Beyond Literal Meaning

A Quantitative Meta-Analysis of N400 and Post-N400 Effects in Figurative Language through Waveform Digitization

P. Canal[1], F. Vespignani[2], V. Mangiaterra[1,3],
F. Luciani[1], F. Frau[1], L. Bischetti[1] & V. Bambini[1]

[1] *Laboratory of Neurolinguistics and Experimental Pragmatics (NEPLab), University School for Advanced Studies IUSS, Pavia, Italy*

[2] *Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Università di Padova, Padova, Italy*

[3] *Laboratoire de sciences cognitives et psycholinguistique, École normale supérieure ENS, Paris, France*

June 19, 2026

Motivation

Two routes of comprehension emerge from ERP research (Kuperberg, 2007; see also Brouwer & Hoecks, 2013) - at least:

- N400 $\rightarrow$ **semantic memory-based** stream comparing lexical information about the incoming words with information from semantic memory.
- P600 $\rightarrow$ **combinatorial** stream integrating words that build up the propositional meaning on the basis of multiple constraints and determine the final interpretation.

I am a P600 hunter

- My career was devoted to searching post-N400 effects in figurative language
- The size of these effects is relatively tiny

Motivation

Two routes of comprehension emerge from ERP research (Kuperberg, 2007; see also Brouwer & Hoecks, 2013) - at least:

- N400 $\rightarrow$ **semantic memory-based** stream comparing lexical information about the incoming words with information from semantic memory.
- P600 $\rightarrow$ **combinatorial** stream integrating words that build up the propositional meaning on the basis of multiple constraints and determine the final interpretation.

I am a P600 hunter

- My career was devoted to searching post-N400 effects in figurative language
- The size of these effects is relatively tiny

Motivation

Two routes of comprehension emerge from ERP research (Kuperberg, 2007; see also Brouwer & Hoecks, 2013) - at least:

- N400 $\rightarrow$ **semantic memory-based** stream comparing lexical information about the incoming words with information from semantic memory.
- P600 $\rightarrow$ **combinatorial** stream integrating words that build up the propositional meaning on the basis of multiple constraints and determine the final interpretation.

I am a P600 hunter

- My career was devoted to searching post-N400 effects in figurative language
- The size of these effects is relatively tiny

Motivation

Two routes of comprehension emerge from ERP research (Kuperberg, 2007; see also Brouwer & Hoecks, 2013) - at least:

- N400 $\rightarrow$ **semantic memory-based** stream comparing lexical information about the incoming words with information from semantic memory.
- P600 $\rightarrow$ **combinatorial** stream integrating words that build up the propositional meaning on the basis of multiple constraints and determine the final interpretation.

I am a P600 hunter

- My career was devoted to searching post-N400 effects in figurative language
- The size of these effects is relatively tiny

ERPs in figurative language

From a pragmatic perspective (Relevance Theory) figurative language involves a) conceptual adjustments, and b) in the derivation of the implicature. We like to tie conceptual adjustments to the N400 and implicature derivation to post-N400 effects.

- **Irony vs. Metaphor:**
  - In metaphor both conceptual adjustment and inference making are needed.
  - In irony the need of deriving the correct implicature is prominent.
- The interplay of N400 and P600 components dominates the literature, however, the P600 is less ubuquitous than the N400: available counts say it amounts to 20% of metaphor studies, and post-N400 effects may also take the shape of sustained negativities (Baiocco, Kielar & Lai, 2026).

ERPs in figurative language

From a pragmatic perspective (Relevance Theory) figurative language involves a) conceptual adjustments, and b) in the derivation of the implicature. We like to tie conceptual adjustments to the N400 and implicature derivation to post-N400 effects.

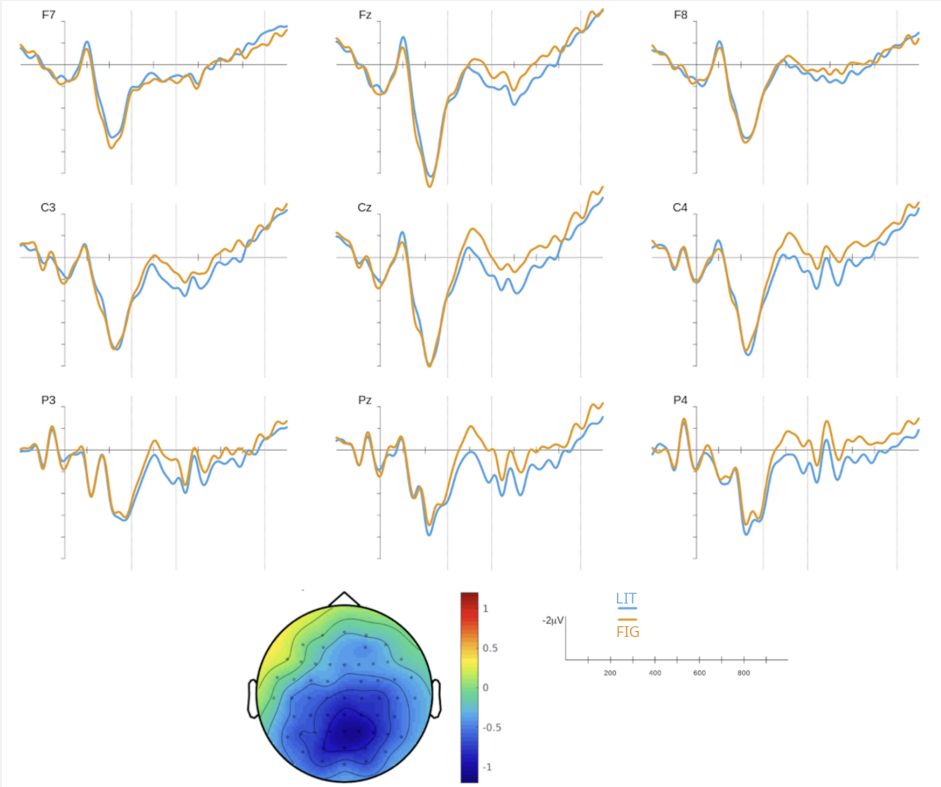
- **Irony vs. Metaphor:**
  - In metaphor both conceptual adjustment and inference making are needed.
  - In irony the need of deriving the correct implicature is prominent.
- The interplay of N400 and P600 components dominates the literature, however, the P600 is less ubuquitous than the N400: available counts say it amounts to 20% of metaphor studies, and post-N400 effects may also take the shape of sustained negativities (Baiocco, Kielar & Lai, 2026).

ERPs in figurative language

From a pragmatic perspective (Relevance Theory) figurative language involves a) conceptual adjustments, and b) in the derivation of the implicature. We like to tie conceptual adjustments to the N400 and implicature derivation to post-N400 effects.

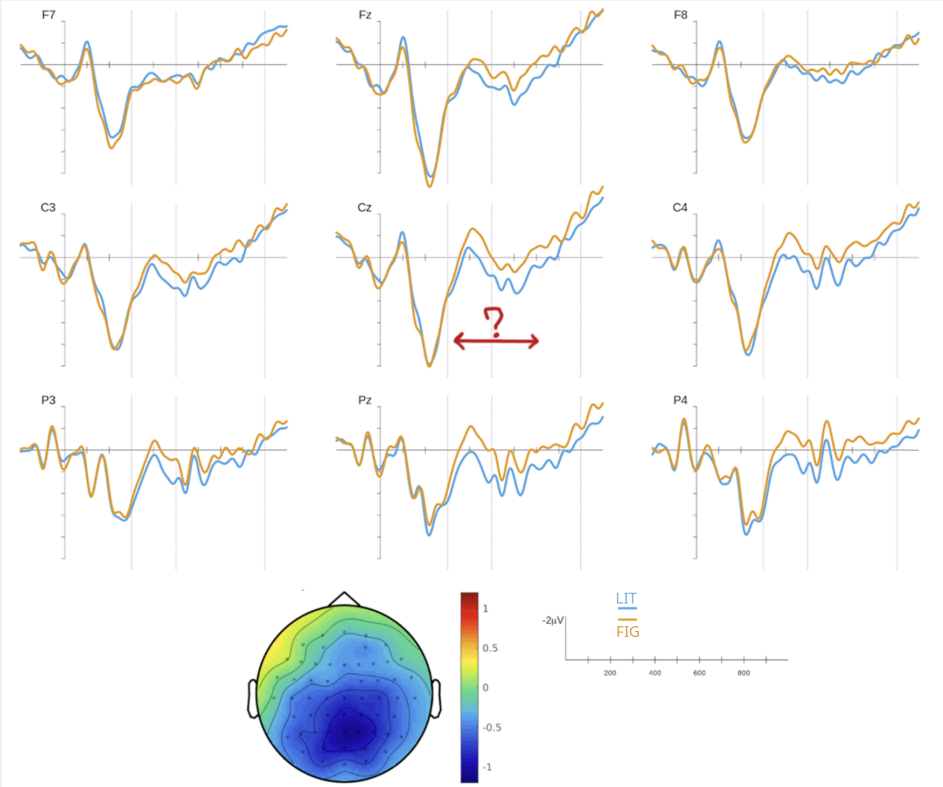
- **Irony vs. Metaphor:**
  - In metaphor both conceptual adjustment and inference making are needed.
  - In irony the need of deriving the correct implicature is prominent.
- The interplay of N400 and P600 components dominates the literature, however, the P600 is less ubuquitous than the N400: available counts say it amounts to 20% of metaphor studies, and post-N400 effects may also take the shape of sustained negativities (Baiocco, Kielar & Lai, 2026).

Why quantitative meta-analyses on ERPs may fail



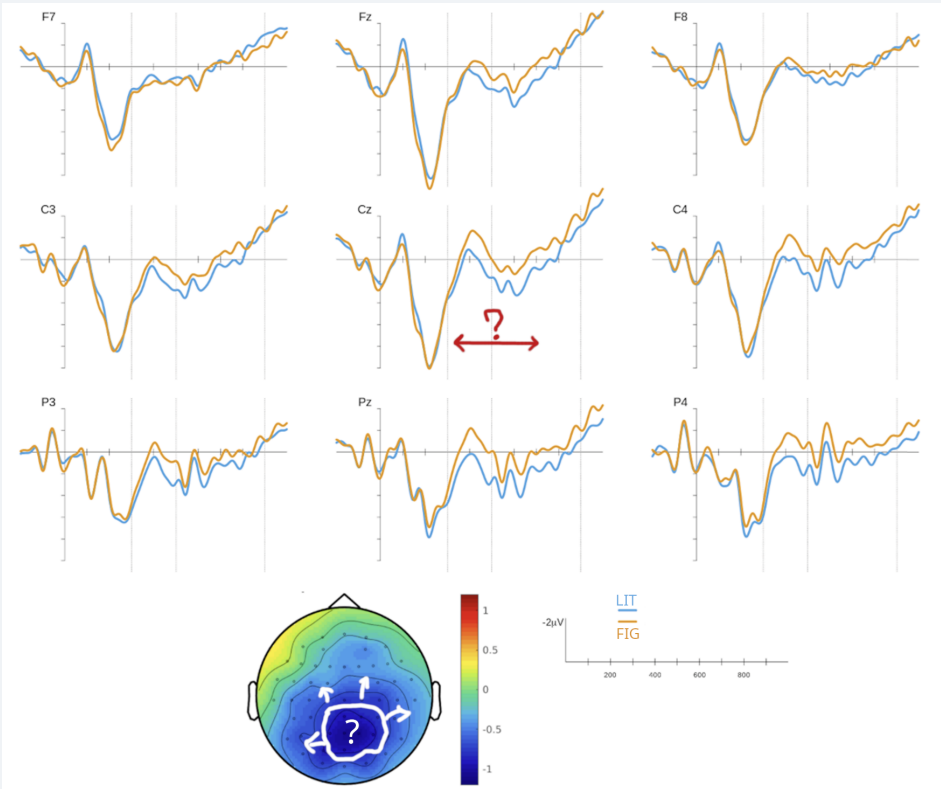
Why quantitative meta-analyses on ERPs may fail

- **Time Windows:** Highly variable across studies based on authors' choices.
- **Missing Data:** Depending on the component of interest, some time windows may be untested.



