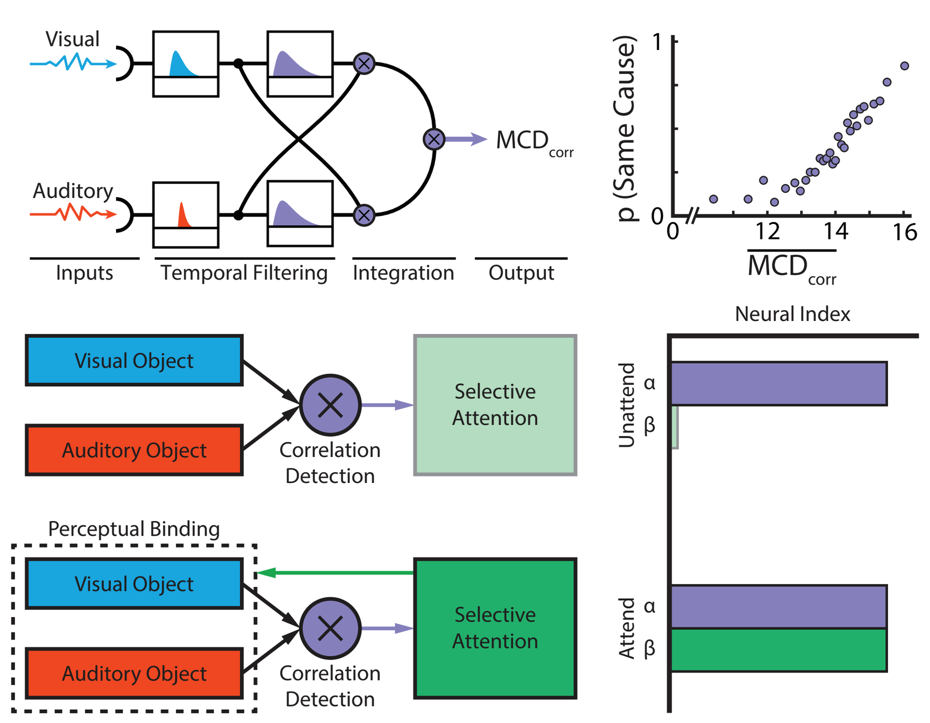
ALPHA AND BETA BAND BRAIN ACTIVITY TRACKS THE TEMPORAL CORRELATION BETWEEN AUDITORY AND VISUAL STIMULI

Aaron R. Nidiffer[1] and Edmund C. Lalor[1,2]

[1]Department of Biomedical Engineering and Department of Neuroscience, University of Rochester, Rochester, NY 14627, USA
[2]School of Engineering, Trinity Centre for Bioengineering and Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin 2, Ireland.

Introduction and conceptual framework

Stimulus features that correspond to a common event in the environment are correlated over time and space. For example, mouth movements of a speaker are correlated with the amplitude envelope of the accompanying speech. The brain uses these correlations to bind these features within and across sensory systems which, in the case of multisensory binding, results in enhancements in speech comprehension and selective attention in multi-speaker situations. Though this much is known, little is known about the underlying mechanism by which the brain generates its representation of that correlation and how it affects binding. Based on previous findings suggesting interactions between multisensory integration, attention, and binding[1-3], we propose a framework where temporal correlation is computed by a previous described multisensory correlation detector (MCD)[4] and serves as a scaffold supporting these processes. We propose that the attentional system represents temporal correlation in alpha activity[5] as a means to track potential correspondence for selection. When selected, representations are then bound via beta band activity[6].

