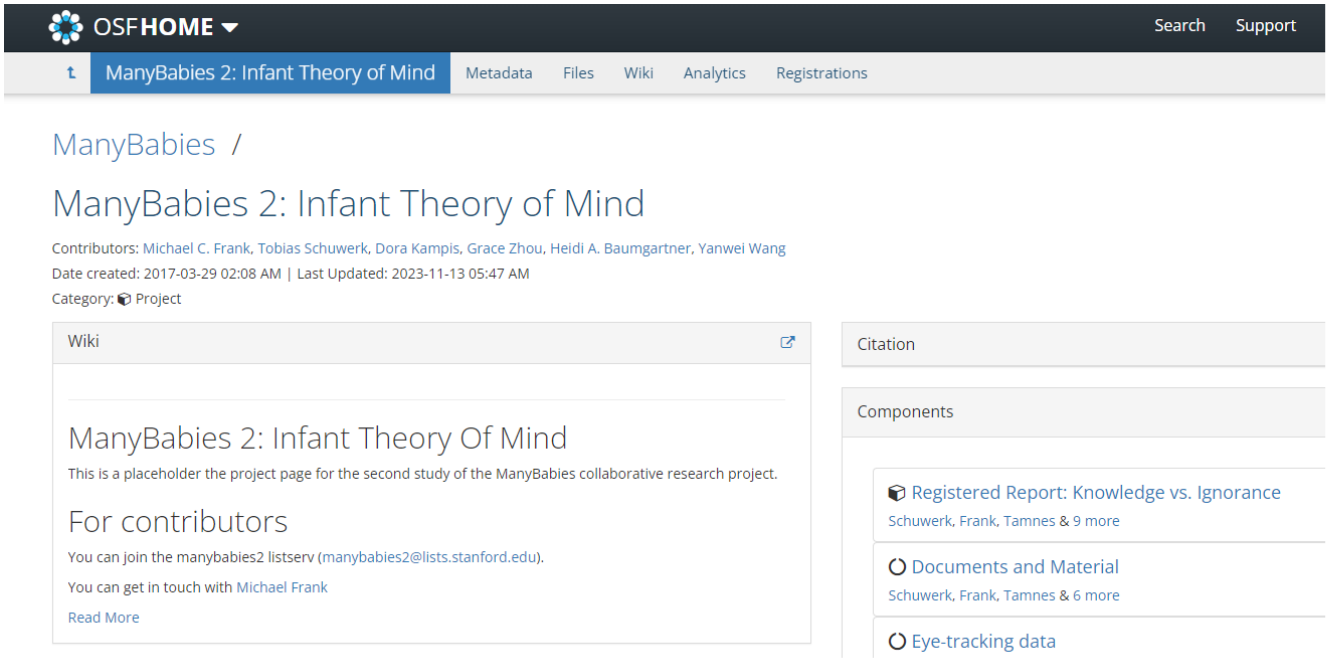
Schematic overview of stimuli and conditions of the test trials.

