During natural vision, semantic novelty modulates fixation-related processing in primate cortex

Vinay S Raghavan[1], Jens Madsen[1], Maximilian Nentwich[2], Marcin Leszczynski[3,4,5], Arnaud Falchier[4,6], Stephan Bickel[2,4,7], Brian E Russ[4,6], **Lucas C Parra**[1]

1. Dept. of Biomedical Engineering, City College of New York, 2. Feinstein Institutes of Medical Research, 3. Dept. of Psychiatry, Columbia, 4. CBIN, Nathan Kline Institute, 5. Cognitive Science Center, Jagiellonian University, 6. Dept. of Psychiatry, NYU, 7. Depts. of Neurology, Neurosurgery, Hofstra/Northwell

Introduction

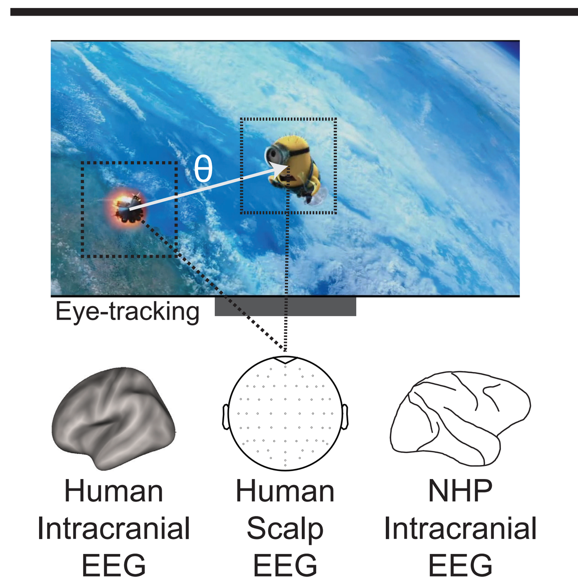
- Hierarchical processing in the primate visual system resembles hierarchical processing in CNNs[1]
- Past work considers visual processing a single fixation
- However, vision proceeds through a **series of fixations**
- Similarly, self-supervised learning uses a series of crops from an image to predict the upcoming embedding[2,3,4]

- Does the brain compare semantic information from foveal vision across saccades?

We use Temporal Response Functions (TRFs) to measure scalp and intracranial EEG responses to novelty features computed from different levels of the ResNet50 hierarchy time-locked to fixation during movie- and image-viewing.[5]

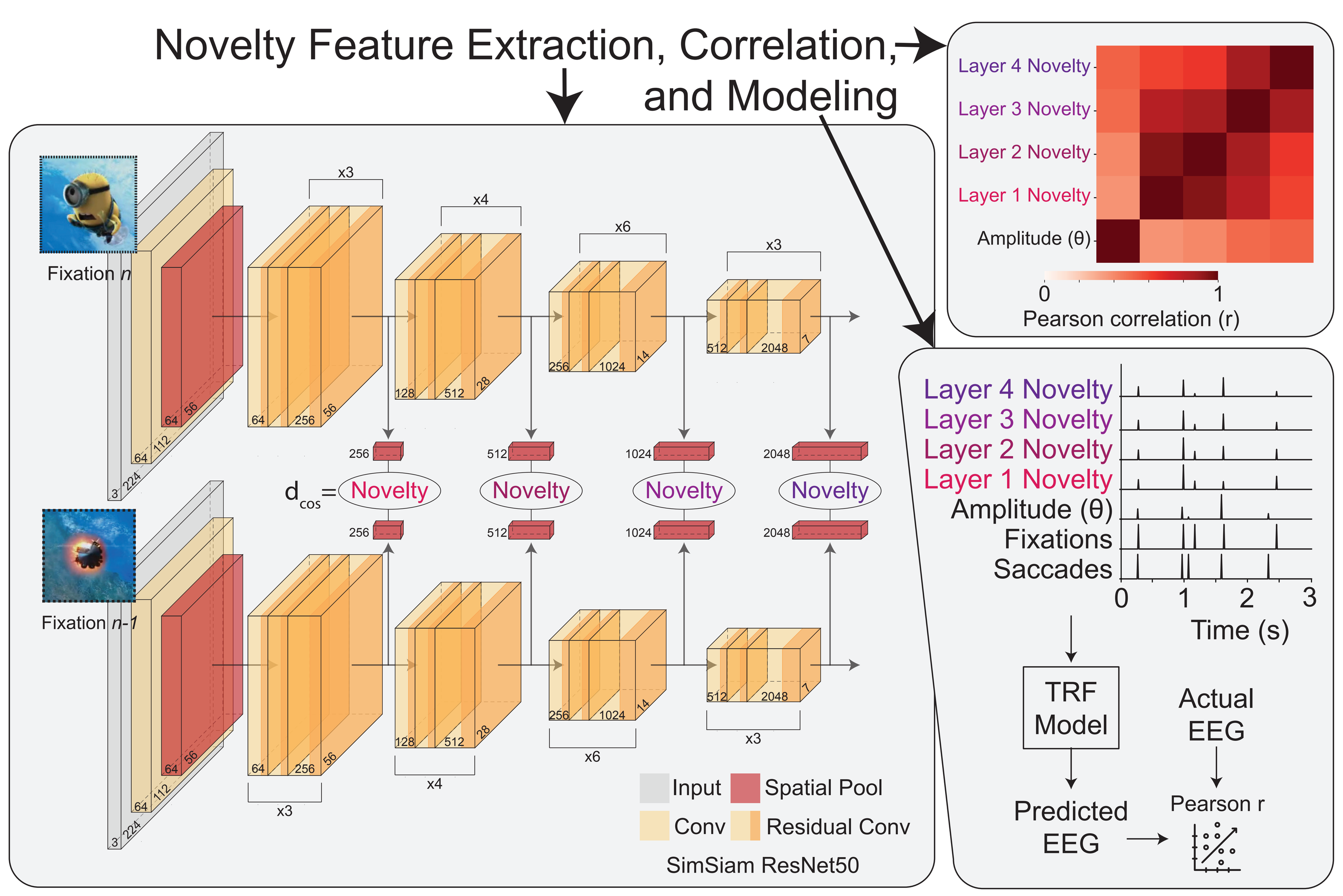
We find...
- Only semantic novelty modulates brain activity
- Scalp EEG modulations resemble MMN and N400
- iEEG modulations occur in ventro-medial visual stream
- Modulations greater during movie- vs image-viewing
- Shows semantic integration of foveal vision