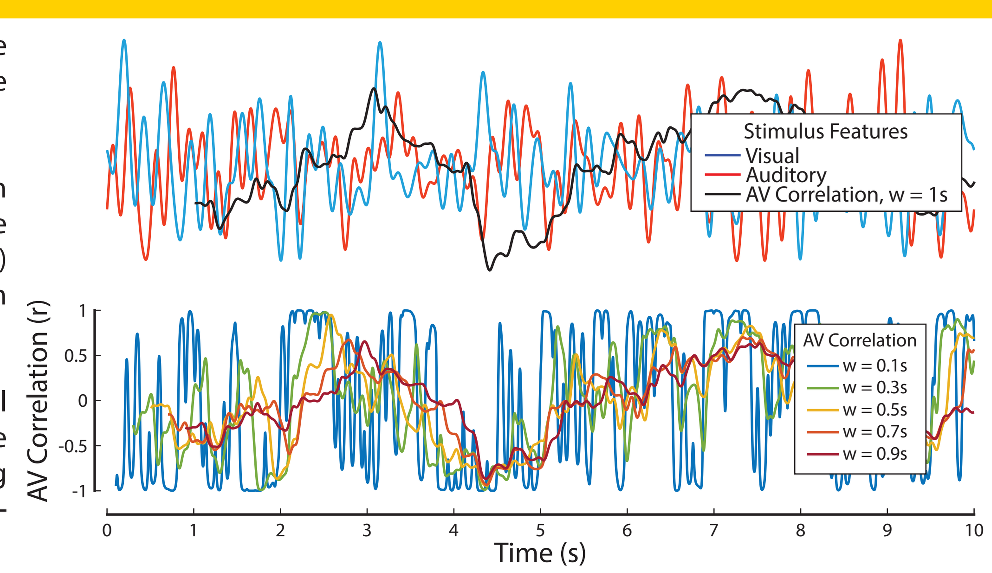


Stimulus, EEG Methods, and TRF Analysis

Continuous streams (30 s) of auditory (broadband white noise) and visual (Gaussian blob) stimuli were amplitude modulated by independent arrhythmic signals[7,8].

AM signals were constructed in the frequency domain from components in the 3-8 Hz range having a flat amplitude profile and random phases. Brief (500 ms) rhythmic (6 Hz) targets were inserted into one or both streams at random times during a second block.

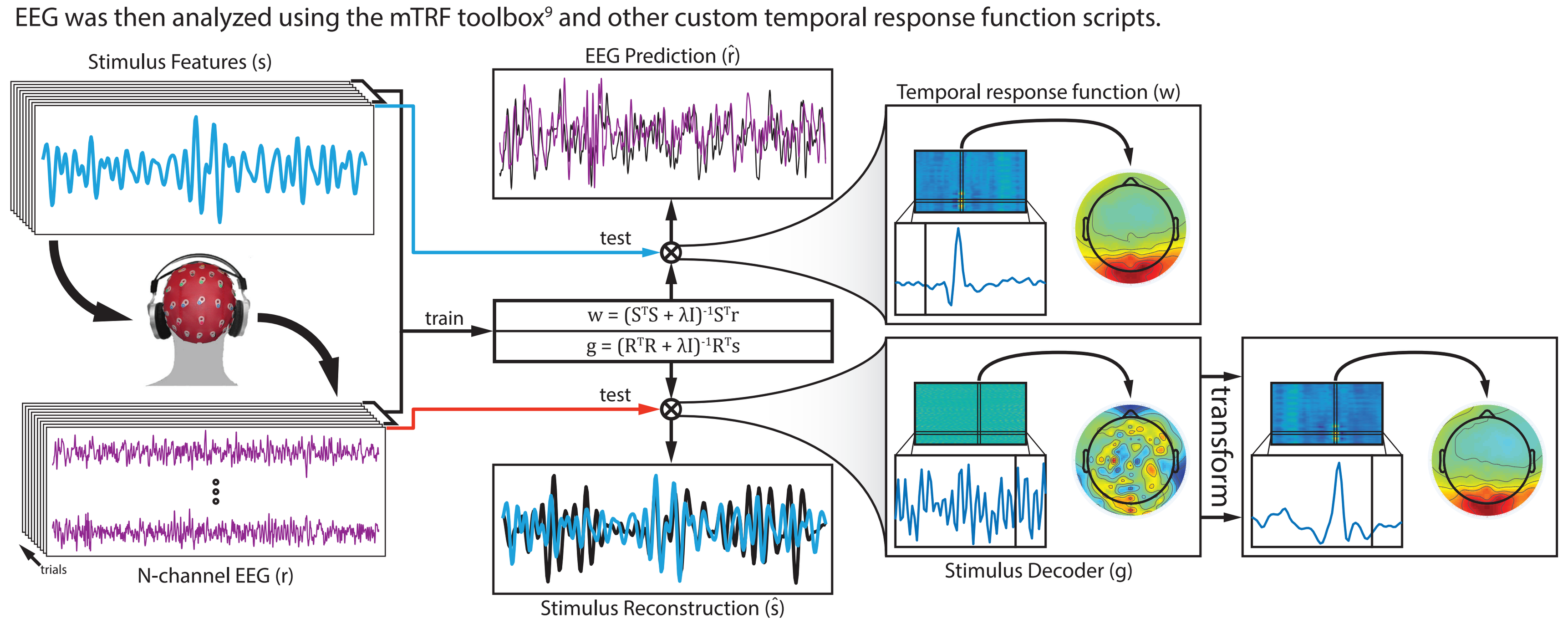
By the nature of their independence, auditory and visual AM signals had a long-term correlation of ~0 but in the short term, correlations computed over time (sliding boxcar function) revealed widely varying correlation envelopes that were unique across time windows.



