# **Supplemental Methods**

Canonical correlation analysis (CCA) was used as a blind source separation technique to remove broadband or electromyographic noise from single trial electroencephalographic (EEG) data, generating de-noised EEG epochs. Our approach is similar to the CCA method described by others(De Clercq et al., 2006, Ries et al., 2013) and previously applied to other studies(Kort et al., 2017, Ford et al., 2016), with some important differences. The method is based on the concepts that true EEG data tend to show high auto-correlation and exhibit power-law scaling (i.e., power is proportional to 1/frequency), but that high frequency random noise in EEG (e.g., muscle artifact, electromyographic (EMG)) tends to show low auto-correlation and violates power-law scaling (i.e., inappropriately high power at higher frequencies relative to low frequencies). The CCA de-noising procedure is performed separately for each subject on the single trial EEG epoch data. For each 3 second epoch, the (S x X) matrix containing the time series of S= 3072 EEG samples (s1, s2, s3 …s3072) at each of X=64 scalp electrodes (x1, x2, x3…x64) is subjected to a CCA with the (S x Y) matrix containing the s+1 time-lagged series of 3072 EEG samples (s2, s3, s4 …s3000, s3001, where s3073=0) at each of the same 64 electrodes (y1, y2, y3…y64). This is the multivariate equivalent of auto-regressive time series correlation. Since both the X and Y vectors each contain 64 electrodes, a total of 64 canonical correlations can be extracted. Each canonical correlation coefficient expresses the correlation of a time series of values representing the weighted sums of the X electrodes with a s+1 time series of values representing the weighted sums of the Y electrodes, with weights chosen to yield the largest canonical correlation that accounts for variance independent of the variance accounted for by all previously extracted canonical correlations. Thus, each canonical correlation coefficient has an associated time series of values that constitutes the canonical variate, X (i.e., each time point has a value that is a linear function of the canonical weights and raw data associated with the 64 electrodes), as well as a similar canonical variate, Y. The current CCA de-noising method only makes use of the set of 64 canonical X variates, one for each of the 64 extracted canonical correlations. When the time series represented by a canonical variate is subjected to a fast Fourier transformation (FFT), the resulting power spectrum can be evaluated to determine whether the canonical variate conforms to the power-law expected from EEG data, in which case it should be retained, or whether it violates the power law as would be expected for high-frequency noise (e.g., EMG contamination), in which case it should be excluded. With this approach, the retained canonical variates are those showing the strongest canonical correlations, whereas the rejected canonical variates are those showing the weakest canonical correlations. The specific criterion used to make these retain/reject decisions is where our CCA denoising approach differs from previously published approaches (e.g., (Ries et al., 2013)).

Previously described CCA denoising approaches have made decisions about which canonical variates to retain or reject by taking the ratio of high frequency (e.g., 15 to 30Hz) to low frequency (e.g., <15Hz) power, rejecting canonical variates with ratios greater than a pre-determined limit (e.g., if high/low > 1/5 in (Ries et al., 2013)). Such ratios provide a very rough heuristic for determining whether a canonical variate’s power spectrum has power-law scaling (i.e., 1/fβ or fα, where -β=α) where log-transformed power decreases linearly with increasing log-transformed frequency. Previous studies (Freeman et al., 2003, Pereda et al., 1998) have suggested that the power-law exponent, α, which corresponds to the slope of the linear regression of log-power on log-frequency, is -1 or less for human EEG, whereas white noise or EMG has an exponent of approximately zero. In order to develop a precise and conservative criterion for deciding which canonical variates to retain, we used simple linear regression to estimate α by regressing log-power on log-frequency. However, rather than running this regression once using all of the available log power and log frequency values obtained from the FFT of the canonical variate, thereby yielding a single estimate of α that could be unduly influenced by contamination from a few frequencies (e.g., 60 Hz or alpha band), a bootstrap procedure was used to generate a distribution of α estimates. Specifically, for each canonical variate and each trial epoch, the linear regression of log-power on log-frequency used to estimate α was repeated 1000 times by randomly sampling, without replacement, half of the frequency bins between 1-125Hz from the FFT. If the distribution of 1000 α estimates, which constitutes a 99.9% bootstrap confidence interval for α, contains a value of -1 or smaller, then the canonical variate was retained, whereas if the smallest value in the bootstrap distribution is greater than -1, the canonical variate is rejected. All rejected canonical variates are then algebraically removed during back-projection to the original EEG epoch space, generating a de-noised EEG epoch.

**References**

**De Clercq W, Vergult A, Vanrumste B, Van Paesschen W, Van Huffel S** (2006). Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. *Institute of Electrical and Electronics Engineers Transactions in Biomedical Engineering,* 53**,** 2583-7.

**Ford JM, Roach BJ, Palzes VA, Mathalon DH** (2016). Using concurrent EEG and fMRI to probe the state of the brain in schizophrenia. *Neuroimage Clinical,* 12**,** 429-41.

**Freeman WJ, Holmes MD, Burke BC, Vanhatalo S** (2003). Spatial spectra of scalp EEG and EMG from awake humans. *Clinical Neurophysiology,* 114**,** 1053-68.

**Kort NS, Ford JM, Roach BJ, Gunduz-Bruce H, Krystal JH, Jaeger J, Reinhart RM, Mathalon DH** (2017). Role of N-Methyl-D-Aspartate Receptors in Action-Based Predictive Coding Deficits in Schizophrenia. *Biological Psychiatry,* 81**,** 514-524.

**Pereda E, Gamundi A, Rial R, Gonzalez J** (1998). Non-linear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. *Neuroscience Letters,* 250**,** 91-4.

**Ries S, Janssen N, Burle B, Alario FX** (2013). Response-locked brain dynamics of word production. *Public Library of Science One,* 8**,** e58197.