Neural encoding of temporal and spectral statistical regularities of reverberant enviroments

SMITH-KETTLEWELL EYE RESEARCH INSTITUTE

Haydée García-Lázaro*, Yavin Alwis*, and Santani Teng*

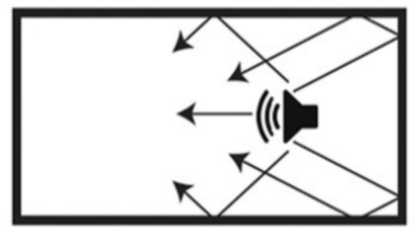
*The Smith-Kettlewell Eye Research Institute, San Francisco, CA, USA

Contact: haydee@ski.org

Poster 2653

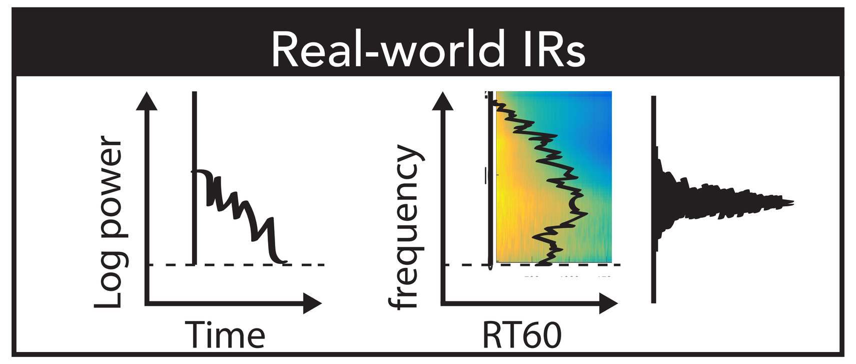
1. BACKGROUND

Auditory scene analysis is the process of perceptually segregating multiple sound sources and other signals such as **reverberation**, the aggregated acoustic reflections from multiple nearby surfaces.



Reverberation

Human listeners are perceptually sensitive to the **temporal and spectral statistical regularities** of reverberation, facilitating perceptual segregation in scene analysis[1], and coding sound sources and their reverberant impulse responses (IRs) separably in the brain[2].



Example properties of reverb

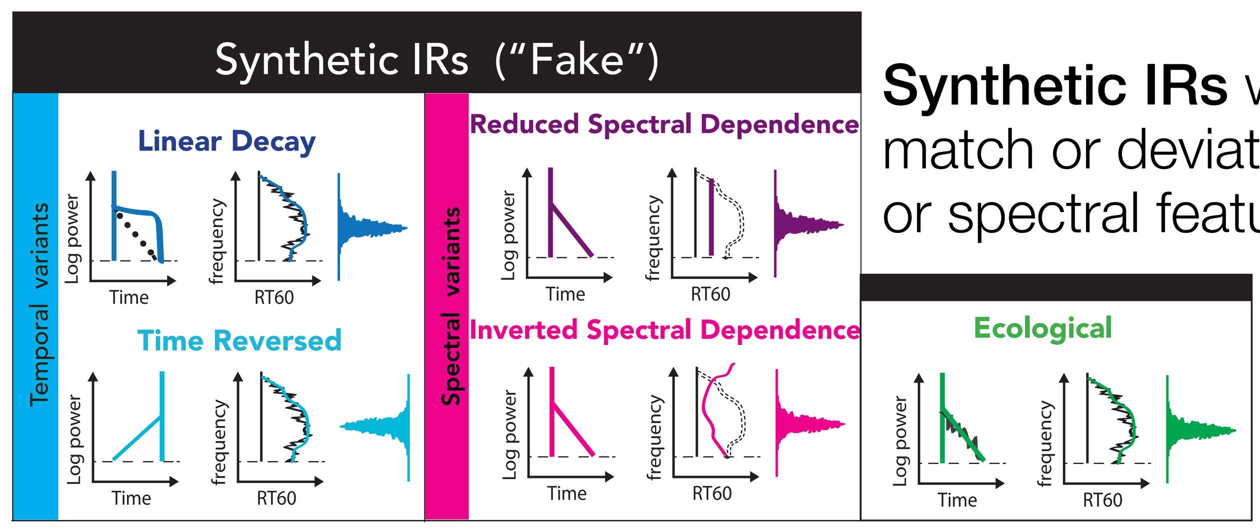
- exponential decay
- a frequency-dependent decay profile.