Methods

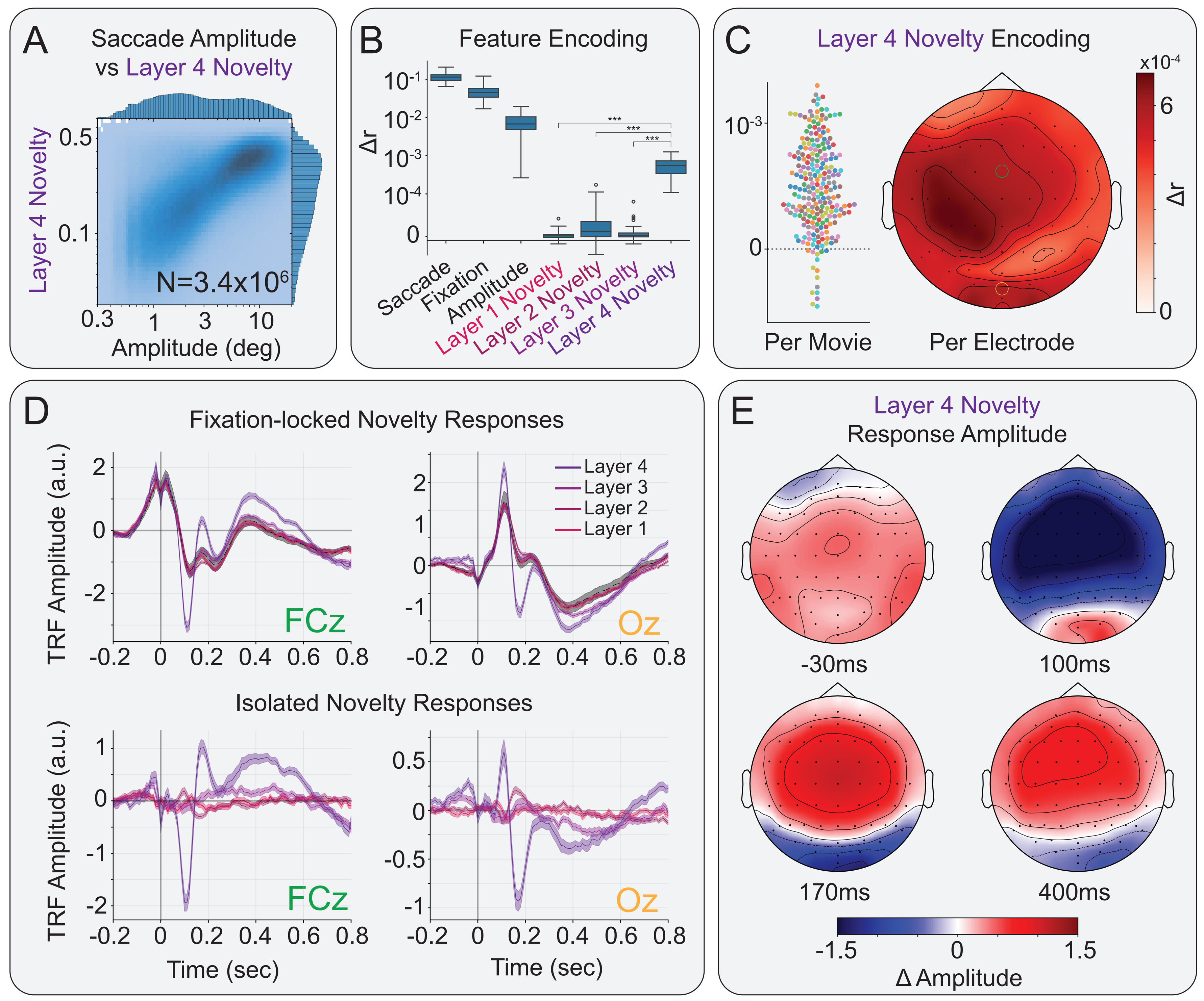


- Human EEG: 79 participants, 64 ch, 231 full-length movies
- Human iEEG: 23/31 participants, 6328 electrodes, 29-44 min each; 8522 electrodes, 6-28 min each
