Bayesian Weighted Averaging Improves Signal-to-Noise Ratio of Human Frequency-Following Responses Without Altering Phase Locking

Lizzie DeLeone, Jadyn Shipman, Fuh-Cherng Jeng, and Sydney W. Bauer

Hearing, Speech and Language Sciences, Ohio University, Athens, Ohio, USA

INTRODUCTION

- **Frequency-following responses (FFRs)** reflect sustained phase-locked neural activity to periodic sounds and provide an objective index of temporal precision and neural synchrony in the auditory pathway. FFRs are widely used to study speech and music encoding, plasticity, and auditory dysfunction (Krizman & Kraus, 2019; Skoe & Kraus, 2010).
- Despite their value, FFRs are small in amplitude and embedded in ongoing EEG activity. Response quality is influenced by neural variability, background noise, recording duration, and participant state, making reliable quantification challenging (Jeng et al., 2025).
- Conventional FFR acquisition improves signal-to-noise ratio (SNR) by averaging a fixed number of stimulus-locked sweeps. However, simple averaging assumes equal sweep quality, even though individual trials often differ substantially in noise contamination.
- **Bayesian weighted averaging** offers a principled alternative: when noise varies across observations, optimal estimation theory supports **inverse-variance weighting**, allowing cleaner sweeps to contribute more strongly than noisier ones (Bishop, 2006; McKearney et al., 2023).
- Applying Bayesian weighted averaging to FFRs may enhance SNR by reducing the influence of contaminated epochs while preserving the underlying temporal structure and phase locking of the neural response.
- This study tests whether Bayesian weighting:
  1. Increases SNR relative to simple averaging
  2. Relates SNR improvement to intrinsic response reliability
  3. Preserves waveform morphology and phase-locking strength.

METHODS

Participants

- 96 college students (24.1 ± 3.5 years old) with normal hearing

Acoustic Stimulus

- An English vowel /i/ with a rising frequency contour (F0 ranging from 102 to 140 Hz), with a duration of 150 ms.
- The stimulus was presented at 3.333 stimuli/s at 70 dB SPL to the right ear.

EEG Recordings

- 3 gold-plated surface recording electrodes
- 8000 accepted sweeps for each recording

Two Averaging procedures

- *Simple Averaging* = $\frac{\sum_{i=1}^{k}(EEG_i)}{k}$
- *Bayesian weighted Averaging* = $\frac{\sum_{i=1}^{k}(EEG_i\,\omega_i)}{\sum_{i=1}^{k}(\omega_i)}$

Outcome Measures

- SNR Improvement
- Effective Sweep Ratio
- FFR Reliability Index
- Phase-Locking Value (PLV) Difference

Statistical Analyses

- Shapiro–Wilk normality test for distributions of SNR Improvements and PLV differences
- Wilcoxon signed-rank test between Simple and Bayesian averaging procedures

RESULTS

EEG Recording

Figure 1. Gold plate electrodes were placed on the high forehead, right mastoid, and low forehead to pick up neural activity elicited by speech stimuli through an insert earphone in the right ear. Participants were encouraged to remain relaxed and still throughout testing.

