Multi-lab and Multiverse

From Preprocessing to Data Analysis

Giulia Calignano and Livio Finos

4Ms WS - February 2025

UNIVERSITÀ DEGLI STUDI DI PADOVA

Contents

My Academic Journey

Theoretical Introduction to Multiverse Analysis

Concrete Example of RDF

MB2 Pupillometry Spin Off

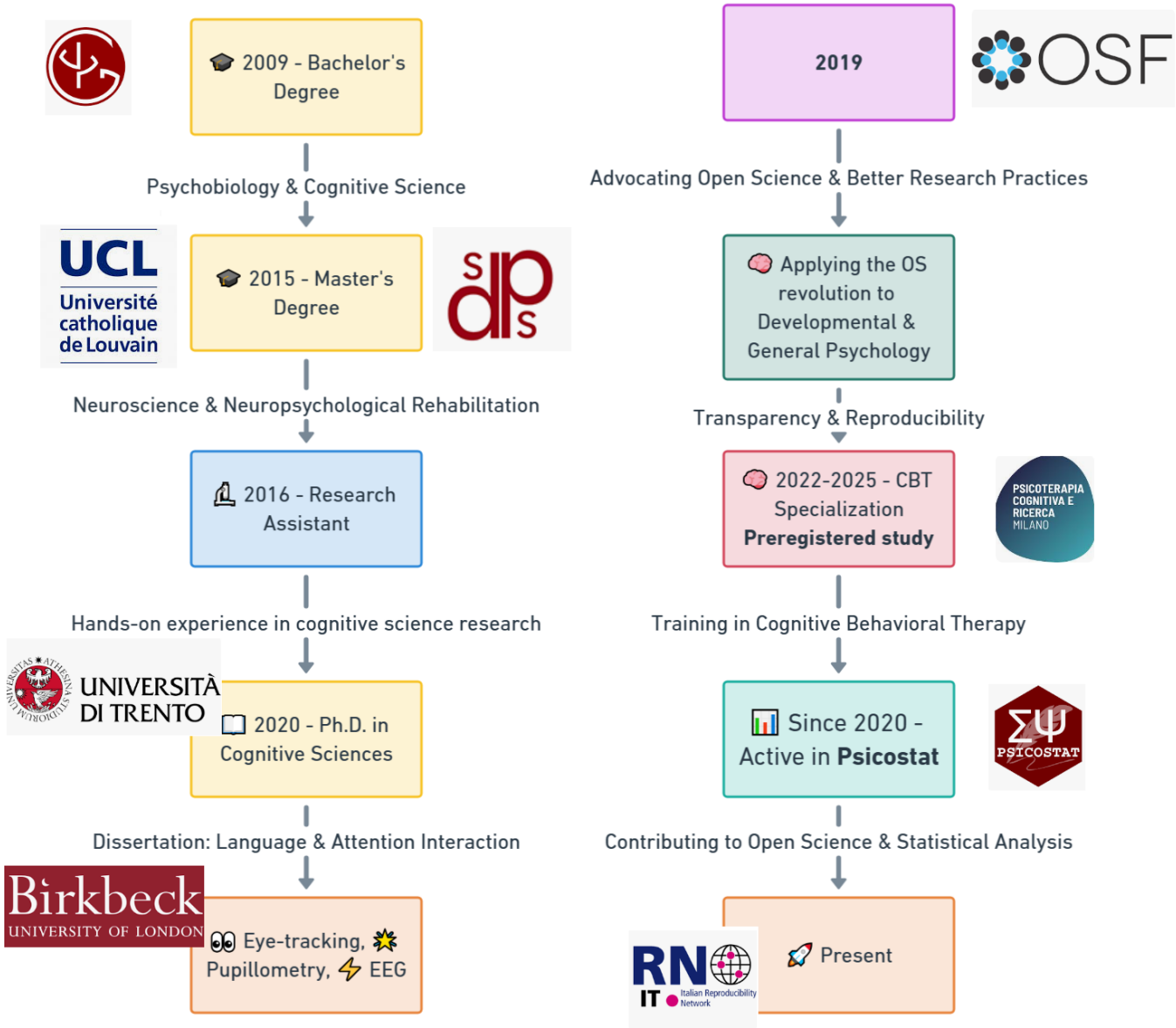
The MB2P Preregistration and Multiverse Forking Paths

Robustness in Multilab Research

Resources

My Academic Journey

Academic Journey



fun facts: I am in love with generalized non-linear mixed effects models and N-1

what I got

1. Secondary data analysis and preregistration: How far can we push simulation (data + code)? *Thanks to Daniel.*

what I got

2. Until what point is the measure robust enough to be considered valid? *Thanks to Jessica.*

what I got

3. Bringing together not only many people but also multiple traditions of data management and processing in a reproducible and informative fashion *Cumulative (multiplicative) knowledge?*
Thanks to Lisa and Filippo

what I got

4. Organized skepticism makes research meaningful *Thanks to Ettore*

Theoretical Introduction to Multiverse Analysis

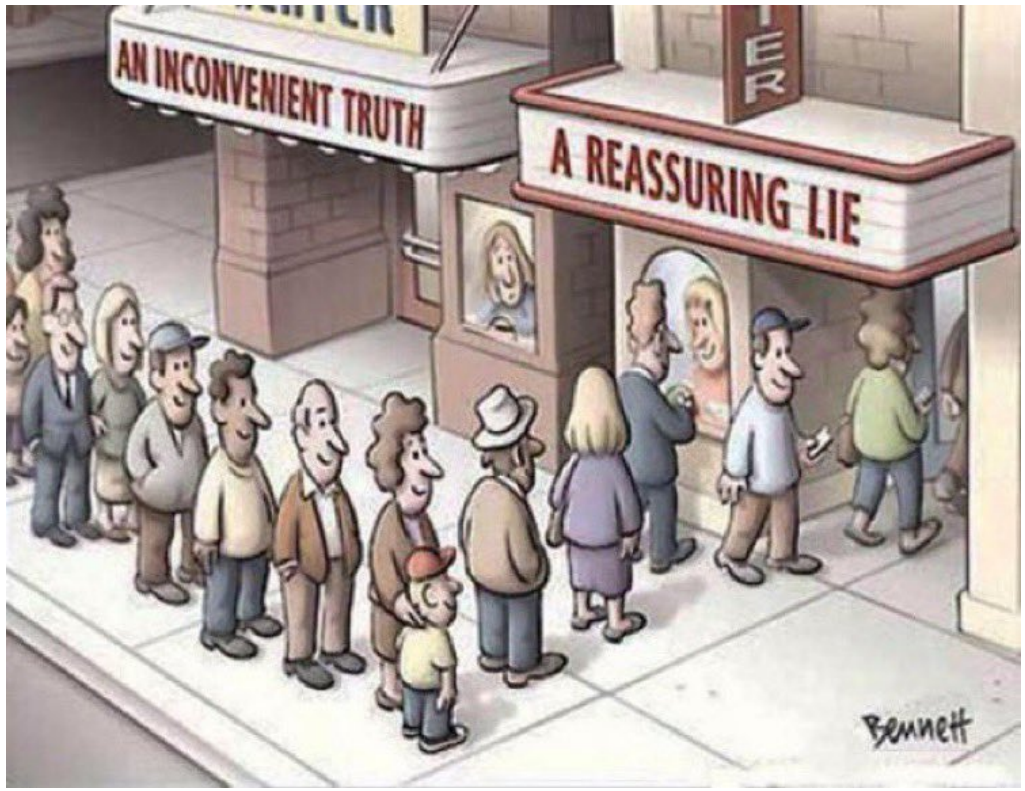
Any datum is not for granted

data

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

which means the 'thing given'

We rarely find data, we actively construct datasets



- A single data collection = a multiverse of possible datasets.

What is Multiverse Analysis?

Definition: A systematic approach to exploring all reasonable analytical paths in research.

- Addresses variability in research decisions.
- Enhances transparency and robustness.
- Provides a structured way to handle uncertainty.

Why Multiverse Analysis Matters

Traditional Analysis

- Focuses on a single analytical path.
- High risk of selective reporting.
- Limited transparency in decision-making.

Multiverse Analysis

- Explores multiple plausible paths.
- Reduces the risk of biased conclusions.
- Provides a more complete picture of data.