- NHP iEEG: 2 rhesus macaques, 27+23 electrodes, 133+73 min; 27+23 electrodes, 56+13 min

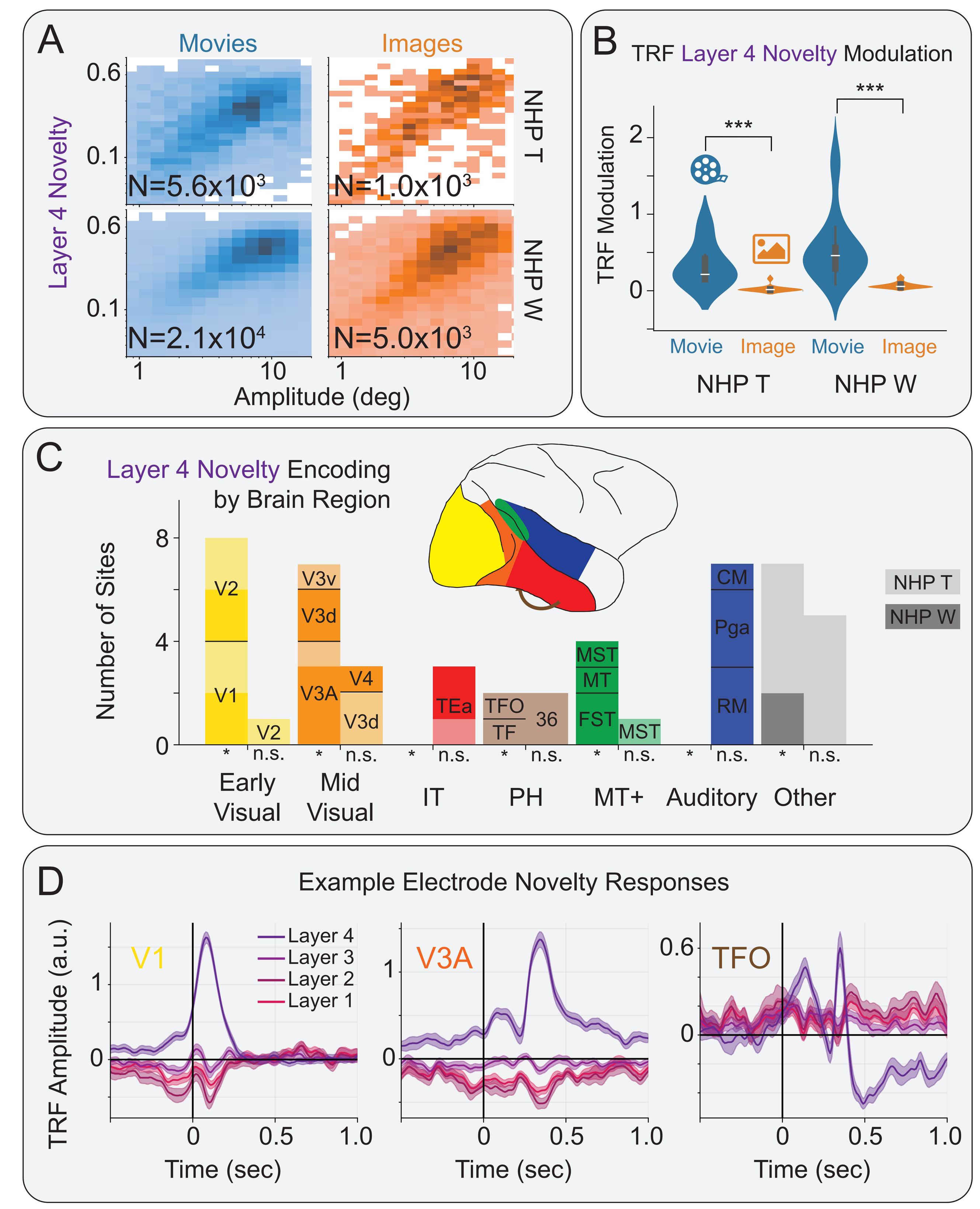
Novelty Feature Extraction, Correlation, and Modeling



