Automated EEG mega-analysis I: Spectral and amplitude characteristics across studies

Nima Bigdely-Shamlo,[a] Jonathan Touryan,[b] Alejandro Ojeda, [a] Christian Kothe, [a] Tim Mullen,[a] Kay Robbins[c*]

[a] Intheon Labs, San Diego, CA, USA
[b] U.S. Army Research Laboratory, Aberdeen Proving Grounds, MD, USA
[c] University of Texas at San Antonio, San Antonio, USA

Abstract
Significant achievements have been made in the fMRI field by pooling statistical results from multiple studies (meta-analysis). More recently, fMRI standardization efforts have focused on enabling the combination of raw fMRI data across studies (mega-analysis), with the hope of achieving more detailed insights. However, it has not been clear if such analyses in the EEG field are possible or equally fruitful. Here we present the results of a large-scale EEG mega-analysis using 18 studies from six sites representing several different experimental paradigms. Our results show that EEG mega-analysis is possible and can provide unique insights unavailable in single studies. Standardized EEG was subjected to a fully-automated pipeline that reduces line noise, interpolates noisy channels, performs robust referencing, removes eye-activity, and further identifies outlier signals. We then define channel dispersion measures to assess the comparability of data across studies and observe the effect of various processing steps on dispersion measures. Using ICA-based dipolar sources, we also observe consistent differences in overall frequency baseline amplitudes across brain areas. For example, we observe higher alpha in posterior *vs* anterior regions and higher beta in temporal regions. We also observe consistent differences in the slope of aperiodic portion of the EEG spectrum across brain areas. This work demonstrates that EEG mega-analysis can enable investigations of brain dynamics in a more generalized fashion, opening the door for both expanded EEG mega-analysis as well as large-scale EEG meta-analysis. In a companion paper, we apply mega-analysis to assess commonalities in event-related EEG features across studies.

Key words: EEG/MEG; Meta-analysis; Mega-analysis; Signal statistics; Large-scale; Neuroinformatics

1 Introduction

Data-pooling of fMRI (functional magnetic resonance imaging) neuroimaging studies has proven extremely valuable for increasing statistical power, assessing inter-subject variability, and determining the generalizability of predictions (Costafreda, 2009). fMRI *meta-analyses* typically combine studies based on coordinates of peak activations (coordinate-based meta-analysis, CBMA) or on 3D statistical images of the activations (image-based meta-analysis, IBMA) (Salimi-Khorshidi et al., 2009) (Gorgolewski et al., 2015). The BrainMap project represents a pioneering effort to establish standardized methods for spatial normalization, referencing to common coordinates, and linking coordinates and images to associated literature (Fox and Lancaster, 2002) (Fox et al., 2014). These and other fMRI data-sharing efforts (Maumet et al., 2016) (Reid et al., 2016) have resulted in an explosion of fMRI meta-analyses (Costafreda et al., 2008) (Hart et al., 2013) (Barquero et al., 2014) that have increased understanding of basic brain function as well as provided tools for diagnosis and evaluation of treatments for mental diseases. With the availability of platforms for sharing raw data such as OpenfMRI (Poldrack et al., 2013), fMRI *mega-analyses* (pooling of raw data rather than meta-data) is also becoming possible. A recent review article by Poldrack et al. (Poldrack et al., 2017) lays out a comprehensive strategy for transparent and reproducible fMRI imaging and makes a strong case for the need for standardized automated processing, standardized benchmarks, and large-scale data sharing.

All of the arguments concerning the need for large-scale meta-analysis and mega-analysis in fMRI apply to electroencephalography (EEG) neuroimaging. However, data-sharing and standardization efforts are much less mature in the EEG arena for a number of reasons. EEG is subject to significant variability from artifacts due to subject motion, muscle activity, eye activity, headset placement, as well as recording issues such as loose detectors. Standards for preprocessing have not yet been established across the EEG community, in part because effective, automated methods for completely removing artifacts are not available. Further, EEG measurements are made through hair, scalp, and skull, so registration to brain features can only be made at a coarse level even when calibration MRI is acquired. Other concerns include variability across recording systems and variability across experimental paradigms (Melnik et al., 2017).