Why the Multiverse Matters

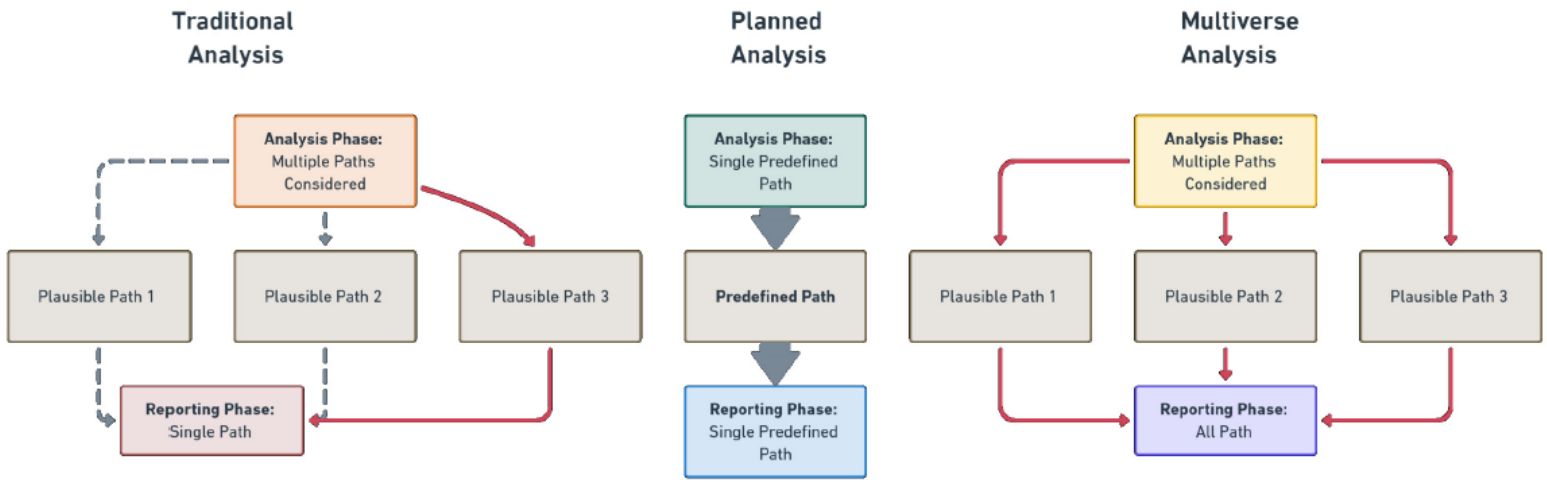


Fig. 1.1 Adapted from Dragicevic et al. (2019), this figure illustrates the relationship between analysis and reporting across three strategies: (1) Traditional Analysis considers multiple branches during analysis but reports only a single selected path; (2) Planned Analysis focuses exclusively on a single predefined branch for both analysis and reporting; and (3) Multiverse Analysis includes plausible branches in the analysis and reports the outcomes for all branches, providing the most transparent approach. The branching structure emphasizes the extent of choices considered and their inclusion in the reported outcomes.

Historical Context: Key Developments

- **2015:** Vibration of Effects (VoE) –Patel et al.
- **2016:** Multiverse Analysis introduced –Steegen et al.
- **2024:** Post-selection Inference in Multiverse Analysis (PIMA) – Girardi et al.

Methodological Approaches

Exploratory Multiverse Analysis

- Identifies and explores patterns.
- Generates hypotheses.
- Flexible and adaptive.

Inferential Multiverse Analysis (PIMA)

- Rigorously tests hypotheses.
- Controls error rates.
- Ensures robustness in statistical inference.

Key Statistical Concepts

Vibration of Effects (VoE)

- Summarizes variability across specifications.
- Uses graphical representations.
- Identifies unstable findings.

Post-selection Inference (PIMA)

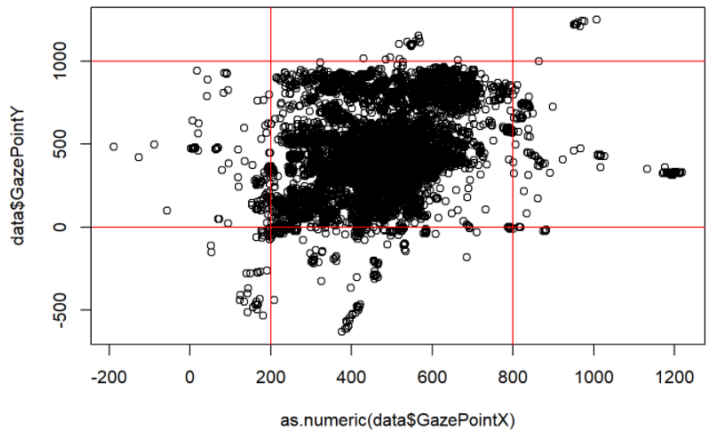
- Controls Type I error rates.
- Uses sign-flipping score tests.
- Balances statistical power and error control.

Degrees of Freedom and Forking Paths

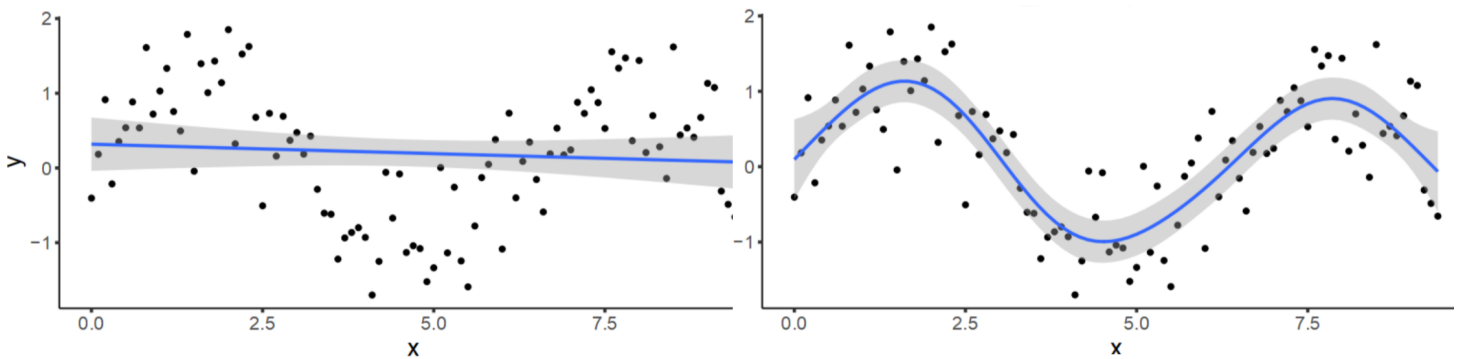
Every decision (data filtering, AOI selection, statistical model choice) introduces degrees of freedom that can significantly affect results.

Degrees of freedom in pupil analysis

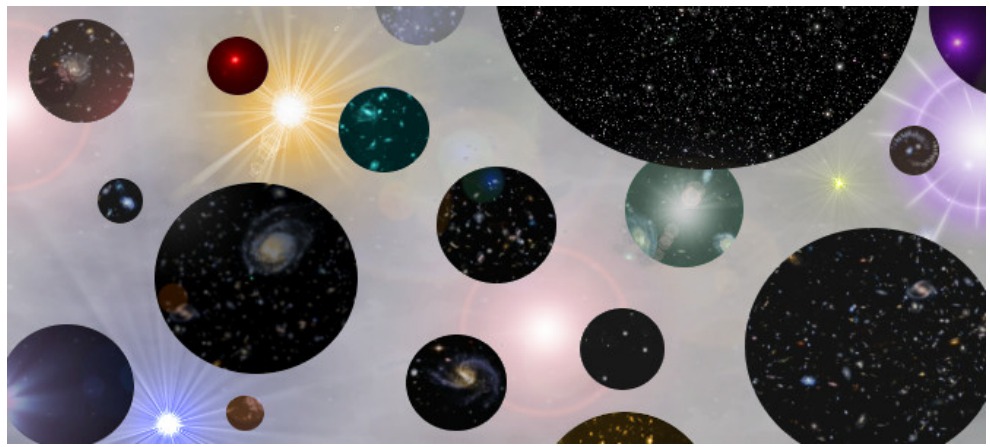
- Area of Interest (AOI) and implausible values e.g. outliers



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



A Multiverse approach



- **a philosophy of statistical reporting** in the manuscript (or 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).