Bayesian Weighted Averaging

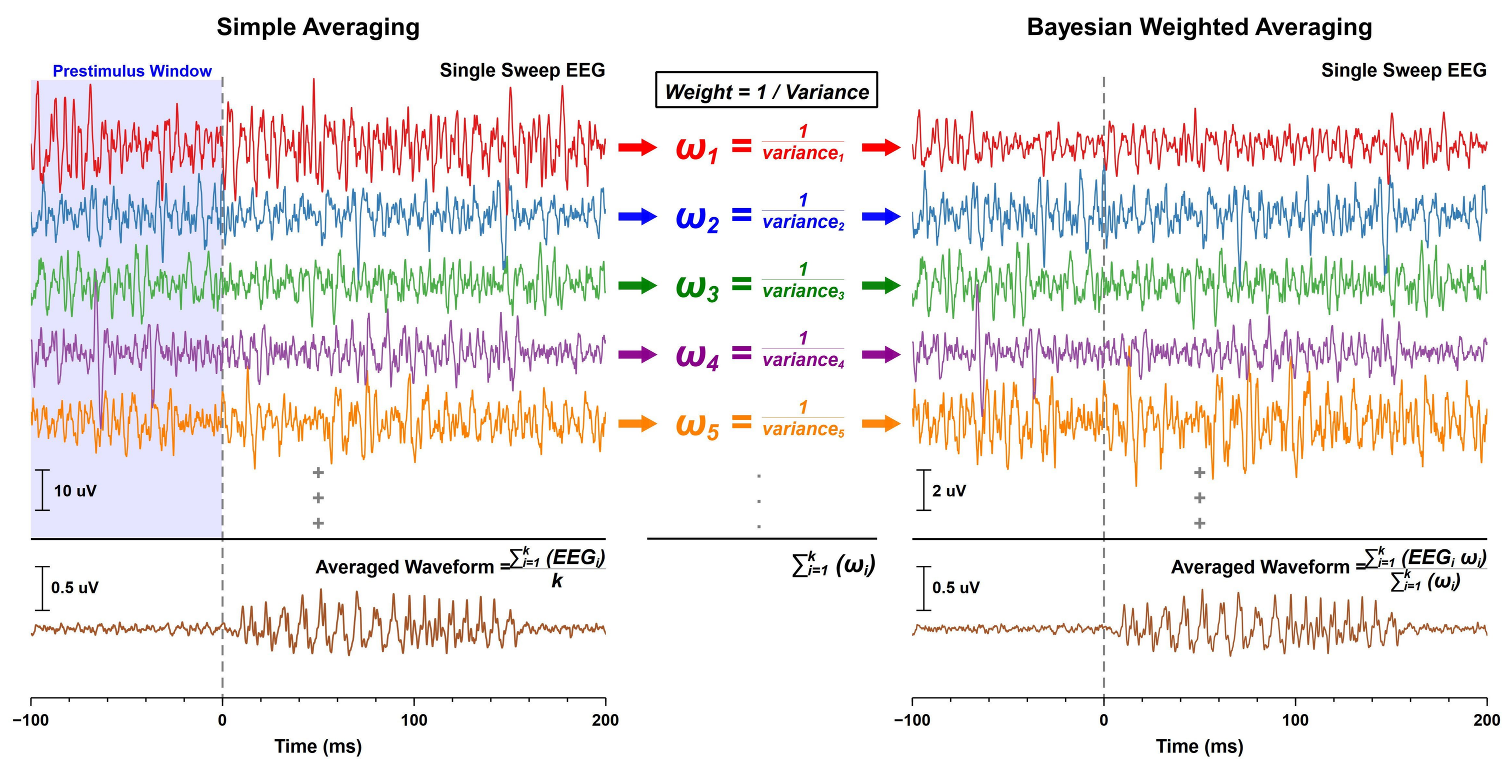

Figure 2. Conceptual illustration of simple averaging and Bayesian weighted averaging procedures. EEG recordings were obtained from a representative participant.

- *Simple Averaging* = $\frac{\sum_{i=1}^{k}(EEG_i)}{k}$
- *Bayesian weighted Averaging* = $\frac{\sum_{i=1}^{k}(EEG_i\,\omega_i)}{\sum_{i=1}^{k}(\omega_i)}$