In addition to issues of artifact handling and standardization of preprocessing, deeper questions concerning EEG mega-analysis exist. How do the basic statistical properties of EEG data acquired from different studies vary across sites/headsets/subjects/paradigms? Can EEG data from different studies be normalized or co-registered in a meaningful way? Can insights about responses in one study be generalized to many contexts? This paper deploys a standardized processing methodology to provide preliminary answers to some of these questions. We present the results of an EEG mega-analysis of 18 studies performed at six different experimental sites. Several of these studies use the same paradigm, for example RSVP (Rapid Serial Visual Presentation) (Bigdely-Shamlo et al., 2008; Sajda et al., 2003), but differ in protocol details such as type of image, presentation rate, or whether manual response was requested when targets were detected. Other studies use tasks such as lane-keeping with vehicle perturbations in a driving simulation.

After placing the data in a standardized form and applying automated standardized preprocessing pipelines, we employ a variety of methods to obtain insights into the statistical properties of continuous EEG activity and sources of variability across studies.

The paper is organized as follows. The Methods section briefly describes the experimental data, provides an overview of the automated processing pipeline used to perform analysis, and introduces a class of robust measures to explore normalization and statistical properties of EEG across heterogeneous studies. The Results section follows a similar organization with a presentation of results on statistical differences and similarities of EEG signals across heterogeneous studies. The Discussion section offers a perspective and directions for the future.

2 Methods

2.1 Experimental data

As summarized in Table 1, this paper includes data for 18 studies from six sites associated with four institutions: the Army Research Laboratory (ARL), the National Chiao Tung University (NCTU), the Netherlands Organization for Applied Scientific Research (TNO), and the University of California at San Diego (UCSD). All studies were conducted with voluntary, fully-informed subject consent and approved by the Institutional Review Boards of the respective institutions.

Table 1: Summary of datasets used. The Rec/Sub column gives the number of recordings and the number of unique subjects within a study. Subjects may overlap across studies at the same site. Headset (B = BioSemi, N = Neuroscan) specifies headset type and number of channels.

Site	Study	Description	Rec/Sub	Headset
ARLT	Cue	Lane-keeping with random or perturb-linked cue.	33/17	B-64
ARLS	GuardA	Visual verification of IDs based on image and ID.	29/27	B-256
ARLS	GuardB	Visual verification of IDs based on image and ID.	22/21	B-256
ARLA	LKBase	Long-running lane-keeping baseline task.	130/108	B-64/256
ARLA	LKCal	Short lane-keeping task for calibration.	245/155	B-64/256
ARLT	Speed	Lane-keeping with manual and auto speed control.	61/32	B-64
ARLT	Traffic	Lane-keeping in visually complex environment.	30/29	B-64
ARLT	Mind	Lane-keeping with audio distractions.	59/20	B-64
ARLS	RSVPB	RSVP with common object images (target varies).	31/27	B-256
ARLH	RSVPC	RSVP with rare non-targets and rare targets.	51/17	B-64
ARLS	RSVPE	RSVP with common object images (target same).	59/10	B-256
ARLH	RSVPI	RSVP with visually similar targets and non-targets.	72/18	B-64
ARLH	VEP	Visual oddball task with images displayed at 0.5 Hz.	18/18	B-64
NCTU	DAS	Lane-keeping with auditory or visual distractions.	201/103	N-62
NCTU	DD	Lane-keeping with math problems.	60/15	N-30
NCTU	LKwAF	Lane-keeping with auditory feedback.	11/10	N-30
TNO	ACC	Driving with faulty adaptive cruise control.	47/15	B-64
UCSD	RSVPU	RSVP with satellite clips of London shown at 12-Hz.	14/8	B-256

