Imaging of brain electric field networks

Lawrence Frank

lfrank@ucsd.edu

UC San Diego https://orcid.org/0000-0001-7235-587X

Vitaly Galinsky

UC San Diego

Olave Krigolson

University of Victoria

Susan F. Tapert

UC San Diego

Stephan Bickel

Feinstein Institutes for Medical Research

Antigona Martinez

The Nathan Kline Institute for Psychiatric Research

Physical Sciences - Article

Keywords: EEG, fMRI, brain connectivity, brain waves, neuroimaging

Posted Date: April 12th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-2432269/v2

Additional Declarations: The authors declare no competing interests.

Imaging of brain electric field networks

Lawrence R. Frank,[1,2*] Vitaly L. Galinsky,[1] Olave Krigolson[3],
Susan F. Tapert,[4] Stephan Bickel[5,6], Antigona Martinez[5]

[1]Center for Scientific Computation in Imaging, UC San Diego, La Jolla, CA, USA
[2]Center for Functional MRI, UC San Diego, La Jolla, CA, USA
[3]Centre for Biomedical Research, University of Victoria, Victoria, BC, Canada
[4]Dept of Psychiatry, UC San Diego, La Jolla, CA, USA
[5]Nathan Kline Institute, Orangeburg, NY, USA
[6] The Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY, USA
*To whom correspondence should be addressed; E-mail: lfrank@ucsd.edu

We present a method for direct imaging of the electric field networks in the human brain from electroencephalography (EEG) data with much higher temporal and spatial resolution than functional MRI (fMRI), without the concomitant distortions. The method is validated using simultaneous EEG/fMRI data in healthy subjects, intracranial EEG data in epilepsy patients, and in a direct comparison with standard EEG analysis in a well-established attention paradigm. The method is then demonstrated on a very large cohort of subjects performing a standard gambling task designed to activate the brain's 'reward circuit'. The technique uses the output from standard EEG systems and thus has potential for immediate benefit to a broad range of important basic scientific and clinical questions concerning brain electrical activity, but also provides an inexpensive and portable alternative to function MRI (fMRI).

1 Main

The human brain communicates internally through exceedingly complex spatial and temporal patterns of electrical signals. Although these signals can be measured using electrodes placed on the surface of the scalp (electroencephalography, or EEG), the ability to reconstruct the spatial and temporal patterns within the brain has been thwarted by the complexity of the inverse problem: What time- (or frequency-) dependent volumetric electrical signals throughout the brain are consistent with the signal measured on the two-dimensional surface of the scalp (*1, 2*)? There is a long-standing belief that it is not possible to detect and reconstruct electrical activity in sub-cortical regions deep within the brain from EEG due to inherent limitations of "volume conduction" (*3*). However, this is not actually a physical limitation, but rather a consequence of the incomplete nature of the standard model used to characterize the EEG signal. Despite the obviously highly dynamical nature of the electrical activity that occurs within the very inhomogeneous and anisotropic composition of brain tissue, current EEG data analysis methods are still based on the assumption that the average tissue bioelectric properties (e.g., the average permittivity $\bar{\epsilon}_0$ and conductivity $\bar{\sigma}$) are sufficient to describe the electric fields E in the brain. This leads to the approximation $|\bar{\epsilon}_0 \partial E/\partial t| \ll |\bar{\sigma} E|$ (*4*) which in turn leads to the assumption that the time depedence $\partial E/\partial t$ can be ignored in the "typical" frequency range of brain signals (*5*). This is the ubiquitous so-called "quasi-static" approximation (*6*).

In reality, it is precisely the anisotropic and inhomogeneous nature of brain tissue that must be taken into account in order to develop an accurate physical model of brain electromagnetic behavior, as we have described in our recently developed universal theory of brain waves called *weakly evanescent transverse cortical waves* (WETCOW) (*7–9*). The surprising consequence of this theory is the existence of electric field waves generated as a consequence of the complex tissue boundaries (e.g., surface waves) that permeate throughout the brain and are in precisely

the frequency range of observed brain electrical activity. This theory explains the broad range of observed but seemingly disparate brain spatiotemporal electrical phenomena from extracellular spiking to cortical wave loops, all of which are predicated on the time dependence of the electric fields within the complex architecture of anisotropic and inhomeneous tissue within the brain. This theory is necessary to provide a solution to the EEG inverse problem which, as shown below, produces a reconstruction of brain electrical activity with high temporal resolution and spatial resolution that is comparable (or even exceeding that of) functional MRI.

2 Methods

The essence of our method (detailed in the Supplementary Material) is to reconstruct the electric field potential throughout the entire brain using our recently developed WETCOW wave model (*7–9*) constrained by MRI-defined tissue properties. This method is called *SPatially resolved EEG Constrained with Tissue properties by Regularized Entropy (SPECTRE)*. This approach solves a dynamic time-dependent construction of Maxwell's wave equations of electromagnetism in an inhomogeneous and anisotropic medium and thus is distinctly different than standard so-called 'source localization' methods, as discussed in Section 6.3.

3 Validation

Validation of any neuroimaging methods is problematic because it is not possible to directly measure brain activity at every location in the brain. Nevertheless, three methods are obvious candidates for assessment of SPECTRE's validity.

The first is comparison with functional MRI (fMRI), the current method of choice for whole brain spatial localization of brain activity. However, association of fMRI with a "standard" for EEG is problematic because it is not measuring electrical activity, but the magnetization changes in hemoglobin as blood becomes deoxygenated during brain activity. The timescale and location

of these changes can be vastly different than those produced by EEG signals. Nevertheless, its capability of spatially localizing activated brain regions merits a comparison. The most direct comparison is between fMRI and EEG data collected simultaneously, which guarantees that the brain activity measured is identical in both experiments. Such "simultaneous fMRI/EEG" experiments are not particularly common as collecting EEG data within an MRI scanner during imaging is notoriously difficult, and the MR imaging procedure significantly distorts the EEG signal. However, a recent open-source study provides such data which is sufficient for our purposes.

A more direct method for validating the ability of SPECTRE to reconstruct localized electrical activity can be constructed from intra-cranial EEG (iEEG) recordings collected during epilepsy studies. Such measurements consist of specially designed EEG sensors distributed linearly along a probe that is inserted deep within a brain that has been exposed by surgical removal of a portion of the skull. By selecting only these electrodes near the brain surface from the full array of electrodes, we can synthesize an artificial surface distribution of electrodes to mimic a standard non-invasive EEG experiment (albeit with a limited coverage of the brain). We have access to such data through an ongoing study which enabled this method of validation as well.

Lastly, a comparison with current "source localization" methods would seem to be in order (*10*). This comparison turns out to be the most problematic as these methods all employ a very different, and quite limited, physical model for the EEG signal, and suffer from computational limitations as well. Despite attempts to make a reasonably valid comparison, it was determined that this was not possible, as described below.

3.1 Validation with simultaneous fMRI/EEG visual task

It is notoriously difficult to get high quality EEG data in simultaneous fMRI/EEG studies as the presence of the rapidly varying magnetic fields present in an fMRI acquisition distort the EEG signal. However, one recent open-source simultaneous fMRI/EEG study of a well controlled visual task (the periodic flashing checkerboard) on multiple subjects ((*11*), available from the Nathan Kline Institute) provides important data to address this question.

The ability of SPECTRE to faithfully reconstruct the spatial distribution similar to fMRI is shown in Fig. 1. Importantly, this comparison was performed on data from a single subject, since brain activity patterns can vary significantly between individual and averaging over multiple subjects obscures specific spatial variations important for validation. In the top rows of Fig. 1 is shown the fMRI EFD mode that automatically detects the activation in the primary visual cortex. In the middle row are shown the SPECTRE modes reconstructed using the 2mm MNI anatomical atlas, chosen because it was closest in resolution ($2mm^3$) to the fMRI data ($\sim 3mm^3$). The very close correspondence between the spatial patterns is evident.

The bottom rows in Fig. 1 clearly demonstrate one of the most compelling, and perhaps surprising, aspects of SPECTRE - its ability to reconstruct activation at spatial resolution *significantly higher* than fMRI. This is a consequence of the SPECTRE reconstruction being based on the solution of the propagation of electromagnetic wave through specific tissue morphologies and bioelectric properties, provided by arbitrary resolution anatomical MRI data. The finer the resolution of the MRI scans, the more details can be available for the reconstruction. This is of course dependent upon the number and distribution of the EEG sensors, but certainly holds for the standard array configurations used in this paper.

Although it is an almost universally believed notion that EEG and fMRI are complementary because EEG has excellent temporal resolution but poor spatial resolution, while fMRI has poor temporal resolution but good spatial resolution, in fact SPECTRE EEG reconstructions

can achieve much higher *intrinsic* temporal *and* spatial resolution. Moreover, because there are no spatial distortions in SPECTRE, this mitigates one of the aspects of fMRI that most confounds spatial localization through signal loss and non-linear geometric distortions. This is shown in Fig.2.

It should be noted that the 'simple' periodic flickering checkerboard stimulus not only activates the primary visual cortex but activates other visual and supplementary fields as well, as is evident from the activity patterns in Fig. 1. A simple stimulus does not imply a simple activation pattern. This notion was a primary motivation for our development of the EFD method for fMRI (*12*). The activation mode reconstructions for both the fMRI and SPECTRE data are based on the EFD which detects complex non-linear interacting spatial-temporal modes of activity (*12*). Thus although the task is a 'simple' visual stimulation, our analysis is not expected to simply detect activity in only the visual cortex, as would be produced by a more standard regression approach (*11*), but in a more complex set of brain networks. Indeed, multiple EFD modes are produced, though we have only shown the one incorporating the primary visual cortex. As we have argued previously (*12*), EFD analysis is more sensitive than simple regression techniques to the complex brain activation patterns predicted by neuroscience, and less sensitive to erroneous identification of noise or non-independent modes than the independent component analysis (ICA) (*12*). Indeed, one of our observations from both the fMRI and EEG data used in this study (*11*) is the appearance of PFC activations associated with visual stimulation, which has been suggestive of conscious visual perception (*13, 14*). Addressing this question is beyond the scope of the current paper.

3.2 Validation with simultaneous fMRI/EEG attention paradigm

Simultaneous EEG/fMRI were collected from subjects within a standard clinical 3T MRI scanner (see Section 6.4.1 for details). The stimuli and paradigm are described in detail in (*15*).

Briefly, bimodal stimuli consisting of short ($\sim 1s$) streams of simple tones (600 and 1000 Hz) alternating at 10 Hz were delivered concurrently with phase-reversing (6Hz) checkerboard patterns presented at fixation. Participants were instructed to selectively attend to either the visual or auditory aspect of the bimodal stimulus and respond when the stream of stimuli in the attended modality ends.

SPECTRE processing was performed in the alpha band. The appearance of visual stimuli elicited a reduction of ongoing alpha (7-14Hz) activity ("event-related desynchronization", ERD) over occipital cortex, believed to occur when cortical regions are brought "on-line" for information processing (*16*). As in previous studies, e.g. (*17*), attended visual stimuli elicited increased (more negative) amplitude of the alpha ERD compared to unattended stimuli (Fig. 3A,B). In contrast, unattended, compared to attended, visual stimuli elicited a greater reduction in ongoing spectral activity within the 5-15Hz frequency range over bilateral middle frontal cortex (Fig. 3C,D). We estimated the neural sources of these attention-related modulations of oscillatory activity across the 8-12Hz frequency band which encompassed both the occipital and frontal activities (Fig.3E). Their anatomical localization was remarkably consistent across several individuals (Fig.4).