Why quantitative meta-analyses on ERPs may fail

- **Time Windows:** Highly variable across studies based on authors' choices.
- **Missing Data:** Depending on the component of interest, some time windows may be untested.
- **Topography:** Different subsets of electrodes are selected, hindering direct comparison.
- **Effect Sizes:** Rarely reported [Clayson, Carbine, Baldwin & Larson, 2019]



A Novel Approach: The Great Grand Average

We applied a novel methodology (Vespignani, 2020):

- Similar to the **Great Grand Average** approach (Sambrook & Goslin, 2015; Moran et al., 2017).
  - There are only three works using this approach [Sambrook & Goslin, 2015; Stewardson & Sambrook, 2020;Liu, Wang, Gozli,Xiang & Jackson, 2020]
- **Core Concept:** Digitizing the ERP waveforms depicted in the published figures of eligible papers.
- **Result:** Extraction of coordinates for each waveform, allowing for a coherent re-analysis of ERP data across studies using a uniform standard [consistent selection of temporal and spatial points].

A Novel Approach: The Great Grand Average

We applied a novel methodology (Vespignani, 2020):

- Similar to the **Great Grand Average** approach (Sambrook & Goslin, 2015; Moran et al., 2017).
  - There are only three works using this approach [Sambrook & Goslin, 2015; Stewardson & Sambrook, 2020;Liu, Wang, Gozli,Xiang & Jackson, 2020]
- **Core Concept:** Digitizing the ERP waveforms depicted in the published figures of eligible papers.
- **Result:** Extraction of coordinates for each waveform, allowing for a coherent re-analysis of ERP data across studies using a uniform standard [consistent selection of temporal and spatial points].

A Novel Approach: The Great Grand Average

We applied a novel methodology (Vespignani, 2020):

- Similar to the **Great Grand Average** approach (Sambrook & Goslin, 2015; Moran et al., 2017).
  - There are only three works using this approach [Sambrook & Goslin, 2015; Stewardson & Sambrook, 2020;Liu, Wang, Gozli,Xiang & Jackson, 2020]
- **Core Concept:** Digitizing the ERP waveforms depicted in the published figures of eligible papers.
- **Result:** Extraction of coordinates for each waveform, allowing for a coherent re-analysis of ERP data across studies using a uniform standard [consistent selection of temporal and spatial points].

A Novel Approach: The Great Grand Average

We applied a novel methodology (Vespignani, 2020):

- Similar to the **Great Grand Average** approach (Sambrook & Goslin, 2015; Moran et al., 2017).
  - There are only three works using this approach [Sambrook & Goslin, 2015; Stewardson & Sambrook, 2020;Liu, Wang, Gozli,Xiang & Jackson, 2020]
- **Core Concept:** Digitizing the ERP waveforms depicted in the published figures of eligible papers.
- **Result:** Extraction of coordinates for each waveform, allowing for a coherent re-analysis of ERP data across studies using a uniform standard [consistent selection of temporal and spatial points].

Systematic review

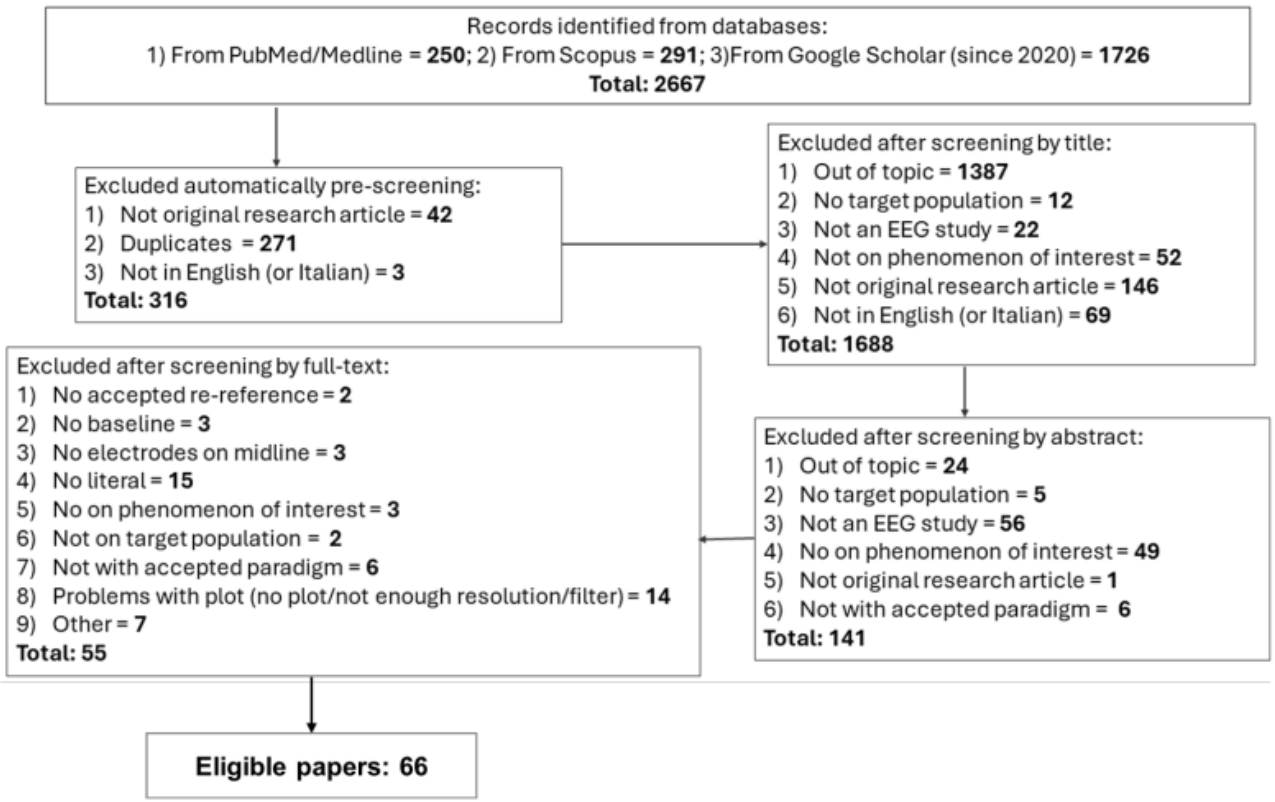
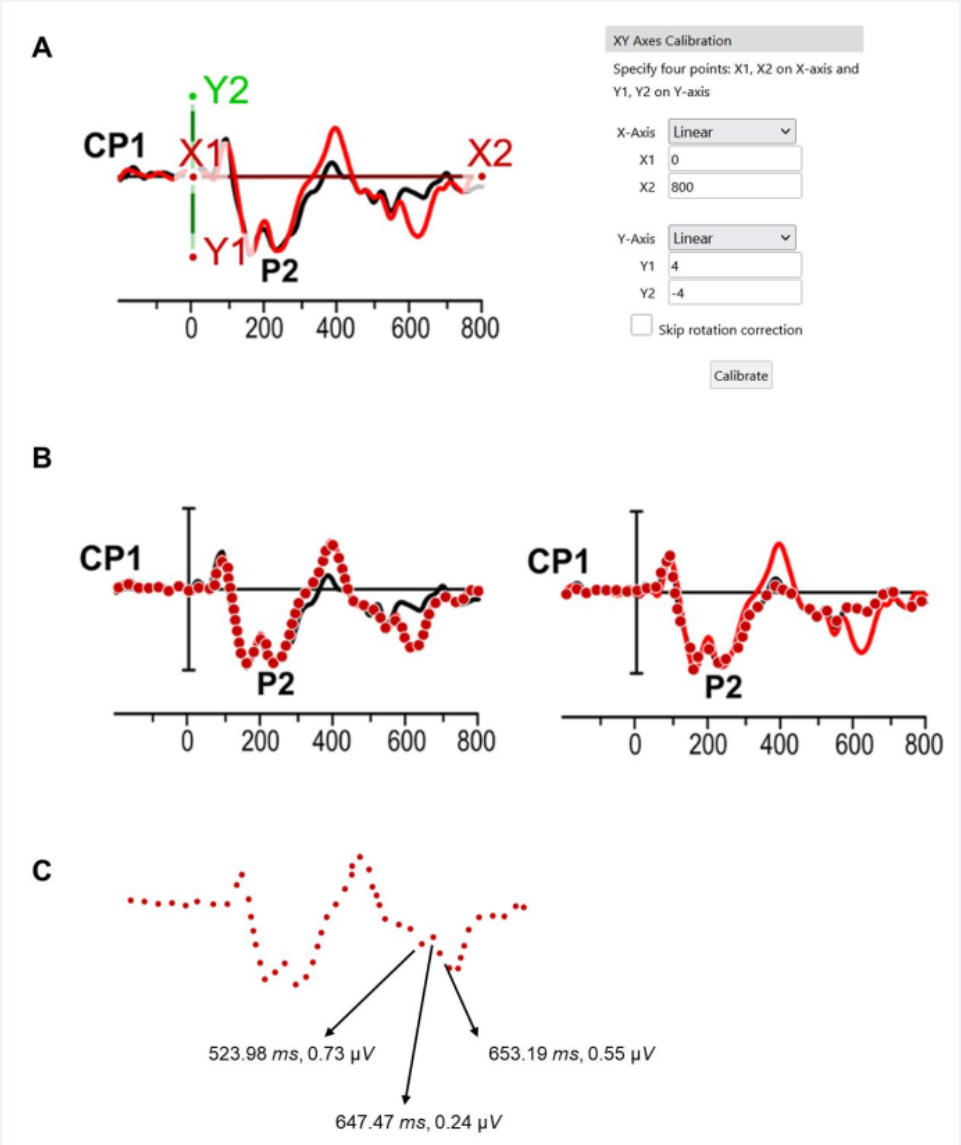


Figure 5.2. Steps of the systematic review. The figure reports the flowchart of paper selection for inclusion into the systematic review.

This amounts to 66 papers, with 73 experiments, on 1837 participants, all with digitizable ERP waveforms and literal controls.

Digitization Process in Web Plot Digitizer



Data Preprocessing: Interpolation & Spatial ROIs

- **Interpolation:** Linear interpolation of voltage between digitized sampling points [to avoid different sampling densities].
- **Spatial Filtering:** Data mapped to discrete Regions of Interest (ROI) in a common 10-05 bidimensional space.
  - **Frontal ROI:** 0.3 to 1.3 z points in the Y-axis.
  - **Posterior ROI:** -0.1 to -1.5 z points in the Y-axis.
  - **Lateral Spread:** Restricted to $|x| < 0.75$ to focus on midline activity.

Data Preprocessing: Interpolation & Spatial ROIs

- **Interpolation:** Linear interpolation of voltage between digitized sampling points [to avoid different sampling densities].
- **Spatial Filtering:** Data mapped to discrete Regions of Interest (ROI) in a common 10-05 bidimensional space.
  - **Frontal ROI:** 0.3 to 1.3 z points in the Y-axis.
  - **Posterior ROI:** -0.1 to -1.5 z points in the Y-axis.
  - **Lateral Spread:** Restricted to $|x| < 0.75$ to focus on midline activity.

Accounting for Study Uncertainty

- Standard ERP studies lack consistent effect size descriptions.
- To perform statistical modeling, we needed to account for the relative uncertainty of each original study.
- **Formula for the contrast between two conditions:**

$$SE_{reconstructed} = \sqrt{2 \cdot \left(\frac{V_{subj}}{N} + \frac{V_{item}}{K} + \frac{V_{resid}}{N \cdot K} \right)}$$

- **Uncertainty:** $SE_{reconstructed}$ used variance and residual from Nieuwland et al (2018) multilab study. We run a sensitivity analysis testing different ratios between by-subj and by-item variance, and the amount of residual error.

Accounting for Study Uncertainty

- Standard ERP studies lack consistent effect size descriptions.
- To perform statistical modeling, we needed to account for the relative uncertainty of each original study.
- **Formula for the contrast between two conditions:**

$$SE_{reconstructed} = \sqrt{2 \cdot \left(\frac{V_{subj}}{N} + \frac{V_{item}}{K} + \frac{V_{resid}}{N \cdot K} \right)}$$

- **Uncertainty:** $SE_{reconstructed}$ used variance and residual from Nieuwland et al (2018) multilab study. We run a sensitivity analysis testing different ratios between by-subj and by-item variance, and the amount of residual error.

Accounting for Study Uncertainty

- Standard ERP studies lack consistent effect size descriptions.
- To perform statistical modeling, we needed to account for the relative uncertainty of each original study.
- **Formula for the contrast between two conditions:**

$$SE_{reconstructed} = \sqrt{2 \cdot \left(\frac{V_{subj}}{N} + \frac{V_{item}}{K} + \frac{V_{resid}}{N \cdot K} \right)}$$

- **Uncertainty:** $SE_{reconstructed}$ used variance and residual from Nieuwland et al (2018) multilab study. We run a sensitivity analysis testing different ratios between by-subj and by-item variance, and the amount of residual error.