How do the neurodynamics of reverberant scene analysis in humans track acoustics vs. perceptual properties?

2. METHODS

Subjects: 14 adults (mean age = 32.3, SD = 5.7 years, 9 male).

Stimuli: 600 **reverberant sound sources** created by convolving a **sound source** (spoken sentence) with a **reverberant impulse response (IR)** characterizing the structure of a sampled space.

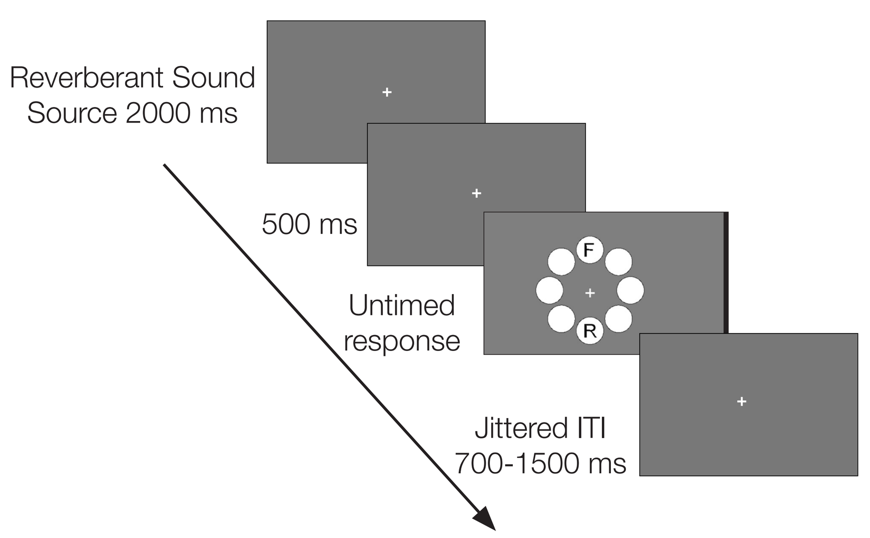


Synthetic IRs were generated to match or deviate from the temporal or spectral features of real-world IRs[1].

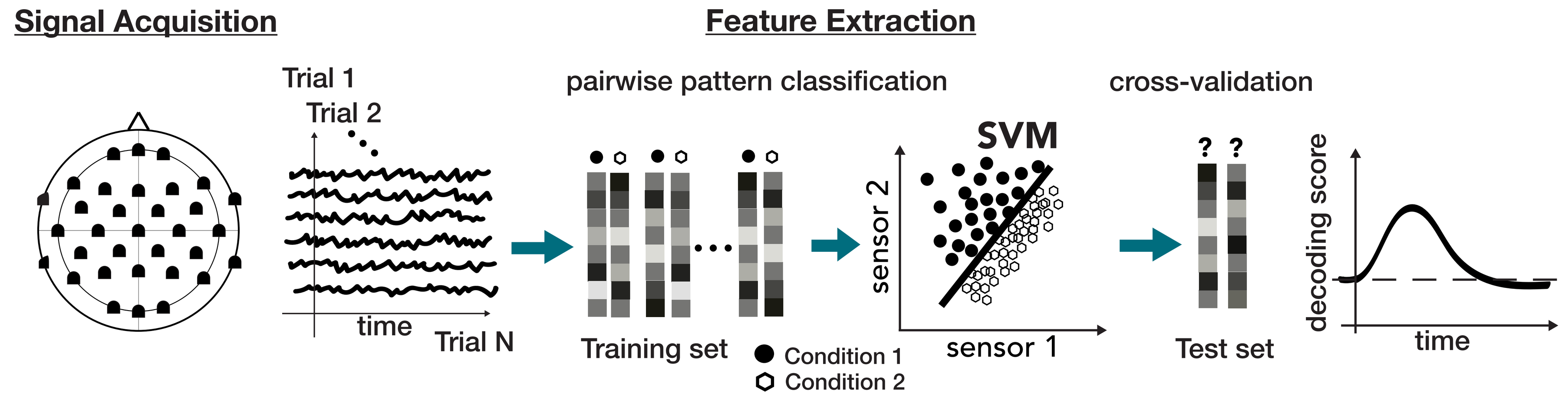
Auditory perceptual task

Judge IR: Real (R) or "Fake" (F) environment?