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

The Multiverse Multilab is a Thing

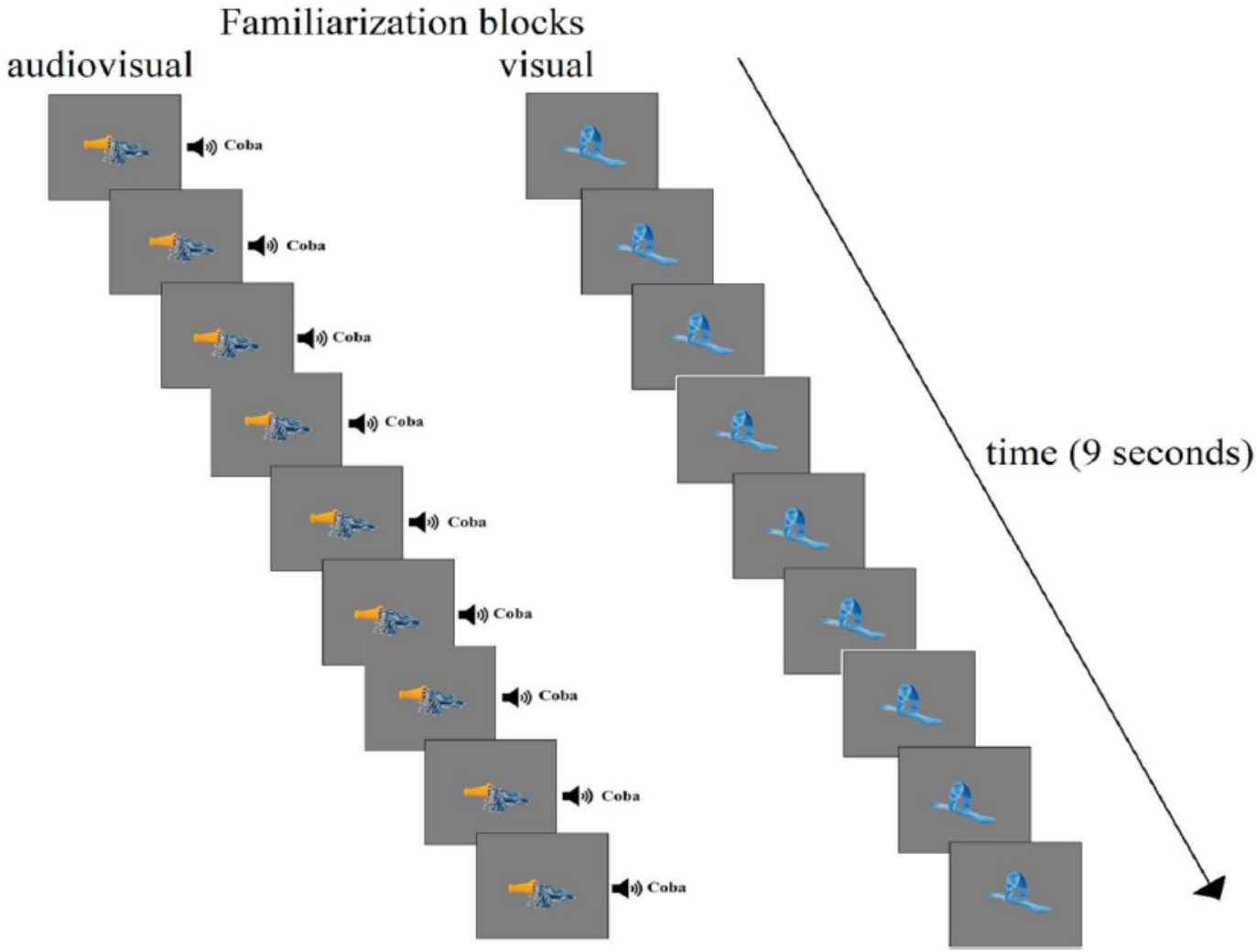
Concrete Example of RDF

The Challenges of Infancy research

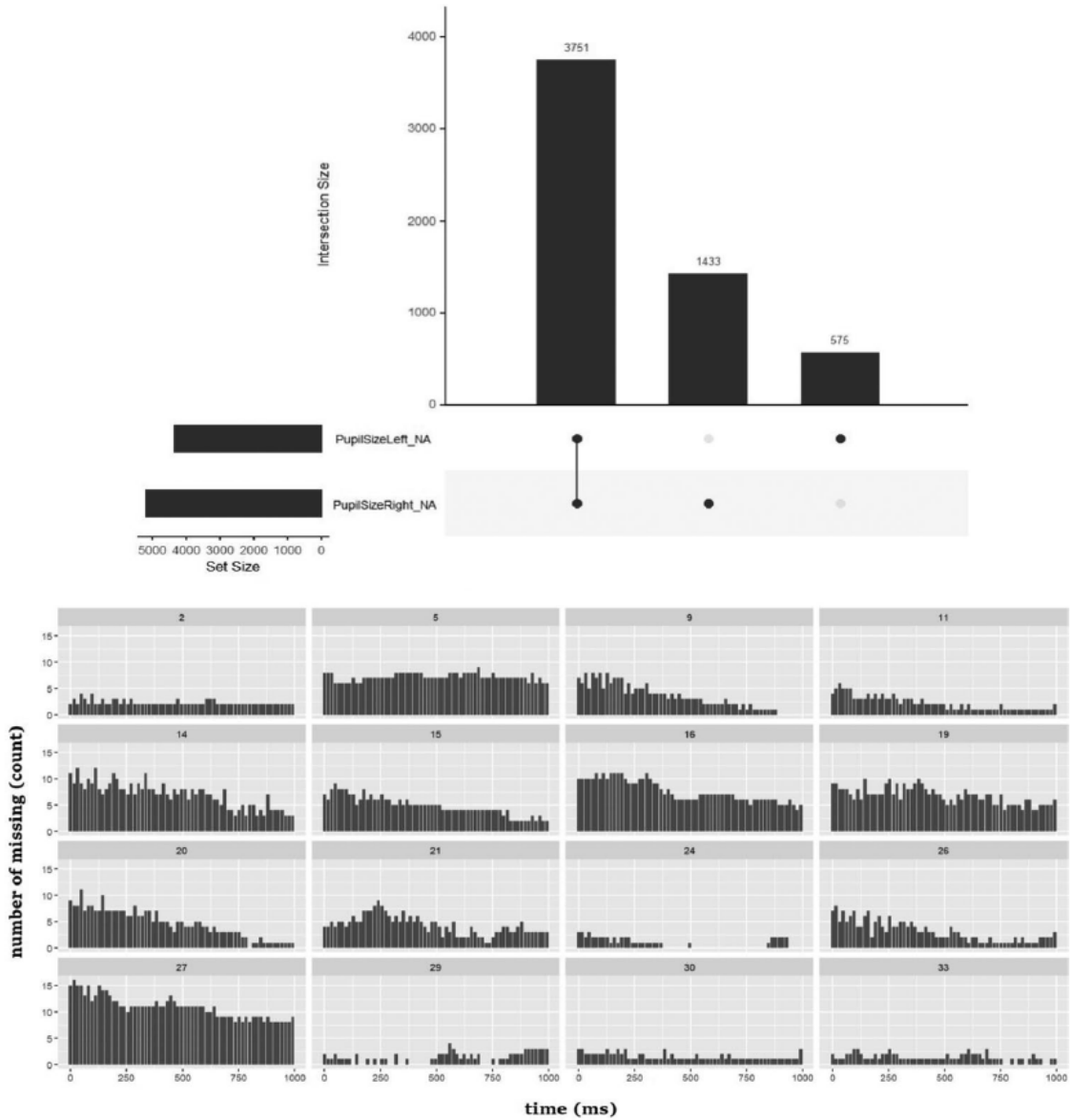
Case Study Pupil measurements within a multiverse framework involves systematic variations in pre-processing and statistical modeling.

The garden of forking paths

An illustrative example



Are missing... data?