Accounting for Study Uncertainty

- Standard ERP studies lack consistent effect size descriptions.
- To perform statistical modeling, we needed to account for the relative uncertainty of each original study.
- **Formula for the contrast between two conditions:**

$$SE_{reconstructed} = \sqrt{2 \cdot \left(\frac{V_{subj}}{N} + \frac{V_{item}}{K} + \frac{V_{resid}}{N \cdot K} \right)}$$

- **Uncertainty:** $SE_{reconstructed}$ used variance and residual from Nieuwland et al (2018) multilab study. We run a sensitivity analysis testing different ratios between by-subj and by-item variance, and the amount of residual error.

Bayesian Modeling: Key Design Choices

- **Robust likelihood (Student-t):** if a number of studies report extreme effects, the model down-weights their influence.
- **Precision weighting:** Each study contributes proportionally to its reliability — larger samples and more items carry more weight.
- **Between-study heterogeneity (τ):** The model estimates *how much* studies genuinely disagree, beyond what sampling error alone explains.
- **Weakly informative priors:** We assume effects are unlikely to exceed $\pm 8\,\mu$V.

Bayesian Modeling: Key Design Choices

- **Robust likelihood (Student-t):** if a number of studies report extreme effects, the model down-weights their influence.
- **Precision weighting:** Each study contributes proportionally to its reliability — larger samples and more items carry more weight.
- **Between-study heterogeneity (τ):** The model estimates *how much* studies genuinely disagree, beyond what sampling error alone explains.
- **Weakly informative priors:** We assume effects are unlikely to exceed $\pm 8\,\mu$V.

Bayesian Modeling: Key Design Choices

- **Robust likelihood (Student-t):** if a number of studies report extreme effects, the model down-weights their influence.
- **Precision weighting:** Each study contributes proportionally to its reliability — larger samples and more items carry more weight.
- **Between-study heterogeneity (τ):** The model estimates *how much* studies genuinely disagree, beyond what sampling error alone explains.
- **Weakly informative priors:** We assume effects are unlikely to exceed $\pm 8\,\mu$V.

Bayesian Modeling: Key Design Choices

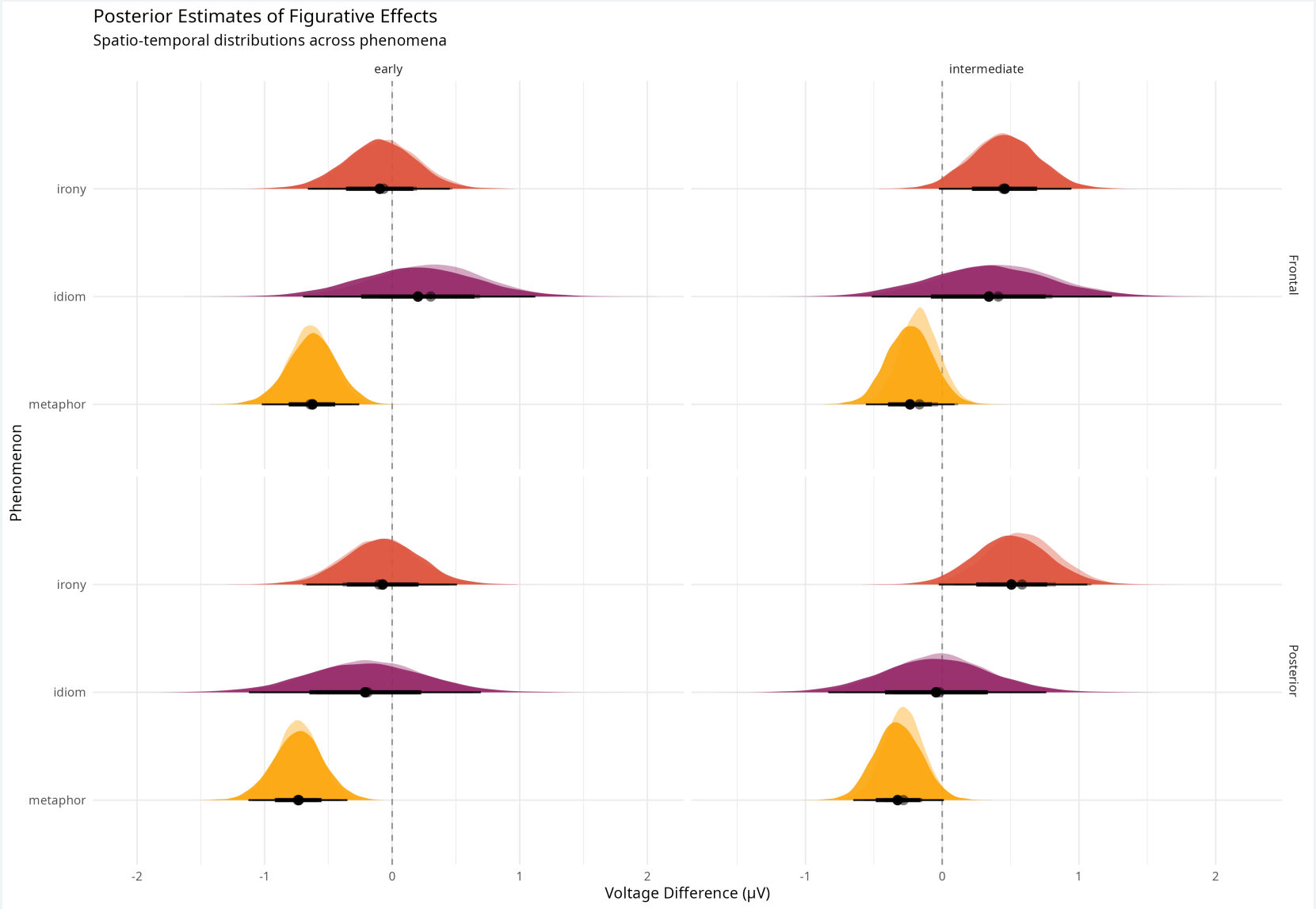
- **Robust likelihood (Student-t):** if a number of studies report extreme effects, the model down-weights their influence.
- **Precision weighting:** Each study contributes proportionally to its reliability — larger samples and more items carry more weight.
- **Between-study heterogeneity (τ):** The model estimates *how much* studies genuinely disagree, beyond what sampling error alone explains.
- **Weakly informative priors:** We assume effects are unlikely to exceed $\pm 8\,\mu$V.

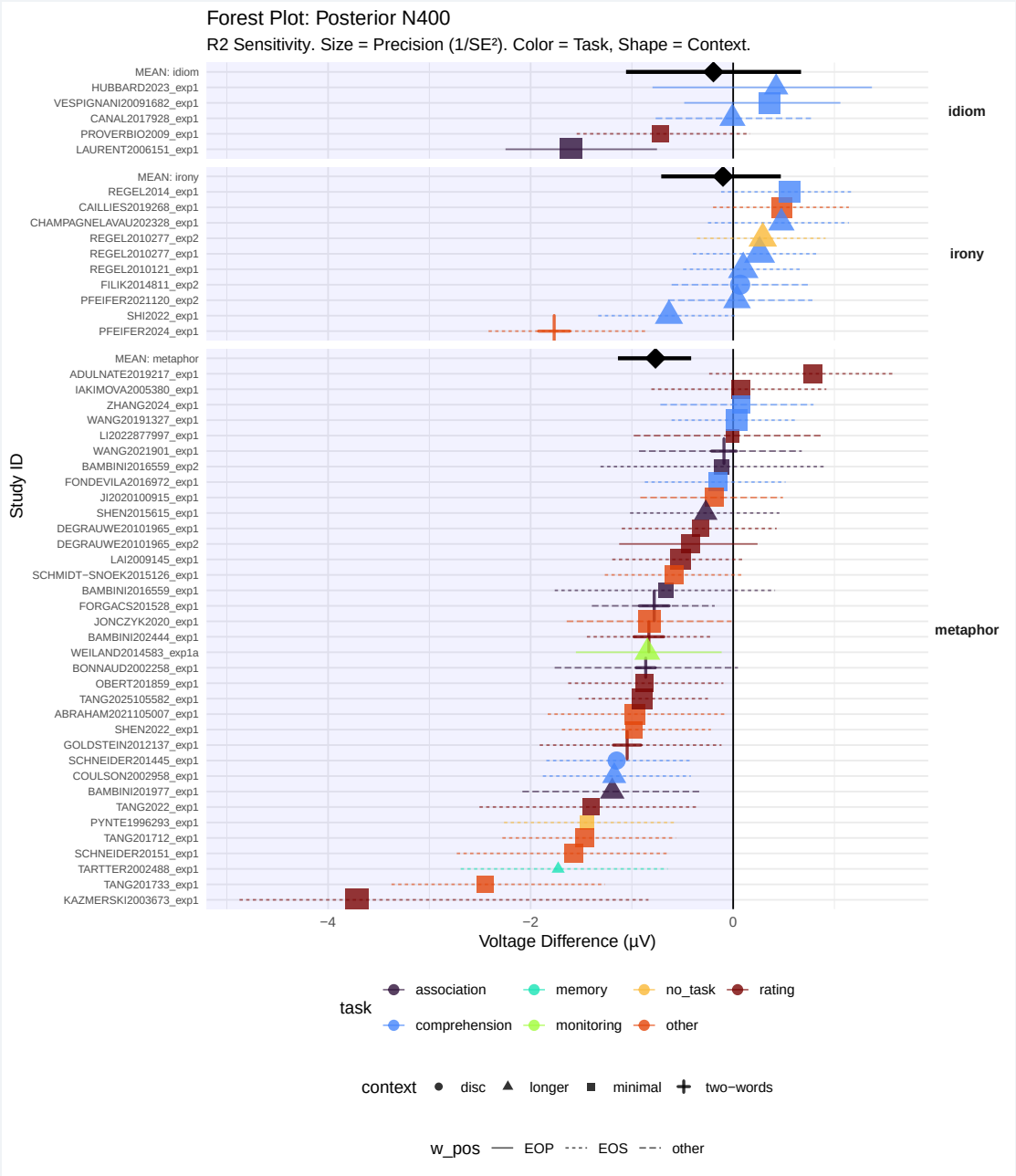
Bayesian Modeling: Key Design Choices

- **Robust likelihood (Student-t):** if a number of studies report extreme effects, the model down-weights their influence.
- **Precision weighting:** Each study contributes proportionally to its reliability — larger samples and more items carry more weight.
- **Between-study heterogeneity (τ):** The model estimates *how much* studies genuinely disagree, beyond what sampling error alone explains.
- **Weakly informative priors:** We assume effects are unlikely to exceed $\pm 8\,\mu$V.

Results: Bayesian modeling [early:300-500ms; intermediate:600-800ms]

Results: Bayesian modeling [early:300-500ms; intermediate:600-800ms]




Forest Plot: Posterior N400
R2 Sensitivity. Size = Precision (1/SE²). Color = Task, Shape = Context.
MEAN: idiom
HUBBARD2023_exp1
VESPIGNANI2091682_exp1
CANAL2017928_exp1
PROVERBIO2009_exp1
LAURENT2006151_exp1
idiom
MEAN: irony
REGEL2014_exp1
CAILLIES2019268_exp1
CHAMPAGNELAVAU202328_exp1
REGEL2010277_exp2
REGEL2010277_exp1
REGEL2010121_exp1
FILIK2014811_exp2
PFEIFER2021120_exp2
SHI2022_exp1
PFEIFER2024_exp1
irony
MEAN: metaphor
ADULNATE2019217_exp1
IAKIMOVA2005380_exp1
ZHANG2024_exp1
WANG20191327_exp1
LI2022877997_exp1
WANG2021901_exp1
BAMBINI2016559_exp2
FONDEVILA2016972_exp1
JI2020100915_exp1
SHEN2015615_exp1
DEGRAUWE20101965_exp1
DEGRAUWE20101965_exp2
LAI2009145_exp1
SCHMIDT−SNOEK2015126_exp1
BAMBINI2016559_exp1
FORGACS201528_exp1
JONCZYK2020_exp1
BAMBINI202444_exp1
WEILAND2014583_exp1a
BONNAUD2002258_exp1
OBERT201859_exp1
TANG2025105582_exp1
ABRAHAM2021105007_exp1
SHEN2022_exp1
GOLDSTEIN2012137_exp1
SCHNEIDER201445_exp1
COULSON2002958_exp1
BAMBINI201977_exp1
TANG2022_exp1
PYNTE1996293_exp1
TANG201712_exp1
SCHNEIDER20151_exp1
TARTTER2002488_exp1
TANG201733_exp1
KAZMERSKI2003673_exp1
metaphor
Study ID
−4
−2
0
Voltage Difference (μV)
task
association
memory
no_task
rating
comprehension
monitoring
other
context
disc
longer
minimal
two−words
w_pos
EOP
EOS
other

Ad interim consideration

Such windowed Bayesian modeling approach answers *is there an effect in window X?*

- We solved the spatial side of the researcher-degrees-of-freedom problem with digitization, but the temporal side is still hand-picked.
- Every waveform is now digitized and interpolated onto a common time grid, so we can model time as continuous rather than discretized.