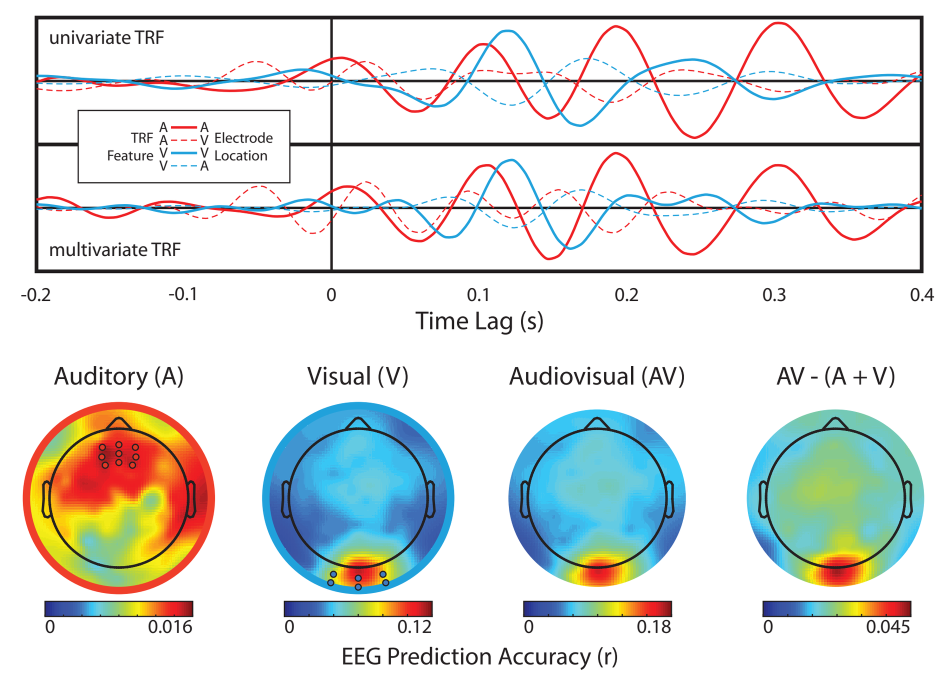
We recorded 128-channel EEG from fifteen participants as they viewed and heard 20 unique audiovisual streams - twice each in two blocks (40 trials per block). First, participants passively observed the stimuli. Second, to facilitate engagement with the correlation, participants were asked to respond to the occurrence of embedded targets in both streams. Participants performed the task successfully (d' = 2.95).