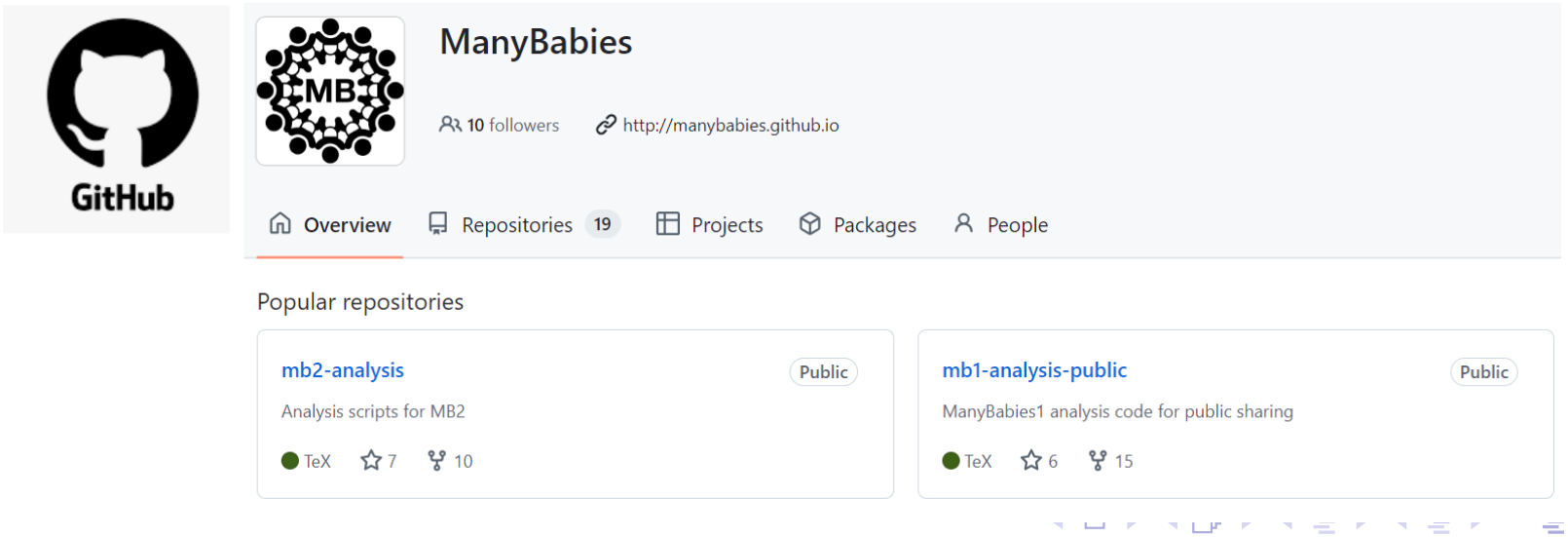


MB2 - ToM



MB2 - ToM


OSF HOME
Search
Support
ManyBabies 2: Infant Theory of Mind
Metadata
Files
Wiki
Analytics
Registrations
ManyBabies /
ManyBabies 2: Infant Theory of Mind
Contributors: Michael C. Frank, Tobias Schuwerk, Dora Kampis, Grace Zhou, Heidi A. Baumgartner, Yanwei Wang
Date created: 2017-03-29 02:08 AM | Last Updated: 2023-11-13 05:47 AM
Category: Project
Wiki
ManyBabies 2: Infant Theory Of Mind
This is a placeholder the project page for the second study of the ManyBabies collaborative research project.
For contributors
You can join the manybabies2 listserv (manybabies2@lists.stanford.edu).
You can get in touch with Michael Frank
Read More
Citation
Components
Registered Report: Knowledge vs. Ignorance
Schuwerk, Frank, Tamnes & 9 more
Documents and Material
Schuwerk, Frank, Tamnes & 6 more
Eye-tracking data


GitHub
MB
ManyBabies
10 followers
http://manybabies.github.io
Overview
Repositories 19
Projects
Packages
People
Popular repositories
mb2-analysis
Public
Analysis scripts for MB2
TeX
7
10
mb1-analysis-public
Public
ManyBabies1 analysis code for public sharing
TeX
6
15