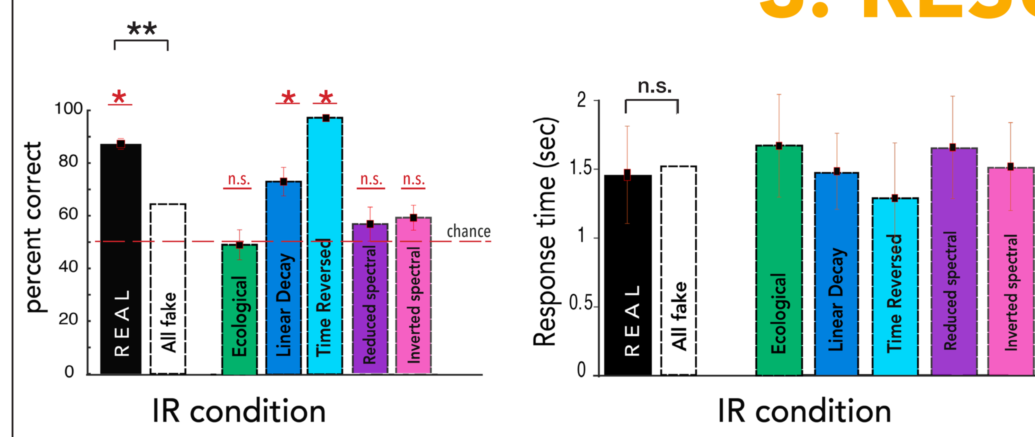
IRs: Real-world IRs 50%; Synthetic IRs 50%