Degrees of freedom (1) Data filtering

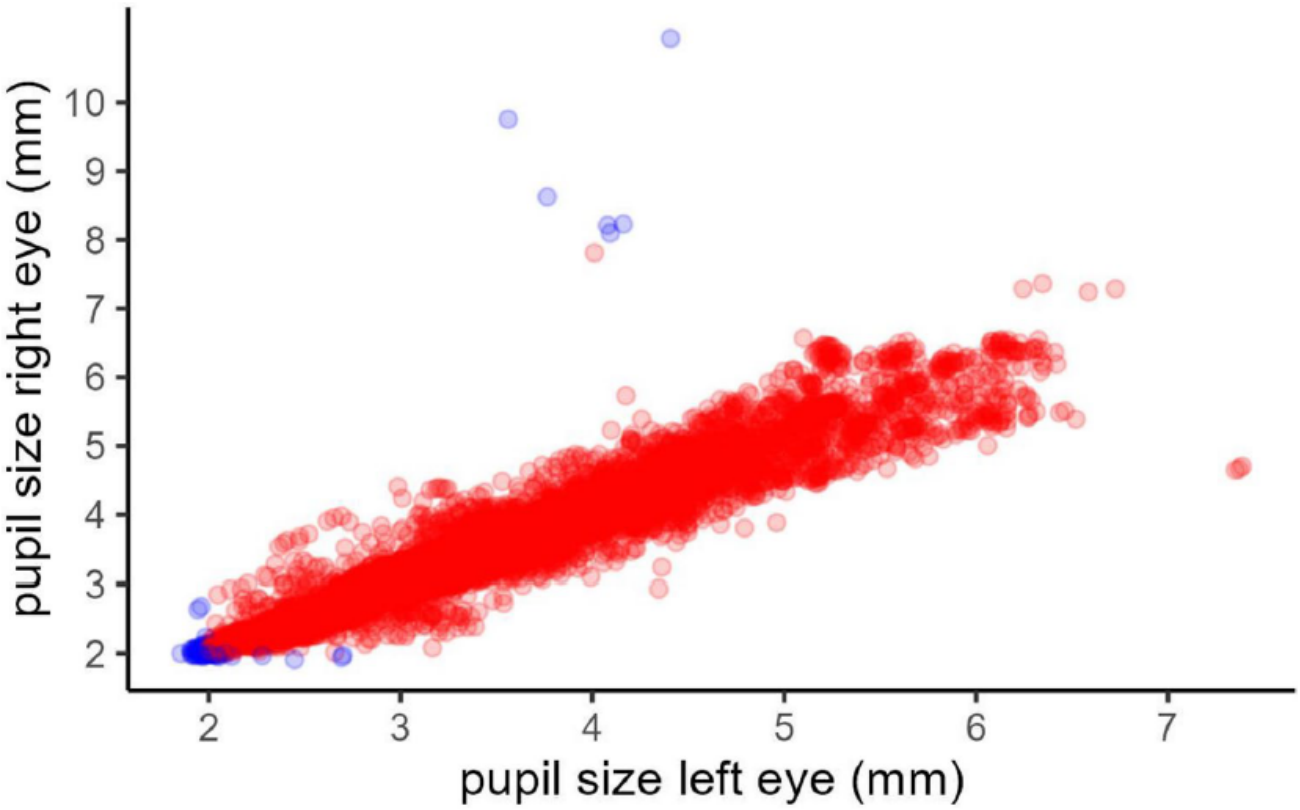
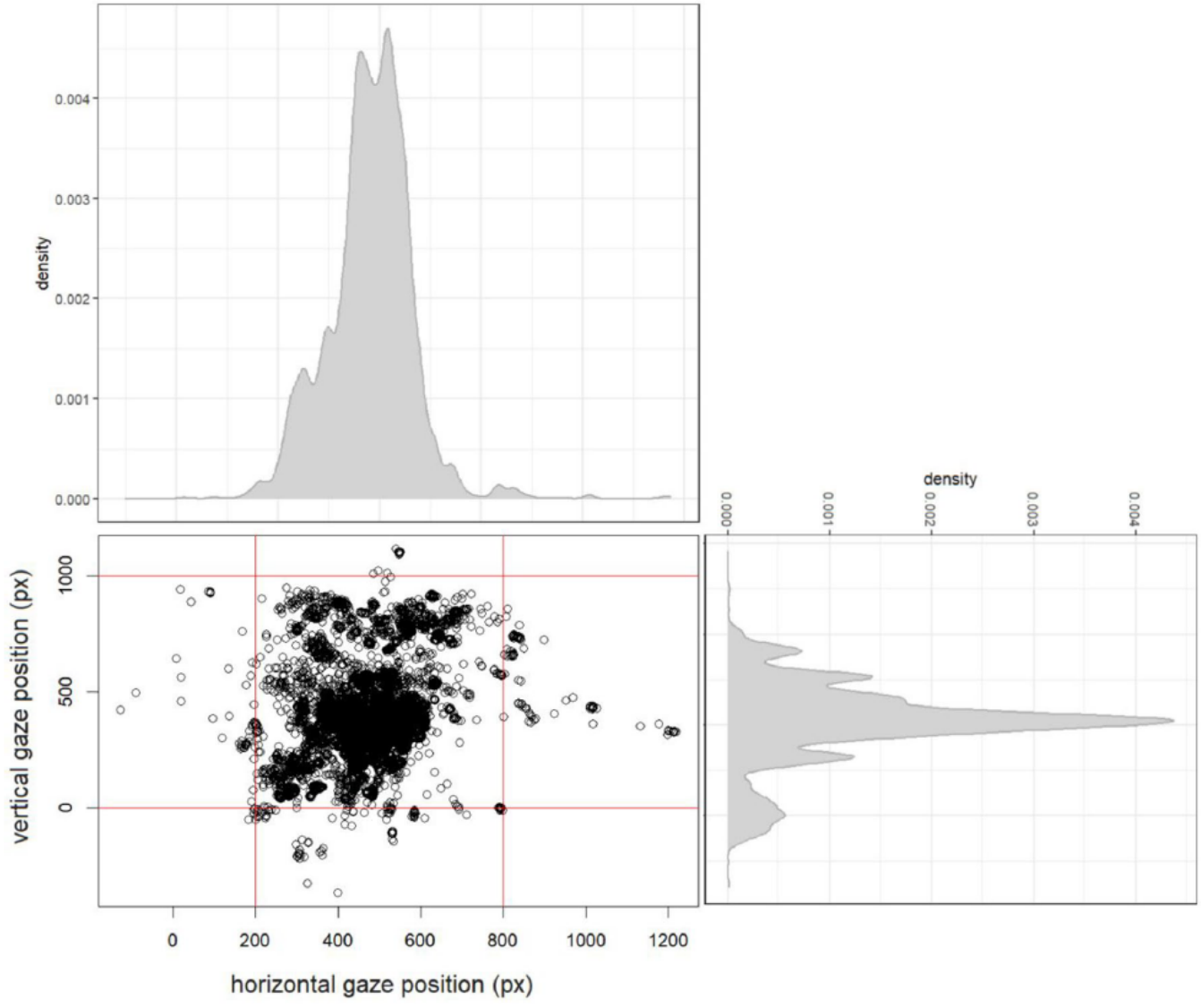


Fig. 3 Scatter plot correlating left and right eye's pupil size. *Blue points* indicate the values excluded in the second filtered dataset (trimmed dataset)