MB2 Pupillometry Spin Off

MB2P - Secondary Data Analysis

MB2P

Measuring Pupil Dilation in Response to Expected and Unexpected Events

MB2P - Secondary Data Analysis

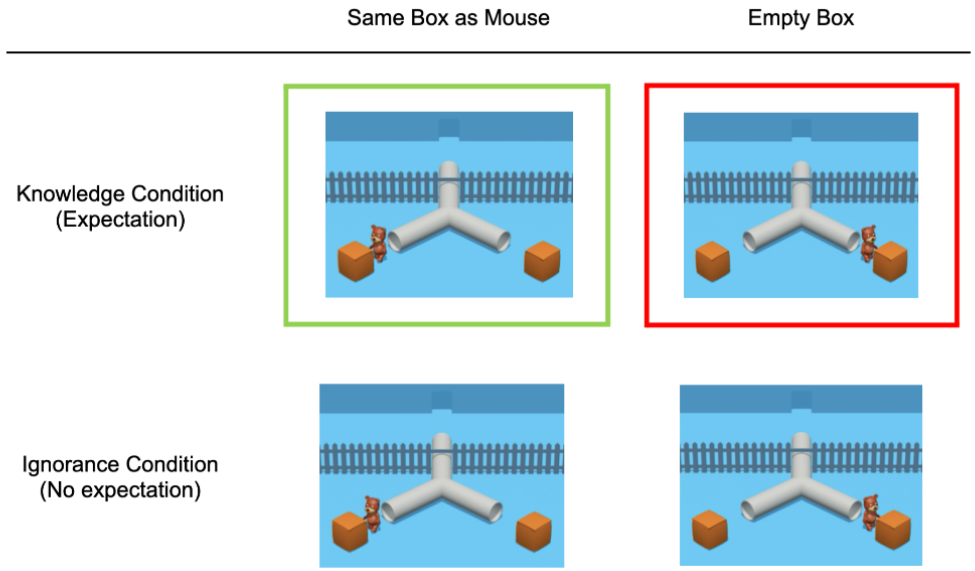
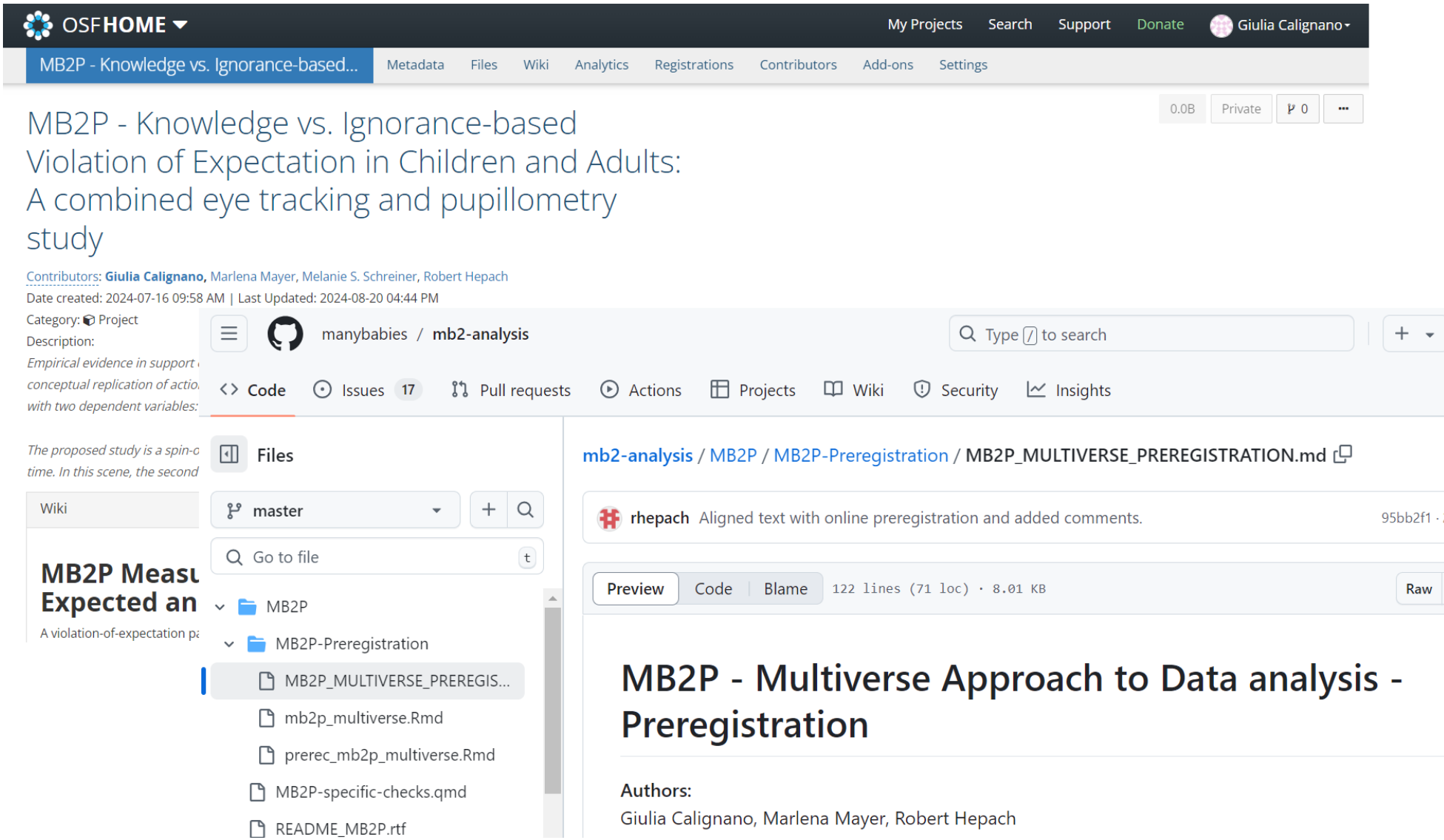


Table 1. Illustrations of the outcomes within the VOE-design. In this example, the mouse is hiding in the left box. The green box marks the congruent outcome while the red box marks the incongruent outcome.