EEG Multivariate Pattern Analysis



3. RESULTS

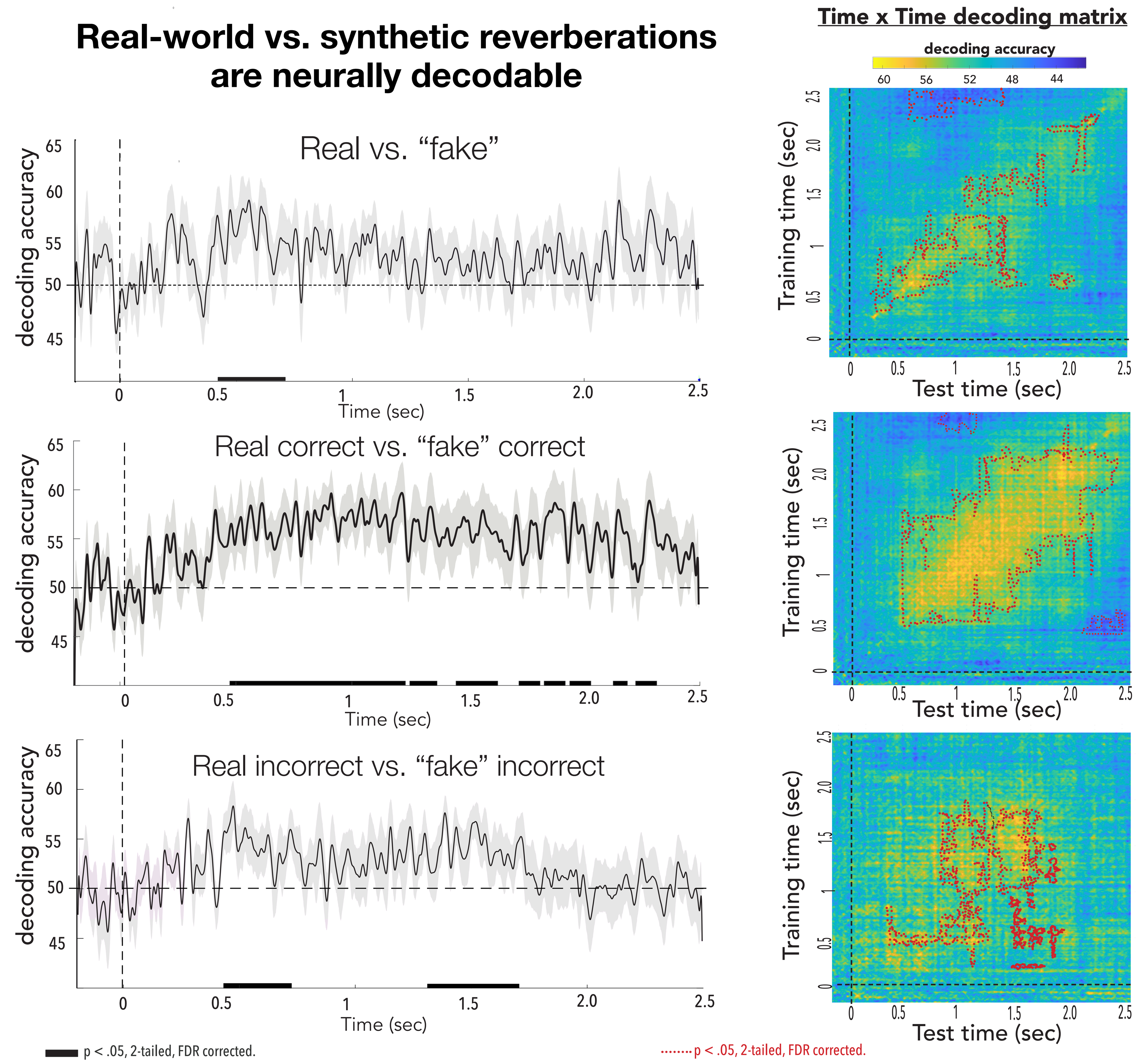


** t= 4.12, df=26, p < 0.001

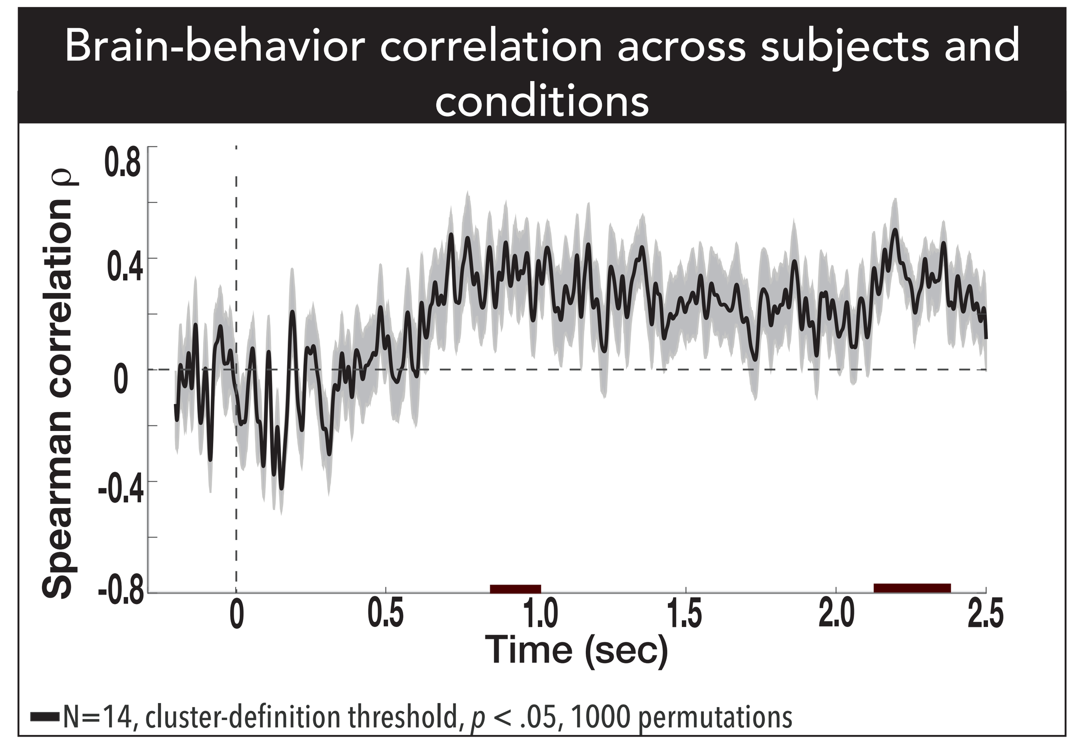
* Wilcoxon, signed-rank test median=50, p < .01

- Overall task performance was **above chance**; it was more accurate to real than fake conditions. Across fake conditions, performance varied widely, resembling priors results[1]
- Response time did not differ across conditions.