A direct comparison of the activation maps derived from both fMRI and EEG using SPECTRE for single study within two subjects (i.e. without any average over studies or subjects) is shown in Fig.5. The comparison is made by choosing specific regions of interest defined in the MNI atlas (occipital cortex and cerebellum) and correlating the activation maps derived from EFD for fMRI and SPECTRE from EEG. Comparison of the similarity of activated regions in individual subjects is generally a non-trivial problem. This is particularly true in the current case where the spatial distortions in fMRI (and lack of them in SPECTRE) make measures such as mean-squared error difficult to interpret. Therefore computation of the correlation coefficient over a predefined atlas ROI is a reasonable conservative meausure. The high correlation coef-

ficients between the maps Fig. 5 are therefore indicative of the consistency between the fMRI and SPECTRE results in the ROI. Note that this does *not* imply similarity over the entire region shown. Indeed, the SPECTRE results show enhanced sensivity to activation in regions not seen in the fMRI.

3.3 Validation through Intra-Cranial EEG recordings

While comparison with fMRI can validate the correct detection of activated brain regions and networks, as shown in the previous section, it cannot inform the question of correct detection of electrical signals, since fMRI is based on a completely different contrast mechanism related to blood oxygenation. A direct validation of SPECTRE's ability to faithfully reconstruct deep electromagnetic activity is, to our knowledge, only achievable with one type of data: intra-cranial EEG (iEEG) recordings such as those used in medically refractory epilepsy patients for seizure onset localization where the electrodes are known to be adjacent to the site of electrical activity (*18, 19*). We analyzed an iEEG recording of a seizure localized in the left mesial temporal region acquired at Northwell Health, NY. All implanted electrodes are shown in Fig.6 (Top row) with each yellow dot depicting one recording contact. Comparing the SPECTRE reconstruction using all of the sensor data with one using only a subset of the data comprised of only the sensors on the surface of the brain (red dots in Fig.6 (Top row)) allows the quantitative assessment of how closely the results from a set of surface electrodes correspond to those produced by intra-cranial measurements recording signal very close to the sources. The results are shown for the alpha frequency band in Fig. 6 and reveal a very close correspondance between the SPECTRE mode reconstruction.

4 Investigation of the 'reward circuit'

Having validated the SPECTRE method directly with simultaneous fMRI/EEG, iEEG, and an attention paradigm, we investigated the ability of SPECTRE to faithfully reconstruct the well known neural 'reward circuit' that is one of the most important in understanding human cognition, emotion, and behavior (*20–22*) and is of great clinical significance in the understanding of addiction (*23, 24*), mood disorders (*25, 26*), and a variety of other conditions (*27*).

We demonstrate that SPECTRE using standard EEG data can accurately map human reward pathways akin to results previously only seen via fMRI. Indeed, fMRI results have highlighted a reward system within the brain that includes midbrain dopamine producing regions (the substansia nigra pars compacta, the ventral tegmental area), the ventral striatum, and multiple regions within the human prefrontal cortex (*28*). Other research using fMRI and source localization of EEG data suggests that the anterior cingulate cortex also plays a key role in reward processing (*29*). In a unifying theory, it has been proposed that all the aforementioned regions work together as a neural system for the optimization of reward driven behavioral change (i.e., reinforcement leaning; (*30*).)

This is of particular clinical significance because addictive behaviors have long been known to be subserved by specific brain regions operating in concert as the reward circuit (*31–35*). The reward circuit is involved in processing rewarding stimuli of any sort and in drug addiction, substances of abuse (e.g., amphetamine) increase dopamine release in a protracted and less regulated manner as compared to typical stimuli, resulting in synaptic plasticity, and altered functioning of this circuit over time.

For our analysis we used large gambling task dataset that includes 500 participants available for download from www.osf.io/65x4v/. The details of the dataset and an extensive analysis using standard EEG analysis methodologies are presented in (*36*). The relevant information

from this study is presented in Supplementary Materials Section 6.4.2.

For each subject trial $n = 10$ power modes were calculated and summed to form the single space-time SPECTRE mode $\tilde{\psi}_{10}$ (see Section 6.1.2). Fig. 7 shows three orthogonal slices of the difference in EFD power summed over all modes between conditions, averaged over all subjects. Activation in key regions of the reward circuit, including the frontal lobes, anterior cingulate gyrus, accumbens, and amygdala are clearly evident. Strong negative activation (i.e., deactivation) is evident in several structures, including the supplementary motor cortex, and the parietal operculum cortex. Activation is also apparent in the lingual gyrus and around the calcarine fissure and, as expected, in bilateral subcortical structures.

In Fig.8 is shown the power per brain regions as defined by the Harvard-Oxford 2mm cortical (top) and subcortical (bottom) atlases. In the cortical regions (top), strong activation is apparent in the frontal cortex (medial, orbital, operculum), cingulate gyrus, paracingulate gyrus, and unsular cortex. Activation in the accumbens is apparent from the data in the sub-cortical atlas Fig. 8 (bottom). These activated regions are consistent with the known elements of the human brain reward circuitry.

Images of statistical significance ($p < .0001$) are shown in Fig. 9. It should be noted that the determination of statistical significance with SPECTRE by 'traditional' methods is potentially misleading as they will tend to *underestimate* activation significance. The estimation of the modes in SPECTRE employs EFD (*12, 37*) which is a probabilistic formulation that *by construction* incorporates space-time neighborhood connectivity so that spatially and temporally coherent patterns ("clusters") are more probable. Traditional methods have the option for "clustering" regions of activation *post-hoc* into their general class of techniques called "bootstrapping" or "permutation inference". Cluster post-detection of an activation is incommensurate with our view of the estimation process wherein the clustering in space-time is a key component indicator of high-probability regions of space-time. Spatially and temporally coher-

ent patterns maybe of low amplitude with apparent low significance by traditional means, but those intensities are within a mode that contains very high significance in cortical regions (e.g., Fig.9) which is predicted by the WETCOW model. Thus the entire SPECTRE mode, including the somewhat diffuse lower intensity regions, is significant.

5 Discussion and Conclusion

The ultimate goal of functional neuroimaging is to non-invasively detect and quantify the spatial and temporal variations in brain activity. Functional MRI (fMRI) and EEG have emerged as the most ubiquitous methods because they offer two important and complementary outcomes. fMRI can non-invasively detect complex spatial and relatively low frequency temporal patterns of activity related to local BOLD changes while EEG can directly detect electrical activity but without the capability of accurate spatial localization.

Ideally, one would measure the electrical activity of the brain at high temporal resolution, as done by EEG, but with the spatial localization capabilities of fMRI. However, there has been a long-standing belief that this is not possible, due to the 'volume-conductance' problem (*3*). In this paper we have shown that this is not an actual physical limitation, as it is always presented, but an artificial constraint stemming from model for how electromagnetic wave propagate within the brain that is overly simplified by neglecting local tissue anisotropy and inhomogeneity. Our recently developed WETCOW model incorporates these tissue characteristics into the electromagnetic field equations (i.e., Maxwell's equations), which predicts the existence of previously undiscovered waves generated precisely as a consequence of the tissue anisotropy and inhomogeity. Unlike standard electromagnetic waves characterized by a frequencies ω that are proportional to their spatial frequency (or wavenumber) k, (i.e., $\omega \sim k$) the WETCOW waves have an inverse relationship $\omega \sim 1/k$. This results in waves that can permeate throughout the brain, not necessarily along neuronal pathways, that are in precisely

the spatial and temporal (frequency) range of observed brain electrical activity. This model provides the physical model upon which the SPECTRE is based, using tissue models derived from standard high resolution anatomical MRI data to reconstruct the modes of brain electrical activity throughout the entire brain volume.

The SPECTRE reconstruction of EEG data provides obvious significant advantages over fMRI in temporal resolution, since EEG data has very high intrinsic temporal resolution ($\sim 1ms$) necessary to capture rapidly varying electric field variations. Moreover, the SPECTRE algorithm can specify what frequency ranges to interrogate, providing a highly flexible analysis framework for focussed investigation of particular frequency bands of interest. On the contrary, even rapid fMRI acquisition is intrinsically limited by the temporal evolution of the contrast mechanism, the BOLD signal, which is related to blood flow and thus of quite low frequency ($\sim 1Hz$).

While the advantages of SPECTRE over fMRI in temporal resolution are clear, what is perhaps surprising is its advantages in *spatial* resolution. The inverse solution that estimates the electric field potential from the EEG data is based on a physical model of wave propagation from tissues whose composition and geometry are derived from high resolution anatomical MRI data. The final resolution of the SPECTRE electric field modes is that of the anatomical data which is typically significantly higher ($\sim .5 - 1mm$)) than the resolution of an fMRI image ($\sim 2mm$). (There are, of course, limitations depending on the number of electrodes in the EEG system.)

But it is also important to recognize that the question of resolution in fMRI is not just a question of the prescribed image resolution of the acquisition. The BOLD physical mechanism that generates the fMRI contrast is a subtle variation in the magnetic susceptibility which causes variations in the local magnetic field, that in turn alters the local signal. fMRI acquisitions are specifically designed to accentuate this effect in order to make is observable. Unfortunately,

local magnetic field variations unrelated to the BOLD mechanism, in particular strong magnetic susceptibility variations due to air/tissue boundaries such as those in the sinus cavities, cause severe non-linear image distortions that effectively alter the location and shape of the affected image volume elements (voxels). This makes even the definition of 'resolution' problematic, as it is essentially a spatially non-linearly varying function. Such effects are absent from EEG, which is simply a set of receiving electrodes (albeit not without its own source of artifacts) (*3*). The SPECTRE reconstruction uses high resolution MRI data acquired with techniques specifically designed to be insensitive to these magnetic susceptibility distortions and thus of very high spatial fidelity.

In this paper, we have successfully validated the SPECTRE method using simultaneous fMRI/EEG experiments. The results not only affirmed SPECTRE's capability to accurately reconstruct spatial distributions of neural activity from EEG data, in alignment with the concurrently acquired fMRI data, but also revealed its efficacy in identifying robust activations across subjects that were not detectable with fMRI alone. These findings underscore SPECTRE's potential to significantly enhance the sensitivity and scope of neuroimaging analyses. Further validation was performed using intra-cranial EEG measurements from an epilepsy study with reconstruction of data from a subset of sensors on the surface of the brain were shown to be consistent with the reconstruction from all the sensors, including those directly next to the activity source. The application of SPECTRE to high resolution EEG data during a gambling task demonstrated its ability to reconstruct a well-known and important brain circuit (*20–22, 24, 31, 32, 38*) that has previously only been detected using fMRI. The analysis revealed significant differences in the brain networks in the alpha range $8 - 12Hz$, consistent with previous spatially resolved fMRI experiments but the analysis is easily carried out in any user-defined frequency ranges of interest (*39–42*), which will be the subject of future work. The SPECTRE methodology is applicable to any EEG study and thus holds promise for a wide range of ongoing studies of

basic neuroscience of reward mechanisms and in clinical applications such as addiction.

The implications for spatially resolved EEG are important not only from a scientific perspective, but from a practical perspective as well. fMRI is a much more involved and expensive procedure, requiring highly trained research or clinical applications specialists in specially designed facilities, and subjecting the subjects to a much more claustrophobic and restricted environment, with the safety concerns always present in MRI imaging experiments. On the contrary, the portability, safety, and relative ease of EEG experiments, which can be carried out in a standard research or clinical office, makes it very attractive. The high spatial and temporal resolution capabilities provided by SPECTRE to standard EEG data offer the possibility of more detailed investigations of brain activity in a wide range of both basic research and clinical settings. This method also has important implications for the democratization of medicine worldwide where there are many populations for which advanced technologies such as fMRI are prohibitive because of cost, citing issues for large specialized equipment, and lack of highly trained personnel.

Human Subjects

All participants provided informed consent approved by the University of Victoria's Human Research Ethics Board. The iEEG data was recorded in drug-resistant epilepsy patients undergoing invasive EEG monitoring at the North Shore University Hospital (Manhasset, NY 11030, USA) for seizure onset localization. All patients provided informed written consent according to a protocol approved by the Institutional Review Board of the Feinstein Institutes for Medical Research in accordance with the Declaration of Helsinki.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code Availability

Software supporting this research is subject to UCSD Technology Transfer restrictions and available with appropriate agreement after reasonable request.