- MB2P is interested in responses to seeing a **goal-congruent vs. a goal-incongruent outcome**
- MB2P asks whether children are **surprised** if the protagonist (the bear) responds in a way that is incongruent with the goal (to follow the mouse)
- Measures: **pupil dilation** and **looking time** in response to the congruent and incongruent events

The MB2P Preregistration and the Multiverse approach

The MB2P Preregistration and the Multiverse approach



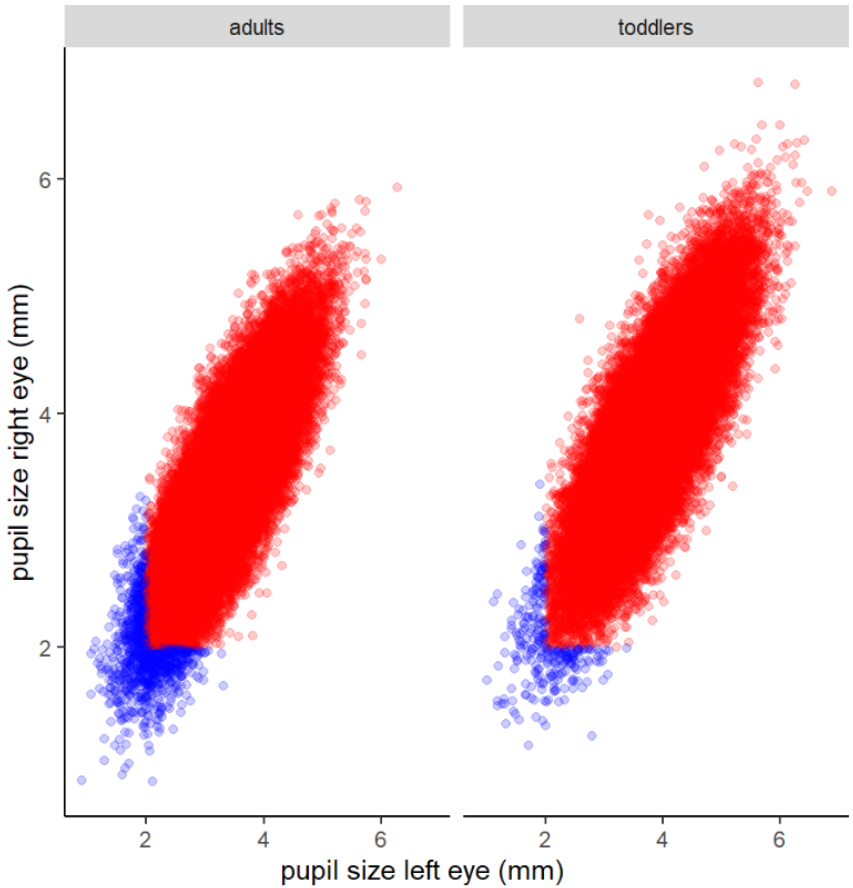
The MB2P Preregistration and Data simulation

- **Data Collection** by MB2 completed in 2023. MB2 will provide the raw data for the present study **after** the Stage 1
- **The analyses** will be based on the second test trial of the main study.
- **Data Simulation** mirroring anticipated data characteristics, **to pre-register preprocessing and analysis strategies**
- The simulated data includes: **participant id**, two **age cohort**, **Timestamp or duration**, **x and y: Coordinates or measurements**, **pupil left and pupil right sizes**, **lab id**, **conditions and outcomes**