Table 1 provides a brief description of each study along with basic metrics (number of recordings, channels, etc.) with more detailed information and corresponding citations appearing in Appendix A. The studies included variations of tasks that fall into two general categories: visual target detection and driving with/without motion platform, traffic, speed control, distraction, auditory feedback, or adaptive cruise control. The visual target detection tasks included simulated guard duty with ID checks (GUARDA and GUARDB), rapid serial visual presentation (RSVPB, RSVPC, RSVPE, RSVPI, RSVPU), and visual oddball (VEP). The other tasks were all based on variations of lane-keeping in a driving simulator, with the exception of ACC, which involved driving in an actual vehicle on a test track. In all, the corpus includes 1,173 EEG recordings running 633 hours. The NCTU, TNO and UCSD experiments were all performed at a single site at the respective institutions, while the ARL experiments were performed at three different sites. ARLH designates experiments performed at the Human Research and Engineering Directorate at Army Research Laboratory (Aberdeen, MD), ARLS designates experiments performed at SAIC Laboratories (Louisville, CO), and ARLT designates experiments performed at Teledyne Corporation (Durham, NC). ARLA designates tasks (LKBase and LKCal) that were performed at all three sites using identical experimental setups. Each EEG dataset was converted to an EEGLAB `EEG` structure (Delorme and Makeig, 2004) (Delorme et al., 2011) prior to analysis.

2.2 Data organization and preprocessing

As mentioned in the Introduction, standardization of EEG for data sharing is less mature than for fMRI. We have developed a standardized data format (EEG Study Schema or ESS) and an extensive collection of support tools in MATLAB to enable large-scale EEG processing (Bigdely-Shamlo et al., 2016b) (*BigEEG Workflow*, 2018). Figure 1 summarizes the analysis pipeline used in this paper, starting in the lower left corner with raw data. The individual containerized EEG recordings are converted to EEGLAB set format and the study-specific event codes are mapped to Hierarchical Event Descriptor (*HED*) strings (Bigdely-Shamlo et al., 2016a). The remainder of the process is completely automated.

Data cleaning proceeds in several stages. The first step is to apply the *PREP* pipeline to all studies (Bigdely-Shamlo et al., 2015). *PREP* is an open-source tool that we have developed to perform robust average referencing. After removing line noise from each data recording, *PREP* identifies and interpolates noisy channels to form a robust average reference. We also removed non-EEG channels from each recording and assigned 10-20 channel labels to channels that do not have standard labels (e.g., 256-channel BIOSEMI caps) based on channel distances to standard 10-20 locations (Klem et al., 1999) (Jurcak et al., 2007). After selecting a maximum of 64 channels based on the 64-channel 10-20 configuration for this analysis, we resampled the data at 128 Hz. Since all 18 studies used in this paper were acquired at rates higher than 128 Hz, resampling is actually down-sampling. We applied a 1690-point high-pass filter at 1 Hz with the *pop_eegfiltnew()* function of EEGLAB to remove low-frequency drift.