References

1. D. Marinazzo, *et al.*, *Brain Topography* **32**, 527 (2019).
2. C. M. Michel, D. Brunet, *Frontiers in Neurology* **10**, 1 (2019).
3. P. L. Nunez, *et al.*, *Electroencephalography and Clinical Neurophysiology* **103**, 499 (1997).
4. S. Taulu, E. Larson, *IEEE Trans Biomed Eng* **68**, 992 (2021).
5. M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, O. V. Lounasmaa, *Rev. Mod. Phys.* **65**, 413 (1993).
6. G. Gaugain, *et al.*, *J Neural Eng* **20** (2023).
7. V. L. Galinsky, L. R. Frank, *Phys. Rev. Res.* **2**, 023061 (2020).
8. V. L. Galinsky, L. R. Frank, *J. Cogn. Neurosci.* **32**, 2178 (2020).
9. V. L. Galinsky, L. R. Frank, *Phys. Rev. Lett.* **126**, 158102 (2021).
10. A. Biasiucci, B. Franceschiello, M. Murray, *Curr Biol* **19**, R80 (2019).
11. Q. K. Telesford, *et al.*, *Scientific Data* **10**, 554 (2023).
12. L. R. Frank, V. L. Galinsky, *Neural Comput.* **28**, 1769–1811 (2016).
13. C. Libedinsky, M. Livingstone, *Journal of Neuroscience* **31**, 64 (2011).
14. S. Paneri, G. G. Gregoriou, *Frontiers in Neuroscience* **11**, 1 (2017).
15. J. Grinband, J. Steffener, Q. R. Razlighi, Y. Stern, *Human Brain Mapping* **38**, 3538 (2017).
16. W. Klimesch, *Trends in Cognitive Sciences* **16**, 606 (2012).
17. J. Foxe, A. Snyder, *Frontiers in Psychology* **2** (2011).
18. G. Ramantani, L. Maillard, L. Koessler, *Seizure: European Journal of Epilepsy* **41**, 196 (2016).
19. P. van Mierlo, B. J. Vorderwülbecke, W. Staljanssens, M. Seeck, S. Vulliémoz, *Clinical Neurophysiology* **131**, 2600 (2020).
20. W. Schultz, *Physiological Reviews* **95**, 853 (2015).
21. S. Haber, *Decision Neuroscience*, J.-C. Dreher, L. Tremblay, eds. (Academic Press, San Diego, 2017), chap. 1, pp. 3–19.

22. M. T. Banich, S. Floresco, *Cognitive, Affective, & Behavioral Neuroscience* **19**, 409 (2019).

23. G. F. Koob, A. J. Roberts, *CNS Spectrums* **4**, 23–37 (1999).

24. E. L. Gardner, *Chronic Pain and Addiction* **30**, 22 (2011).

25. C. A. Naranjo, L. K. Tremblay, U. E. Busto, *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **25**, 781 (2001).

26. S. J. Russo, E. J. Nestler, *Nature Reviews Neuroscience* **14**, 609 (2013).

27. R. G. Lewis, E. Florio, D. Punzo, E. Borrelli, *Advances in Experimental Medicine and Biology* **1344**, 57 (2021).

28. S. McClure, M. York, P. Montague, *The Neuroscientist* **10**, 260 (2004).

29. C. B. Holroyd, M. G. Coles, *Cortex* **44**, 548 (2008). Special Issue on "Selection, preparation, and monitoring: Current approaches to studying the neural control of action".

30. C. B. Holroyd, M. G. Coles, *Psychological Review* **109**, 679–709 (2002).

31. G. F. Koob, N. D. Volkow, *Neuropsychopharmacology* **35**, 217 (2010).

32. A. I. Leshner, *Science* **278**, 45 (1997).

33. S. F. Tapert, *et al.*, *Archives of General Psychiatry* **60**, 727 (2003).

34. S. M. Tyree, L. de Lecea, *Frontiers in Systems Neuroscience* **11**, 50 (2017).

35. A. M. Kallen, C. J. Patrick, B. D. Bartholow, G. Hajcak, *Biological Psychology* **181**, 108597 (2023).

36. C. C. Williams, T. D. Ferguson, C. D. Hassall, W. Abimbola, O. E. Krigolson, *Psychophysiology* **58**, e13722 (2021).

37. L. R. Frank, V. L. Galinsky, *J. Phys. A* **49**, 395001 (2016).

38. J. Olds, P. Milner, *Journal of Comparative and Physiological Psychology* **47**, 419 (1954).

39. M. R. Hagerty, *et al.*, *Neural Plasticity* **2013** (2013).

40. A. Cantisani, *et al.*, *European archives of psychiatry and clinical neuroscience* **266**, 755 (2016).

41. K. Stavropoulos, L. Carver, *Mol Autism* **9**, 1 (2018).

42. C. H. Leung, E. Pang, *International Journal of Technology, Knowledge and Society* **17**, 33 (2021).

43. V. L. Galinsky, A. Martinez, M. P. Paulus, L. R. Frank, *Neural Computation* **30**, 1725 (2018).

44. D. Gottlieb, S. A. Orszag, *CBMS-NSF Regional Conference Series in Applied Mathematics: Numerical Analysis of Spectral Methods: Theory and Applications* (Society for Industrial and Applied Mathematics, 1977).

45. V. Fonov, A. Evans, R. McKinstry, C. Almli, D. Collins, *NeuroImage* **47**, S102 (2009). Organization for Human Brain Mapping 2009 Annual Meeting.

46. V. L. Galinsky, L. R. Frank, *Magn. Reson. Med.* **81**, 1225 (2018).

47. L. R. Frank, V. L. Galinsky, *Phys. Rev. E* **89**, 11 (2014).

48. L. R. Frank, V. L. Galinsky, L. Orf, J. M. Wurman, *J. Atmos. Sci.* **75**, 709 (2018).

49. T. A. Enßlin, M. Frommert, F. S. Kitaura, *Phys. Rev. D* **80**, 105005 (2009).

50. E. T. Jaynes, *Phys. Rev.* **106**, 620 (1957).

51. E. T. Jaynes, *Phys. Rev.* **108**, 171 (1957).

52. V. L. Galinsky, L. R. Frank, *Neural Computation* **29**, 1441 (2017).

53. R. B. Buxton, L. R. Frank, *J. Cerebr. Blood Flow Metab.* **17**, 64 (1997).

54. P. Jezzard, *NeuroImage* **62**, 648 (2012). 20 YEARS OF fMRI.

55. L. R. Frank, R. B. Buxton, E. C. Wong, *Magn. Reson. Med.* **45**, 635 (2001).

56. R. Pascual-Marqui, C. Michel, D. Lehmann, *International Journal of Psychophysiology* **18**, 49 (1994). Available at https://www.uzh.ch/keyinst/NewLORETA/Software/Software.htm.

57. H. Hallez, *et al.*, *Physics in Medicine & Biology* **50**, 3787 (2005).

58. H. Hallez, *et al.*, *Journal of Neuroengineering and Rehabilitation* **4**, (2007).

59. S. Dattola, F. C. Morabito, N. Mammone, F. La Foresta, *Electronics* **9** (2020).

60. S. Taulu, E. Larson, *IEEE transactions on bio-medical engineering* **68**, 992 (2021).

61. A. Gramfort, T. Papadopoulo, E. Olivi, M. Clerc, *Biomed Eng Online* **9**, 45 (2010).

62. J. Kybic, *et al.*, *IEEE Trans Med Imaging* **24**, 12 (2005).

63. N. von Ellenrieder, P. A. Valdes-Hernandez, C. H. Muravchik, *Med Biol Eng Comput* **47**, 1083 (2009).

64. D. Gutierrez, A. Nehorai, *IEEE Trans Biomed Eng* **55**, 1103 (2008).

65. P. H. Schimpf, C. Ramon, J. Haueisen, *IEEE Trans Biomed Eng* **49**, 409 (2002).

66. J. J. Ermer, J. C. Mosher, S. Baillet, R. M. Leah, *Phys Med Biol* **46**, 1265 (2001).

67. J. C. Mosher, R. M. Leahy, P. S. Lewis, *IEEE Trans Biomed Eng* **46**, 245 (1999).

68. C. M. Michel, *et al.*, *Clinical Neurophysiology* **115**, 2195 (2004).

69. A. Ojeda, K. Kreutz-Delgado, J. Mishra, *Neural Computation* **33**, 2408 (2021).

70. M. J. Rosa, J. Kilner, F. Blankenburg, O. Josephs, W. Penny, *NeuroImage* **49**, 1496 (2010).

71. M. Schultze-Kraft, R. Becker, M. Breakspear, P. Ritter, *Progress in Biophysics and Molecular Biology* **105**, 67 (2011).

72. J. Jorge, W. van der Zwaag, P. Figueiredo, *NeuroImage* **102**, 24 (2014).

73. R. Abreu, *et al.*, *Brain topography* **35**, 282 (2022).

74. R. Niazy, C. Beckmann, G. Iannetti, J. Brady, S. Smith, *NeuroImage* **28**, 720 (2005).

75. P. Lakatos, *et al.*, *Journal of Neurophysiology* **94**, 1904 (2005).

76. A. Martinez, *et al.*, *Front Hum Neurosci* **9**, 371 (2015).

77. A. Martínez, *et al.*, *Biological Psychiatry* **86**, 557 (2019). Schizophrenia: Genomics to Therapeutics.

Acknowledgments

LRF and VLG were supported by NSF grant ACI-1550405 and NIH grant R01 AG054049. ST was supported by NIH U24 AA021695, NIH U01 AA021692, NIH U01 DA041089, NIH R01 DA057567, OK was supported by NSERC Discovery Grant RGPIN 2016-0943. AM was supported by (NIMH) R21MH123875. SB was supported by NIDCD R01DC019979. The technology in this paper is covered under US Patents 10909414, 10789713, 11270445.