The MB2P - Preregistration and Multiverse Forking Paths

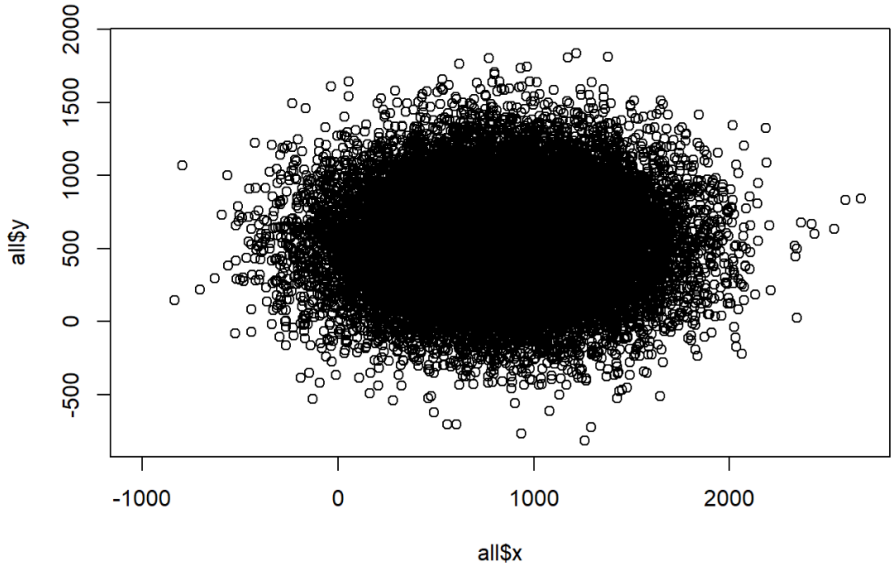
First Degree of Freedom: *Filtering Extreme Pupil Values*

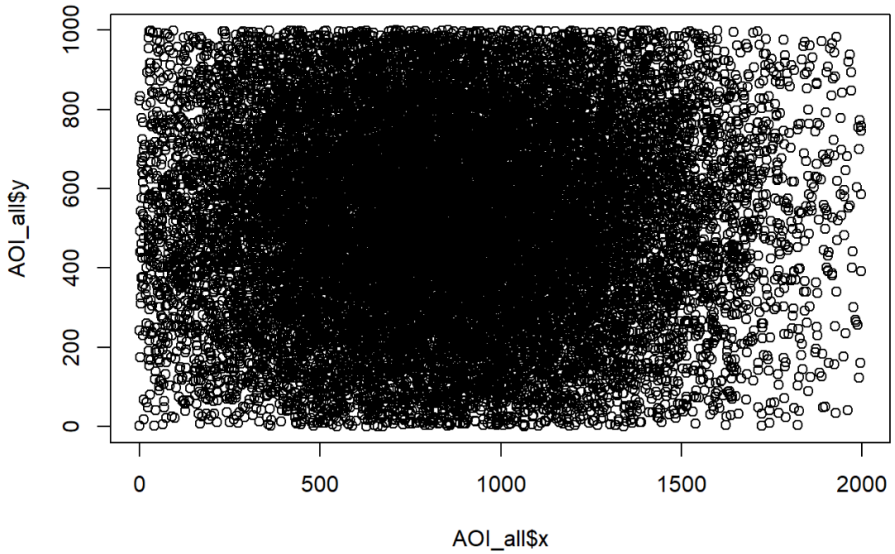
- Step 1: Filter implausible pupil values (obtaining 2 datasets)



The MB2P - Preregistration and Multiverse Forking Paths

Second Degree of Freedom: *Spatial filtering*





The MB2P - Preregistration and Multiverse Forking Paths

Third Degree of Freedom: *Moving average filtering*

- A moving average is a method to smooth data by calculating the average of a set number of consecutive data points, which then shifts forward over time
- For example, in pupillometry, a 5-second moving average can be used to track pupil diameter changes by averaging measurements from each 5-second window, reducing the influence of short-term fluctuations like blinking or minor eye movements

The MB2P - Preregistration and Multiverse Forking Paths

Fourth Degree of Freedom: *Moving average filtering Baseline Correction*

- Method 1: 5 seconds before the bear resolution
- Method 2: 300 milliseconds after the bear resolution
- Method 3: 500 milliseconds after the bear resolution.