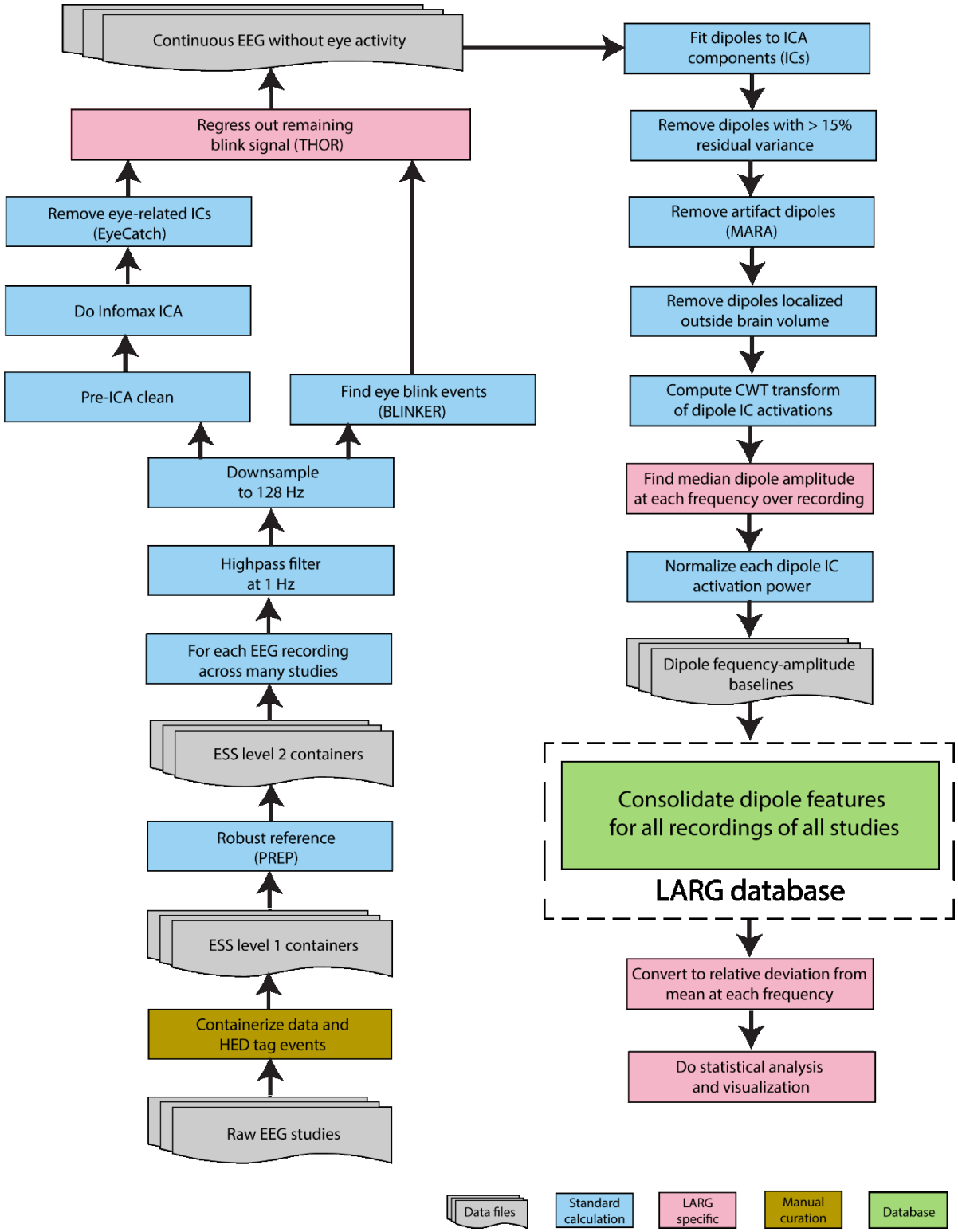

Figure 1. Pipeline for automated large-scale analysis of EEG.

We removed eye-activity from EEG using independent component analysis (ICA) followed by removal of residual blink activity using regression. After performing Infomax ICA (Bell and Sejnowski, 1995) using the portion of the data without outlier amplitudes as described in Mullen et al. (Mullen et al., 2015), we applied a modified version of *EyeCatch* (Bigdely-Shamlo et al., 2013a) to identify independent components (ICs) associated with eye activity. *EyeCatch*, which contains a database of 3,453 manually identified eye-related ICA scalp maps, flags "eye" ICs as those that either have a similar topography (correlation > 0.94) to at least one scalp map from the database or are moderately similar (correlation > 0.85) and have a power ratio of at least 100 for at least 1% of data frames. To compute the power ratio, *EyeCatch* performs a discretized continuous wavelet transform (Lilly and Olhede, 2012; Mallet, 2008) using the MATLAB *cwt()* function on "suspect" ICs and calculates the power in the range [1, 3] Hz divided by the power in the range [3, 15] Hz at each time point. We then removed the subspace spanned by the ICs identified by *EyeCatch* as containing eye activity from the channel data.

