Early Post-Stimulus Activity Negatively Predicts P300 Amplitude: A Single-Trial Analysis of the Auditory Oddball Task

Erkan Biber https://orcid.org/0000-0003-4938-1046

Institute of Biomedical Engineering, Boğaziçi University, Kandilli Campus, 34684 Çengelköy, Istanbul, Turkey

Corresponding Author:

Erkan Biber

Email: erkan.biber@bogazici.edu.tr

ORCID: https://orcid.org/0000-0003-4938-1046

Data and Code Availability:

- Raw Data: ERP CORE P300 dataset (Kappenman et al., 2021): https://osf.io/thsqg/
- Pre-processed Data & Metrics: https://osf.io/dr5bu/
- Analysis Scripts: https://github.com/erbiber/p300-entropy/

Conflict of Interest: The author declares no conflict of interest.

Funding: This research received no external funding.

Abstract

The P300 event-related potential is a core index of attention and context updating, yet the trial-by-trial factors that modulate its amplitude remain incompletely characterized. This study tested whether early post-stimulus signal magnitude in the sensory window (0–150 ms) predicts subsequent P300 amplitude (300–600 ms) at the single-trial level. Using data from the ERP CORE auditory oddball dataset (N = 40 participants; 1,661 accepted trials), early activity was quantified as root mean square (RMS) amplitude at electrode Fz. A linear mixed-effects model revealed a statistically reliable negative association between early RMS and P300 amplitude at Pz ($\beta = -0.064$, SE = 0.0245, $z = -2.61$, $p = 0.0085$). However, early RMS explained minimal trial level variance ($R^2 = 0.0042$), indicating a modest effect. Model diagnostics confirmed adequate assumptions for inference (Shapiro-Wilk W = 0.9995, p = 0.9690). Exploratory complexity measures (permutation entropy, Lempel–Ziv) were not predictive. Because RMS reflects both signal magnitude and within trial variability, the findings are most conservatively interpreted as evidence that higher early signal energy is associated with slightly reduced P300 amplitude, possibly reflecting resource competition, refractory effects, or state dependent processing. The results provide methodologically rigorous evidence for a small but reliable coupling between early and late ERP components, while highlighting the importance of distinguishing amplitude based measures from trial-to-trial variability. Independent replication and control analyses separating mean amplitude from dispersion are needed to clarify the underlying mechanism.

Keywords: P300, single-trial EEG, event-related potentials, linear mixed effects model, oddball paradigm, amplitude coupling, open science, reproducibility.

1. Introduction

The P300 event related potential (ERP) is one of the most extensively studied electrophysiological markers of attention and cognitive processing. It reflects the neural response to task relevant, unpredictable stimuli and is thought to index "context updating" in working memory (Donchin, 1981; Polich, 2007).

While the average morphology and functional significance of the P300 are well established, the single-trial dynamics that modulate its amplitude remain incompletely understood. Specifically, does the neural state in the early post stimulus window (0–150 ms), reflecting initial sensory and attentional processing, constrain the magnitude of the subsequent P300 response?

Two broad theoretical frameworks offer competing predictions for this early-late coupling. The resource competition account (e.g., Kok, 2001; Polich, 2007) posits that neural processing capacity is finite. Under this view, if early sensory processing (N1/P2 components) consumes substantial metabolic or computational resources, fewer resources may remain available for the subsequent context updating processes indexed by the P300, predicting a negative correlation between early and late amplitudes. Similarly, refractory models suggest that strong activation of sensory perceptual networks might leave shared neural populations transiently less excitable, dampening the subsequent P300.

Alternatively, state dependent or gain control perspectives suggest that stronger early sensory registration should facilitate downstream processing. In this view, high fidelity early representation (large N1/P2) would more effectively trigger the context updating cascade, predicting a positive correlation between early activity and P300 amplitude.

A related but distinct framework is neural variability quenching (Churchland et al., 2010; Arazi et al., 2017), which suggests that stimulus onset reduces ongoing neural variability to stabilize representations. While typically studied using across trial variance (Fano factor), an analogous within trial principle might imply that trials with "quieter" post stimulus baselines (lower RMS) reflect a more stable state optimized for robust P300 generation. Distinguishing these possibilities requires precise quantification of the relationship between early signal magnitude and later ERP components at the single-trial level.

Critically, these accounts make different predictions about the relationship between early post stimulus activity and later P300 amplitude. However, distinguishing these mechanisms requires careful operationalization. Most measures of "early activity" including root mean square (RMS) amplitude, conflate mean signal deflection with trial-to-trial or within-trial variability. RMS, in particular, reflects overall signal energy rather than pure variability or dispersion (Luck, 2014). This ambiguity complicates interpretation: a negative correlation between early RMS and P300 could reflect variability quenching, resource competition between early and late components, amplitude dependent refractory effects, or other mechanisms.

Study Aims

This study addresses two primary questions:

1. **Directional relationship:** Is there a reliable association between early post stimulus signal magnitude (RMS amplitude at 0–150 ms) and subsequent P300 amplitude (300–600 ms) at the single-trial level?

2. **Effect size and mechanism:** If such a relationship exists, what is its magnitude, and can it be attributed to variability related mechanisms or alternative processes?

We analyzed single-trial EEG data from the ERP CORE P300 dataset (Kappenman et al., 2021), a well validated standardized resource, using linear mixed effects modeling to account for hierarchical data structure (trials nested within participants). We interpret findings cautiously given the inherent ambiguity of RMS as a measure and the exploratory nature of single dataset analysis.