Degrees of freedom (2) Area of interest



Degrees of freedom (3) Dealing with blinks

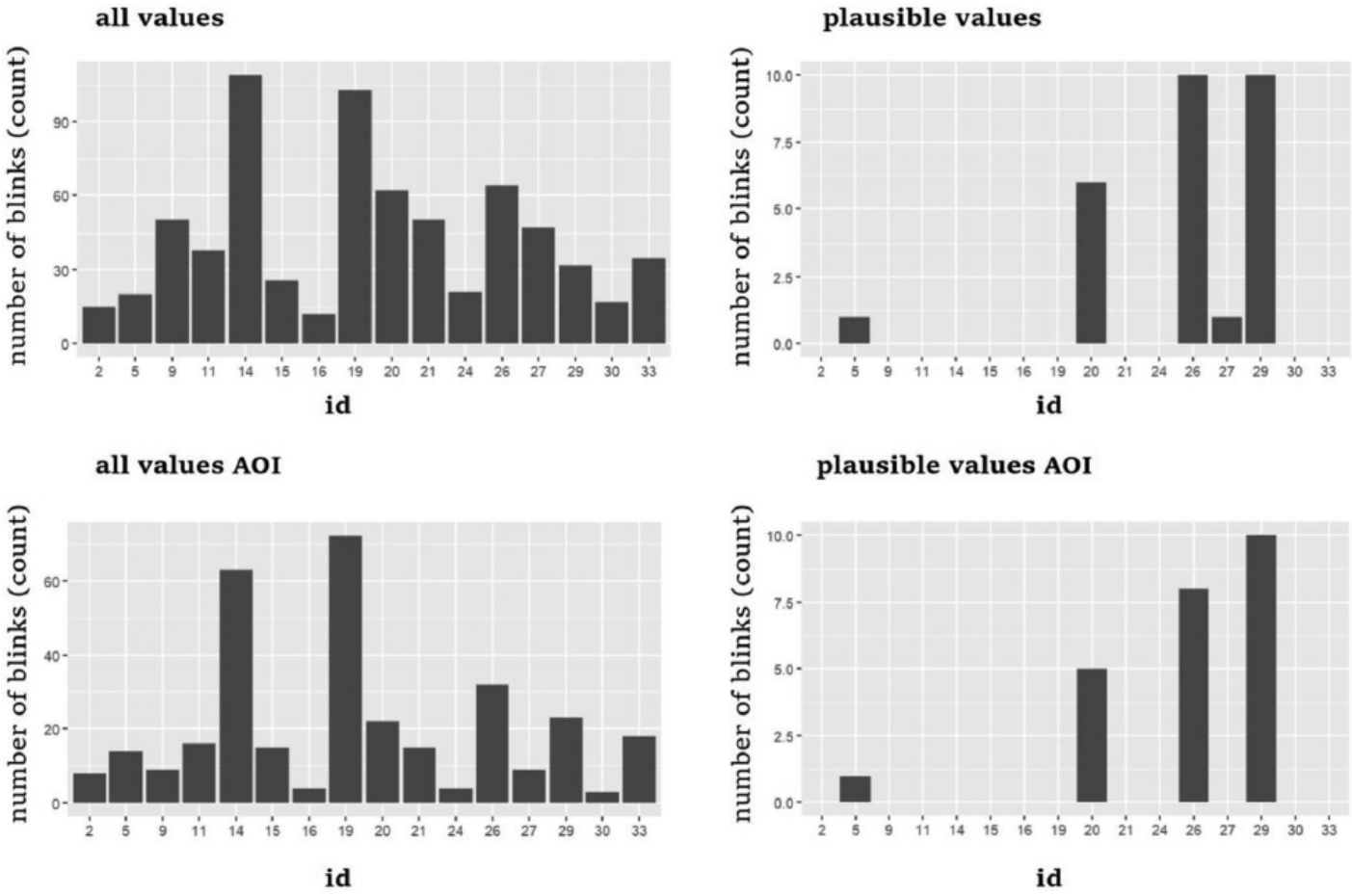


Fig. 5 Distribution of blink detected across the participants (id) in the four datasets, i.e., all values in the whole screen, plausible values in the whole screen, all values within the area of interest (AOI), plausible values within the AOI

Degrees of freedom (4) Baseline correction

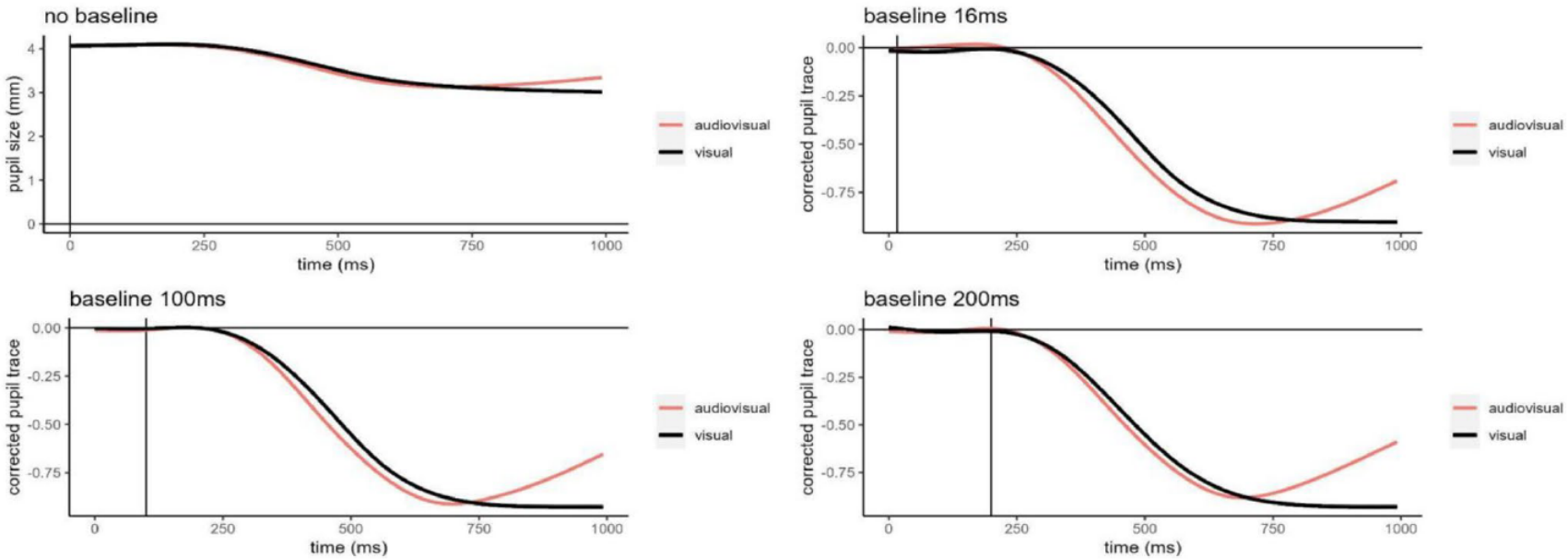


Fig. 6 Average pupil size variation (no baseline) and pupil changes relative to baseline (16, 100, and 200 ms) smoothed across time, we used the dataset with trimmed values filtered by the AoI and interpolated blinks for illustrative purposes. The *red* and *black lines* represent the audiovisual and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

Degrees of freedom (5) Participants inclusion

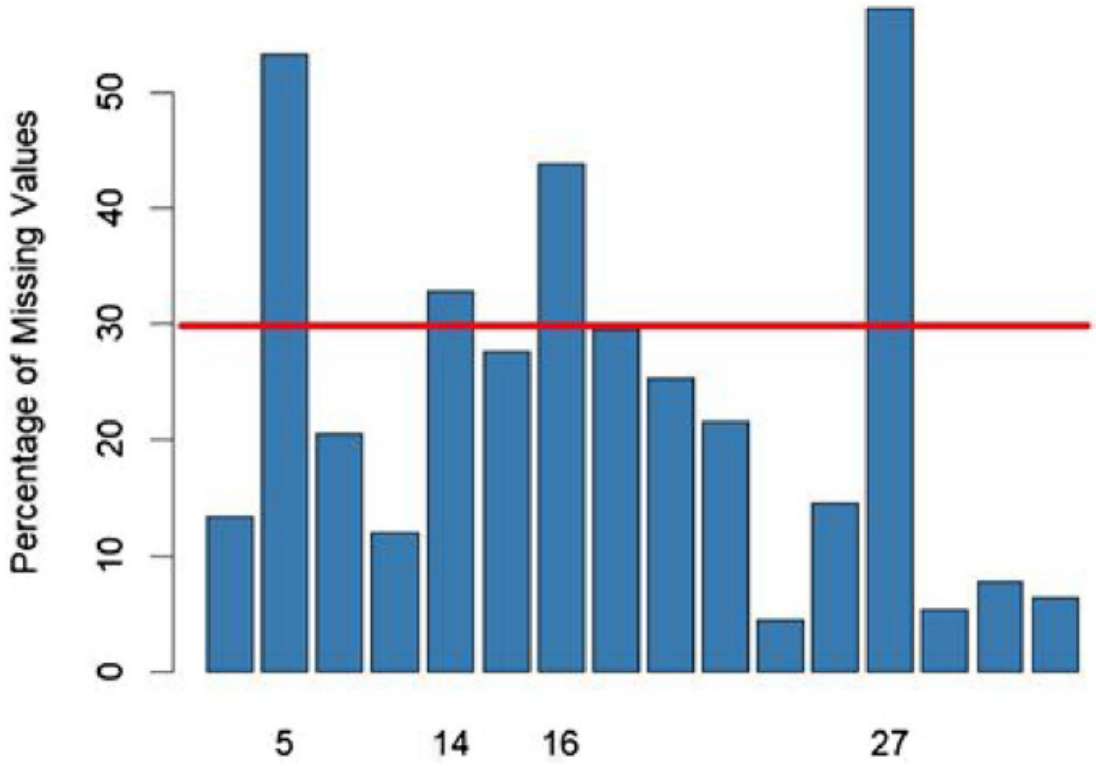


Fig. 7 Percentage of missing values by subject (ID). The *red line* indicates the cut-off value. Note that all that only ID participants above the cut-off are shown in the *x*-axis