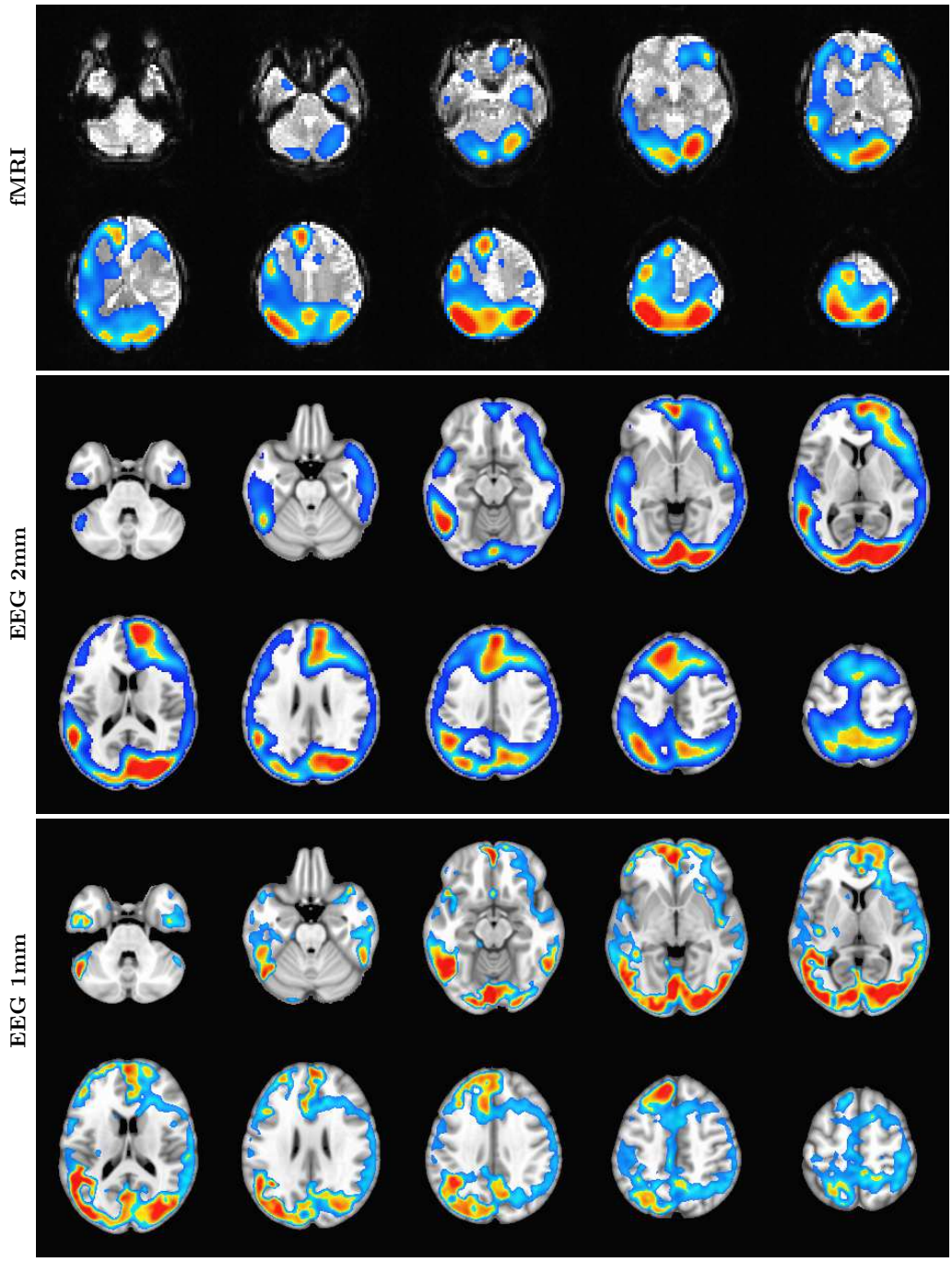

Figure 1: Comparison of EFD reconstructed fMRI activity (top) with SPECTRE reconstructed EEG at both 2mm (middle) and 1mm (bottom) spatial resolution (axial view) from a single representative subject from an open-source study with simultaneous fMRI and EEG (*11*). In both cases, the weighted sum of the power over all modes is shown. The task was a simple 8Hz flashing checkerboard with 4 on/off cycles.

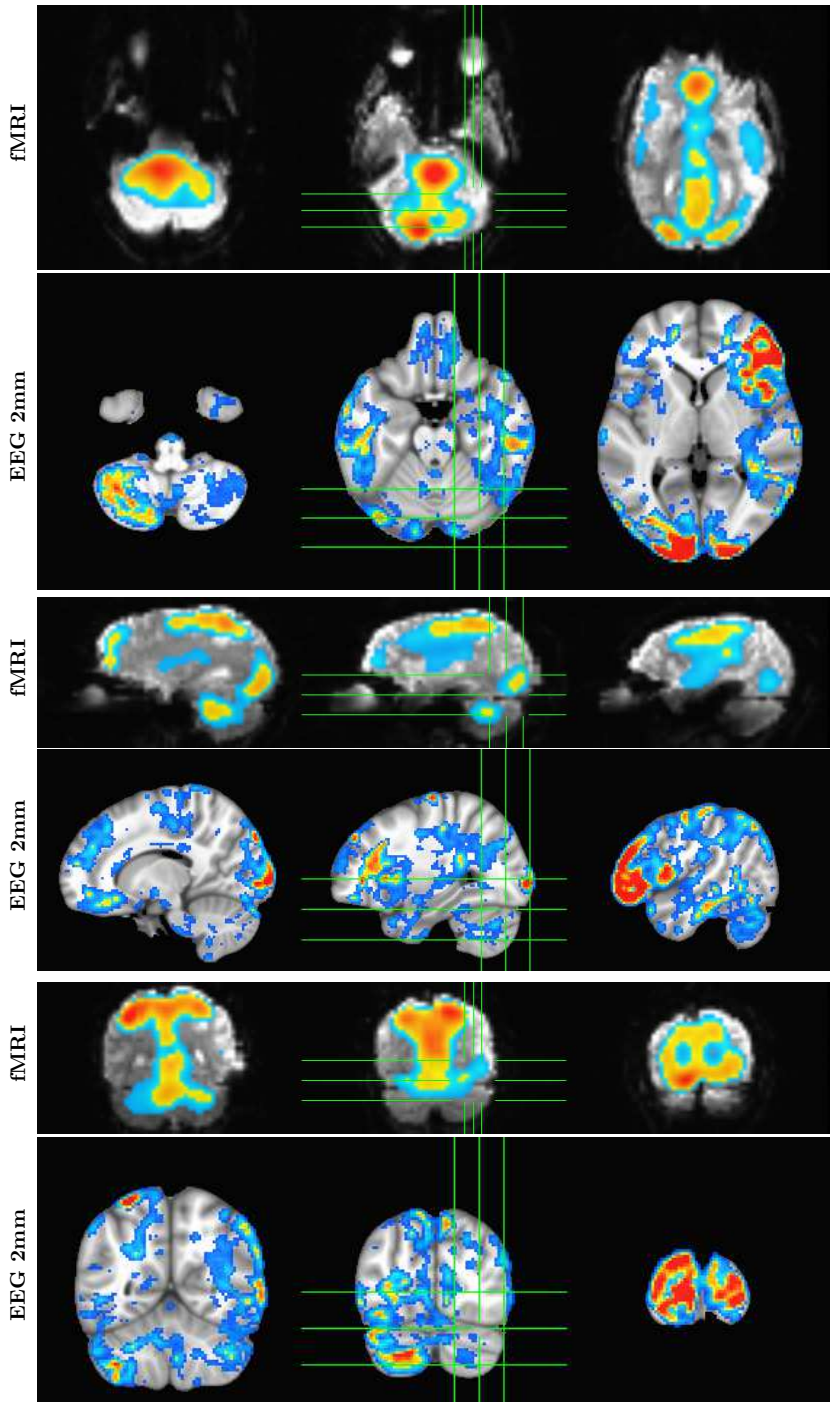

Figure 2: Comparison of EFD reconstructed fMRI activity (top) with SPECTRE reconstructed EEG at 2mm (bottom) from a single subject with simultaneous fMRI and EEG. (Same data source but different subject as in Fig. 1) demonstrating the fine spatial resolution produced by SPECTRE, and the ability to reconstruct activations in regions prone to severe distortions in fMRI, such as the frontal lobes.

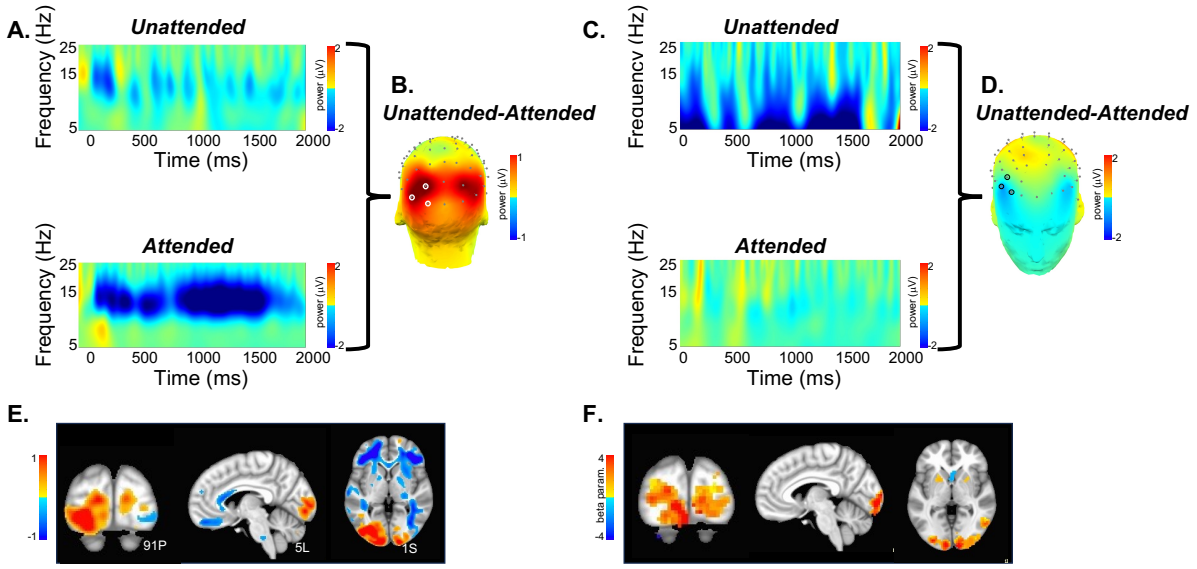

Figure 3: A. Baseline-corrected EEG activity from a single subject elicited by unattended (top) and attended (bottom) visual stimuli averaged across the cluster of 3 occipital electrode sites (PO7, PO3, O1) denoted in B. by white circles. Over the broad alpha frequency band (7-16Hz) there was a reduction in total power (from the pre- to post-stimulus latency interval) which was greater for attended, compared to unattended, visual stimuli. B. Scalp-topography of the mean difference in oscillatory (8-12Hz) activity for unattended minus attended visual stimuli across the 0-2000ms latency interval. As expected, attention modulated (reduced) the power of these oscillations over the visual cortex. C. As in A. for three frontal electrode sites (F6, F8, AF6) denoted in D. by black circles. In contrast to visual cortex, in bilateral frontal regions, unattended visual stimuli elicited a greater reduction of oscillatory activity between 5-10Hz (theta-alpha frequency). D. Frontal view of the unattended minus attended difference topography between 0-2000ms in the 8-12Hz frequency band. E. Source estimates derived from mean (baseline-corrected) oscillatory power between 0-2000ms and across 8-12Hz for the same subject shown in panels A-D, superimposed on the MNI template brain. Hot colors (yellow to red) indicate greater attention-related modulation (reduction) of activity and the inverse for warm colors (light to dark blue). F. BOLD signal (beta parameter estimate) contrasting activation to visual stimuli when attended versus activation to the same stimulus when unattended. Attention-related enhancement of the BOLD signal in visual cortex mirrors the reduction in alpha power obtained in the same subject using EEG.

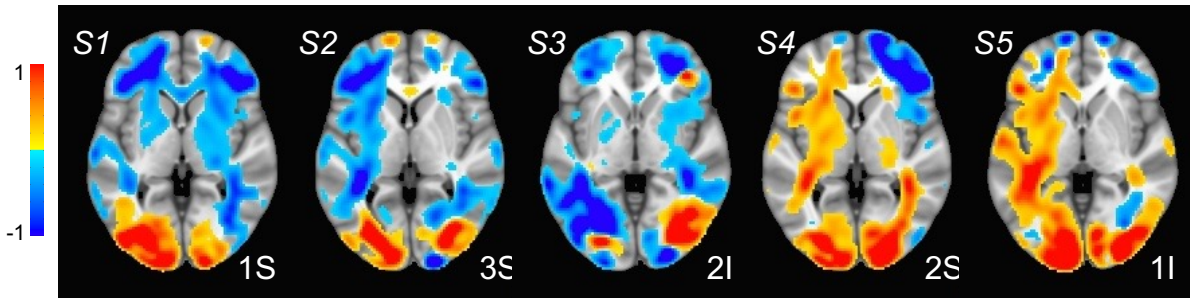

Figure 4: Estimated localization of neural sources for 8-12Hz oscillatory activity (unattended minus attended; 0-2000ms) for five participants (S1-S5). Colors are as in 1E. A prominent bilateral occipital source associated with increased attentional modulation is observable in all participants. A bilateral source localized in middle frontal cortex and indicating less modulation is also consistently observed across participants.