GAMs put between-study heterogeneity into the time domain directly, providing complementary evidence.

```
bam(uV_effect ~  s(time_int, k=30) +
    s(time_int, unique_exp_id, bs="fs", m=1, k=15),
    data = df_ROI, weights = w_precision,
    rho = 0.35, AR.start = start_event)
```

Ad interim consideration

Such windowed Bayesian modeling approach answers *is there an effect in window X?*

- We solved the spatial side of the researcher-degrees-of-freedom problem with digitization, but the temporal side is still hand-picked.
- Every waveform is now digitized and interpolated onto a common time grid, so we can model time as continuous rather than discretized.

GAMs put between-study heterogeneity into the time domain directly, providing complementary evidence.

```
bam(uV_effect ~  s(time_int, k=30) +
    s(time_int, unique_exp_id, bs="fs", m=1, k=15),
    data = df_ROI, weights = w_precision,
    rho = 0.35, AR.start = start_event)
```

Ad interim consideration

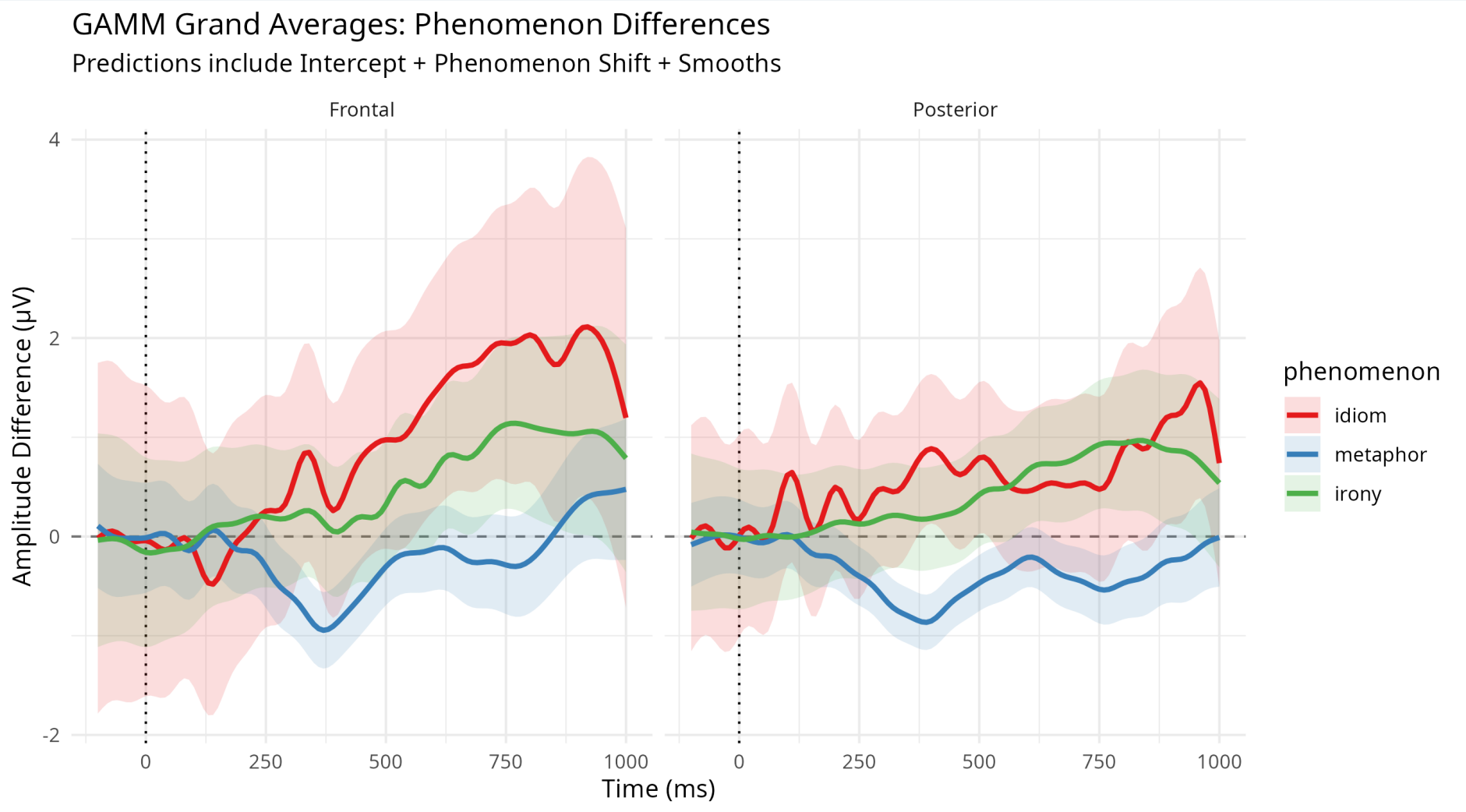
Such windowed Bayesian modeling approach answers *is there an effect in window X?*

- We solved the spatial side of the researcher-degrees-of-freedom problem with digitization, but the temporal side is still hand-picked.
- Every waveform is now digitized and interpolated onto a common time grid, so we can model time as continuous rather than discretized.

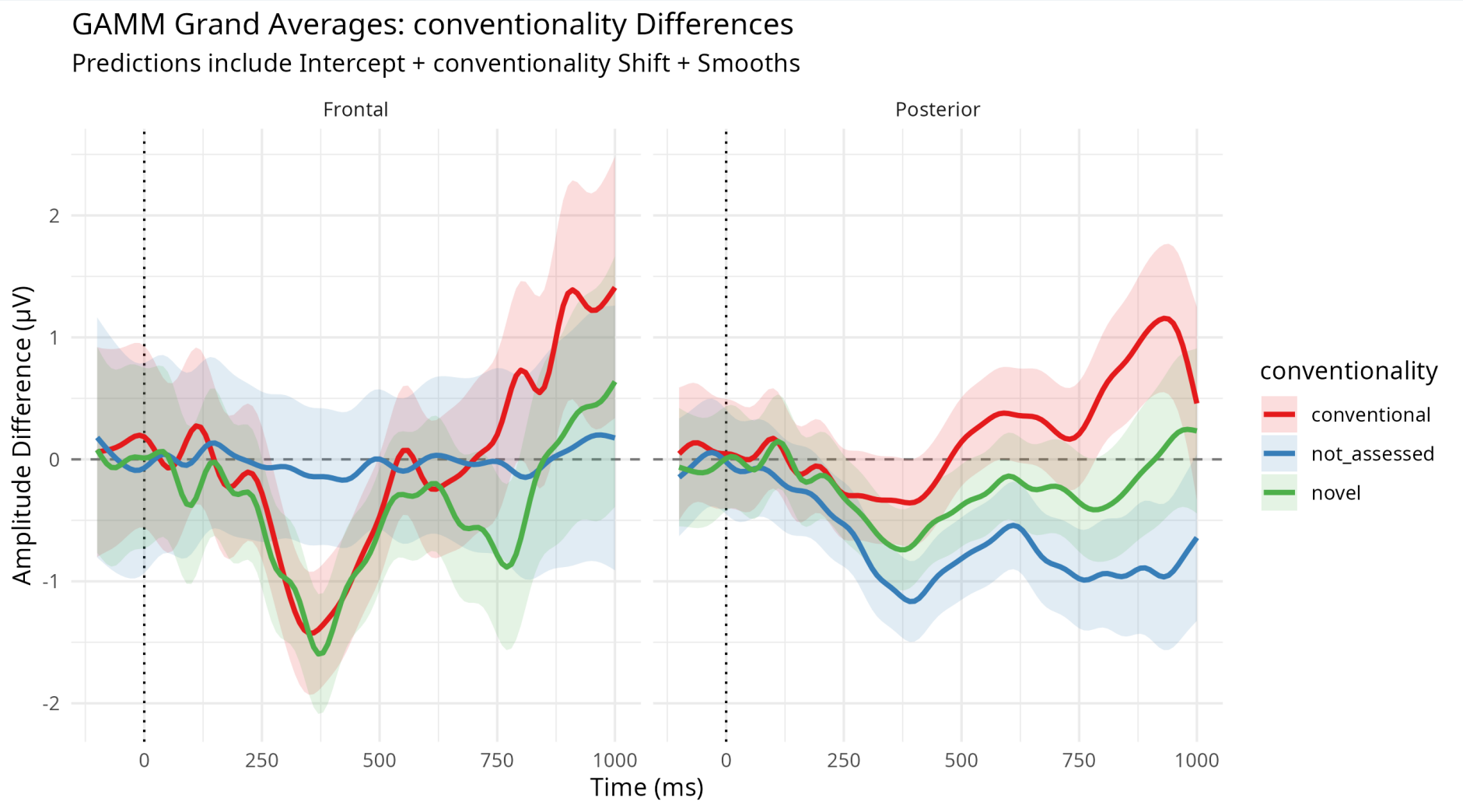
GAMs put between-study heterogeneity into the time domain directly, providing complementary evidence.

```
bam(uV_effect ~  s(time_int, k=30) +
    s(time_int, unique_exp_id, bs="fs", m=1, k=15),
    data = df_ROI, weights = w_precision,
    rho = 0.35, AR.start = start_event)
```

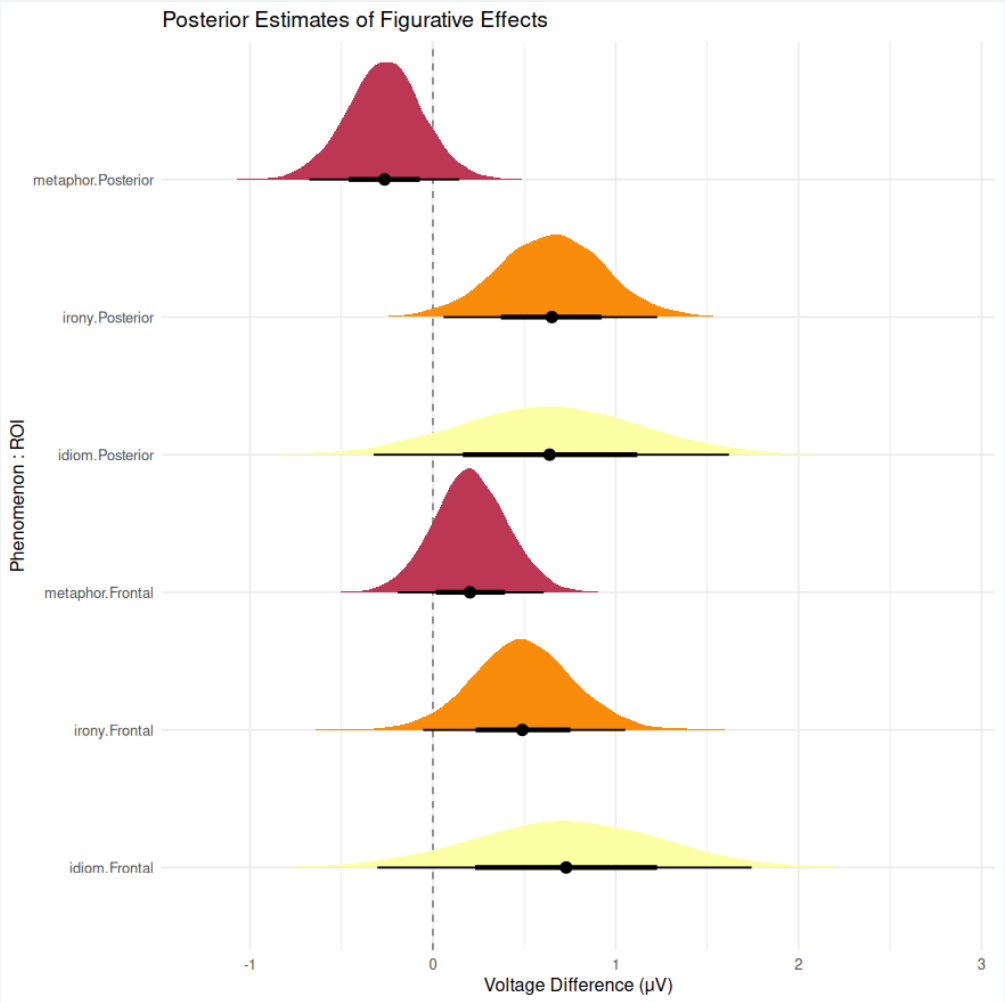
Results: GAM Phenomenon



Results: GAM type



Figurative effects in the late time window



Conclusions: Methodological Contributions

- The digitization of waveforms provides a uniform standard, bypassing idiosyncratic choices in time-window and electrode selection.
- The use of `brms` with Student-t likelihood and precision weighting (N/K) looks a promising approach even with heterogeneous study designs.