Degrees of freedom (6) A multiverse of models

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

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

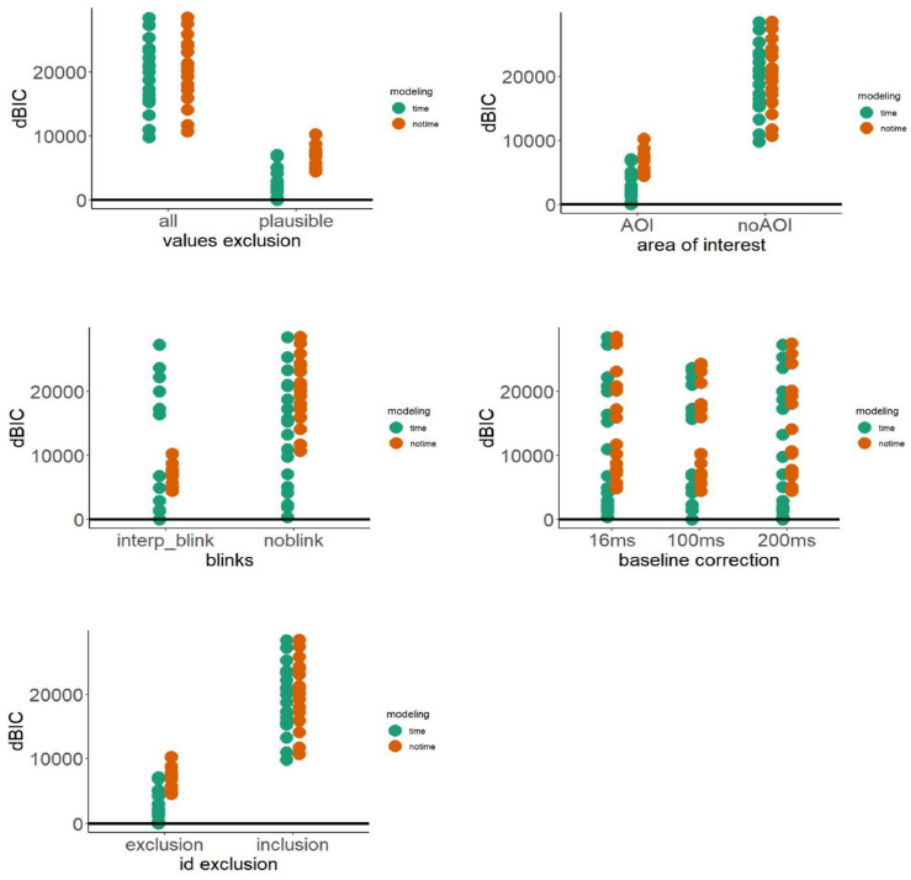
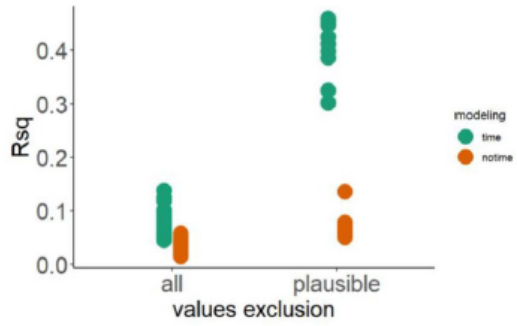
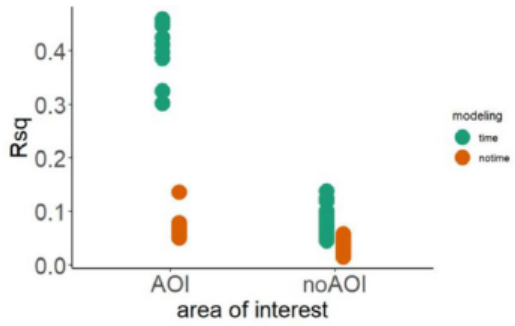
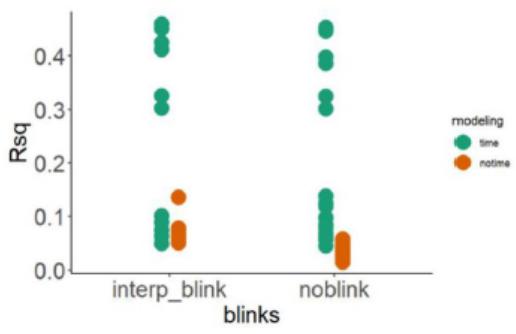


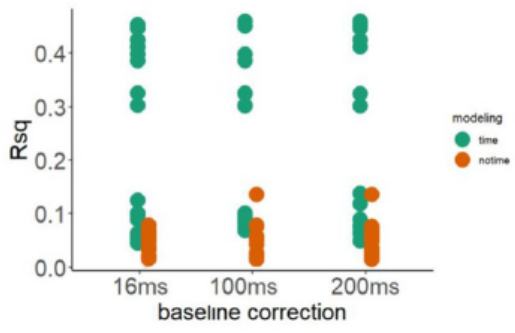
Fig. 8 The figure shows the delta BIC (the lower the better) and *R* squared of the two models (i.e., with and without smoother terms for time) for the 48 datasets. Plots are split by the first (extreme vs. trimmed values), the second (no AoI vs. AoI), the third (no blink vs. interpolated blink), the fourth (16-, 100-, and 200-ms baseline) and the fifth (participant inclusion and exclusion) degrees of freedom

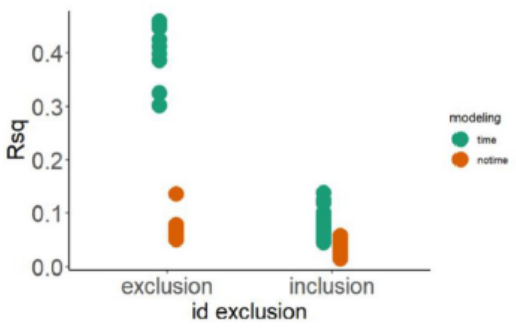
Degrees of freedom (6) A multiverse of models