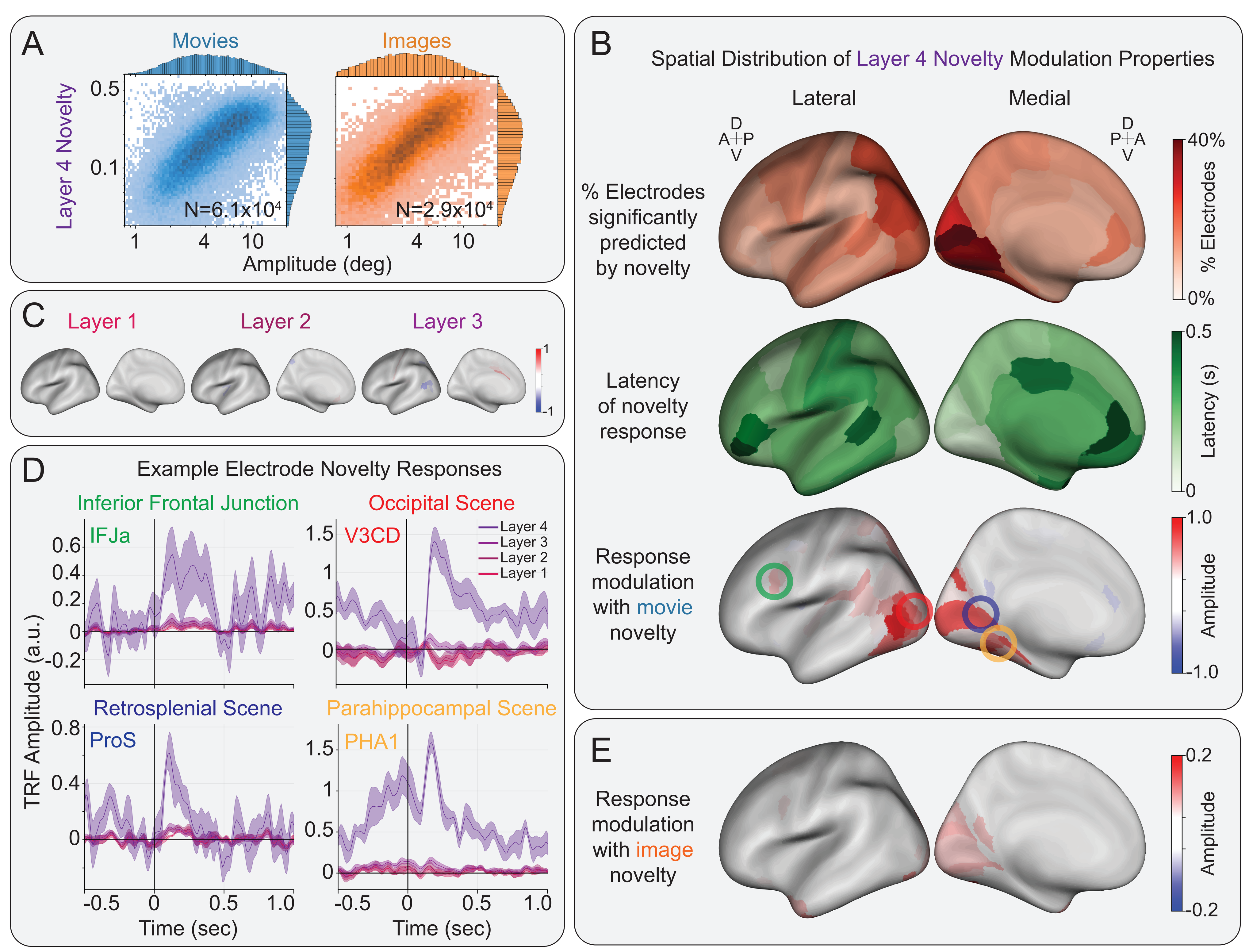
Human Scalp EEG Responses



NHP iEEG Responses



Human iEEG Responses



References

1. Yamins, D. L., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. Nature neuroscience
2. Chen, T., Kornvblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. In *International conference on machine learning*
3. Chen, X., & He, K. (2021). Exploring simple siamese representation learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*
4. Wen, Z., & Li, Y. (2021). Toward understanding the feature learning process of self-supervised contrastive learning. In *International Conference on Machine Learning*
5. Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. *Frontiers in human neuroscience*

Paper