Conclusions: Methodological Contributions

- The digitization of waveforms provides a uniform standard, bypassing idiosyncratic choices in time-window and electrode selection.
- The use of `brms` with Student-t likelihood and precision weighting (N/K) looks a promising approach even with heterogeneous study designs.

Conclusions: Methodological Contributions

- The digitization of waveforms provides a uniform standard, bypassing idiosyncratic choices in time-window and electrode selection.
- The use of `brms` with Student-t likelihood and precision weighting (N/K) looks a promising approach even with heterogeneous study designs.

Conclusions: Theoretical Insights

- **Phenomenon-Specific Dynamics:** Metaphor and Irony show distinct "neural signatures":
  - **Metaphor:** Massive early N400, possibly carrying over to the intermediate time window.
  - **Irony:** Minimal N400, followed by a robust, sustained Late Positivity (> 800 ms).
- **Beyond Fixed Windows (GAMs):** The Generalized Additive Models revealed a **growing positivity after 750 ms** that was partially masked by traditional 600–800 ms windows.
- **The Frontal Shift in Metaphor (800–1000 ms):** Metaphor processing does not trigger a canonical P600. Rather, it engages a late frontal component (84% of likelihood), which can be linked to the derivation of the implicated meaning. Previously attested 20% probability of having a P600 in scoping review.

Conclusions: Theoretical Insights

- **Phenomenon-Specific Dynamics:** Metaphor and Irony show distinct "neural signatures":
  - **Metaphor:** Massive early N400, possibly carrying over to the intermediate time window.
  - **Irony:** Minimal N400, followed by a robust, sustained Late Positivity (> 800 ms).
- **Beyond Fixed Windows (GAMs):** The Generalized Additive Models revealed a **growing positivity after 750 ms** that was partially masked by traditional 600–800 ms windows.
- **The Frontal Shift in Metaphor (800–1000 ms):** Metaphor processing does not trigger a canonical P600. Rather, it engages a late frontal component (84% of likelihood), which can be linked to the derivation of the implicated meaning. Previously attested 20% probability of having a P600 in scoping review.

Conclusions: Theoretical Insights

- **Phenomenon-Specific Dynamics:** Metaphor and Irony show distinct "neural signatures":
  - **Metaphor:** Massive early N400, possibly carrying over to the intermediate time window.
  - **Irony:** Minimal N400, followed by a robust, sustained Late Positivity ($>$ 800 ms).
- **Beyond Fixed Windows (GAMs):** The Generalized Additive Models revealed a **growing positivity after 750 ms** that was partially masked by traditional 600–800 ms windows.
- **The Frontal Shift in Metaphor (800–1000 ms):** Metaphor processing does not trigger a canonical P600. Rather, it engages a late frontal component (84% of likelihood), which can be linked to the derivation of the implicated meaning. Previously attested 20% probability of having a P600 in scoping review.

Conclusions: Theoretical Insights

- **Phenomenon-Specific Dynamics:** Metaphor and Irony show distinct "neural signatures":
  - **Metaphor:** Massive early N400, possibly carrying over to the intermediate time window.
  - **Irony:** Minimal N400, followed by a robust, sustained Late Positivity (> 800 ms).
- **Beyond Fixed Windows (GAMs):** The Generalized Additive Models revealed a **growing positivity after 750 ms** that was partially masked by traditional 600–800 ms windows.
- **The Frontal Shift in Metaphor (800–1000 ms):** Metaphor processing does not trigger a canonical P600. Rather, it engages a late frontal component (84% of likelihood), which can be linked to the derivation of the implicated meaning. Previously attested 20% probability of having a P600 in scoping review.

Conclusions: Theoretical Insights

- **Phenomenon-Specific Dynamics:** Metaphor and Irony show distinct "neural signatures":
  - **Metaphor:** Massive early N400, possibly carrying over to the intermediate time window.
  - **Irony:** Minimal N400, followed by a robust, sustained Late Positivity ($>$ 800 ms).
- **Beyond Fixed Windows (GAMs):** The Generalized Additive Models revealed a **growing positivity after 750 ms** that was partially masked by traditional 600–800 ms windows.
- **The Frontal Shift in Metaphor (800–1000 ms):** Metaphor processing does not trigger a canonical P600. Rather, it engages a late frontal component (84% of likelihood), which can be linked to the derivation of the implicated meaning. Previously attested 20% probability of having a P600 in scoping review.

Thanks! *I've got backup slides.....*

We thank Psicostat crew for previous feedback on the preliminary analyses.

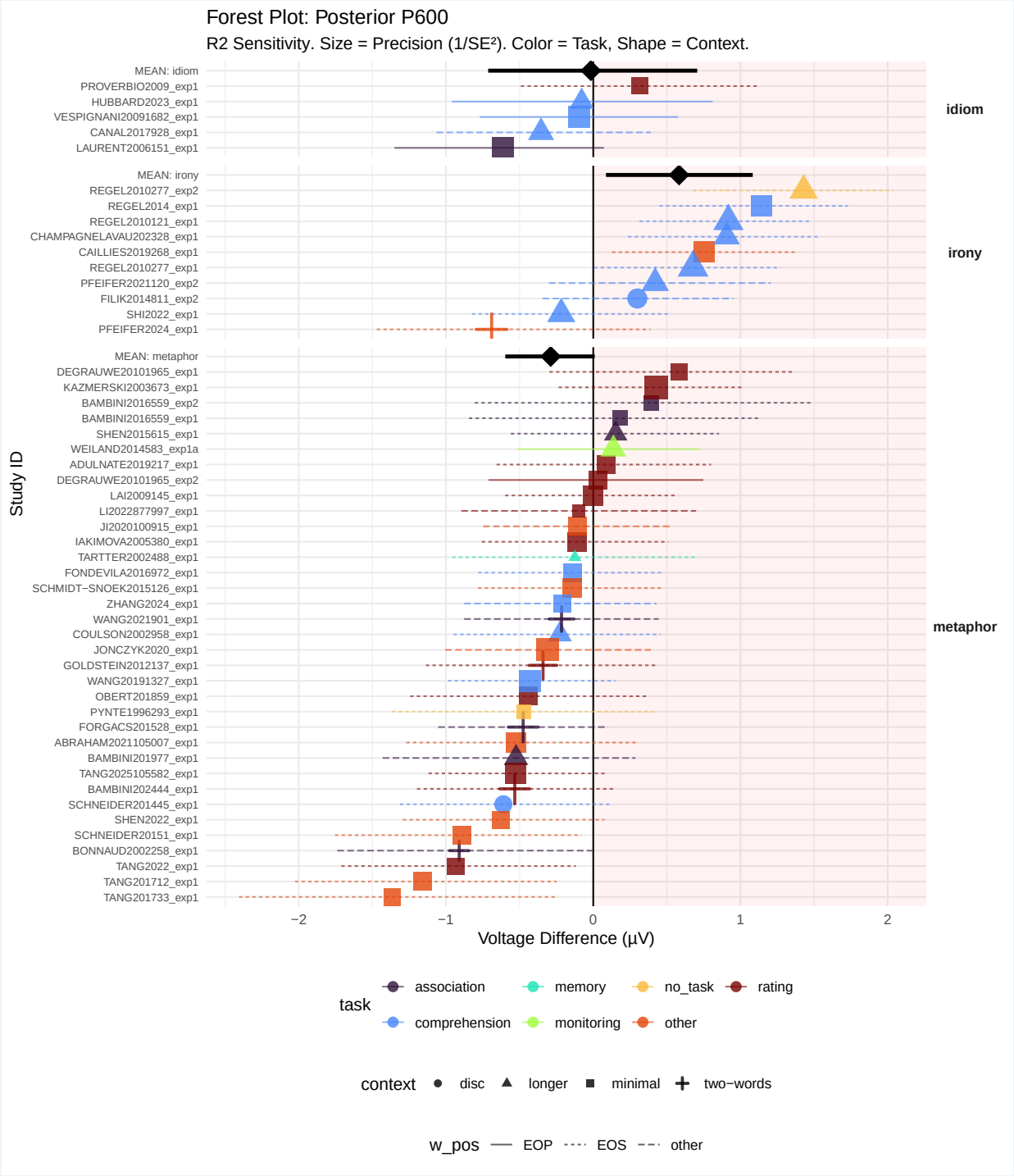
We thank Claude and Gemini for assistance with code debugging and discussion of model diagnostics during the development of the analyses; all analytic decisions and interpretations remain the responsibility of the authors

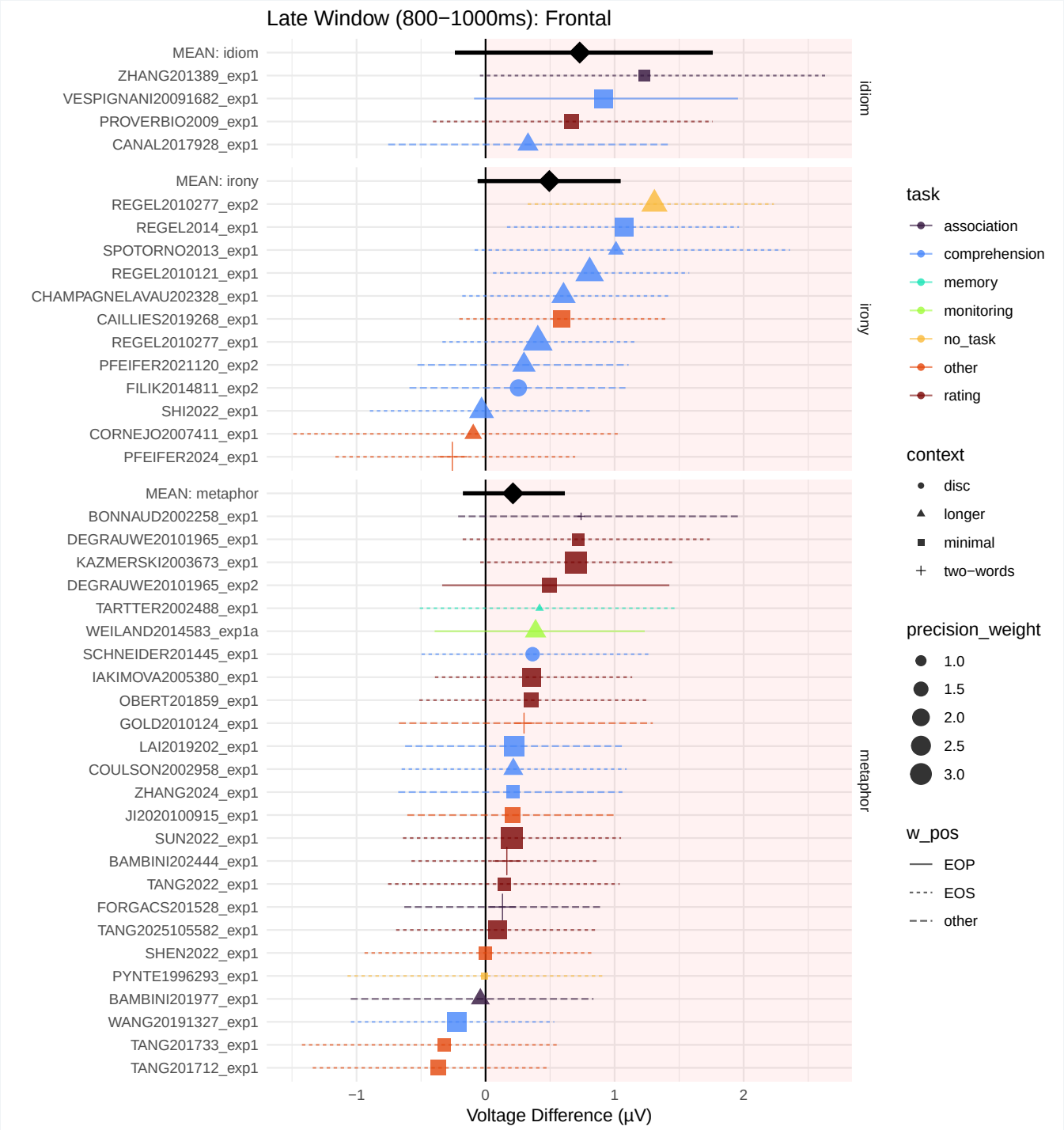
Eligibility & Inclusion Criteria

- Focus on language comprehension (written/auditory) targeting a specific word.
- Sample of healthy adults (16–60 years), L1 speakers.
- Must include at least one ERP plot of a single midline channel (Oz, Pz, CPz, Cz, FCz, Fz) or a cluster including one.
- Plot must show:
  - Pre-stimulus baseline $\geq$ 100ms.
  - Post-stimulus activity $\geq$ 800ms.
- Must have a literal control condition.
- Standard referencing (mastoids, ear lobes, average).