SNR Improvement

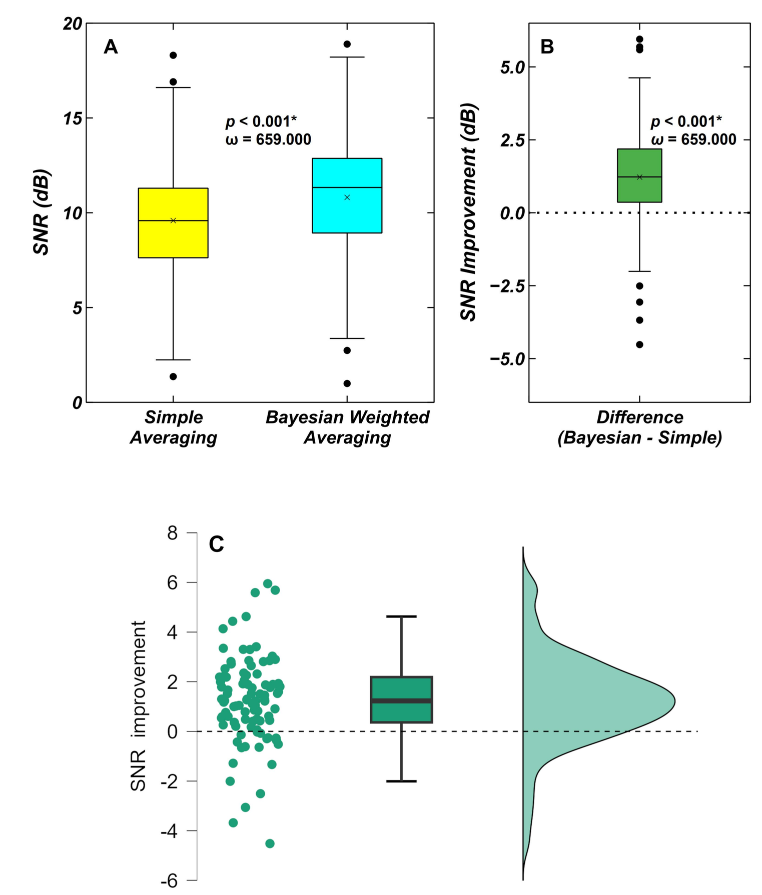

Figure 3. Comparison of SNR values (A) and SNR Improvement (B & C). A dotted horizontal line in panels B and C shows a reference line where SNR Improvement is zero.

ELECTROPHYSIOLOGICAL ASSESSMENT

Differential Improvement (Noisy Recordings Benefits More)

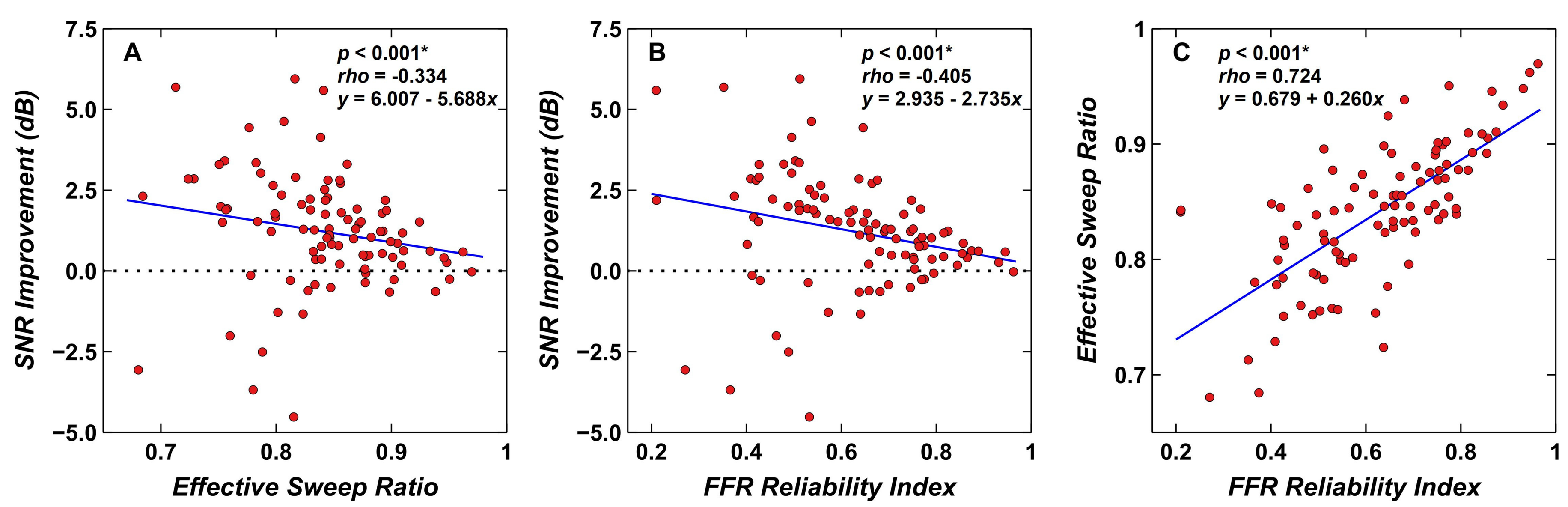

Figure 5. Associations among SNR Improvement, Effective Sweep Ratio, and FFR Reliability Index. Scatter plots depicting pairwise relationships among SNR Improvement, Effective Sweep Ratio, and the FFR Reliability Index. Spearman correlation coefficients (rho) and associated p values are shown in each panel. Solid lines represent simple linear regression fits. Dotted horizontal lines indicate a reference line with zero SNR Improvement.