Degrees of freedom (6) A multiverse of models

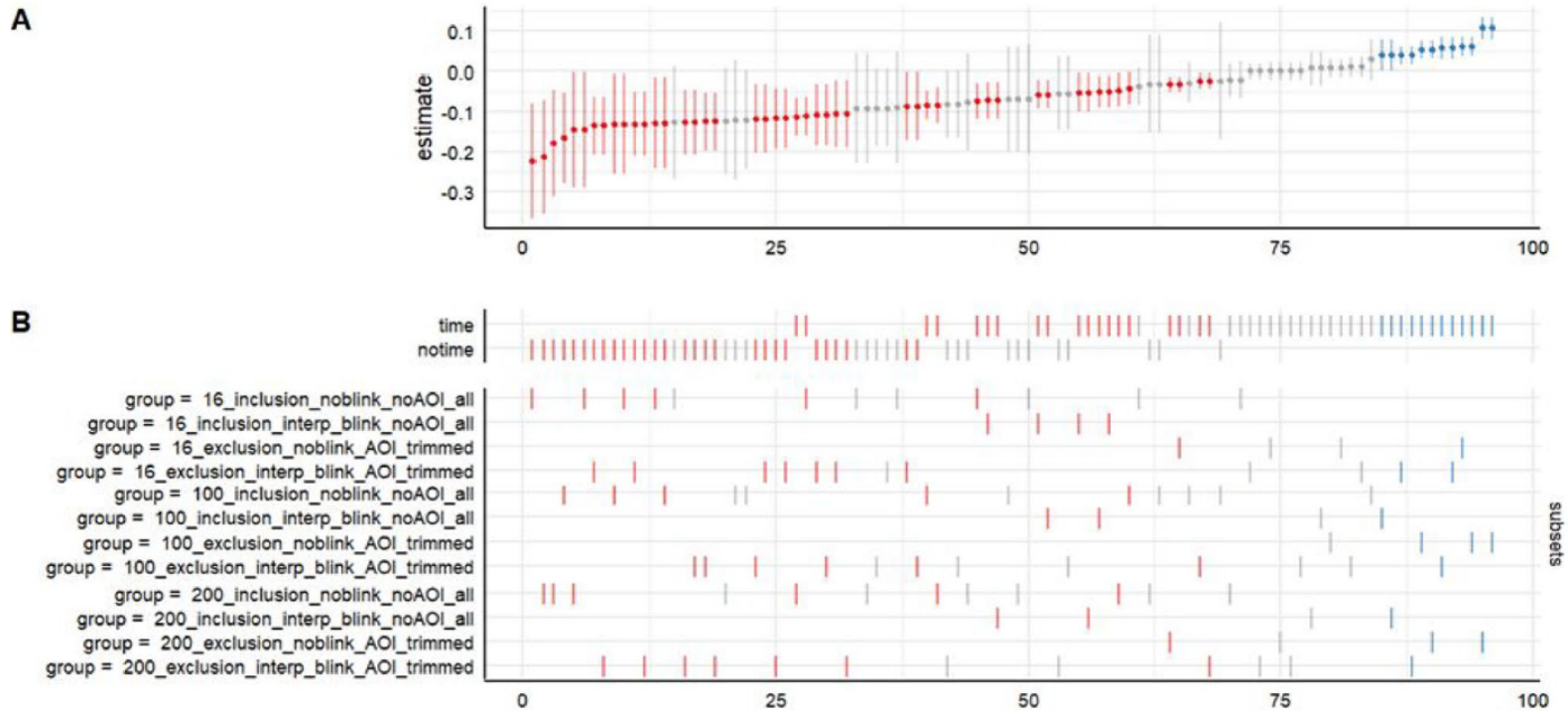
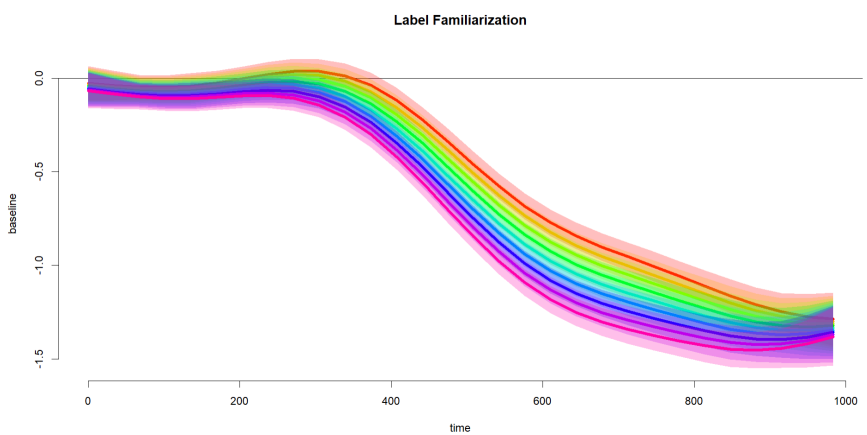
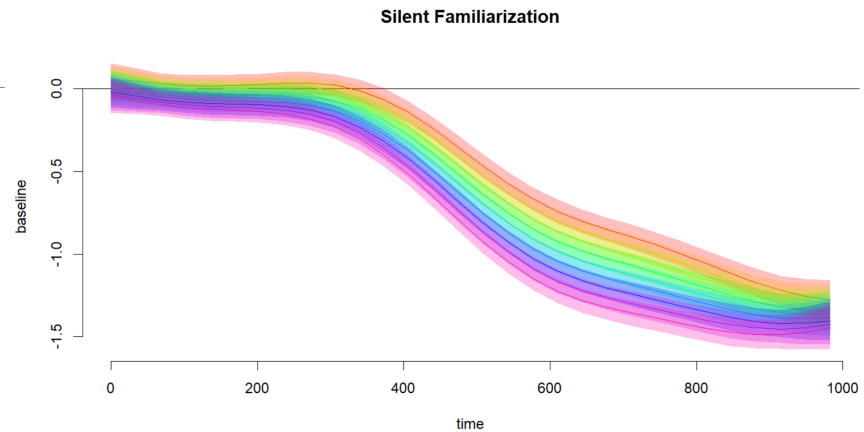


Fig. 9 **A** The 96 coefficient's estimates and relative 95% CI related to the Visual vs. Audiovisual regressor. **B** Relative combinations by the six degrees of freedom of the multiverse analysis. The direction of the significant results are highlighted (negative = *red*, positive = *blue*, *gray* = 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





Multiverse: Robustness and Generalizability

Nosek's (2018) Approach:

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

MB2 Pupillometry Spin Off

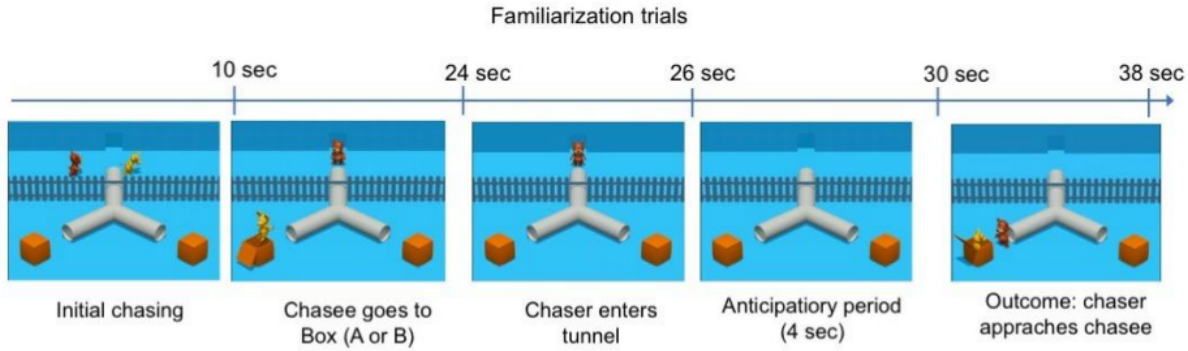
MB2 - Theory of Mind (ToM)

Stage 1 Registered Report

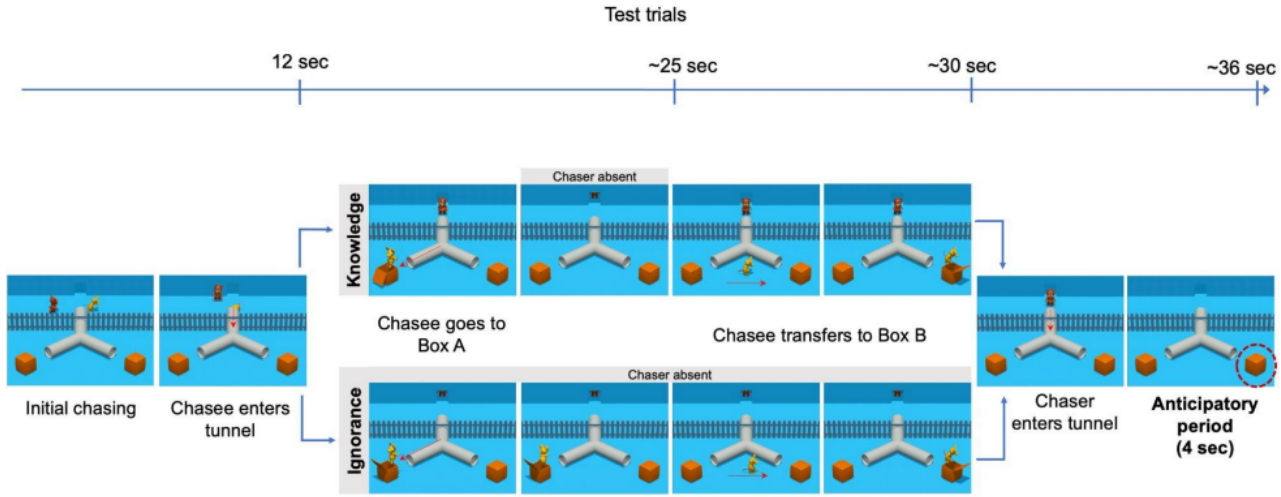
Schuwerk, T., Kampis, D.**, Baillargeon, R., Biro, S., Bohn, M., Byers-Heinlein, K., Dörrenberg, S., Fisher, C., Franchin, L., Fulcher, T., Garbisch, I., Geraci, A., Grosse Wiesmann, C., Hamlin, K., Haun, D. B. M., Hepach, R., Hunnius, S., Hyde, D. C., Karman, P., ... Rakoczy, H. (accepted pending data 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 - Paradigm

Timeline of the familiarization trials.