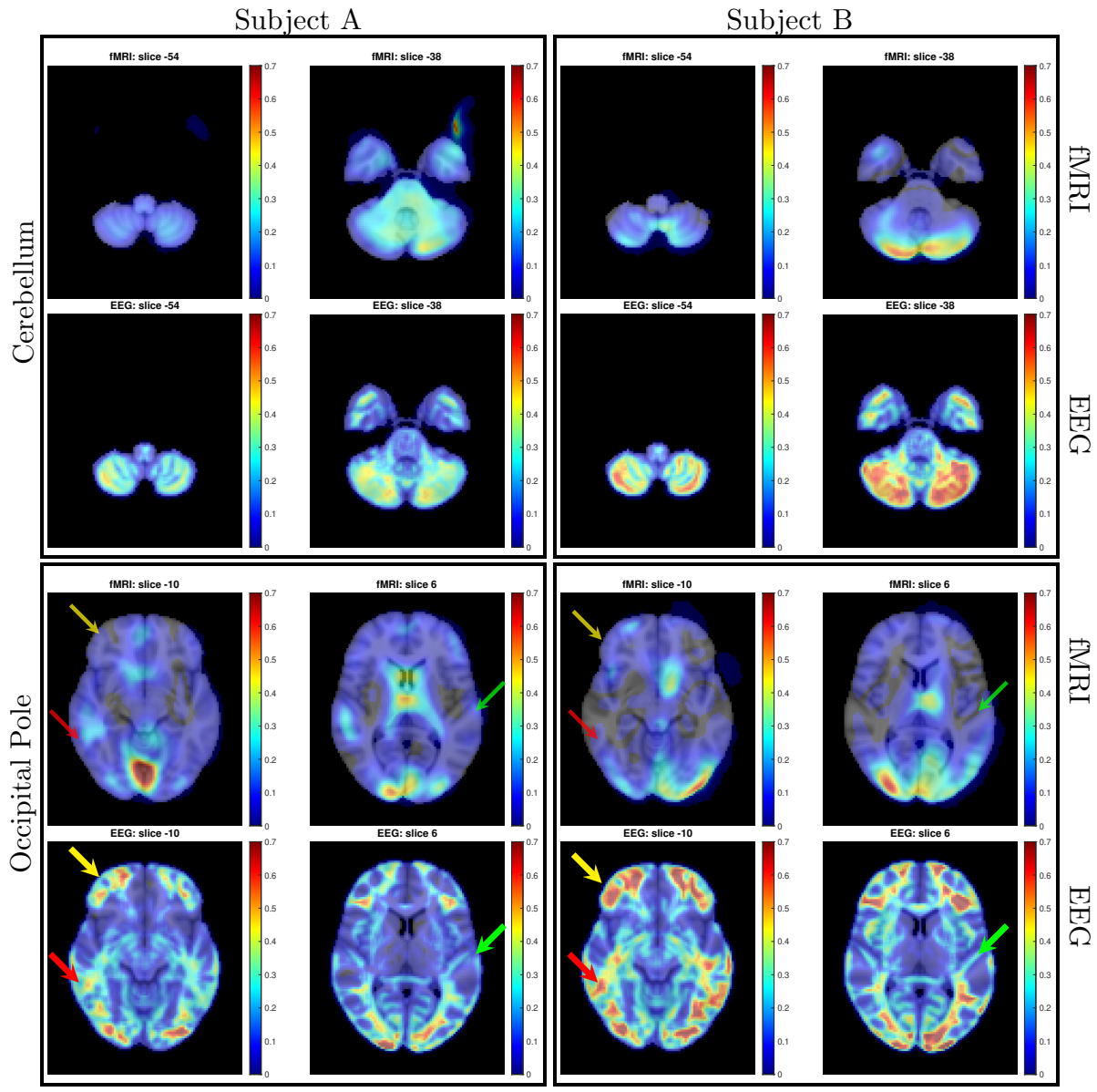

Figure 5: Direct comparison of activation maps from two participants (Subject A, left; Subject B, right) in the bimodal (auditory + visual) stimulation paradigm described for Figures 3 and 4. In each subject, two brain regions – the cerebellum and the occipital pole (top and bottom rows, respectively), were delineated based on the MNI atlas and EFD activation maps were correlated across these entire regions. Correlation coefficients were as follows: for Subject A, cerebellum=0.74, occipital pole=0.70; for Subject B, cerebellum=0.70, occipital pole=0.84. Correlations were computed only for regions exhibiting activation levels above 0.1. In contrast to fMRI, the SPECTRE technique identified robust activations in bilateral middle and inferior frontal cortex (indicated by yellow arrows) and middle temporal cortex (red arrows). It also discerned activations along the superior temporal cortex, including areas encompassing the primary auditory cortex (green arrows).

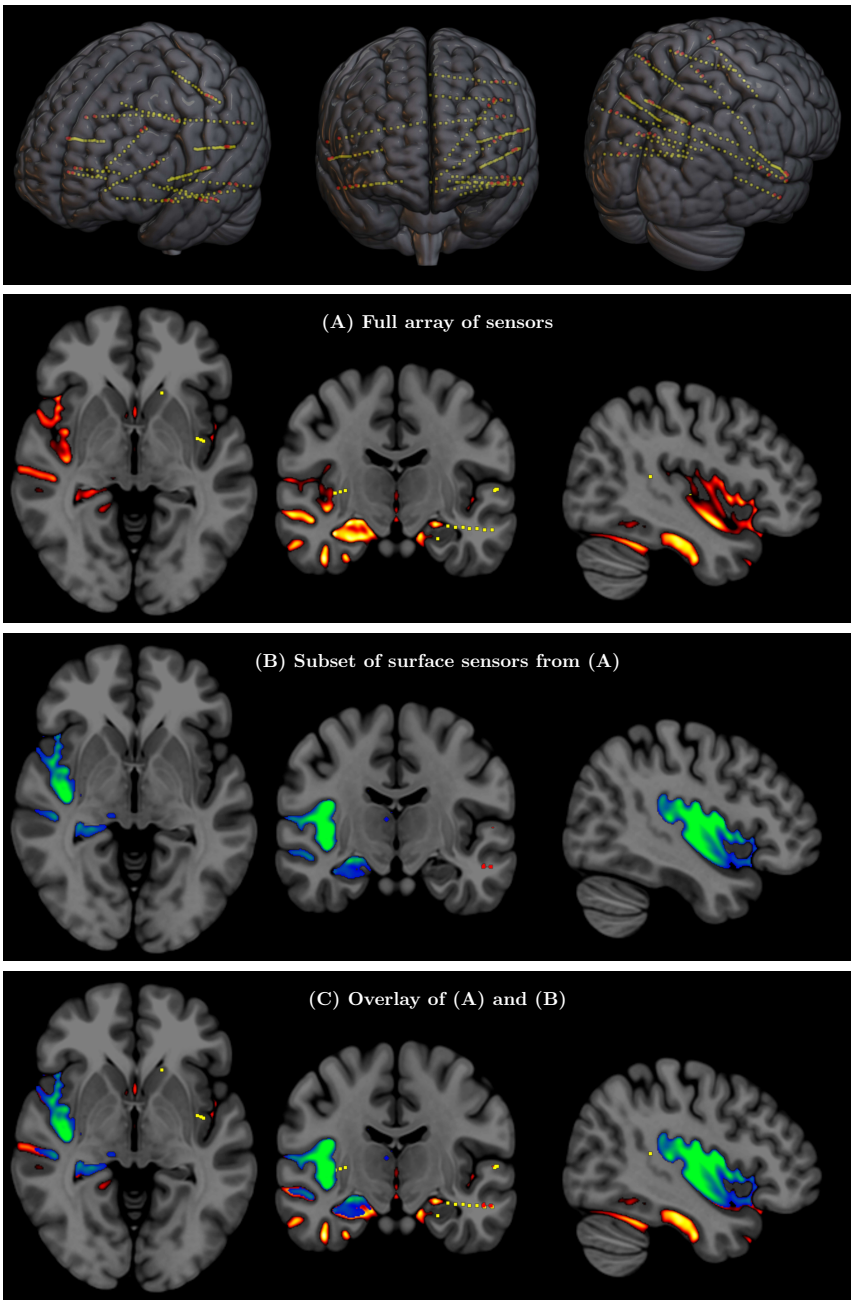

Figure 6: (Top row) Full array of intra-cranial EEG contacts from a recording in a medically refractory epilepsy patient (yellow dots). Red dots indicate subset of surface-only electrodes to mimic a standard non-invasive (i.e., extra-cranial) EEG study. SPECTRE α band reconstruction from (A) full array of intra-cranial EEG sensors from an epilepsy study (yellow dots) in top row and (B) from subset of surface electrodes (red dots) in top figure. (C) Overlay of (A) and (B) validating that the surface based is correctly reconstructing the local electric field potential detected by the intra-cranial electrodes.

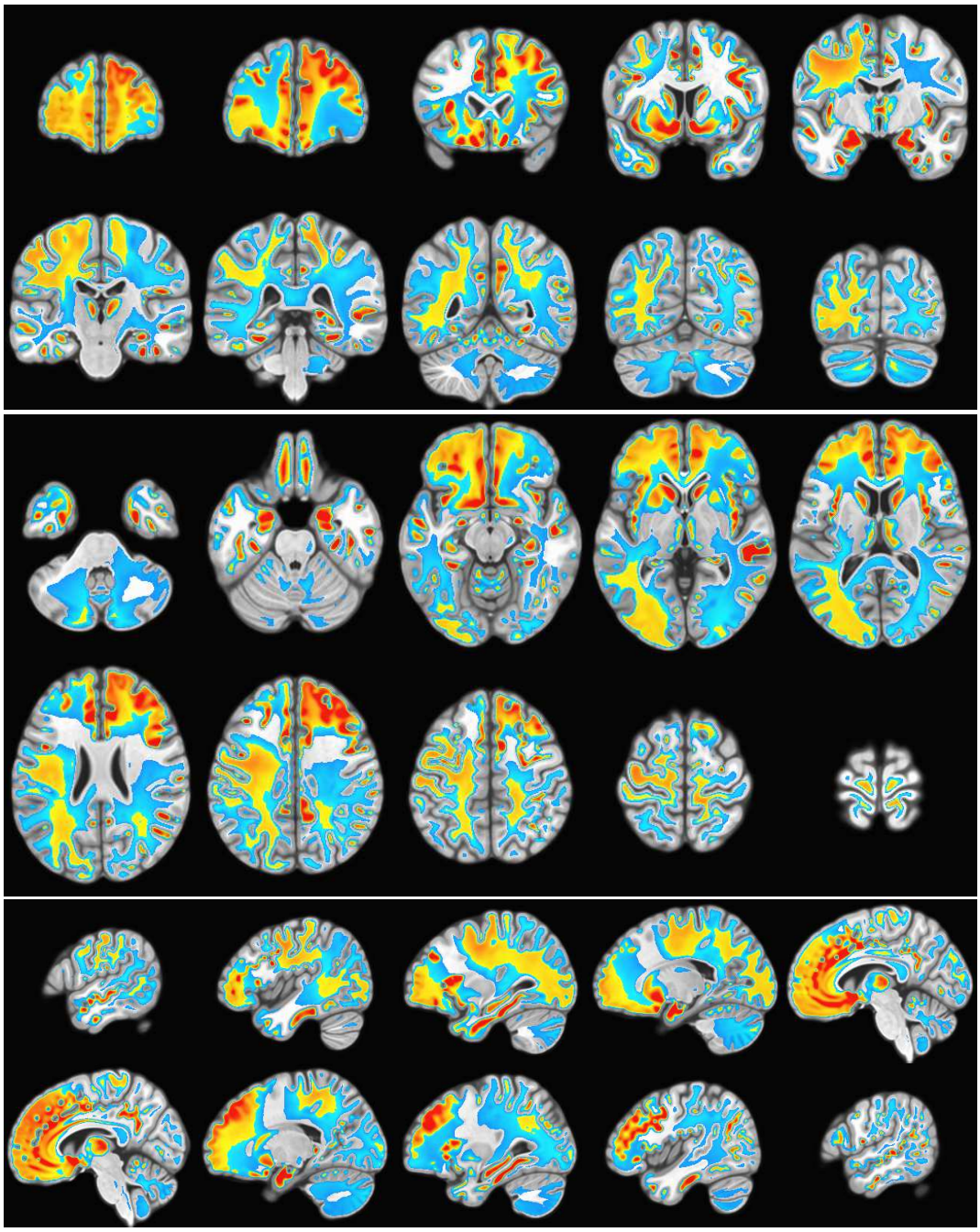

Figure 7: Gambling task EEG from 500 subject cohort. Alpha power of the weighted summed over the first $n = 10$ SPECTRE modes $\tilde{\psi}_{10}$. Activation in key regions of the reward circuit, including the frontal lobes, paracingulate gyrus, accumbens, and amygdala are clearly evident. Negative activation (i.e., deactivation) is evident in the supplementary motor cortex and the left temporal-parietal regions.