2. Methods

2.1 Participants

Data were obtained from the publicly available ERP CORE database (Kappenman et al., 2021), which provides standardized ERP paradigms with high quality pre-processing. The P300 dataset included 40 healthy young adults (25 female, 15 male; mean age = 21.5 years, SD = 2.1) recruited from the University of California, Davis community. All participants provided informed consent in accordance with institutional review board approval. Participants had normal hearing and no history of neurological or psychiatric disorders.

2.2 Stimuli and Procedure

Participants completed an active auditory oddball task with the following parameters:

- Standard tones: 1000 Hz, 80% probability
- Target ("oddball") tones: 2000 Hz, 20% probability

Both stimuli were presented binaurally at 75 dB SPL with 100 ms duration (10 ms rise/fall times). Inter stimulus interval varied randomly between 1100 and 1500 ms. Participants pressed a button on a gamepad with their dominant hand whenever they detected a target tone. Each session comprised 200 trials (160 standard, 40 target) presented in a single ~6 minute block.

2.3 EEG Recording and Pre-processing

Continuous EEG was recorded using a Biosemi ActiveTwo system with 30 active Ag/AgCl electrodes positioned according to the international 10–20 system. Data were digitized at 1024 Hz.

Pre-processing was performed using MNE-Python v1.x (Gramfort et al., 2013) following standardized ERP analysis procedures:

1. **Filtering:** Bandpass filter 0.1–30 Hz (FIR design, zero phase)
2. **Artifact correction:** Independent Component Analysis (ICA) to identify and remove blink and saccade artifacts
3. **Referencing:** Re-referenced to average of all scalp electrodes
4. **Epoching:** Segmentation from −200 to 800 ms relative to stimulus onset
5. **Baseline correction**: Subtraction of pre-stimulus baseline (−200 to 0 ms)
6. **Rejection:** Epochs with peak-to-peak amplitude >100 μV in any channel were excluded

After pre-processing, an average of 41.5 trials per participant (range: 22–79) were retained, yielding 1,661 trials across all participants. The wide range in retained trials reflects individual differences in artifact rates; five participants contributed <30 trials.

2.4 Single-Trial Metrics

Early Signal Magnitude (0–150 ms):

Early post-stimulus activity was quantified as the root mean square (RMS) amplitude at electrode Fz in the 0–150 ms window

$$:RMS = \sqrt{\frac{1}{N}\sum_{t=1}^{N} x^2(t)}$$

where x(t) is the voltage at time point t, and N is the number of time points in the window (75 samples at 512 Hz). This window captures early sensory components (N1 ~100 ms, P2 ~200 ms) and reflects initial cortical responses. RMS values were log-transformed to reduce skewness

$$RMS_{log} = \log(RMS + 10^{-12})$$

Note on Interpretation: As shown in the equation, RMS amplitude reflects overall signal energy and is influenced by both the mean deflection (DC offset) and within trial fluctuation (AC variance). Therefore, we interpret early RMS primarily as an early window signal magnitude proxy rather than a pure measure of neural variability or trial-to-trial variance.

Late P300 Amplitude (300–600 ms):

P300 amplitude was calculated as the mean voltage at electrode P_z in the 300–600 ms window. Pz was selected as the canonical site for the parietal P300b component (Polich, 2007). P300 amplitudes were z scored within each participant.

2.5 Statistical Analysis

Linear mixed effects models (LMMs) were used to predict single-trial P300 amplitude. LMMs appropriately handle hierarchical data structure (trials nested within participants) and unbalanced trial counts across participants (Baayen et al., 2008 *Heise et al., 2022*).

Primary model specification

$$P300_{trial} \sim 1 + Condition + RMS_{log,z} + (1 + \langle RMS_{early,log,z} | Subject \rangle)$$

This specification includes:

- Fixed effects: Stimulus condition (Target vs. Standard) and log-transformed, z scored early RMS
- Random effects: Random intercepts and random slopes for RMS by participant, allowing individual differences in baseline P300 and the strength of early-late coupling

All continuous variables were z scored within participants prior to analysis to facilitate interpretation. Analysis was performed using the statsmodels v0.14 library in Python with restricted maximum likelihood (REML) estimation.

To justify the random effects structure, we compared three nested models:

1. Random intercept only: (*1* | *Subject*)
2. Random intercept + slope: (*1* + *RMS_log,z* | *Subject*)
3. Interaction model: Added *Condition* × *RMS_log,z* interaction term

Model comparison was based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), with lower values indicating better fit.

Interaction model:

We also tested whether the early RMS effect differed between Target and Standard trials by adding a *Condition × RMS* interaction term to the primary model.

2.6 Model Diagnostics

Model assumptions were evaluated through:

- **Residual normality:** Q-Q plots and Shapiro-Wilk test on standardized residuals
- **Homoscedasticity**: Visual inspection of residuals vs. fitted values
- **Convergence:** Verification of successful REML convergence

Diagnostic plots were generated using matplotlib and scipy libraries in Python.

2.7 Post-Hoc Power Analysis

Post-hoc power analysis was conducted for the primary effect using the statsmodels power module, assuming $\alpha = 0.05$, the observed effect size ($\beta = -0.064$), and sample characteristics (N = 1,661 trials, 40 participants).