As a second prong in the effort to remove the effects of eye activity in the EEG time signal, We applied *BLINKER* (Kleifges et al., 2017) to channel data to identify latencies associated with different phases of eye blinks. *BLINKER* inserts five events into the `EEG.event` structure (*Left base*, *Left zero*, *Blink peak*, *Right zero*, *Right base*) for each blink. *BLINKER* also extracts a continuous "blink" signal that follows the blink-induced EEG activity between each pair of [*Left zero*, *Right zero*] markers and is zero outside these intervals. We then applied temporal regression using a method similar to that proposed by (Kristensen et al., 2017) to regress out both the effects of blink events and the continuous blink signal produced by *BLINKER*, assuming blinks are fixed patterns in an interval [−1, 1] seconds time-locked to the *Blink peak* events.

All of the analyses reported in this paper use data processed by the pipeline of Figure 1. Several analysis steps normalize the data by removing the median and dividing by the robust standard deviation. We refer to this normalization as a "robust z-score" data transformation. The robust standard deviation is defined as 1.4826 times the median absolute deviation (*MAD*). In this paper, we refer to the robust standard deviation of an EEG channel as the "robust channel amplitude."

2.3 Variation of channel amplitudes across scalp positions and recordings

To investigate whether EEG signals have consistently different amplitudes across channel locations and recordings, we low-pass filtered each recording at 20 Hz and selected the subset of 26 channels common across all of the recordings: Fp1, Fp2, F3, Fz, F4, F7, F8, FC3, FCz, FC4, FT7, FT8, C3, Cz, C4, TP7, TP8, CP3, CPz, CP4, P3, Pz, P4, O1, Oz, and O2. Each recording can then be represented by a single 26×1 vector of its robust channel amplitudes, consisting of the robust z-scores of the individual channels. We refer to this positive 26×1 vector as the recording's "amplitude vector". To perform analysis across R recordings, we formed the "amplitude matrix", a $26 \times R$ matrix of these column vectors.

We visualized the overall dependence of robust amplitude on channel position by applying the median function to the rows of the amplitude matrix and then plotting the resulting 26×1 vector using the EEGLAB *topoplot()* function. To visualize potential dependencies of the amplitude

vectors on study identity, we applied *t-SNE* (t-distributed stochastic neighbor embedding) (Van Der Maaten and Hinton, 2008) (Van Der Maaten, 2014) to project the columns (recordings) of the amplitude matrix to a 2-D space. We used the MATLAB *tsne()* function with a perplexity parameter of 20 for the visualizations displayed in this paper.

A known difficulty with joint analysis of EEG recordings is that the variability introduced by headset technology and sensor placement as well as subject hair, scalp and skull characteristics causes statistical properties of data to vary across recordings. A simple approach for addressing this problem is to model the variability by a recording-specific scaling factor and to "normalize" the data in each recording by removing the effect of such scaling. We investigated five (5) different ways of computing a recording-specific normalization factor: mean, median, Huber mean (Huber, 1964), Euclidean (L^2) norm, and robust standard deviation. The normalization factors, except for robust standard deviation, were computed column-wise on the amplitude matrix (across channels) as summarized in Figure 2 for the Huber mean. Unlike the other normalization methods, robust standard deviation normalization used the robust standard deviation of all values in the recording, concatenated across all channels. Regardless of the normalization method, the resulting amplitude matrix was divided column-wise by the $1 \times R$ vector of resulting recording normalization factors to produce a normalized amplitude matrix for each normalization method.