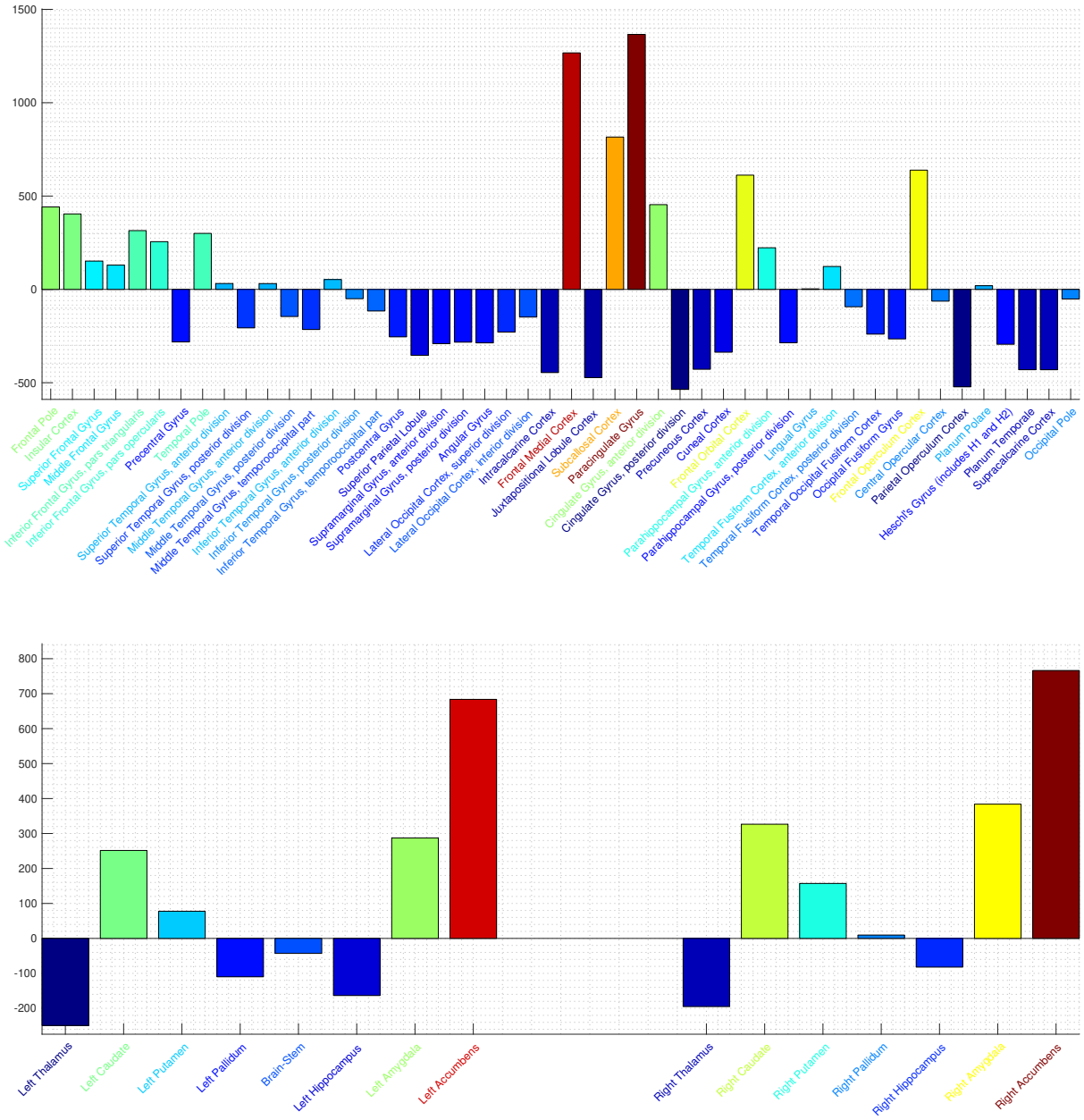

Figure 8: SPECTRE power per brain region in the Harvard-Oxford 2mm cortical (top) and subcortical (bottom) atlases. Colormap is from hot/yellow (activated) to blue (de-activated). Activation in key regions of the reward circuit, including the frontal lobes, paracingulate gyrus, subcallosal cortex/nucleus accumbens, and amygdala are clearly evident. Negative activation (i.e., deactivation) is evident in the supplementary motor area, posterior cingulate, and thalamus. Activation of the important reward element accumbens is evident in the bottom plot. Also of note is the relatively similar activation in the bilateral subcortical elements.

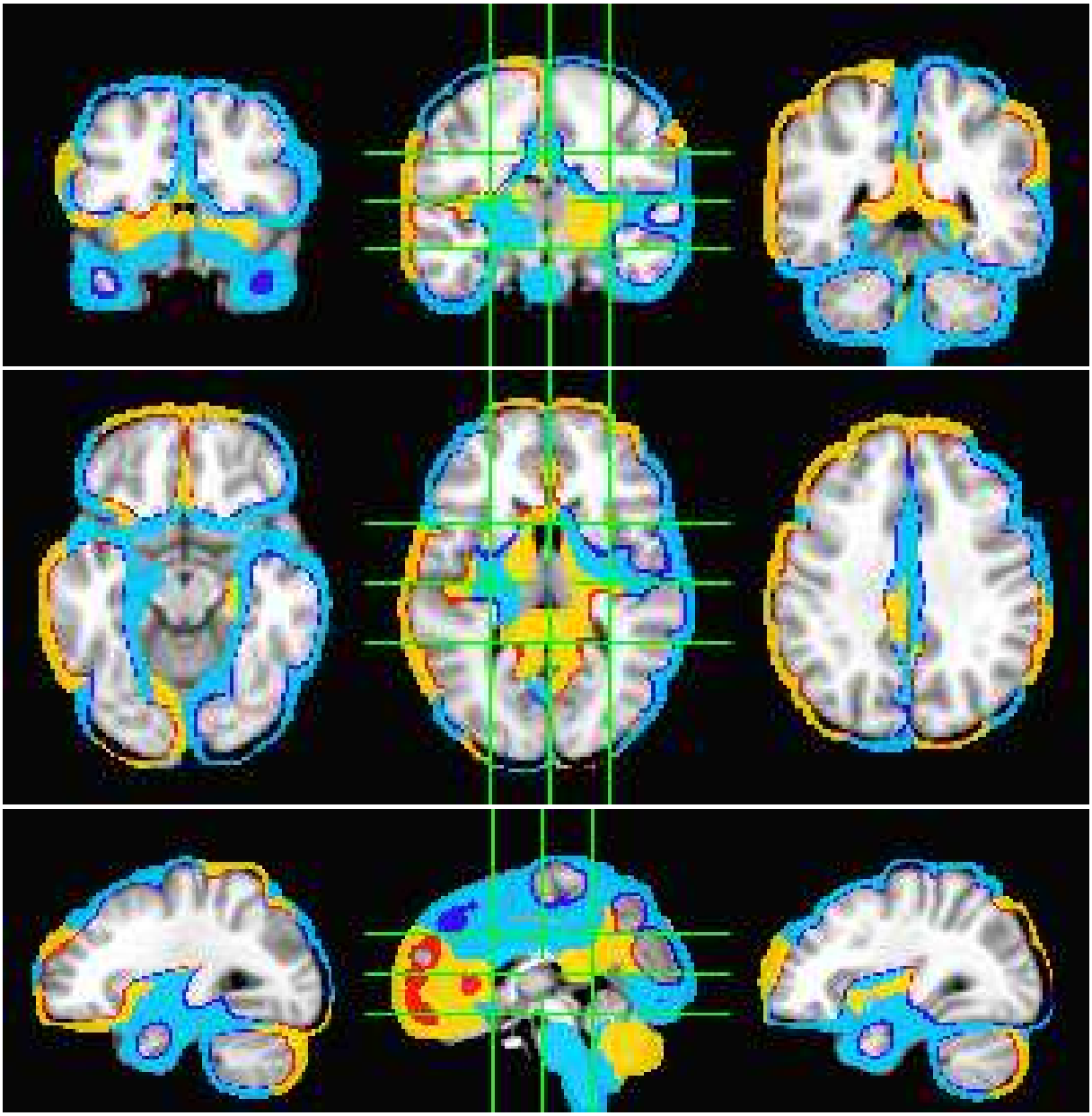

Figure 9: Statistical significance. t-statistic between the SPECTRE power modes pre- and post-stimulus reward experiment. Calculations were performed using the standard AFNI 3dttest++ algorithm. Yellow/red color reflects positive changes, blue color reflects negative changes. Significance threshold was $p = 10^{-8}$, indicating strong statistical significance.

6 Supplementary Material: Data and Methods

6.1 SPECTRE details

6.1.1 Solution to the inverse EEG problem

Our method for solving the inverse EEG problem can be summarized as follows. Given a standard EEG dataset from N electrodes and a high-resolution anatomical (HRA) MRI dataset with high contrast between gray matter (GM) and white matter (WM), the solution to the inverse EEG problem can be formulated as an approximation for the volumetric distribution of electrostatic potential inside the complex inhomogeneous and anisotropic tissues and complicated morphology of the MRI domain (*43*). A general form of this approximation is based on Maxwell's equations in an anisotropic and inhomogeneous medium expressed as charge continuity. This is used to derive the frequency dependent electrostatic field potential ϕ that is dependent upon the electrical properties of the tissue permittivity, permeability, and conductivity. These parameters can be estimated from the HRA data. In the Fourier (i.e., frequency) domain, the electrostatic potential satisfies the equation

$$\left(\Sigma^{ij} - I\omega\varepsilon\delta^{ij}\right)\partial_i\partial_j\phi_\omega = \left[I\omega(\partial_i\varepsilon)\delta^{ij} - (\partial_i\Sigma^{ij})\right]\partial_j\phi_\omega + \mathscr{F}_\omega, \tag{1}$$

which is written in tensor form where a summation is assumed over repeated indices. This can be expressed in the form $\hat{L}\phi_\omega = \hat{R}\phi_\omega + \hat{\mathscr{F}}_\omega$ in terms of the operators $\hat{L} \equiv \partial^i\partial_i$, a frequency dependent source term $\hat{\mathscr{F}}_\omega$ and the operator

$$\hat{R} \equiv \frac{\sigma + I\omega\varepsilon}{\sigma^2 + \omega^2\varepsilon^2}\left[I\omega(\partial_i\varepsilon)\delta^{ij} - (\partial_i\Sigma^{ij}) - \left(\Sigma^{ij} - \sigma\delta^{ij}\right)\partial_i\right]\partial_j$$

where $\mathbf{\Sigma} = \{\Sigma^{ij}\}$ is a local tissue conductivity tensor, $\sigma = Tr\mathbf{\Sigma}/3 = \Sigma^i_i/3$ is an isotropic local conductivity. Terms in square brackets show that the parts of $\hat{R}\phi_\omega$ can be interpreted in terms of different tissue characteristics and may be important for understanding the origin of sources of the electro-/magnetostatic signal detected by the EEG sensors. The first term

$(\omega(\partial^i\varepsilon)(\partial_i\phi_\omega))$ corresponds to areas with sudden change in permittivity, e.g. the WM/GM interface. The second term $((\partial_i\Sigma^{ij})(\partial_j\phi_\omega))$ corresponds to regions where the conductivity gradient is the strongest, i.e. the GM/CSF (cerebral spinal fluid) boundary. Finally, the last term $(\Sigma^{ij}\partial_i\partial_j\phi_\omega - \sigma\partial^i\partial_i\phi_\omega)$ includes areas with the strongest conductivity anisotropies, e.g. input from major WM tracts. The frequency and position dependent internal sources $\hat{\mathscr{F}}_\omega$ can be used to incorporate various nonlinear processes including multiple frequency effects of the efficient synchronization/desynchronization by brain waves or effects of their critical dynamics. This term is ignored in the current paper.