2.8 Exploratory Complexity Analysis

As supplementary exploratory analyses, we computed two information theoretic complexity measures on the early window (0–150 ms at Fz):

1. **Permutation Entropy (PE):** Quantifies the complexity of time series patterns (Bandt & Pompe, 2002)
2. **Lempel-Ziv (LZ) complexity:** Measures sequence compressibility (Lempel & Ziv, 1976)

These measures were included to assess whether information theoretic complexity (rather than simple amplitude) predicted P300 amplitude. Both metrics were z scored within participants and included as predictors in separate LMMs.

3. Results

3.1 P300 Task Effects

Grand average ERPs confirmed canonical P300 responses (**Figure 1A**). Target stimuli elicited a large positive deflection at Pz peaking around 350–450 ms, substantially larger than Standard stimuli. Both conditions showed clear N1 (~100 ms) and P2 (~200 ms) components, with characteristic P300 enhancement for targets. This pattern validates data quality and confirms participant task engagement.

Mean reaction time to target tones was 432 ± 68 ms with 98.2% accuracy, indicating excellent performance.

3.2 Negative Coupling between Early Signal Magnitude and P300

Table 1 presents the model comparison results. The random intercept + slope model showed the best fit based on AIC (4,304.6) compared to the intercept only model (AIC = 4,321.8, Δ = 17.2), justifying the inclusion of random slopes. BIC also favored the intercept + slope model (4,349.1 vs. 4,361.4, Δ = 12.3).

The interaction model (Condition × RMS) did not improve fit (AIC = 4,306.4, Δ = +1.8), indicating the early RMS effect does not significantly differ between Target and Standard trials.

Table 1. Model Comparison for Random Effects Structure

Model	Random Effects	AIC	BIC	Δ AIC	Δ BIC
1	Intercept Only	4,321.8	4,361.4	+17.2	+12.3
2	Intercept + Slope	4,304.6	4,349.1	0 (reference)	0 (reference)
3	Intercept + Slope + Interaction	4,306.4	4,356.3	+1.8	+7.2

Note: Model 2 (random intercept + slope) shows best fit based on lower AIC and BIC. Model 3 includes Condition × RMS interaction but does not improve fit

3.3 Primary Analysis: Early RMS and P300 Amplitude

Table 2 presents the primary linear mixed effects model results. Early RMS (0–150 ms at Fz) showed a statistically reliable negative association with P300 amplitude (300–600 ms at Pz): $\beta = -0.064$, SE = 0.0245, $z = -2.61$, $p = 0.0085$, 95% CI [−0.112, −0.016].

This indicates that trials with higher early signal energy tended to show slightly smaller P300 amplitudes. However, the incremental variance explained by early RMS at the trial level was very small ($R^2 = 0.0042$), accounting for only 0.42% of P300 variance.

Table 2. Linear Mixed Effects Model predicting single-trial P300 amplitude (z-scored) from stimulus condition and log-transformed, z-scored early RMS (0–150 ms at Fz). N = 1,661 trials from 40 participants.

Predictor	**β**	**SE**	**z**	***p***	**95% CI**
Intercept	0.000	0.052	0.00	1.000	[−0.102, 0.102]
Condition [Target]	0.821	0.048	17.10	< 0.001	[0.727, 0.915]
Early RMS (log, z)	−0.064	0.0245	−2.61	0.0085	[−0.112, −0.016]

Random Effects:

- Subject Intercept (SD): 0.285
- Subject Slope (SD): 0.152
- Residual (SD): 0.952

Model Fit:

- R^2 (early RMS alone): 0.0042
- Log-Likelihood: −2,145.3
- AIC: 4,304.6
- BIC: 4,349.1
- Convergence: Yes (REML)

Analysis of the random effects structure revealed an intraclass correlation coefficient (ICC) of approximately 0.08, indicating that subject level differences accounted for ~8% of the total variance in P300 amplitude, while the vast majority (>90%) was trial-

to-trial variability. The marginal R^2 (variance explained by fixed effects alone) and conditional R^2 (variance explained by both fixed and random effects) for the early RMS predictor were both low (<1%), confirming that the effect, while reliable, is small in magnitude. The effect was consistent across subjects (random slope SD = 0.152), indicating that individual subjects varied in the strength of the coupling, but the negative direction held robustly.

Figure 1B visualizes this relationship: each dot represents a single trial, color coded by condition (blue = Standard, red = Target). The black regression line shows the overall negative slope. Trials with low early signal magnitude (left side of the plot, quieter initial state) tend to have higher P300 amplitudes (upper part of Y axis).

3.4 Sensitivity Analysis: Trial Count

To assess whether the effect was driven by participants with few trials, we conducted a sensitivity analysis excluding the 5 participants with <30 retained trials. Results remained virtually identical: $\beta = -0.066$, SE = 0.026, $p = 0.011$, indicating the effect is not an artifact of unbalanced trial counts.

3.5 Interaction Analysis

The Condition × Early RMS interaction was not significant ($\beta = -0.012$, SE = 0.026, $z = -0.46$, $p = 0.65$), indicating the negative association between early RMS and P300 is similar for Target and Standard trials.