EEG data were band pass filtered (0.5-50 Hz), bad channels were rejected, eye blinks were removed via ICA, missing channels were interpolated. The signal was filtered again into 4Hz frequency bands centered at 2.5 - 46.5 Hz (if applicable), epoched into trials, averaged across repetitions, power computed (if applicable), and normalized using z-score.

EEG was then analyzed using the mTRF toolbox[9] and other custom temporal response function scripts.



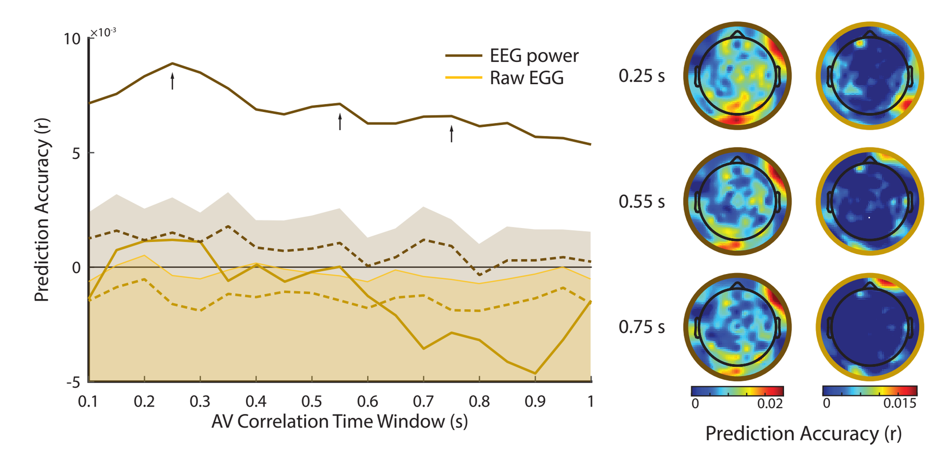
Results - Cortical responses track temporal correlation.



Univariate auditory (red) and visual (blue) TRFs were well fit and significantly predicted EEG responses above a noise floor derived from permuting shuffled trials.

Unique auditory and visual topographies revealed that each stimulus representation predicted EEG over areas of the scalp consistent with auditory and visual generators, respectively.

A multivariate unisensory TRF significantly improved predictions across all channels (but especially auditory and visual channels) compared to an additive model baseline using univariate TRFs (A + V)[10], suggesting some form of cross-modal interaction in sensory cortices.



Unlike unisensory TRFs, correlation TRFs did not significantly predict raw EEG responses across most time-windows (orange). However, TRFs trained to predict EEG power (brown) did so significantly better than shuffled trials (dashed lines).