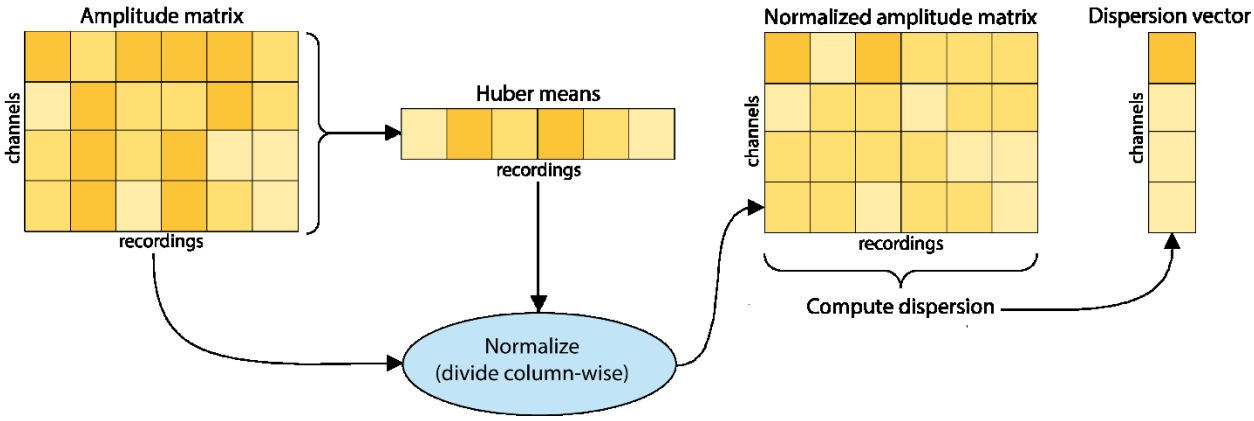

Figure 2. Steps for computing the normalized amplitude matrix and the channel dispersion vector for Huber mean normalization. A similar procedure was used for other normalization methods.

We performed several tests to assess whether normalizing each recording by a recording-specific constant could improve statistical comparability across recordings from different studies. To characterize comparability, we computed the "dispersion vector" by taking the robust standard deviation of each row (across recordings) of an amplitude matrix (after normalization) and dividing by the row median (across recordings) to obtain a 26×1 channel dispersion vector representing each method. We also computed median dispersion across all channels for each channel dispersion vector to find the normalization method that was most effective in reducing overall amplitude dispersion across recordings.

In order to find out whether the differences among these median dispersions were statistically significant for different normalization methods, we applied the pairwise Wilcoxon signed rank test

(using MATLAB *signrank()* function) between dispersion vectors associated with the best normalization method (the method associated with the least median dispersion) and the dispersion vectors associated with other normalization methods. We also plotted the channel dispersion vectors representing different methods using EEGLAB *topoplot()* function.

To investigate how much of channel amplitude variability could be explained by a recording-specific constant scaling of all channel amplitudes in each recording, we plotted channel-pair amplitudes across all recordings and computed adjusted R^2 values for linear fits of data points corresponding to each channel-pair amplitude separately. The significance of the slopes being non-zero was also computed. The same processing steps on channel-pair amplitudes were then repeated after normalizing channel amplitudes in each study by their Huber mean.

2.4 Variation of channel baseline amplitude across scalp positions and frequencies

To investigate whether EEG signals at different frequencies have consistently different amplitudes across channel locations, we analyzed continuous EEG after removal of eye activity (as output from the top left of the pipeline of Figure 1) and normalization. We selected the same subset of 26 channels used in the investigation of channel signal amplitudes and calculated the discretized continuous wavelet transform (using the MATLAB *cwt()* function) for each channel to obtain a time-varying amplitude spectrogram (square root of power of 50 frequencies logarithmically sampled between 2 and 40 Hz). The baseline spectral amplitude vector for each channel is a 50×1 vector of the median over time of the amplitudes at each of the 50 specified frequencies in the channel spectrogram. To focus on the relative amplitudes of different frequencies in each channel, we normalized each of these baseline spectral amplitude vectors to have Euclidean norm of one (1). We refer to the result as the "normalized channel spectrogram".