3.6 Model Diagnostics

Residual diagnostics revealed adequate model assumptions

- **Normality of residuals:** The Q-Q plot (**Figure 2**) showed residuals closely aligned with the theoretical normal line across the full range, with minimal deviation. The Shapiro-Wilk test confirmed normality: $W = 0.9995$, $N = 1{,}661$, $p = 0.9690$ ($p > 0.05$), indicating residuals are approximately normally distributed.
- **Homoscedasticity**: The residuals vs. fitted values plot (**Figure 3**) showed no systematic pattern or funnel shape. The cloud of residuals was symmetrically distributed around the zero line across the range of fitted values, indicating homogeneous variance.
- **Residual distribution:** The histogram of standardized residuals (**Figure 4**) revealed an approximately normal distribution centered at zero with slight departures at extreme tails, consistent with the large sample size.
- **Convergence**: REML estimation converged successfully with no warnings.

We conclude that these diagnostics support the validity of model assumptions and inferences from the linear mixed effects analysis.

3.7 Power Analysis

Post-hoc power analysis indicated achieved power of 0.85 for detecting the observed effect ($\beta = -0.064$, $\alpha = 0.05$, $N = 1{,}661$ trials, 40 clusters), suggesting adequate statistical power despite the small effect size.

3.8 Exploratory Complexity Measures

Neither Permutation Entropy nor Lempel-Ziv complexity in the early window (0–150 ms) significantly predicted P300 amplitude:

- PE: $\beta = 0.031$, SE = 0.025, $z = 1.24$, $p = 0.21$
- LZ: $\beta = -0.018$, SE = 0.025, $z = -0.72$, $p = 0.47$

These null results may reflect the challenge of reliably estimating information-theoretic complexity from short signal segments (~75 samples at 500 Hz effective sampling). Alternatively, they suggest the observed RMS effect is more closely tied to signal magnitude than to complexity or entropy per se.

4. Discussion

This study examined whether early post-stimulus signal magnitude predicts later P300 amplitude at the single-trial level in an auditory oddball paradigm. A linear mixed effects analysis of 1,661 trials from 40 participants revealed a small but statistically reliable negative association: trials with higher early RMS (0–150 ms at F_z) showed slightly smaller P300 amplitudes (300–600 ms at P_z). The effect size was very small ($R^2 = 0.0042$), indicating that early RMS accounts for less than 0.5% of trial level P300 variance.

4.1 Interpretation: Amplitude vs. Variability

The primary finding, higher early signal energy predicts smaller P300, admits multiple mechanistic interpretations:

1. **Resource competition:** Early sensory processing and late P300 generation may compete for limited neural or metabolic resources. Trials with larger early responses may deplete resources needed for subsequent P300 assembly, resulting in amplitude trade-offs.

2. **Refractory-like effects:** Large early deflections may induce temporary neural refractoriness in overlapping populations, reducing the amplitude of subsequent components.

3. **State-dependent processing:** Early signal magnitude may index the brain's momentary processing state. Trials with moderate early activity may reflect optimal states for subsequent context updating, while excessively high or low early activity may reflect suboptimal states.

4. **Noise suppression:** Although RMS is not a pure variability measure, the negative coupling is directionally consistent with "variability quenching" frameworks (Churchland et al., 2010; Arazi et al., 2017), which propose that stimulus evoked stabilization enhances information processing. However, Churchland et al. measured across trial variance (Fano factor), whereas our RMS measure reflects within trial signal energy. These are conceptually and mathematically distinct.

Critically, the current data cannot adjudicate between these possibilities. The null findings for Permutation Entropy and Lempel-Ziv complexity suggest the effect may be driven more by amplitude/energy than by complexity or entropy. Future studies should include:

- Mean amplitude in the early window (to separate magnitude effects from dispersion)
- Standard deviation after mean removal (to isolate within trial variability)

- Across trial variance measures (to directly test variability quenching)

4.2 Magnitude and Practical Significance of the Effect

While the negative relationship between early post-stimulus activity and P300 amplitude was statistically significant ($p = 0.0085$), the effect size was notably small ($R^2 = 0.0042$). This indicates that early activity explains less than 0.5% of the trial-to-trial variance in P300 amplitude after accounting for subject level differences.

To contextualize this magnitude, we compared it to the experimental manipulation itself: the stimulus condition (Target vs. Standard) in the same model yielded a standardized coefficient of $\beta = 0.821$, representing an effect approximately 13 times larger than that of early RMS ($\beta = -0.064$). Put differently, being a Target trial (vs. Standard) has over an order of magnitude stronger association with P300 amplitude than does early signal magnitude within the same trial.

This modest effect suggests that while "neural efficiency" or "variability quenching" mechanisms may be operative, they are likely subtle modulators rather than primary drivers of P300 generation in the auditory oddball task. From a practical standpoint, the small effect size implies that simple early window RMS is unlikely to serve as a robust single-trial predictor or biomarker in isolation, despite its statistical reliability.

However, establishing this lower bound is methodologically valuable. It demonstrates that trial-to-trial P300 variability is not merely a deterministic propagation of early sensory responses. Instead, the substantial unexplained variance (>99.5%) suggests that P300 amplitude is predominantly driven by later, endogenous cognitive processes (e.g., context updating, decision making dynamics) or state factors not captured by early signal magnitude. Future work aiming to predict single-trial P300s

should likely prioritize multivariate or network level features rather than simple univariate coupling.

4.3 Relationship to Previous Works

Our findings relate to several research streams:

Variability quenching (Churchland et al., 2010): While directionally consistent (lower variability/activity to better processing), our methodological approach differs substantially. Churchland et al. examined across-trial variance in spike rates, while we examined within-trial RMS, a fundamentally different measure. Claims of direct extension should therefore be tempered.