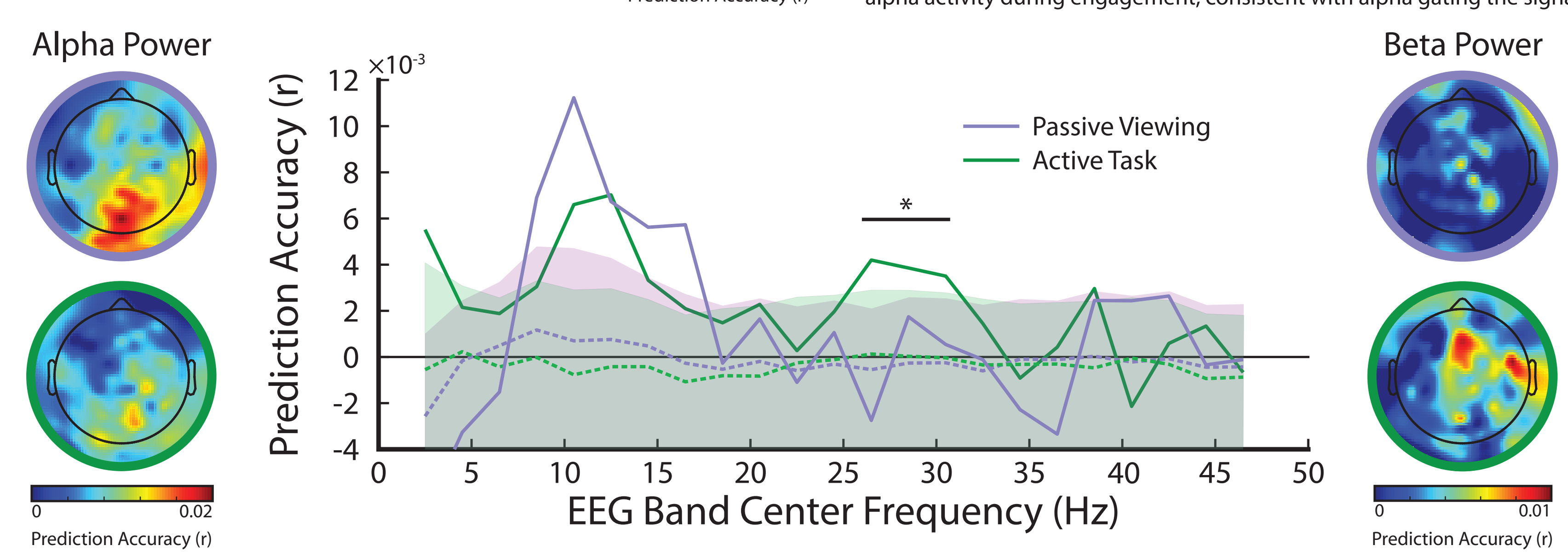
Predictions were best using a time-window of 250 ms. However, local peaks at 550 ms and 750 ms might warrant further investigation. It is likely that the brain tracks a weighted average of multiple time-scales.

Topographies show that the predictions were strong over a large swath of scalp electrodes, peaking at occipital locations.

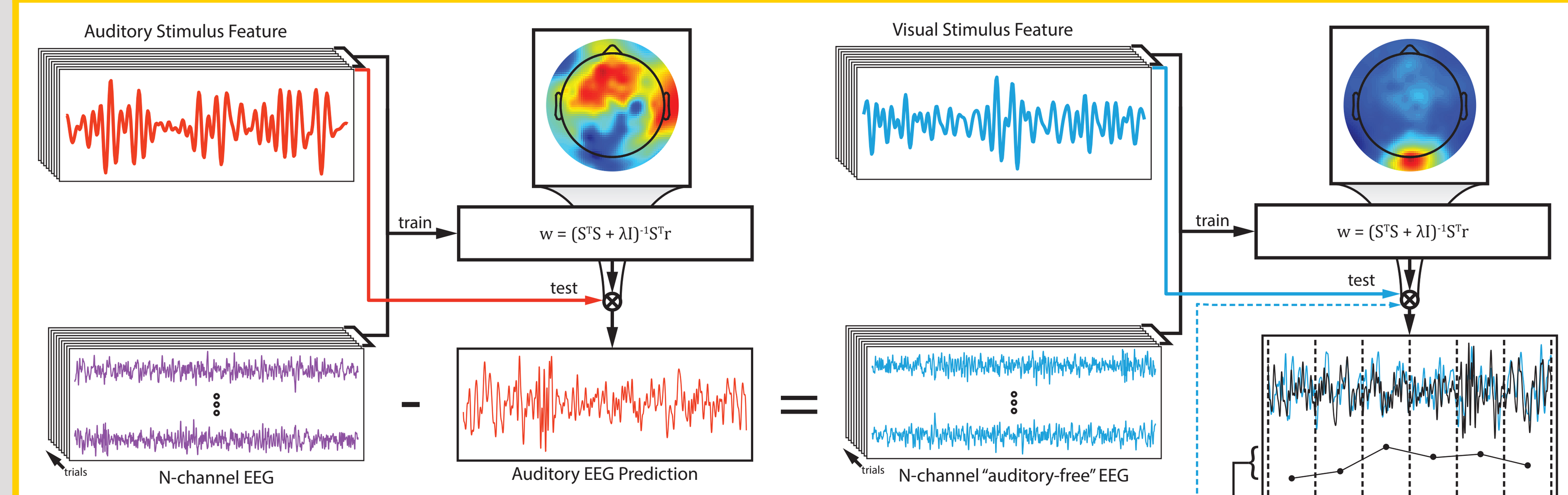
When decomposed into different frequency bands, only alpha and beta power significantly encode temporal correlation.