[Extensive annotation of experiments features was carried out: annotation (PDF)]

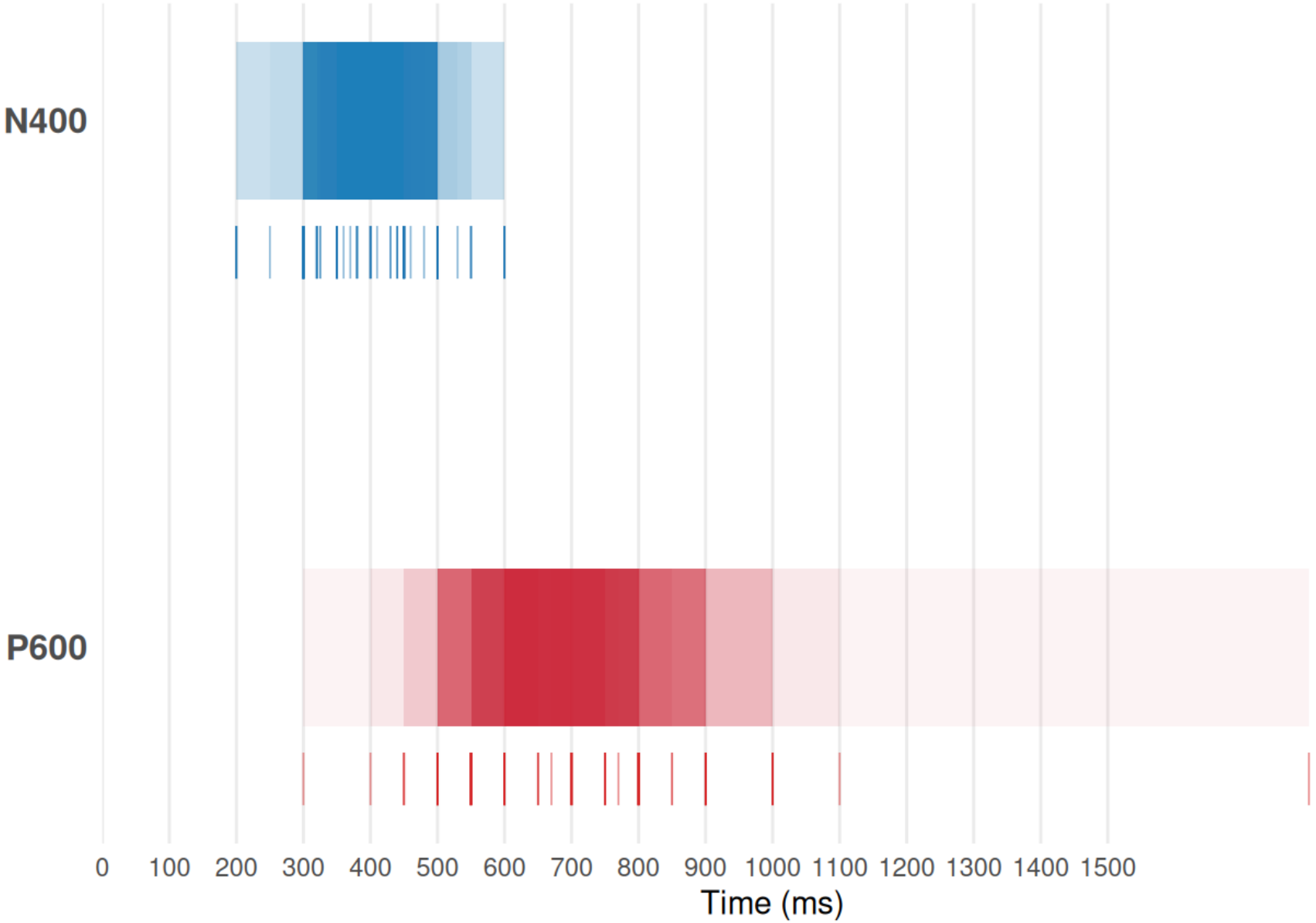
Win	Hyp	ROI	Estimate	95% CI	ER (BF)	Post.Prob
Metaphor						
300-500	Neg	Fron	**-0.63**	[-0.96, -0.31]	1599.0	1*
		Post	**-0.73**	[-1.05, -0.42]	3999.0	1*
600-800	Neg	Fron	**-0.23**	[-0.50, 0.04]	11.9	0.92
		Post	**-0.32**	[-0.60, -0.06]	37.6	0.97*
800-1000	Pos	Fron	**0.20**	[-0.13, 0.54]	5.1	0.84
		Post	-0.26	[-0.61, 0.09]	0.1	0.10
Irony						
300-500	Neg	Fron	-0.09	[-0.57, 0.36]	1.7	0.63
		Post	-0.07	[-0.57, 0.41]	1.5	0.59
600-800	Pos	Fron	**0.46**	[0.04, 0.88]	25.8	0.96*
		Post	**0.51**	[0.05, 0.97]	28.1	0.97*
800-1000	Pos	Fron	**0.49**	[0.04, 0.94]	24.9	0.96*
		Post	**0.64**	[0.16, 1.12]	67.4	0.99*
Idiom						
300-500	Neg	Fron	0.20	[-0.54, 0.93]	0.5	0.33
		Post	-0.21	[-0.96, 0.55]	2.1	0.68
600-800	Pos	Fron	0.34	[-0.38, 1.09]	3.4	0.77
		Post	-0.04	[-0.69, 0.64]	0.9	0.46
800-1000	Pos	Fron	**0.74**	[-0.10, 1.57]	12.8	0.93
		Post	**0.66**	[-0.14, 1.45]	10.4	0.91

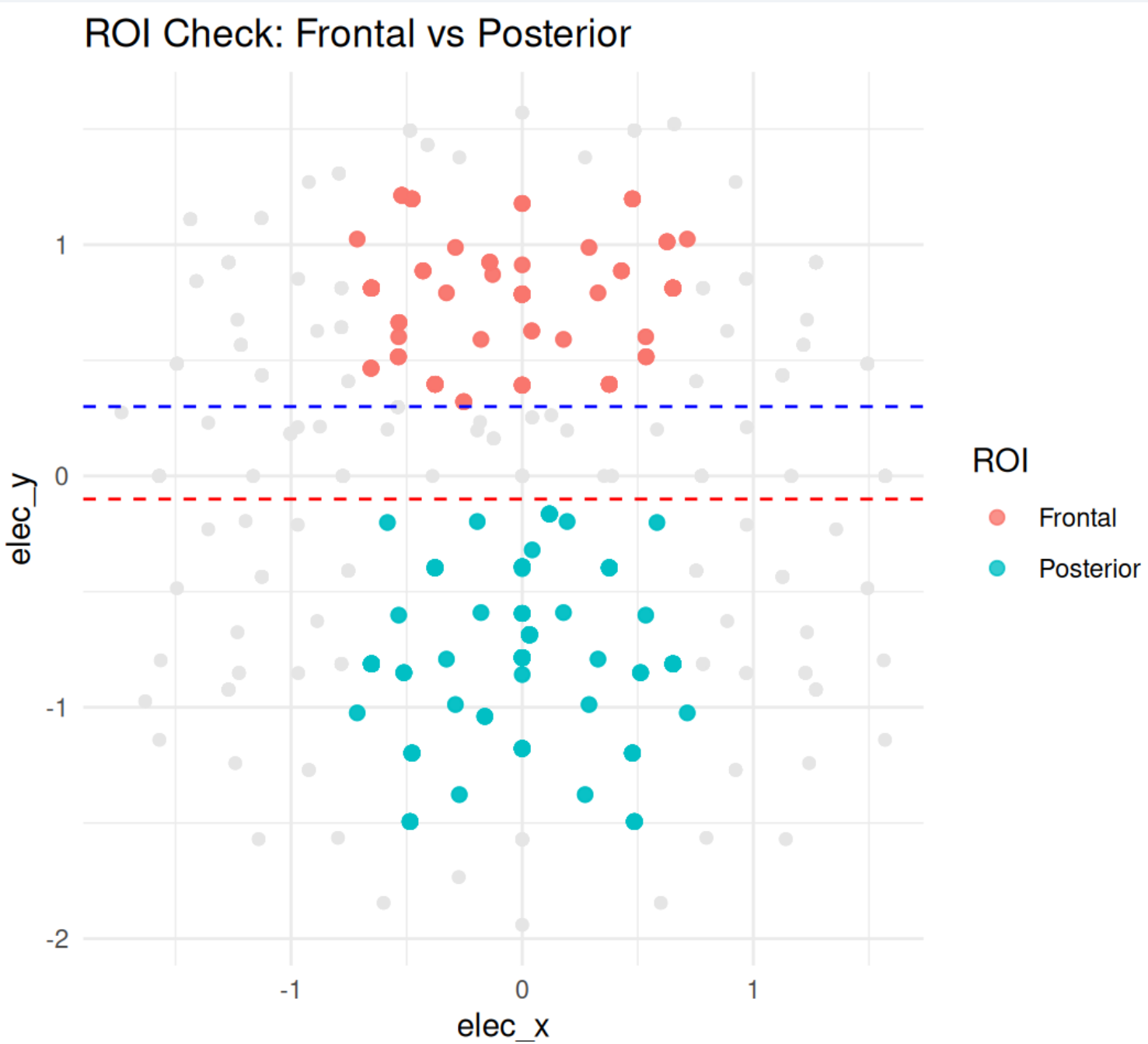

Forest Plot: Posterior P600
R2 Sensitivity. Size = Precision (1/SE²). Color = Task, Shape = Context.
MEAN: idiom
PROVERBIO2009_exp1
HUBBARD2023_exp1
VESPIGNANI20091682_exp1
CANAL2017928_exp1
LAURENT2006151_exp1
idiom
MEAN: irony
REGEL2010277_exp2
REGEL2014_exp1
REGEL2010121_exp1
CHAMPAGNELAVAU202328_exp1
CAILLIES2019268_exp1
REGEL2010277_exp1
PFEIFER2021120_exp2
FILIK2014811_exp2
SHI2022_exp1
PFEIFER2024_exp1
irony
MEAN: metaphor
DEGRAUWE20101965_exp1
KAZMERSKI2003673_exp1
BAMBINI2016559_exp2
BAMBINI2016559_exp1
SHEN2015615_exp1
WEILAND2014583_exp1a
ADULNATE2019217_exp1
DEGRAUWE20101965_exp2
LAI2009145_exp1
LI2022877997_exp1
JI2020100915_exp1
IAKIMOVA2005380_exp1
TARTTER2002488_exp1
FONDEVILA2016972_exp1
SCHMIDT−SNOEK2015126_exp1
ZHANG2024_exp1
WANG2021901_exp1
COULSON2002958_exp1
JONCZYK2020_exp1
GOLDSTEIN2012137_exp1
WANG20191327_exp1
OBERT201859_exp1
PYNTE1996293_exp1
FORGACS201528_exp1
ABRAHAM2021105007_exp1
BAMBINI201977_exp1
TANG2025105582_exp1
BAMBINI202444_exp1
SCHNEIDER201445_exp1
SHEN2022_exp1
SCHNEIDER20151_exp1
BONNAUD2002258_exp1
TANG2022_exp1
TANG201712_exp1
TANG201733_exp1
metaphor
Study ID
−2
−1
0
1
2
Voltage Difference (µV)
task
association
memory
no_task
rating
comprehension
monitoring
other
context
disc
longer
minimal
two−words
w_pos
EOP
EOS
other