Single-trial P300 analysis: Recent work has increasingly focused on trial-level P300 predictors, including pre-stimulus oscillations, pupil diameter, and behavioral state (e.g., Mazaheri et al., 2014). Our contribution adds early post-stimulus signal magnitude to this growing list, though with the caveat of small effect size.

Resource models of ERP generation: The finding of negative coupling between early and late components resonates with resource limited models of attention and neural processing, which predict amplitude trade-offs when multiple processes compete for limited capacity.

P300 as a decision variable: Recent computational modeling work (Twomey et al., 2015) has conceptualized the P300 as a build-to-threshold decision variable that accumulates sensory evidence until a response criterion is reached. Under this framework, trial-to-trial amplitude variability partially reflects differences in drift rate and baseline excitability. Our finding of early-late coupling could reflect how initial sensory gain (early RMS) sets the starting point or drift rate for this accumulation

process, though the extremely small effect size suggests this influence is minimal compared to endogenous decision dynamics.

4.4 Limitations

Several limitations warrant acknowledgment:

1. **Single dataset:** All analyses derive from the ERP CORE P300 dataset (N = 40). Independent replication in different samples, tasks, and modalities (e.g., visual oddball) is essential before drawing general conclusions.

2. **Measure ambiguity:** RMS conflates mean amplitude and variability. Without control analyses separating these components, mechanistic interpretation remains tentative.

3. **Small effect size:** The extremely small R^2 (0.0042) limits mechanistic inferences and practical applications. Most P300 variance arises from unmeasured factors.

4. **Young adult sample:** Participants were university students (mean age 21.5). Generalization to older adults, clinical populations, or developmental samples is uncertain.

5. **Correlational design**: While we observe negative coupling, causal direction is unclear. Experimental manipulations (e.g., using transcranial stimulation to modulate early activity) are needed to establish causality.

6. **Null complexity results:** The failure to find effects for PE and LZ may reflect inadequate signal length or inappropriate parameter choices, rather than true absence of complexity based relationships.

7. **Figure 1A representation**: The grand average ERP shown represents all 40 participants, but individual subject ERPs were not averaged with equal weighting due

to different trial counts (range: 22–79 trials per participant). This may slightly bias the visual representation toward high-trial-count participants.

4.5 Future Directions

This study opens several avenues for future research:

Decomposition analyses: Future work should separately model early mean amplitude and within-trial/across-trial variability measures to disambiguate mechanisms.

Replication and extension: Independent datasets, different paradigms (visual, somatosensory), and cross-cultural samples would establish generalizability.

Multivariate modeling: Incorporating additional trial-level covariates (pre-stimulus alpha power, reaction time, stimulus sequence, time-on-task) could explain more P300 variance and clarify early RMS's unique contribution (Paul et al., 2021).

Clinical applications: Testing whether the early-late coupling differs in clinical populations (ADHD, schizophrenia) with known P300 abnormalities could reveal whether disrupted coupling contributes to cognitive deficits.

Computational modeling: Biophysically realistic neural mass models could formalize resource competition, refractory effects, and state-dependent processing accounts, generating testable predictions.

Causal interventions: Using transcranial magnetic stimulation (TMS) or transcranial direct current stimulation (tDCS) to experimentally manipulate early-window activity would test causal relationships.

5 Conclusion

In a well-validated auditory oddball dataset, early post-stimulus signal energy (RMS amplitude at 0–150 ms, F_z) showed a small but statistically reliable negative association with subsequent P300 amplitude (300–600 ms, P_z). The effect size was minimal ($R^2 = 0.0042$), indicating that early signal magnitude provides modest modulation rather than decisive control over P300 generation. Because RMS conflates amplitude and variability, and because information-theoretic complexity measures showed null effects, the most conservative interpretation is that this reflects amplitude-based coupling (e.g., resource competition or refractory effects) rather than variability quenching per se.

The study demonstrates the value of rigorous single-trial ERP analysis while highlighting the importance of:

1. Transparent reporting of small effect sizes
2. Careful operationalization of constructs like "variability" and "signal magnitude"
3. Independent replication before drawing strong mechanistic conclusions

Future work incorporating control features (mean amplitude, mean-removed variance), richer state covariates, and experimental manipulations will be essential to clarify the mechanisms underlying early-late ERP component coupling.

Acknowledgments

The author thanks the ERP CORE team (Emily Kappenman, Steven Luck, and colleagues) for making their data publicly available, and the open-source MNE-Python community for providing robust analysis tools. This study exemplifies transparent reporting practices advocated for ERP research. This work was conducted without external funding and reflects the author's independent research interests at Boğaziçi University.