Alpha power encodes correlation independent of task demands whereas beta power only encodes correlation when participants are overtly engaged with the stimulus ($p = 0.042$).

Topographies reveal distinct scalp locations encoding correlation in alpha (parieto-occipital) and beta (frontal) and suggest a potential suppression of alpha activity during engagement, consistent with alpha gating the signal.



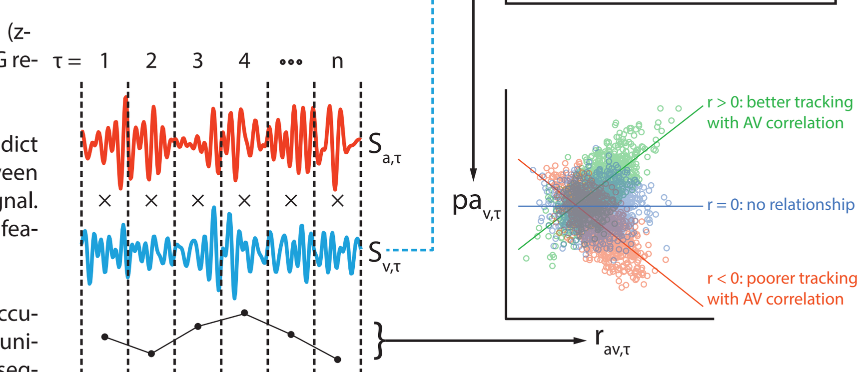
Results - Correlation enhances cortical sensory responses.