Schematic overview of stimuli and conditions of the test trials.



MB2P - Secondary Data Analysis

MB2P

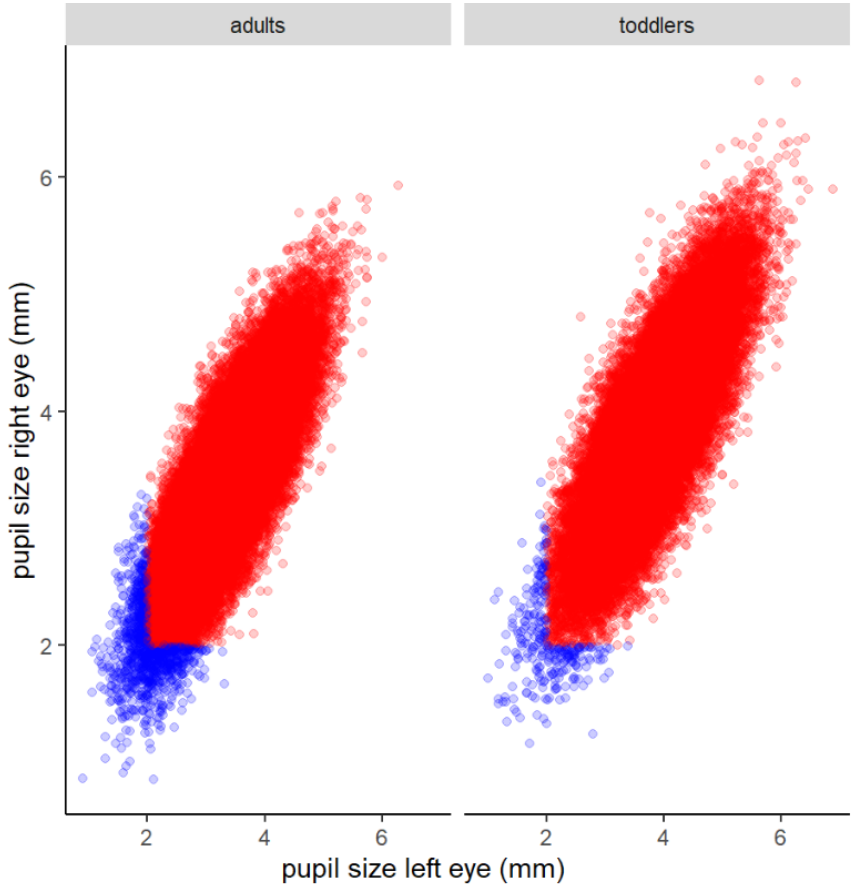
Measuring Pupil Dilation in Response to Expected and Unexpected Events

The MB2P Preregistration and Data Simulation

- **Data Simulation** to pre-register preprocessing and analysis strategies.
- Simulated data includes: **participant ID, age cohorts, timestamp/duration, (x,y) coordinates, pupil left/right size, lab ID, conditions, and outcomes**.

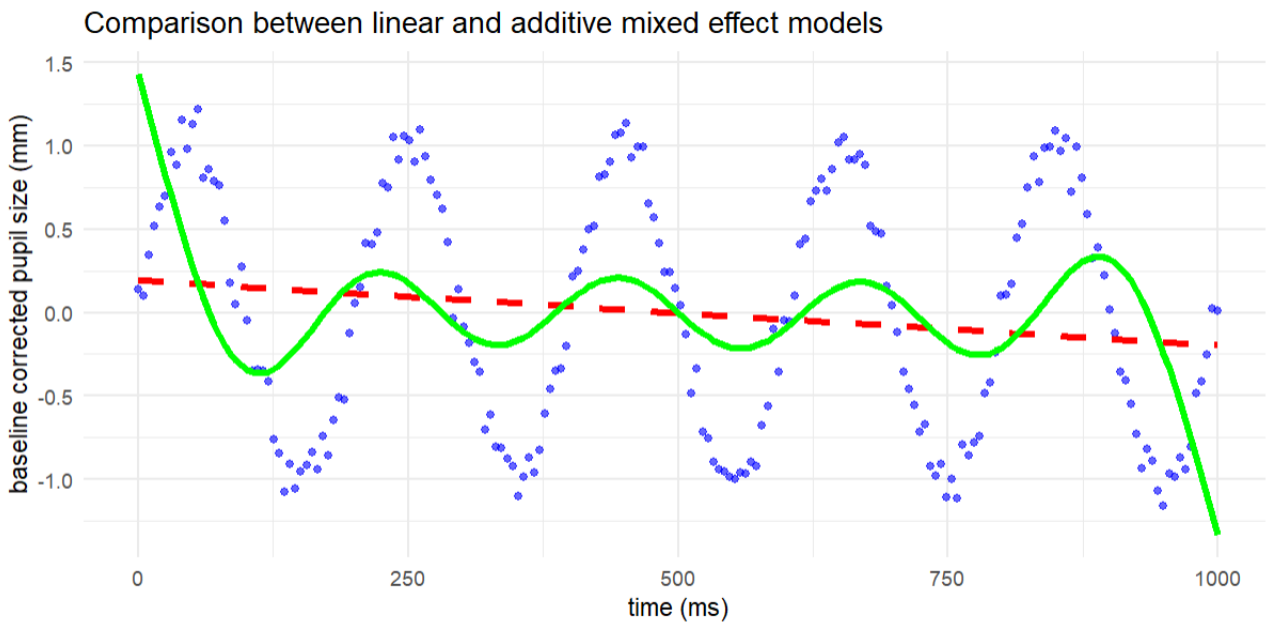
The MB2P Preregistration and Multiverse Forking Paths

Degrees of Freedom: Filtering Extreme Pupil Values



Statistical Modeling

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



Robustness in Multilab Research

Multiverse: Robustness and Generalizability

Nosek's (2018) Approach:

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

Why Robustness Matters

- No dataset is ever "perfect" $\rightarrow$ It is crucial to show the impact of preprocessing on results.
- Statistical models that evaluate *how* and *when* an effect emerges are more informative than a simple p-value.
- The multiverse approach enhances robustness and prevents p-hacking (by discoling it).

Resources

Resources

- Repository Multiverse Multilab Hands on 4Ms
- GitHub ManyBabies
- MB2P

References

- Dragicevic, P., Jansen, Y., Sarma, A., Kay, M., Chevalier, F. (2019). *Increasing the transparency of research papers with explorable multiverse analyses*. CHI Conference on Human Factors in Computing Systems.
- Girardi, P., Vesely, A., Lakens, D., Altoè, G., Pastore, M., Calcagnì, A., Finos, L. (2024). *Post-selection inference in multiverse analysis (PIMA)*. Psychometrika.
- Ioannidis, J. P. (2005). *Why most published research findings are false*. PLoS Medicine.
- Patel, C. J., Burford, B., Ioannidis, J. P. (2015). *Assessment of vibration of effects due to model specification*. Journal of Clinical Epidemiology.
- Steegen, S., Tuerlinckx, F., Gelman, A., Vanpaemel, W. (2016). *Increasing transparency through a multiverse analysis*. Perspectives on Psychological Science.

Thank you

Thank you for your attention!

UNIVERSITÀ DEGLI STUDI DI PADOVA