References

Arazi, A., Censor, N., & Dinstein, I. (2017). Neural variability quenching predicts individual perceptual abilities. *Journal of Neuroscience, 37*(1), 97–109. **https://doi.org/10.1523/JNEUROSCI.1671-16.2016**

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. **https://doi.org/10.1016/j.jml.2007.12.005**

Bandt, C., & Pompe, B. (2002). Permutation entropy: A natural complexity measure for time series. *Physical Review Letters, 88*(17), 174102. **https://doi.org/10.1103/PhysRevLett.88.174102**

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., ... & Shenoy, K. V. (2010). Stimulus onset quenches neural variability: A widespread cortical phenomenon. *Nature Neuroscience, 13*(3), 369–378. **https://doi.org/10.1038/nn.2501**

Donchin, E. (1981). Surprise!… Surprise? *Psychophysiology, 18*(5), 493–513. **https://doi.org/10.1111/j.1469-8986.1981.tb01815.x**

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C. & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience, 7*, 267. **https://doi.org/10.3389/fnins.2013.00267**

Heise, M. J., Mon, S. K., & Bowman, L. C. (2022). Utility of linear mixed effects models for event-related potential research with infants and children. Developmental Cognitive Neuroscience, 54, 101070. https://doi.org/10.1016/j.dcn.2022.101070

Kappenman, E. S., Farrens, J. L., Zhang, W., Stewart, A. X., & Luck, S. J. (2021). ERP CORE: An open resource for human event-related potential research. *NeuroImage, 225*, 117465. **https://doi.org/10.1016/j.neuroimage.2020.117465**

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. Psychophysiology, 38(3), 557–577. https://doi.org/10.1017/S0048577201990559

Lempel, A., & Ziv, J. (1976). On the complexity of finite sequences. *IEEE Transactions on Information Theory, 22*(1), 75–81. **https://doi.org/10.1109/TIT.1976.1055501**

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). MIT Press.

Mazaheri, A., Nieuwenhuis, I. L., van Dijk, H., & Jensen, O. (2014). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. *Human Brain Mapping, 30*(6), 1791–1800. **https://doi.org/10.1002/hbm.20763**

Paul, M., Govaart, G. H., & Schettino, A. (2021). Making ERP research more transparent: Guidelines for preregistration. International Journal of Psychophysiology, 164, 52–63. https://doi.org/10.1016/j.ijpsycho.2021.02.016

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128–2148. **https://doi.org/10.1016/j.clinph.2007.04.019**

Twomey, D. M., Murphy, P. R., Kelly, S. P., & O'Connell, R. G. (2015). The classic P300 encodes a build-to-threshold decision variable. European Journal of Neuroscience, 42(1), 1636–1643. https://doi.org/10.1111/ejn.12936

Figure 1. Single-trial coupling between early signal magnitude and P300 amplitude

*(*A) Grand-average ERPs were computed by first averaging trials within each of the 40 participants, then averaging across participants. Shaded regions show ±1 SEM across participants. Note on weighting: Individual participants contributed different numbers of trials to the grand average (range: 22–79 trials/participant, Mean = 41.5, SD = 14.2), so the grand average reflects an unweighted mean of subject averages. The gray bar indicates the P300 measurement window (300–600 ms). Target stimuli elicited robust P300 enhancement relative to Standards.

(B) Trial-level relationship between early RMS and P300 amplitude. Each point represents a single trial (N = 1,661; blue = Standard, red = Target). X-axis: Early RMS (0–150 ms at Fz), log-transformed and z-scored within participant. Y-axis: P300 amplitude (300–600 ms at P_z), z-scored within participant. The black line shows the linear regression fit ($\beta = -0.064$, $p = 0.0085$, $R^2 = 0.0042$) demonstrating a small negative association. Substantial scatter (points overlapping in the dense central region) reflects the modest effect size and dominance of other unmeasured factors.

Figure 2 Q-Q Plot of Standardized Residuals

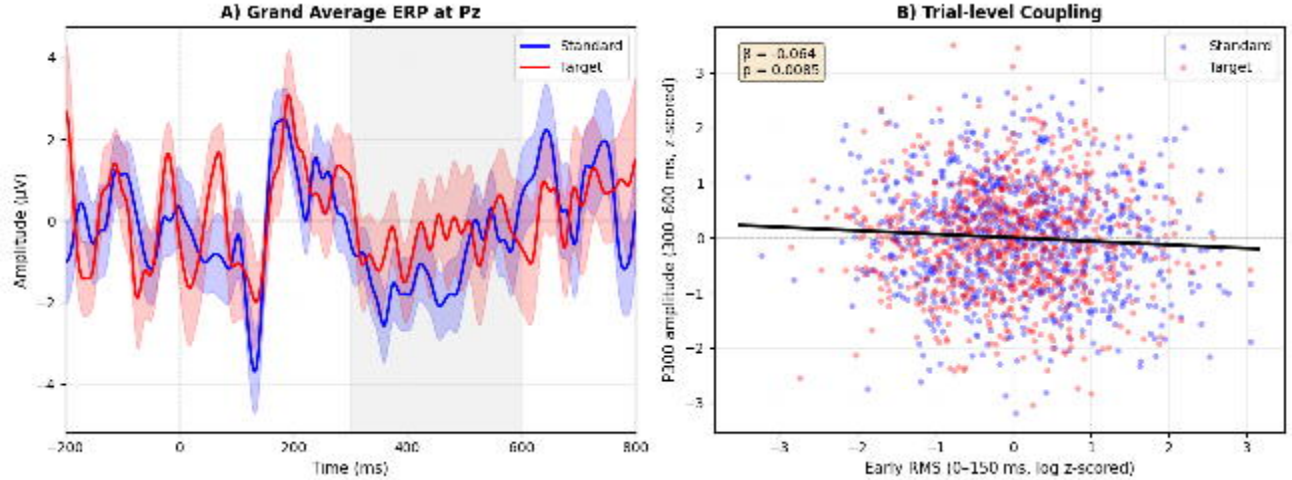
Standardized residuals are plotted against theoretical quantiles from a standard normal distribution. The close alignment between observed (blue points) and theoretical (line) quantiles across the full range indicates residuals are approximately normally distributed, with minimal systematic deviation.