Late Window (800−1000ms): Frontal
MEAN: idiom
ZHANG201389_exp1
VESPIGNANI20091682_exp1
PROVERBIO2009_exp1
CANAL2017928_exp1
idiom
MEAN: irony
REGEL2010277_exp2
REGEL2014_exp1
SPOTORNO2013_exp1
REGEL2010121_exp1
CHAMPAGNELAVAU202328_exp1
CAILLIES2019268_exp1
REGEL2010277_exp1
PFEIFER2021120_exp2
FILIK2014811_exp2
SHI2022_exp1
CORNEJO2007411_exp1
PFEIFER2024_exp1
irony
MEAN: metaphor
BONNAUD2002258_exp1
DEGRAUWE20101965_exp1
KAZMERSKI2003673_exp1
DEGRAUWE20101965_exp2
TARTTER2002488_exp1
WEILAND2014583_exp1a
SCHNEIDER201445_exp1
IAKIMOVA2005380_exp1
OBERT201859_exp1
GOLD2010124_exp1
LAI2019202_exp1
COULSON2002958_exp1
ZHANG2024_exp1
JI2020100915_exp1
SUN2022_exp1
BAMBINI202444_exp1
TANG2022_exp1
FORGACS201528_exp1
TANG2025105582_exp1
SHEN2022_exp1
PYNTE1996293_exp1
BAMBINI201977_exp1
WANG20191327_exp1
TANG201733_exp1
TANG201712_exp1
metaphor
−1
0
1
2
Voltage Difference (μV)
task
association
comprehension
memory
monitoring
no_task
other
rating
context
disc
longer
minimal
two−words
precision_weight
1.0
1.5
2.0
2.5
3.0
w_pos
EOP
EOS
other

Time Windows

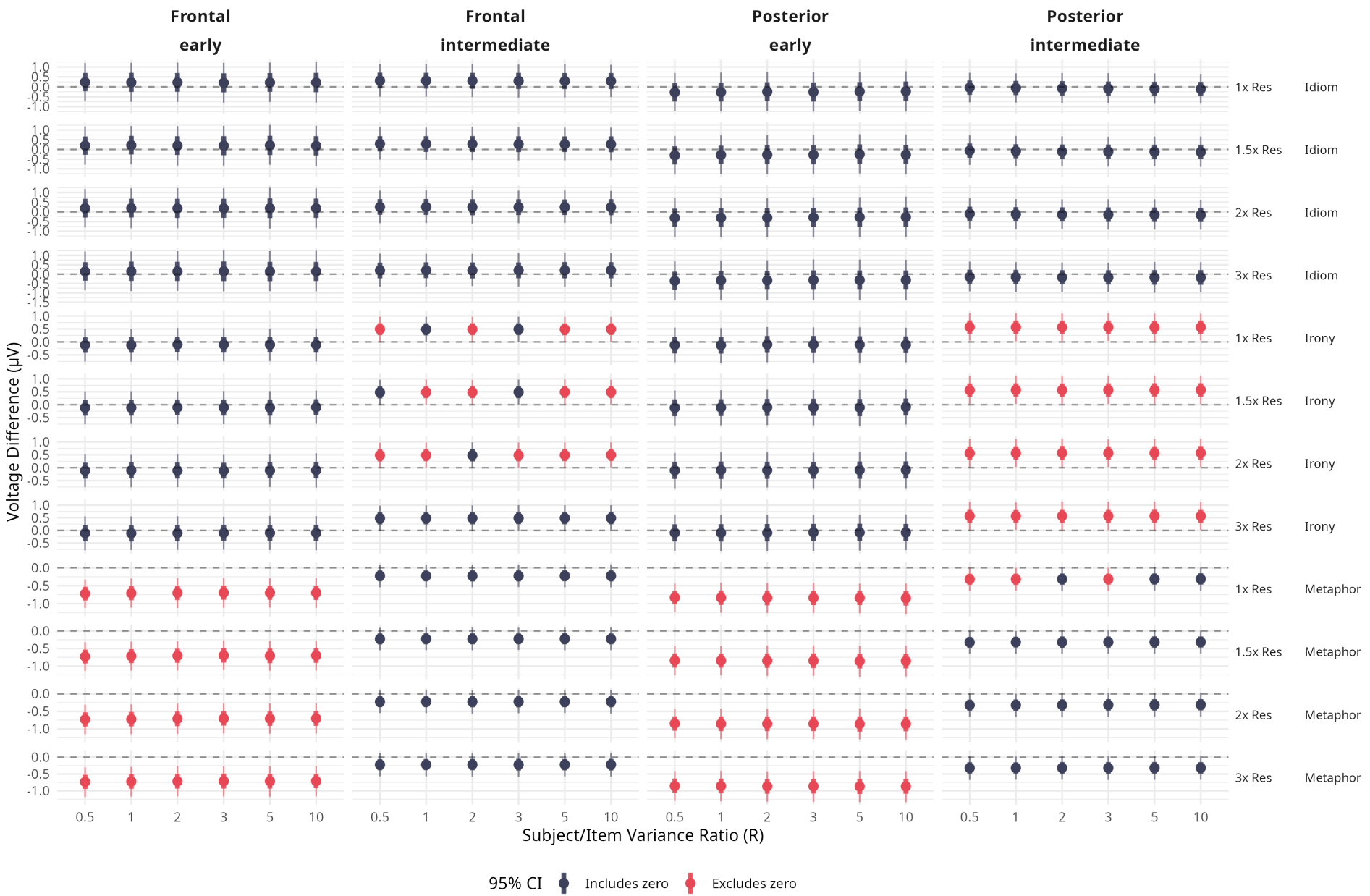
Darker regions indicate high consensus among studies

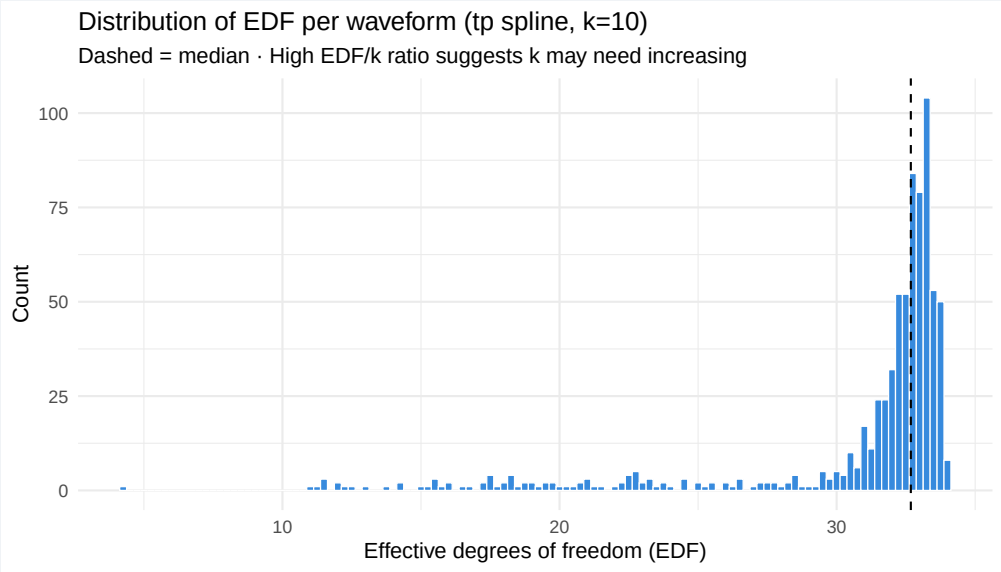


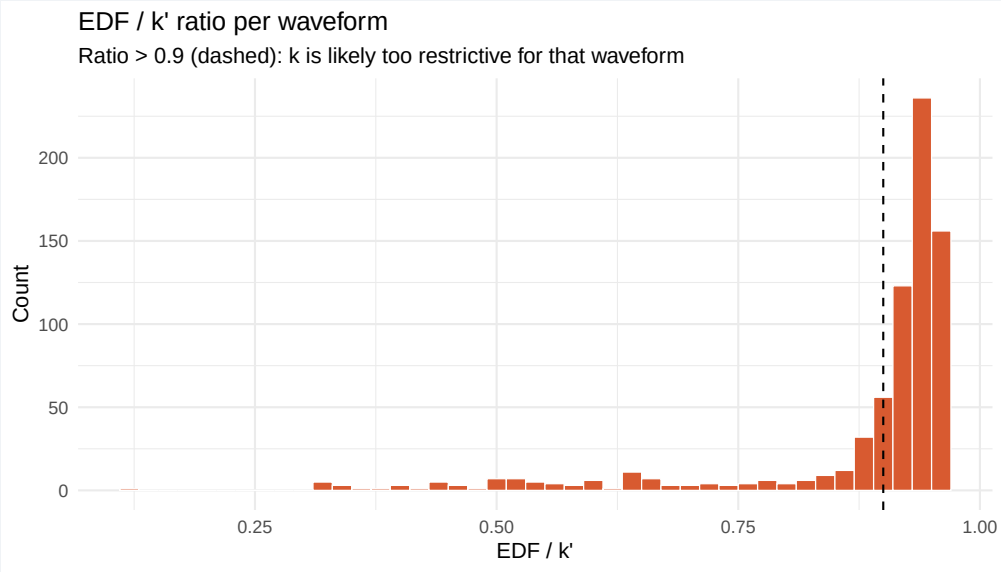

ROI Check: Frontal vs Posterior
elec_y
elec_x
1
0
-1
-2
-1
0
1
ROI
Frontal
Posterior