Real-world vs. synthetic reverberations are neurally decodable

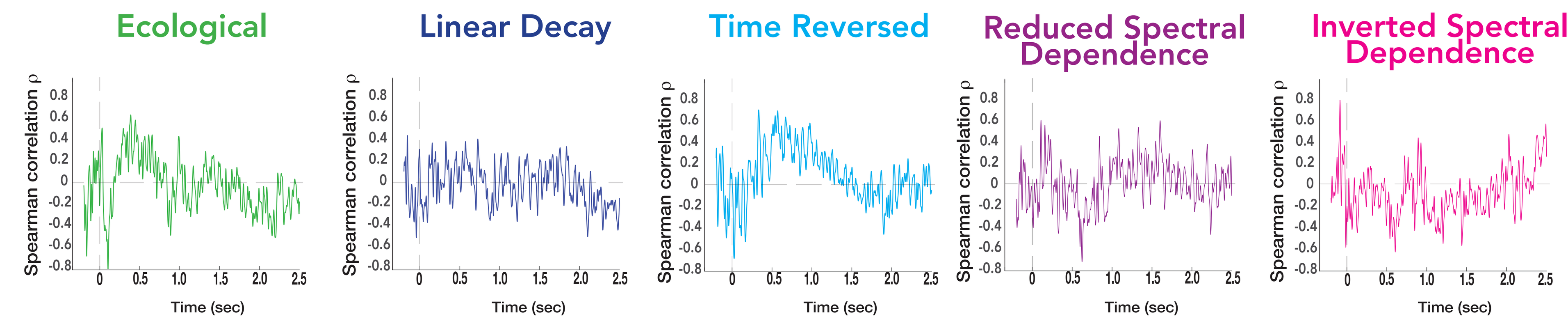


Neural decoding accuracy correlates with task performance



Decoding accuracy and behavioral-performance correlated significantly during two windows time from 800 ms to 1000 ms, and 2100 ms to 2400 ms after stimulus onset.

Decoding-performance correlations varied widely across conditions



4. CONCLUSIONS

- Participants **reliably distinguished real-world from synthetic reverberation** even with stimuli containing complex speech sound sources, with performance mirroring previous work[1].
- Even with a wide variety of IRs used, and speech source sounds unique for each trial, we were able to decode real vs. fake stimulus conditions, suggesting the **neural response pattern captured consistent statistics of auditory scenes** robustly across salient properties of both the varying source and the trial-unique IRs.
- **Stimulus conditions were better decoded when subselecting correct trials** (about 75% overall), suggesting that perceptual report, not just stimulus attributes, is related to classifier performance.
- The differing time courses of the brain-behavior correlations suggest heterogeneous processing of temporal and spectral stimulus cues.

5. REFERENCES

1. Traer, J., & McDermott, J. H. (2016). Statistics of natural reverberation enable perceptual separation of sound and space. Proceedings of the National Academy of Sciences, 113(48). http://mcdermottlab.mit.edu/Reverb/ReverbDemos.html
2. Teng S, Sommer V, Pantazis D, Oliva A. Hearing scenes: A neuromagnetic signature of perceived auditory source and reverberant space separation. eNeuro 4(1). doi: 10.1523/ENEURO.0007-17.2017
3. TIMIT Acoustic-Phonetic Continuous Speech Corpus https://catalog.ldc.upenn.edu/LDC93S1

6. ACKNOWLEDGMENTS

We thank Traer and McDermott (2016) for publicly available code (https://mcdermottlab.mit.edu/Reverb/ReverbSummary.html) and James Traer for generous assistance and commentary in designing and creating stimuli for the present study.

Funding sources: E. Matilda Ziegler Foundation for the Blind (to ST).