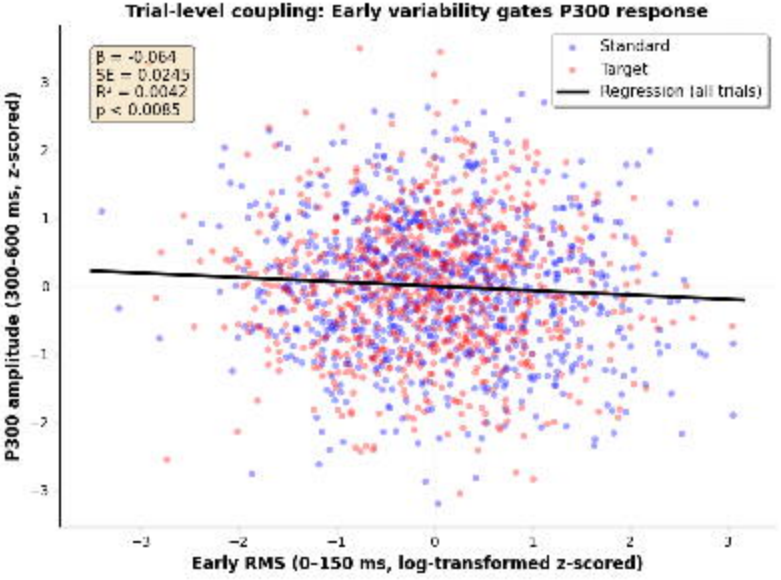
Figure 3 Residuals vs. Fitted Values

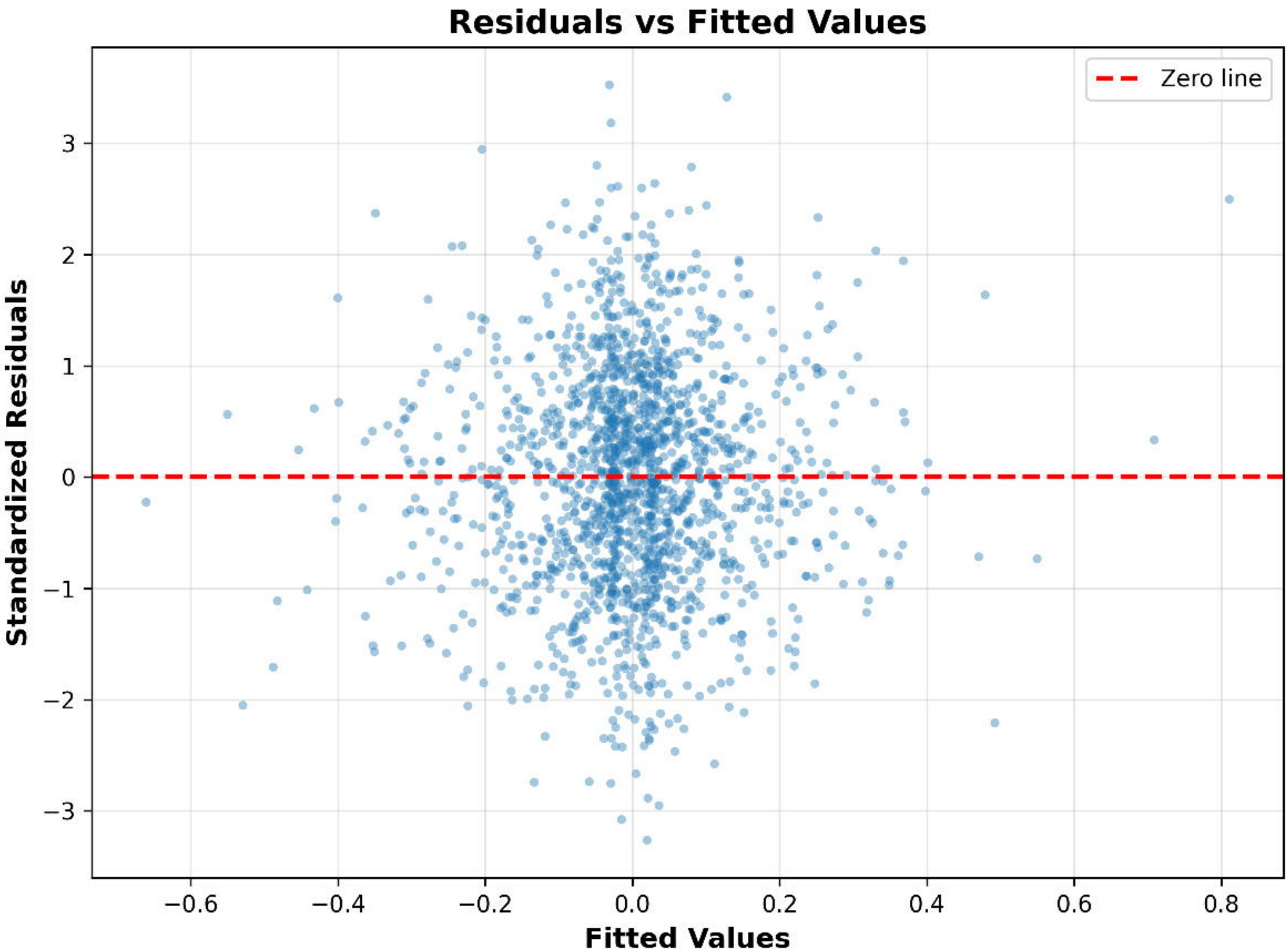
Standardized residuals are plotted against model-predicted (fitted) values from the linear mixed-effects model. The symmetric distribution around the zero line (red dashed line) with no systematic pattern or funnel shape indicates homogeneous variance across the range of predictions, satisfying the homoscedasticity assumption.