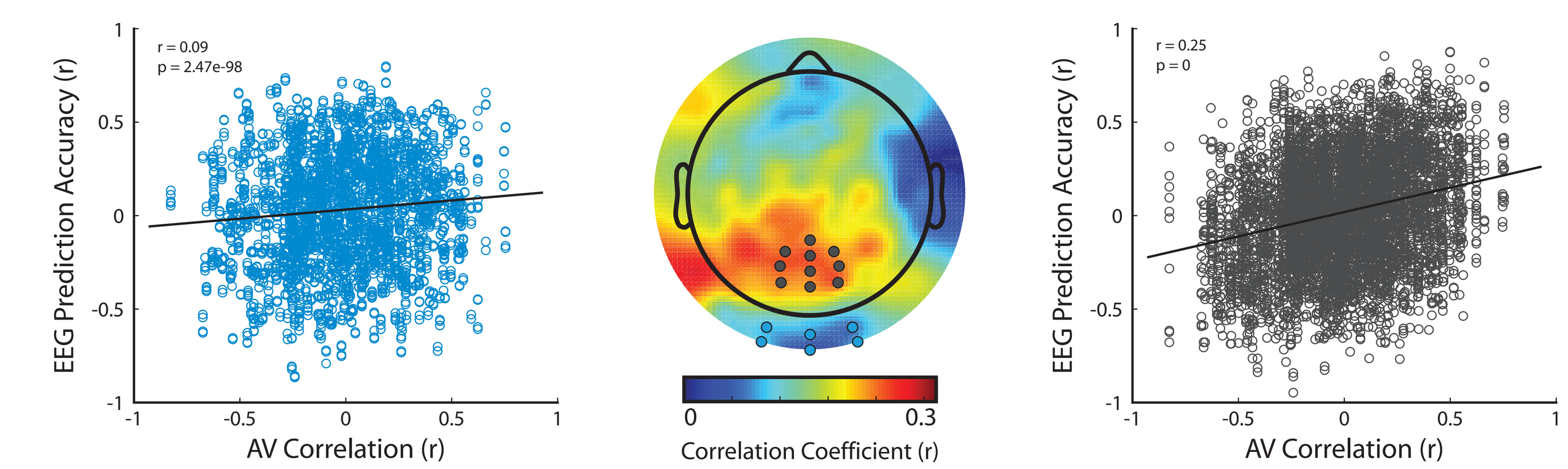


In order to isolate cortical responses to the visual stimulus we removed normalized (z-score) EEG predictions from the univariate auditory TRF from the original processed EEG responses. We trained a new univariate visual TRF using this auditory-free EEG

Using this TRF, one-second segments of the visual stimulus feature were used to predict corresponding one-second segments of EEG. We then measured the similarity between these one-second predictions to the corresponding second of the auditory-free EEG signal. We also assessed the temporal correlation between the auditory and visual stimulus features for each corresponding second.

A correlation between momentary temporal correlation and visual EEG prediction accuracy informs us as to whether the temporal correlation improves (r>0) or disturbs (r<0) unisensory encoding. Correlations were compared to noise floor (permuting shuffled segments and trials)[11].





Although the relationship between temporal correlation and visual encoding is small across participants, it is highly reliable and significant in pre-defined visual channels. Additionally, across those channels this effect is significanly larger than the noise floor in 7-9 of 15 participants, depending on channel ($ps = 1.8\times10^{-7}, 7.4\times10^{-9}, 2.3\times10^{-10}$, binomial test).

Inspection of the topography of these correlations revealed a peak anterior to visual channels over parietal areas. The effect in these channels was significant in more individual participants (12-14 of 15, depending on channel; $ps < 1.1\times10^{-15}$, binomial test). This peak potentially represents higher-order visual processing or and future work will be aimed at determining its functional significance.

Summary and Conclusions

- Audiovisual temporal correlation is tracked at short time scales by cortical responses.
- Correlation is tracked across different frequency bands depending on attentional engagement.
- Correlation improves the representation of sensory stimuli over unisensory cortex.

These results are consistent with a model where the attentional system continually monitors correlation through alpha activity and engages perceptual binding via beta band activity, highlighting the complex interplay between multisensory integration, attention, and binding.

References

1. Treissman and Gelade (1980). A feature-integration theory of attention. Cognitive Psychology, 12.
2. Talsma et al (2010). The multifaceted interplay between attention and multisensory integration. Trends in Cognitive Sciences, 14.
3. Bizley et al (2016). Defining auditory-visual objects: behavioral tests and physiological mechanisms. Trends in Neuroscience, 39(2).
4. Parise and Ernst (2016) Correlation detection as a general mechanism for multisensory integration. Nature Communications, 7.
5. Klimesch (2012). Alpha-band oscillations, attention, and controlled access to stored information. Trends in Cognitive Sciences 16(12).
6. Hipp et al (2011). Oscillatory Synchronization in large-scale cortical networks predicts perception. Neuron, 69.
7. Maddox et al (2015). Auditory selective attention is enhanced by a task-irrelevant temporally coherent visual stimulus in human listeners. eLife.
8. Attilgan et al (2018). Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. Neuron, 97.
9. Crosse et al (2016). The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. Frontiers in Human Neuroscience, 10.
10. Crosse et al (2015). Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. Journal of Neuroscience, 35(42).
11. Broderick et al (2019). Semantic Context Enhances the Early Auditory Encoding of Natural Speech. Journal of Neuroscience.

Supported by NIH grant DC016297

National Institute on Deafness and Other Communication Disorders (NIDCD)

aaron.nidiffer@rochester.edu

arnndffr.us

lalorlab.net