A. ***SNR Improvement negatively correlated with Effective Sweep Ratio.***
B. ***SNR Improvement negative correlated with FFR Reliability Index.***
C. ***Effective Sweep Ratio positively correlated with FFR Reliability Index.***

Phase Locking - Unchanged

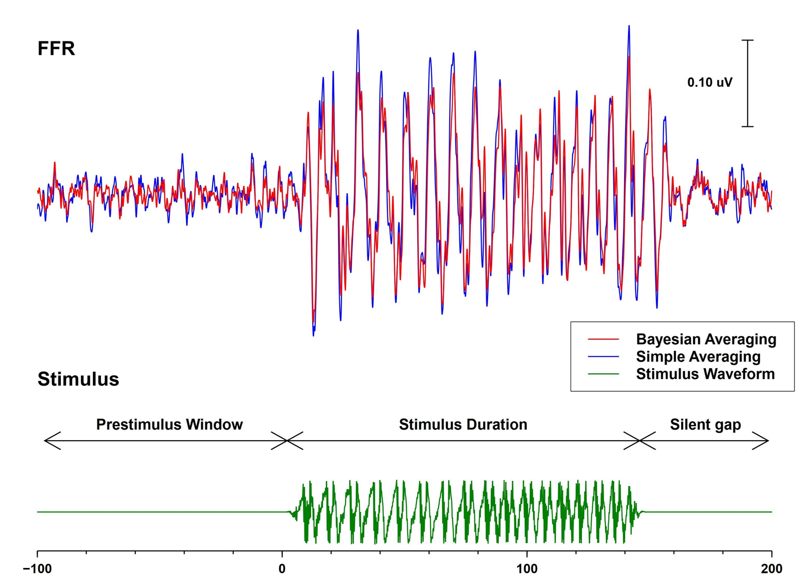

Figure 6. Across all participants, the mean cross-correlation coefficient was 0.926 (standard deviation = 0.066), indicating a high degree of morphological similarity between the two averaged waveforms. Similarly, PLV Difference (i.e., PLV_Bayesian – PLV_Simple) values had a mean of −0.000002180 (standard deviation = 0.00002136), which was effectively zero, indicating ***negligible deviation between the two approaches*** at the individual level.

DISCUSSION

Four Major Findings:

1. Bayesian weighting significantly increased SNR relative to conventional averaging.
2. SNR improvements were largest in recordings with lower effective sweep ratios and reduced reliability.
3. Waveform morphology was preserved across methods.
4. Phase-locking values were statistically unchanged, indicating improved robustness without altering neural synchrony.

Relationship Between SNR Improvement and Data Quality Metrics

SNR improvement was inversely associated with both Effective Sweep Ratio and the FFR Reliability Index, demonstrating that weighting preferentially benefits recordings with lower initial quality. The strong association between these metrics supports their validity as indicators of recording robustness and suggests that Bayesian weighting functions as an adaptive estimator rather than a uniform signal enhancer.

Neurophysiological and Methodological Implications

By preserving morphology and phase locking, Bayesian weighting maintains physiological fidelity while enhancing detectability. The approach may reduce required sweep counts, improve feasibility in clinical populations, and enhance reproducibility.

Limitations and Future Directions

Noise was estimated from prestimulus activity; alternative modeling strategies warrant exploration. Generalization to other stimuli and clinical populations remains to be determined. Future work should examine test–retest reliability, acquisition efficiency, and integration with adaptive stopping rules.

ACKNOWLEDGMENTS

This study was supported in part by a Student Research Grant (FY25-596) to J. S. and L. D. and F.-C. J. Research Funds at Ohio University.

REFERENCES

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.

Jeng, F.-C., Carriero, A. E., & Bauer, S. W. (2025). Detecting Human Frequency-Following Responses Using an Artificial Neural Network. Perceptual and Motor Skills, 00315125251347006. https://doi.org/10.1177/00315125251347006

Krizman, J., & Kraus, N. (2019). Analyzing the FFR: A tutorial for decoding the richness of auditory function. Hearing Research, 382, 107779. https://doi.org/10.1016/j.heares.2019.107779

McKearney, R. M., Bell, S. L., Chesnaye, M. A., & Simpson, D. M. (2023). Optimising weighted averaging for auditory brainstem response detection. Biomedical Signal Processing and Control, 83, 104676. https://doi.org/10.1016/j.bspc.2023.104676

Skoe, E., & Kraus, N. (2010). Auditory brain stem response to complex sounds: A tutorial.