The MB2P - Preregistration and Multiverse Forking Paths

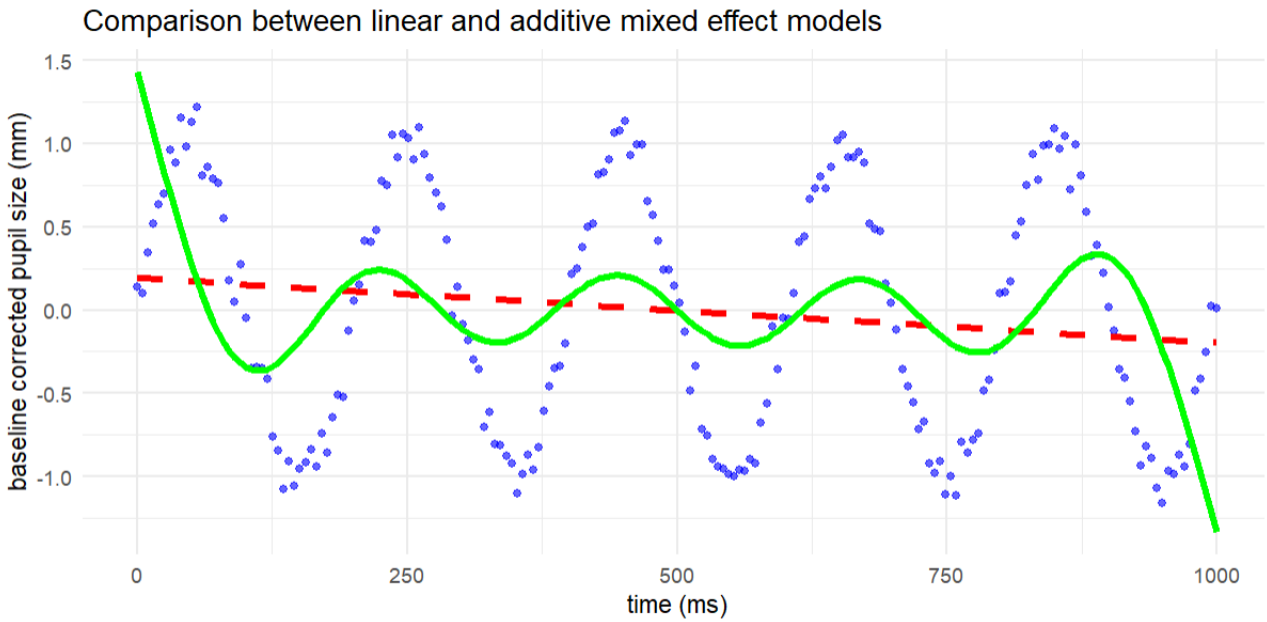
Fifth Degree of Freedom: *Participant Exclusion*

- including only those subjects with two valid test trials
- with one valid test trial (as defined by the main MB2-inclusion criteria),
- with no trial exclusion criteria applied.

The MB2P - Preregistration and Multiverse Forking Paths

Statistical Modeling

- Model 1: Linear mixed-effect model with condition, outcome, and age
- Model 2: Linear mixed-effect model with condition, outcome, and age over time
- Model 3: Generalized Additive mixed-effect with condition, outcome, age, time, and participant-specific effects



Crispy considerations

- **No practice leads to perfectly clean data**, yet it is fundamental to show the impact of preprocessing on statistical results (Steegen et al., 2016 Dragicevic et al., 2019)
- Pupillometry is a useful indirect measure of **the time course** of attention since infancy (Blaser et al., 2014; Brisson et al., 2013; Sirois Jackson, 2012; Tamasi et al. 2016)
- Statistical models that allow to evaluate *how* and *when* an effect emerge is fundamental in cognitive development (more than p-value)
- Statistical significance **is not useful at all when taken alone**
- multiverse analysis address **robustness** and avoid p-hacking
- Embracing (rather than being afraid of) the inherent uncertainty of infant data **increase our understanding of individual differences in developmental pathways of attention, learning processes, and beyond**.

Thank You!

filippo.gambarota@unipd.it, giulia.calignano@unipd.it

UNIVERSITÀ DEGLI STUDI DI PADOVA

Cognitive Science Arena

Brixen (BZ, Italy), February, 13-15th, 2025.