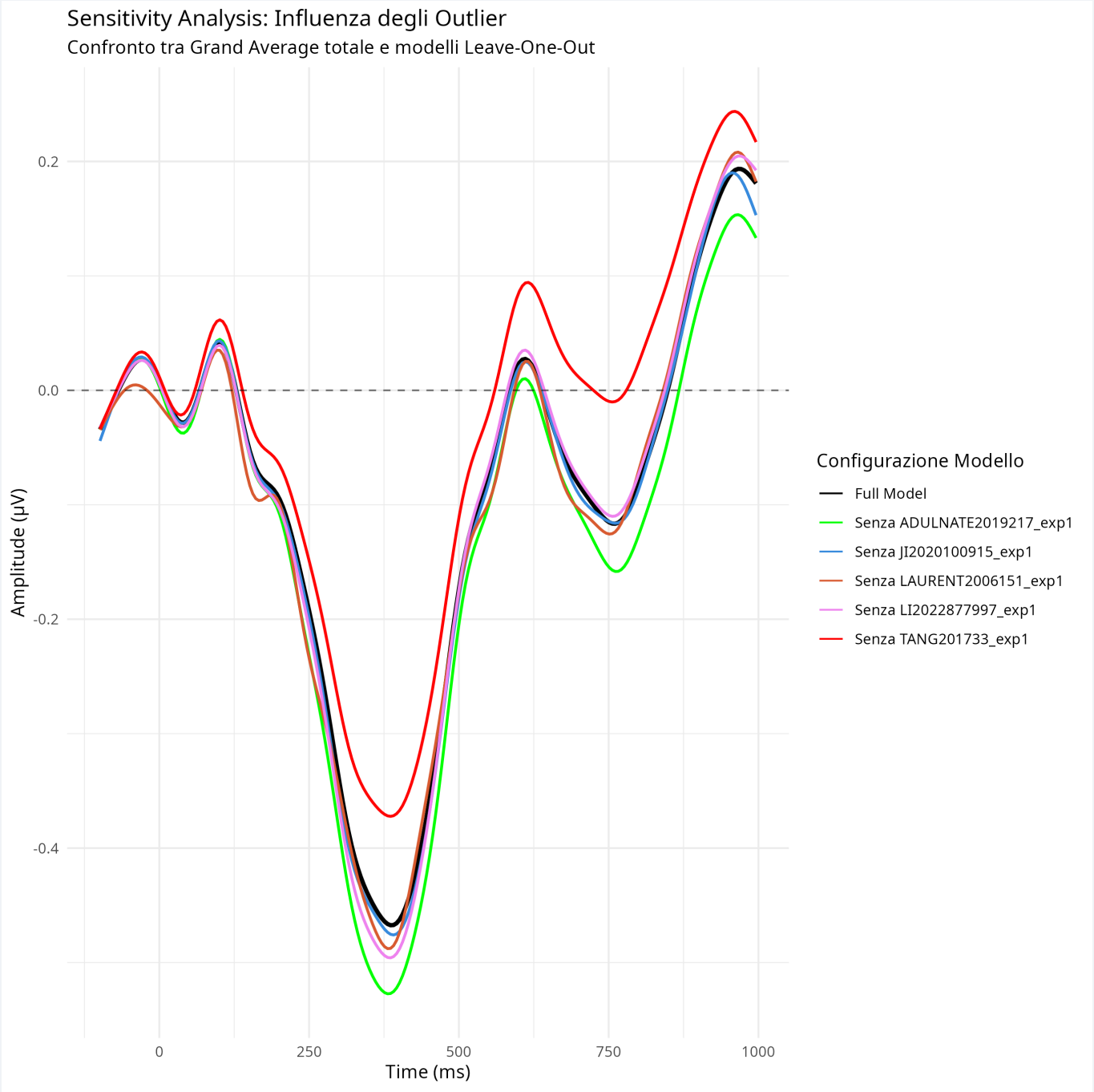
2D Sensitivity Analysis

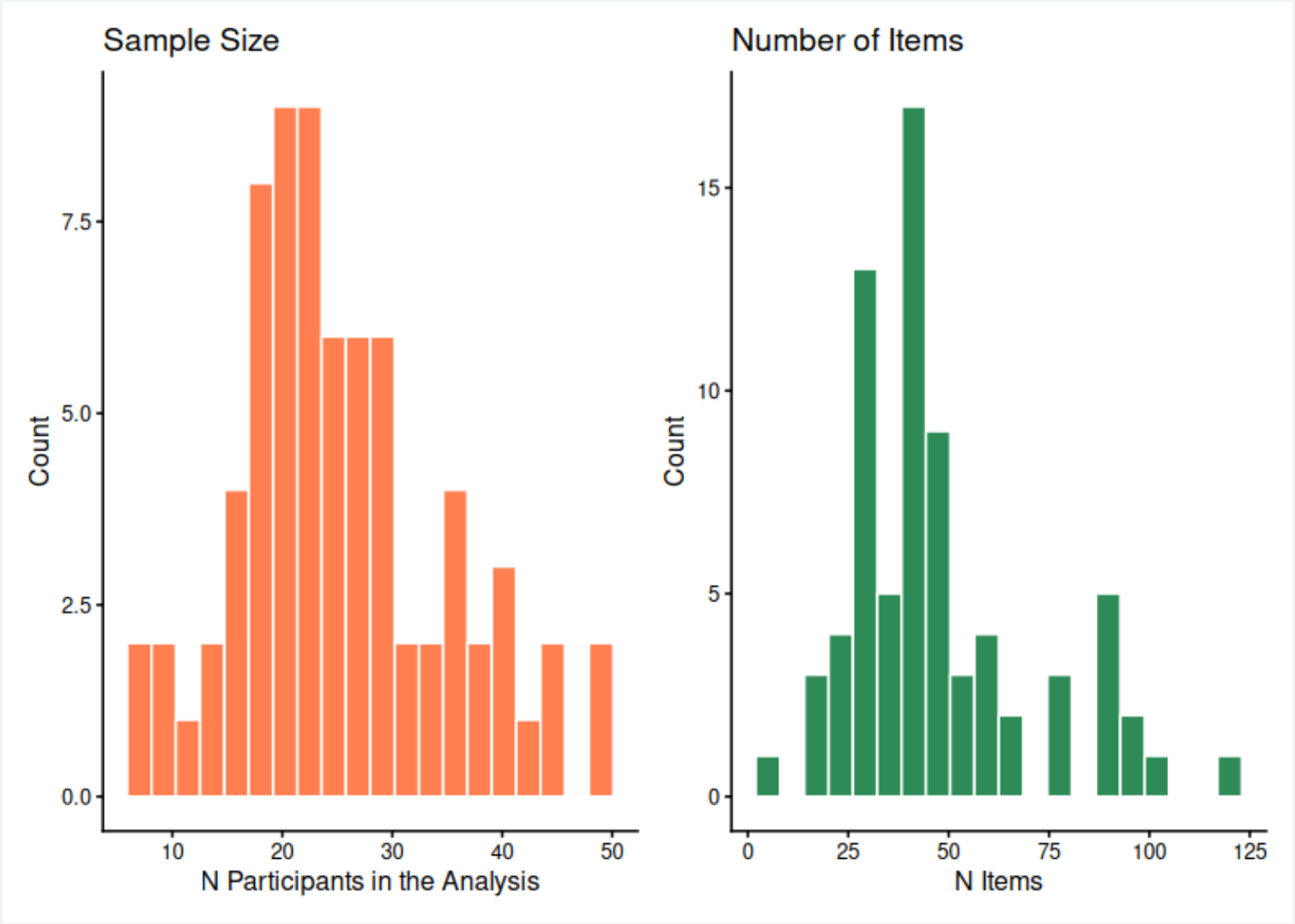
Testing τ^2/σ^2 Ratio (x-axis) across Residual Noise Multipliers (facet rows)

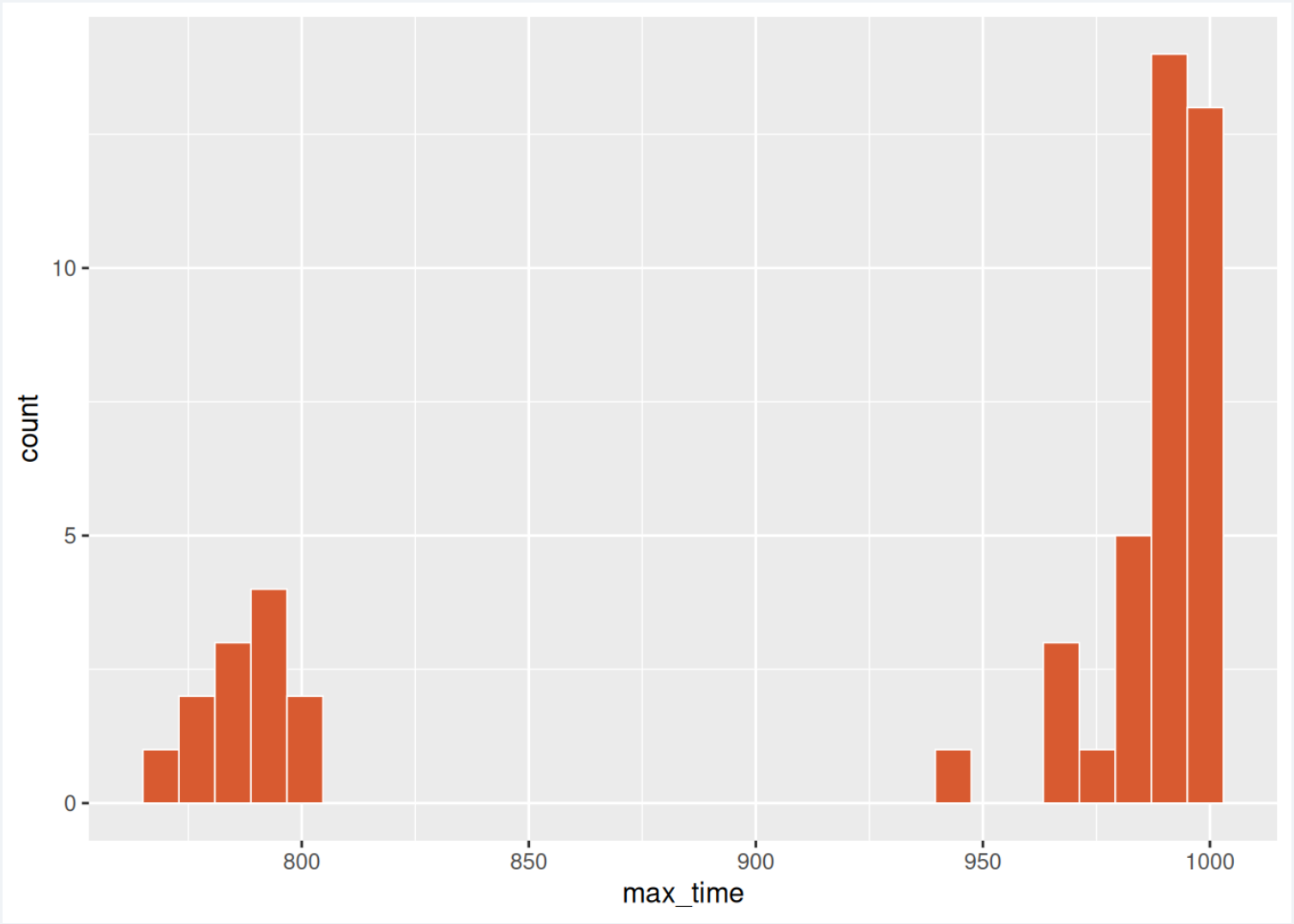


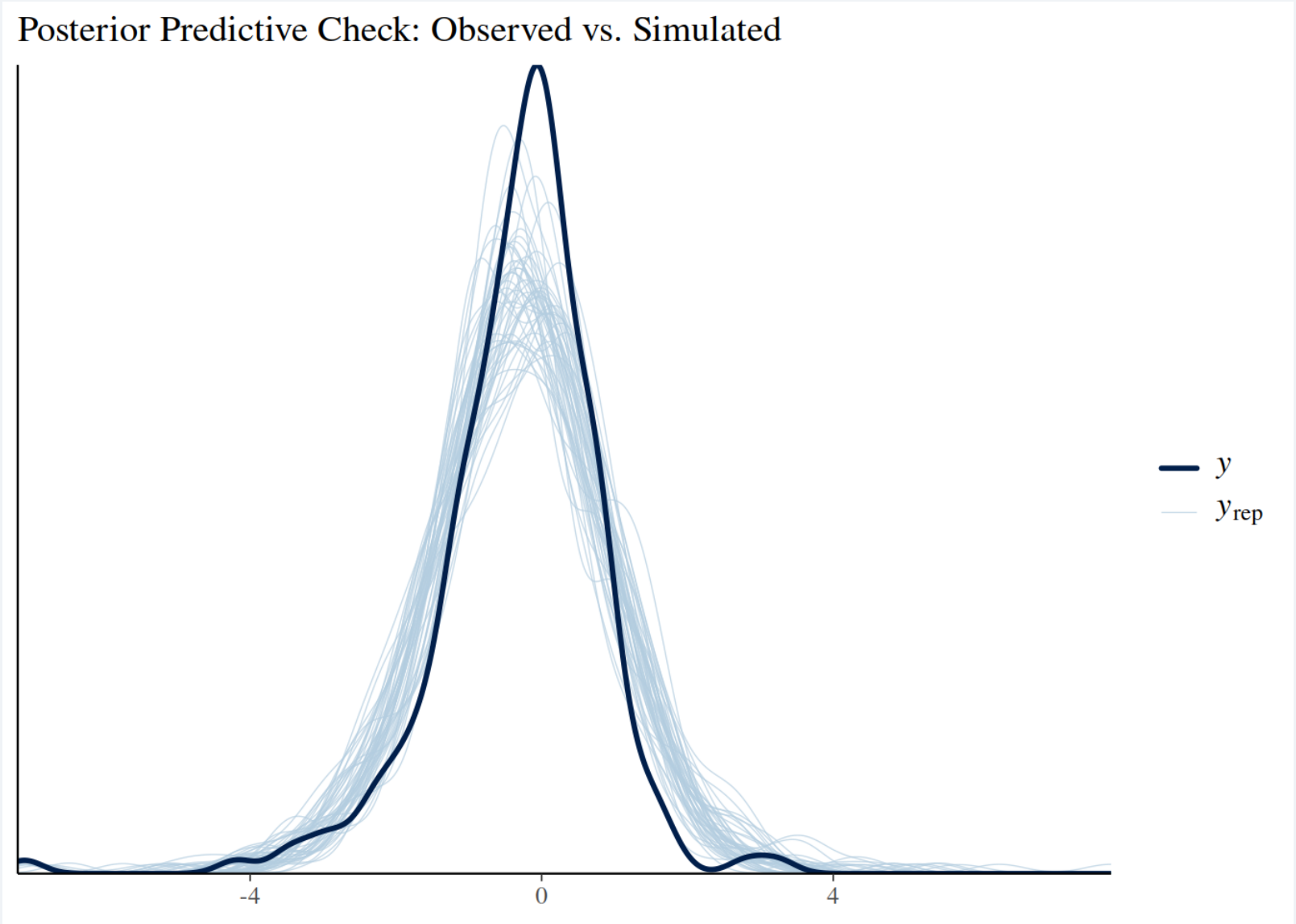

Distribution of EDF per waveform (tp spline, k=10)
Dashed = median · High EDF/k ratio suggests k may need increasing
Count
100
75
50
25
0
10
20
30
Effective degrees of freedom (EDF)


EDF / k' ratio per waveform
Ratio > 0.9 (dashed): k is likely too restrictive for that waveform
Count
200
150
100
50
0
0.25
0.50
0.75
1.00
EDF / k'


Sensitivity Analysis: Influenza degli Outlier
Confronto tra Grand Average totale e modelli Leave-One-Out
Amplitude (μV)
0.2
0.0
-0.2
-0.4
0
250
500
750
1000
Time (ms)
Configurazione Modello
Full Model
Senza ADULNATE2019217_exp1
Senza JI2020100915_exp1
Senza LAURENT2006151_exp1
Senza LI2022877997_exp1
Senza TANG201733_exp1


Sample Size
Count
7.5
5.0
2.5
0.0
10
20
30
40
50
N Participants in the Analysis
Number of Items
Count
15
10
5
0
0
25
50
75
100
125
N Items


count
10
5
0
800
850
900
950
1000
max_time



The model recovers the bulk of the distribution; Student-t absorbs the handful of high-magnitude studies without excluding them