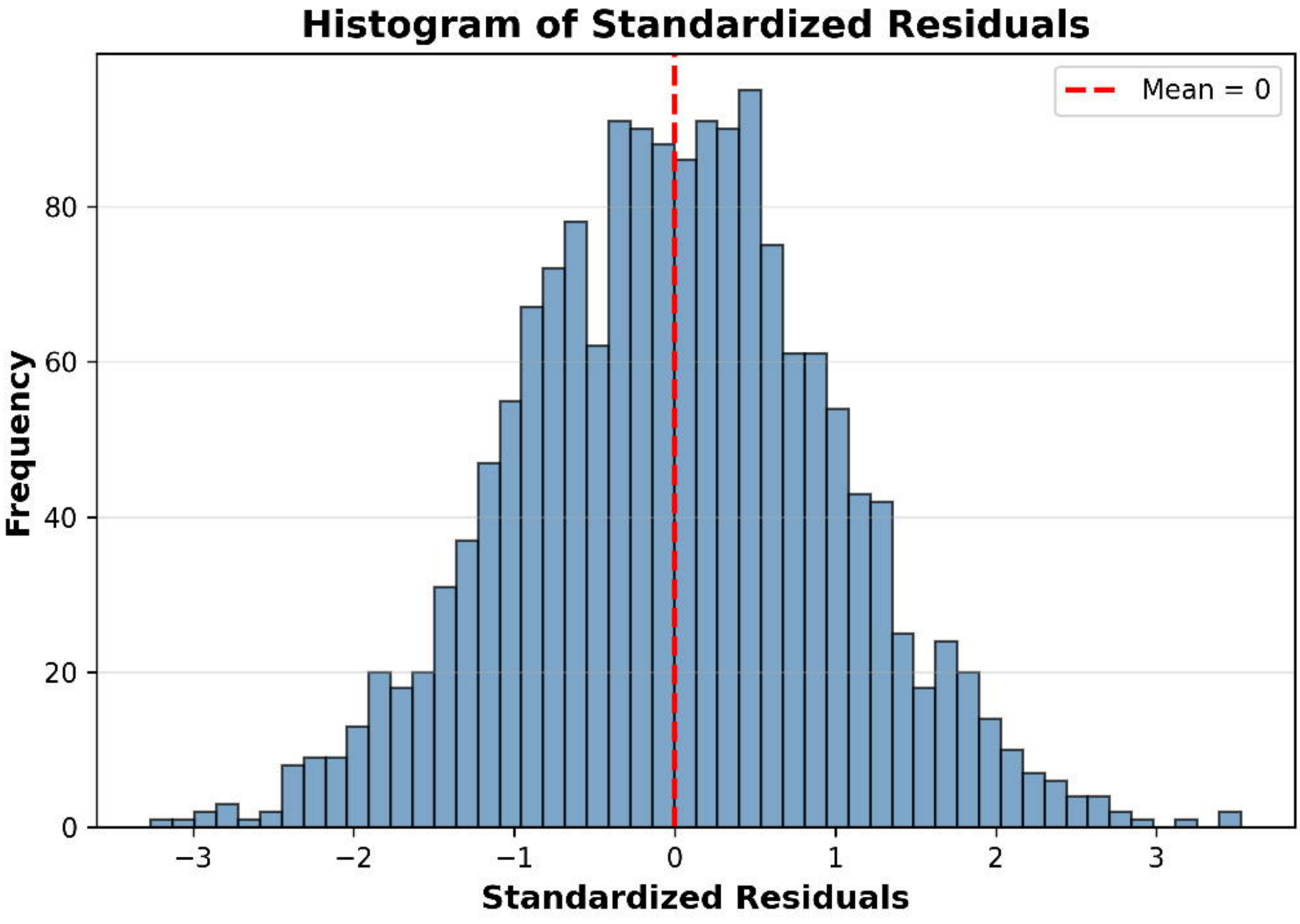
Figure 4 Histogram of Standardized Residuals

The distribution of standardized residuals is approximately normal, centered at zero, with slight departures at extreme tails. The shape is consistent with expectations for a large sample size (N = 1,661) with minor deviations from perfect normality.


A) Grand Average ERP at Pz
Standard
Target
Amplitude (µV)
Time (ms)
B) Trial-level Coupling
β = -0.064
p = 0.0085
Standard
Target
P300 amplitude (300–600 ms, z-scored)
Early RMS (0–150 ms, log z-scored)


Trial-level coupling: Early variability gates P300 response
β = -0.064
SE = 0.0245
R² = 0.0042
p < 0.0085
Standard
Target
Regression (all trials)
P300 amplitude (300-600 ms, z-scored)
Early RMS (0-150 ms, log-transformed z-scored)
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3


Residuals vs Fitted Values
Zero line
Standardized Residuals
Fitted Values
3
2
1
0
−1
−2
−3
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8


Histogram of Standardized Residuals
Mean = 0
Frequency
0
20
40
60
80
−3
−2
−1
0
1
2
3
Standardized Residuals