In order to investigate the statistical significance of the observed differences in channel amplitude at each frequency, we applied Wilcoxon signed rank tests (MATLAB *signrank()* function) on the robust z-scores of the channel spectral amplitudes for each (frequency, recording) pair. The null hypothesis of these statistical tests was that in each (frequency, recording) pair, each channel spectral amplitude could be modeled by a constant plus a random value drawn from a zero-median distribution.

2.5 Variation of channel covariance across studies

To investigate systematic variations of channel covariance across headsets, paradigms, and other study details, we calculated a single covariance matrix for each recording using the same subset of 26 channels used for the investigation of channel signal amplitudes. We used the same segments of the data that were used to compute Infomax ICA, excluding portions with abnormally high power to avoid distortion due to artifacts. We plotted the two-dimensional *t-SNE* projections of distances between recording covariance matrices using both a study/site color scheme and an EEG headset color scheme to visualize the systematic variations in covariance across recordings with respect to these factors. We used the Riemannian distance metric recommended by Förstner and Moonen (Förstner and Moonen, 2003):

$$d(\mathbf{A},\mathbf{B}) = \sqrt{\left(\sum_{i=1}^{n} ln^2(\lambda_i(\mathbf{A},\mathbf{B}))\right)}$$

Here n is the number of channels, and the λ_i are generalized eigenvalues. For comparison, we also applied *t-SNE* to the recording channel correlation matrix and to the channel standard deviation vectors (computed as the square root of the variance rather than as the robust standard deviation used in the amplitude analysis). The correlation features used Riemannian distance and the standard deviation features used Euclidean distance. We examined variations with and without recording-specific Huber normalization of channel amplitudes.

We also applied representational similarity analysis (RSA) as proposed by Kriegeskorte et al. (Kriegeskorte et al., 2008) to determine whether the patterns observed in the *t-SNE* visualizations were statistically significant. Representational analysis determines whether pairwise sample differences associated with two aspects of the data are more correlated than would be expected from chance. This significance is computed by random permutation testing.

We used the Riemannian distance $d(C_i, C_j)$ as one representation in the RSA analysis, where C_i and C_j represent the covariance matrices of recordings i and j, respectively. The second representation is a membership matrix M_{ij}, which has an entry of 1 if recording i and recording j are in the same category and 0 if recording i and recording j are not in the same category. The results presented in the paper use the membership function designating whether the recordings have the same headset/cap type.

2.6 Variation of brain sources across space and frequency

We inferred equivalent dipole locations for maximally-independent sources of EEG activity by applying the EEGLAB *dipfit* plugin to scalp maps of the ICA components (ICs) not identified as associated with the corneo-retinal potential or eye muscles during preprocessing. To further select ICs most likely to be associated with brain sources, we only considered dipoles that matched all the following criteria: (a) the scalp map explained by the dipole accounted for at least 85% of the variance, (b) the EEGLAB *MARA* plugin (Winkler et al., 2014, 2011a) did not identify the component scalp map as an artifact (e.g., muscle or eye), and (c) the dipole was located inside the brain volume, as identified by the *ft_sourcedepth()* function of the EEGLAB *fieldtrip_lite* plugin (Oostenveld et al., 2011). Note that each dipole is associated with exactly one recording.