The inverse problem can be solved by constructing an approximate solution for the potential ϕ across an entire brain volume iteratively as $\hat{L}\phi_\omega^{(k)} = \hat{R}\phi_\omega^{(k-1)}$ and $\tilde{\phi}_\omega^{(K)} = \alpha_K \sum_{k=0}^{K} \phi_\omega^{(k)}$ (*43*): where a single iteration forward solution is found using a Fourier-space pseudo-spectral approach (*44*). Data from MRI can be used to define the complex brain tissue morphology and constrain the tissue specific values of $\mathbf{\Sigma}$ and ε. This procedure of inverting the WETCOW brain wave model constrained by MRI-defined tissue properties is called *SPatially resolved EEG Constrained with Tissue properties by Regularized Entropy (SPECTRE)*.

SPECTRE is flexible in its ability to incorporate relevant prior information from MRI data. HRA data is useful for tissue segmentation and assignment of mean values for permittivity and conductivity. Diffusion MRI (dMRI) data further allows construction of estimates of the conductivity tensor anisotropy. In the present study, only HRA data were used for the estimation procedure. One important practical point is that very often HRA data are not acquired for participants in EEG studies. In this case it is sufficient to use HRA data from a standard atlas, and spatially register the EEG data to the atlas. Indeed, this was case in the present paper, where the standard 2mm, 1mm, and 0.7mm resolution T1-weighted anatomical MRI Montreal Neurological Institute (MNI) (*45*) atlases were used, to which the EEG data were aligned using our non-linear registration algorithm SYMREG (*46*).

6.1.2 Mode reconstruction

The reconstructed volumetric time series of the estimated electrostatic potential field can be thought of as similar to the EM equivalent of an fMRI dataset, containing a multitude of correlated spatiotemporal patterns or "modes" of the system.

The problem then becomes one of detecting the multiple modes in complex non-linear systems. We have addressed this problem previously in our development of the *entropy field decomposition* (EFD) method, which is a probabilistic framework for estimating spatial-temporal modes of complex non-linear systems containing multivariate interacting fields. It is formally based on a field-theoretic mathematical formulation of Bayes' Theorem that enables the hierarchy of multiple orders of field interactions including coupling between fields (*12, 37*). Its practical utility is enabled by incorporation of the theory of entropy spectrum pathways (ESP) (*47*), which uses the space-time correlations in each individual dataset to automatically select the very limited number of highly relevant field interactions. In short, it selects the configurations with maximum path entropy, summarized in the equilibrium (i.e., long time) distribution μ^*.

Our terminology is as follows (see Section 6.1.3 details). The k'th EFD mode of the coupled parameters $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_m\}$ is denoted $\psi_{\boldsymbol{\alpha}}^{(k)}(\boldsymbol{x}, t)$ where $k = 1, \ldots, n$ for some user defined n. While each of these modes provides unique information on coherent spatio-temporal activity, for characterizing the total brain activity it is often most useful and efficient to sum these modes: $\tilde{\psi}_n = \sum_k^n \psi_{\boldsymbol{\alpha}}^{(k)}(\boldsymbol{x}, t)$. In this study we utilize both individual modes and the summed modes, as appropriate.

A strength of the EFD method is that it uses prior information contained in individual datasets - there are no training datasets or averages across datasets - just the prior information contained within the single dataset of interest. This method has shown utility in resting state fMRI data (*12*) and in meteorology in the application to severe local storms, in particular tornadic supercells (*48*). The fact that this method uses prior information embedded within

single datasets without the need for any 'training' is of significance to clinical studies in which important individual variations can be lost in the averaging process. It is also particularly important in the current paper where our validation necessitates comparison with single subject studies.

6.1.3 Entropy Field Decomposition

The *entropy field decomposition* (EFD) is a general probabilistic method for the estimation of spatial-temporal modes of complex non-linear, non-period, non-Gaussian multivariate data (*12, 37*). The goal of EFD is to estimate the field $\psi(\boldsymbol{x}, t)$ that describes a continuous (in both space $\boldsymbol{x}$ and time t) parameter space from which the signal s_l are discrete samples of $s_l = \int \psi(\xi)\delta(\xi - \xi_l)d\xi$. In general, this can be done by constructing the posterior distribution of $\psi(\boldsymbol{x}, t)$ given the data d and any prior information I that is available, via Bayes' Theorem re-expressed in the language of field theory (*49*) as

$$p(\psi|s, I) = \frac{e^{-H(s,\psi)}}{Z(s)} \tag{2}$$

where $Z(d) = p(d|I) = \int d\psi\, e^{-H(s,\psi)}$ is the *partition function* (considered a constant in this application) and $H(s, \psi) = -\ln p(s, \psi|I)$ is the *information Hamiltonian* which takes the form

$$H(s, \psi) = H_0 - j^\dagger\psi + \frac{1}{2}\psi^\dagger D^{-1}\psi + H_i(s, \psi) \tag{3}$$

where H_0 is essentially a normalizing constant that can be ignored, j and D are the information source and propagator, and $\dagger$ means the complex conjugate transpose. H_i is an interaction term (*49*)

$$H_i = \sum_{n=1}^{\infty} \frac{1}{n!} \int \cdots \int \Lambda^{(n)}_{s_1 \cdots s_n} \psi(s_1) \cdots \psi(s_n) ds_1 \cdots ds_n \tag{4}$$

where $\Lambda^{(n)}_{s_1 \cdots s_n}$ terms describe the interaction strength. In highly complex non-linear systems, j, D, and $\Lambda^{(n)}_{s_1 \cdots s_n}$ are often unknown and too complex for deriving effective and accurate approximations. In this case the ESP method (*47*), based on the principal of maximum entropy (*50, 51*),

provides a general and effective way to introduce powerful prior information to find the most significant contributions to $H(d,\psi)$ by using coupling between different spatio-temporal points that is available from the data itself. This is accomplished by constructing a *coupling matrix* that characterizes the relation between locations i and j in the data $Q_{ij} = e^{-\gamma_{ij}}$ where the γ_{ij} are Lagrange multipliers that describe the relations and depend on some function of the space-time locations i and j. The eigenvalues λ_k and eigenvectors $\phi^{(k)}$ of the coupling matrix Q

$$\sum_j Q_{ij}\phi_j^{(k)} = \lambda_k \phi_i^{(k)} \tag{5}$$

then define the transition probability from location j to location i of the k'th mode as

$$p_{ijk} = \frac{Q_{ji}}{\lambda_k}\frac{\phi_i^{(k)}}{\phi_j^{(k)}} \tag{6}$$

For each transition matrix (6) there is a unique stationary distribution associated with each path k:

$$\mu^{(k)} = \left[\phi^{(k)}\right]^2 \qquad \text{where} \qquad \mu_i^{(k)} = \sum_j \mu_j^{(k)} p_{ijk} \tag{7}$$

where $\mu^{(1)}$, associated with the largest eigenvalue λ_1, corresponds to the maximum entropy stationary distribution. Note that (7) is written to emphasize that the squaring operation is performed on a pixel-wise basis.

The essence of the EFD approach (*12, 37*) is to incorporate these coupling matrix priors into the information Hamiltonian (3) by expanding the signal $s(\boldsymbol{x},t)$ into a Fourier expansion using $\{\phi^{(k)}\}$ as the basis functions

$$s_i = \sum_k^K \left[a_k \phi_i^{(k)} + a_k^\dagger \phi_i^{\dagger,(k)}\right] \tag{8}$$

which allows expressing the information Hamiltonian (3) in this ESP basis as

$$H(d,a_k) = -j_k^\dagger a_k + \frac{1}{2}a_k^\dagger \Lambda a_k + \sum_{n=1}^\infty \frac{1}{n!}\sum_{k_1}^K \cdots \sum_{k_n}^K \tilde{\Lambda}^{(n)}_{k_1\cdots k_n} a_{k_1}\cdots a_{k_n} \tag{9}$$

where matrix Λ is the diagonal matrix Diag$\{\lambda_1, ..., \lambda_K\}$, composed of the eigenvalues of the coupling matrix, and $j_k = \int j\, \phi^{(k)} ds$ is the amplitude of kth mode in the expansion of the source j and the new interaction terms $\tilde{\Lambda}^{(n)}$ are

$$\tilde{\Lambda}^{(n)}_{k_1 \cdots k_n} = \int \cdots \int \Lambda^{(n)}_{s_1 \cdots s_n} \phi^{(k_1)} \cdots \phi^{(k_n)} ds_1 \cdots ds_n \tag{10}$$

from which can be derived (*12, 37*) a simple expression for the solution for the amplitudes a_k

$$\Lambda a_k = \left(j_k - \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k_1}^{K} \cdots \sum_{k_n}^{K} \tilde{\Lambda}^{(n+1)}_{k k_1 \cdots k_n} a_{k_1} \cdots a_{k_n} \right) \tag{11}$$

through the eigenvalues and eigenvectors of coupling matrix.

The EFD methods can be extended to multiple modalities by incorporating coupling between different parameters, which we call *Joint Estimation with Entropy Regularization* (JESTER) (*52*). For $m = 1, ..., M$ different modalities $d^{(m)}$ with the coupling matrices $Q^{(m)}$ that all correspond to the same unknown signal s, intermodality coupling matrix can be constructed as the product of the coupling matrices for the individual modalities expressed in the ESP basis and registered to a common reference frame, which we denote $\widetilde{Q}^{(m)}$: That is, the joint coupling matrix is $\mathscr{Q}^{(m)} = \prod_m \widetilde{Q}^{(m)}$. More specifically, the joint coupling matrix $\mathscr{Q}_{ij}$ between any two space-time locations (i, j) can be written in the general (equivalent) form as

$$\ln \mathscr{Q}_{ij} = \sum_{m=1}^{M} \beta^{(m)}_{ij} \ln \widetilde{Q}^{(m)}_{ij} \tag{12}$$

where the exponents $\beta^{(m)}$ can either be some constants or functions of data collected for different modalities $\beta^{(m)}_{ij} \equiv \beta^{(m)}(\widetilde{\boldsymbol{d}}_{\boldsymbol{i}}, \widetilde{\boldsymbol{d}}_{\boldsymbol{j}}), \widetilde{\boldsymbol{d}}_{\boldsymbol{i}} \equiv \{\widetilde{d}^{(1)}_i, ..., \widetilde{d}^{(M)}_i\}$ where $\widetilde{d}^{(m)}_i$ and $\widetilde{Q}^{(m)}_{ij}$ represent, respectively, the data and the coupling matrix of the modality dataset m represented in the ESP basis and evaluated at locations r_i and r_j of a common reference domain R:

$$\widetilde{d}^{(m)}_i = d^{(m)}\left(\psi^{(m)}(r_i)\right), \qquad \widetilde{Q}^{(m)}_{ij} = Q^{(m)}\left(\psi^{(m)}(r_i), \psi^{(m)}(r_j)\right) \tag{13}$$

where $\psi^{(m)} : R \to X$ denotes a diffeomorphic mapping of m-th modality from the reference domain R to an acquisition space X.

6.2 Comparison with functional MRI

Functional MRI (fMRI) has become the *de facto* neuroimaging method for spatial and temporal localization of brain activity. The contrast mechanism that forms the basis of fMRI is the blood oxygenation level dependent (BOLD) variations in the magnetic state of hemoglobin and its influence on the local MRI signal as a function of the local metabolism and hemodynamics (*53*). Consequently, the spatial and temporal characteristics of the fMRI signal are related to blood flow and metabolic dynamics, rather than direct measures of electrical activity. In particular, the signal variations will be spatially localized in vascular pathways and the temporal variations, being related to blood flow effects, are very slow compared to electrical activity. In short, the spatial-temporal dynamics measured by fMRI need not (and, in fact, will not) correspond exactly to the spatial-temporal patterns of electrical activity. Numerous experimental realities also make fMRI problematic as a gold standard. In particular, fMRI is facilitated by enhancing the sensitivity of MRI to the BOLD contrast mechanism, which requires enhancing the sensitivity to local magnetic field variations through the use of T_2-weighted pulse sequences (*54*) which lead to increased geometric distortions, compromising not only spatial resolution but confounding the spatial localization of the activity in a complex, non-linear fashion. Gross distortions can lead to significantly reduced signal-to-noise and even completely unrecoverable signal loss, particularly in regions near air/tissue interfaces, such as in the prefrontal cortex (PFC). Moreover, the complex non-linear interactions between the magnetic fields and physiological variations such as respiration and cardiac pulsations produce a variety of complex spatiotemporal signal distortions (*55*). While mitigating these artifacts is an area of very active research, they remain a serious problem for fMRI.

Nevertheless, certain very simple task-based fMRI experimental paradigms, such as finger tapping or rapidly flickering checkerboard stimuli, repeated at periodic on/off intervals, have been established as experiments that produce repeatable robust activations in known brain net-

works, and are commonly used as basic testbeds for assessment of analysis algorithms. When combined with simultaneous EEG acquisition, such experiments provide two different types of data that can be compared as a form of validation, with the proviso that these two methods are imaging different physical quantities.

6.3 Comparison with state-of-the-art Source Localization methods

There is a long history of attempts to spatially localize EEG activity and these are generally called "source localization" methods (*56–59*). These methods are fundamentally different from the SPECTRE approach as they involve numerous stringent assumptions about brain electrical activity such as a fixed set of static dipole sources, an idealized geometric model of the head reduced to a few (typically 3) shells, that spatially close points are more likely synchronized, and the smoothness of the solution. (see (*60*) and references therein).

These methods all implicitly assume the "quasi-static" approximation to the EM field equation which entails ignoring the time dependent terms in Maxwell's equations, which are dependent on tissue conductivity properties which are themselves frequency dependent. The resulting solutions are therefore static, have no frequency dependence, and are insensitive to the detailed spatially variable electrical properties of the tissues. However, as discussed in detail in the development of the WETCOW model (*7–9*), these assumptions are incompatible with the basic physics of brain electrical activity. The SPECTRE approach is to employ the WETCOW model and solve the actual physical problem of the complete Maxwell's equations in an inhomogeneous and anisotropic medium. The WETCOW theory provides a comprehensive framework for characterizing the propagation of EM fields through the complex brain tissue microstructure and larger scale morphology (e.g., cortical folding), and provides the dynamic solution to the electric potential field necessary to solve the EEG inverse problem.

The pseudo-spectral computational approach used in SPECTRE has some important ad-

vantages over the finite/boundary element approaches typically used for electrostatic modeling of brain activity (*61–67*). It does not use surface meshes and so does not require limiting the location of sources to a small number of surfaces with fixed number of static dipole sources constrained to the surfaces. And the distribution of both electrostatic and geometric properties of the media (conductivity, permittivity, anisotropy, inhomogeneity - derived from the MRI data) are incorporated at every location throughout the volume. It is thus able to find a time dependent spatial distribution of the electrostatic potential at every space-time location of a multidimensional volume as a superposition of source inputs from every voxel of the same volume (*43*). These traits allow it to model wave-like signal propagation inside the volume and can detect and characterize significantly more complex dynamical behavior of the sources of the electrostatic activity recorded at the sensor locations than traditional methods.

To understand intuitively why SPECTRE is capable of reconstructing EM activity through the entire brain, including deep within subcortical structures, a simple idealized example is helpful. Consider two point current sources of different frequencies, one in the cortical layer close to the scalp, the second deep within the subcortical structures of the brain. Consider a single sensor placed on the scalp collinear with the two sources. Standard source localization methods will not see the deep source, since there is no frequency dependence, and the signal falloff is simply a function of the distance from the sensor. Therefore, the close source completely dominates the signal model. Since all tomographic imaging methods (e.g., MRI, CT, etc.) depend strongly on both the spatial and temporal sampling of the measured physical system, this effective invisibility of currents in the standard quasi-static model essentially precludes the solution of the true inverse EEG problem and necessitates the artificial construction of assumed dipole distribution on pre-chosen artificial internal structures. In contrast, in SPECTRE the sources are not dipoles, but frequency sources that extend throughout the entire brain volume subject to the boundary conditions imposed by both the tissues geometry and its spatially and frequency

dependent properties. The surface electrodes are assumed to be sensing EM waves emanating from the entire brain across a broad frequency spectrum limited only by the sensors. Used in conjunction with an HRA MRI data that provides the spatial distribution of the frequency-dependent tissue electrical properties that constrain the possible solution, SPECTRE can invert the wave equations to provide an estimate of the spatiotemporal distribution of the electric field potential.

The problem of spatially localizating the EEG signal involves estimating the most probable distribution of electric field amplitudes given an array of sensors. This is essentially a problem of correctly modeling the physics of how electromagnetic waves propagate through the complex environment of a the convoluted brain tissue morphology and the anisotropic and inhomogeneous nature of brain tissue. The current state-of-the-art approach to this problem, called 'source localization' such as *low resolution electromagnetic tomography* or LORETA algorithm with its many variations (*56–59*), also called 'EEG source imaging' (*68, 69*)) involves using a pre-defined brain atlas, arbitrarily placing dipole source on the surface, and calculating the contribution from these sources. Some methods propose using fMRI as a prior, which has the disadvantage of requiring fMRI acquisitions (*70–73*). The current source localization methods are based on a static model for the electric field caused by a fixed set of pre-defined dipole sources. (see (*2*) for a review of current methods). This model is inherently limited because in reality the brain's electrical field variations are time dependent and generated by an essentially continuous distribution of sources throught the entire brain. This description is the essence of the WETCOW theory (*7–9*) which describes how highly coherent localized electric field phenomena, such as cortical wave loops and synchronized spiking, are produced by the complex non-linear interactions of waves across multiple spatial and temporal scales. In a typical application of the SPECTRE method, we use MNI volumetric grid with 2mm (902629 voxels), 1mm (7221032 voxels), or 0.73mm (11393280 voxels). All voxels in our models are

considered sources of electromagnetic activity consistent with the local intravoxel tissue characteristics (via (1)) rather than an assumed dipolar form. In an attempt to compare SPECTRE with current state-of-the-art source localization method, we downloaded the currently available LoretaKey1 software (*59*), which uses a set of 6239 fixed dipoles. In order to make a fair comparison, we tried to use LoretaKey1 with a number of dipoles comparable to our $2mm$ number of voxels. We found that LoretaKey1 is not able to handle this size due to "out-of-memory" crashes. We scaled down the number of dipoles in 2, 4 , 10 and 20 times. Only with around 45K dipoles were we able to make Loreta run. It ran for approximately 24 hours, but then again crashed due to out-of-memory problems. Our processing with 2mm, requires around 650Mb of memory. At this point it was decided that it was not possible for Loreta to provide a result that would usefully inform the efficacy of the SPECTRE method. Our highest 0.73 resolution processing can be completed on a modern workstation taking 16-20Gb of memory in a matter of minutes.

6.4 Data

6.4.1 Attention paradigm data

Simultaneous EEG/fMRI acquisition: Functional and structural MRI images were acquired on a Siemens 3T TIM-Trio scanner (NKI Center for Biomedical Imaging and Neuromodulation) equipped with a 32-channel phased array head coil. Structural T1 and T2 scans were collected using standard sequences. Whole-brain BOLD data was acquired with a gradient-echo EPI sequence (TR=2000ms;TE=30ms;flip angle=80°). EEG data were acquired concurrently with fMRI using an MR compatible EEG amplifier (BrainVision MR series, Brain Products, Munich, Germany) and a 64-channel MR-compatible ring electrode cap with 10–20 International System electrode placement cap. EEG data was sampled at a rate of 5 KHz EEG data were acquired at a rate of 5 KHz using BrainVision Recorder software (Brain Products). Electrocardiographic data

were captured from electrodes on the backs of subjects. The reference electrode was positioned between Fz and Cz. Scanner and heartbeat artifacts were removed offline from the EEG signal using an average template subtraction procedure (*74*) and the data was resampled to 250Hz.

Traditional EEG analysis: The single-trial EEG signal from each electrode was convolved with a 3-cycle Morlet wavelet computed over a 3 second window centered at the onset of each stimulus and averaged separately for each stimulus type. The averaged spectral amplitude at each time point was then baseline corrected by subtracting the mean spectral amplitude over the -200 to -50 pre-stimulus interval. Further details of post-processing and time-frequency analyses methods are described in (*75–77*).

6.4.2 Reward circuit data

Task and Design Acquisition Participants completed a simple gambling task (*36*). On each trial, they saw a black fixation cross for 500 ms, followed by two colored squares for 500 ms, and then, the fixation cross turned gray (go cue) and participants were to select one of the two squares (square locations—left, right—were randomized on each trial) within a 2,000 ms time limit. They were then presented with a black fixation cross for 300 to 500 ms, and then, simple feedback as to their performance ("WIN" for gain, "LOSE" for loss) for 1,000 ms in black font. If the participants responded before the go cue they were instead delivered "TOO FAST" feedback and if they did not respond before the 2,000 ms time limit, it would be considered a loss. The goal of the participants was to accumulate wins by determining which of the two squares would more often lead to gains (60% vs. 10%). In this task, participants accumulated wins; however, were not paid money. They would see the same pair of colors for one block of 20 trials. They conducted six blocks of unique color pairs.

Participants Five hundred undergraduate students were included and were recruited via the University of Victoria psychology participant pool (see (*36*) for details). The data was collected until 500 participants became available that were not characterized by one of the following a priori criteria: trial count after artifact rejection were less than 15 per condition, total artifact rejection exceeded 40% of trials rejected, FCz (electrode of interest) specific artifact rejection exceeded 40% of trials rejected, or independent component analysis based blink correction failed. These criteria were extremely strict to ensure clean data in the analyses, and as such a total of 637 participants were analyzed before reaching the goal of 500 clean participants. All participants had normal or corrected-to-normal vision and volunteered to take part in the experiment for extra course credit in a psychology course. All participants provided informed consent approved by the University of Victoria's Human Research Ethics Board.

Data Acquisition and Preprocessing EEG data were recorded from either a 64 or 32 electrode (Ag/AgCl) EEG system (ActiCAP, Brain Products, GmbH, Munich, Germany) using Brain Vision Recorder (Version 1.21.0004, Brain Products GmbH, Munich, Germany). During recording, electrodes were referenced to a common ground, impedances were kept below 20 kΩ on average, data were sampled at 500 Hz, and an antialiasing low-pass filter of 245 Hz was applied via an ActiCHamp amplifier (Revision 2, Brain Products GmbH, Munich, Germany). Stimuli and EEG markers were temporally synced using a DataPixx stimulus synchronization unit (VPixx, Vision Science Solutions, Quebec, Canada)

Data were re-referenced to an average mastoid reference and filtered using a 0.1 to 30 Hz passband (Butterworth, order 4) and a 60 Hz notch filter. Correction for eye blinks was performed using EEGLAB's independent component analysis (ICA). Components reflective of blinks were manually identified and removed via topographic maps and component loadings, and data were reconstructed. Data were then segmented from -500 to $1,500ms$ relative to

feedback stimulus onset, baseline corrected using a -200 to $0ms$ window, and run through artifact rejection with $10\mu V/ms$ gradient and $100\mu V$ maximum–minimum criteria. Data were pre-processed to identify noisy or damaged electrodes using artifact rejection trial removal rates for each electrode.

The 1 second of recorded sequence for each ”WIN” or ”LOSE” event were extracted from recordings for each participants (with 22 ms of pre-event sample and 488 ms of post-event sample) and combined together to form separate winning and loosing datasets. Each of those datasets were processed using SPECTRE to construct the approximate inverse solution for the potential ϕ across an entire 2 mm MNI brain volume.