As with analysis in channel space, we calculated the discretized continuous wavelet transform (using the MATLAB *cwt()* function) of the continuous time course of the IC activation associated with each dipole to obtain a time-varying amplitude spectrogram (square root of power of 50 frequencies logarithmically sampled between 2 and 40 Hz). The baseline spectral amplitude vector for each IC activation is a 50 × 1 vector of the median over time of the amplitudes at each of the 50 specified frequencies in the IC spectrogram. To focus on the relative amplitudes of different frequencies and to facilitate cross-recording comparison, we scaled each spectrogram so that its baseline spectral amplitude vector had Euclidean norm one. We refer to the result as the normalized spectrogram for an IC.

To investigate the differences in normalized spectrograms, we converted these values to relative deviations from means computed at each frequency across all ICs. We define the spectral amplitude deviation value at frequency f for the ith IC as:

$$D_i^f = S_i^f - \frac{1}{n}\sum_{j=1}^{n} S_j^f$$

S_i^f is the normalized spectrogram value at frequency f for IC i, and n is the number of ICs. We then analyzed the spatial distribution of these deviations for different frequency bands. Using Measure Projection Analysis (*MPA*) we computed the spatial Gaussian weighted averages (12 mm standard deviation) of IC spectral amplitude deviations for all brain voxels (Bigdely-Shamlo et al., 2013b). Averages were also computed for regions of interest (ROIs) selected from the AAL brain atlas (Tzourio-Mazoyer et al., 2002).

The power spectrum of electrophysiological signals at the channel or source level can be modeled as a combination of periodic and aperiodic components, reflecting narrowband oscillations (spectral "peaks") and approximately $1/f$ "background" activity, respectively (Onton and Makeig, 2006) (Miller et al., 2009). While the properties of periodic/oscillatory activity in various frequency bands have been extensive studied and linked to an array of cognitive and behavioral states and neurological disorders, the aperiodic component of the power spectrum has been less well-studied.

In order to characterize the spatial distribution of the aperiodic spectrum, we computed the Huber mean power spectra across all time points for each IC. We used the Huber mean, instead of median, in order to more closely estimate mean power while ignoring outliers caused by artifacts. We then used the *FOOOF* python toolbox (Haller et al., 2018) (*fooof*, 2018) to model the power spectrum, P, as a linear combination of aperiodic background activity, L, and a weighted sum of N Gaussian functions, G_n, each representing a periodic component (i.e., a "peak" in the spectrum):

$$P = L + \sum_{n=0}^{N} G_n$$

$$G_n = a_n \times exp(\frac{-(F-c)^2}{2w_n^2})$$

Each Gaussian component, G_n, is characterized by its amplitude, a_n, and RMS bandwidth, w_n. F is a vector of frequencies. The aperiodic signal L is modelled as an exponential function in semilog-power space (linear frequencies and log power values) as:

$$L = b - \log(k + F^{\chi})$$

where b is the broadband offset, χ is the slope, and k is the "knee" parameter that determines the inflection point of the aperiodic exponential function (Miller et al., 2009).

We set the maximum number of Gaussians to 10 and their width to be between 3 and 20 Hz. This provided *FOOOF* with ample degrees of freedom to model the spectrum and to separate the periodic and aperiodic activity. After fitting each dipole IC power spectrum using *FOOOF*, we removed ICs with a goodness of fit (R^2) less than 0.95. We then computed the average slope, $\bar{\chi}$, for all ICs and the deviation of each IC's slope from this average: $D_i^{\chi} = \chi_i - \bar{\chi}$.

Similar to the analysis of the spatial distribution of baseline "periodic" amplitude deviations, we analyzed the spatial distribution of aperiodic slope deviations using Measure Projection Analysis (*MPA*), computing the spatial Gaussian weighted averages (12 mm standard deviation) of these values for all brain voxels (Bigdely-Shamlo et al., 2013b).

3 Results

3.1 Division by a recording-specific constant improves recording and study comparability

To investigate whether scaling the data from each recording by a single recording-specific number improves the comparability across recordings, we looked at channel dispersion across studies for different normalization approaches and at different stages in processing. Figure 3 shows the effect on channel dispersion of different types of processing.

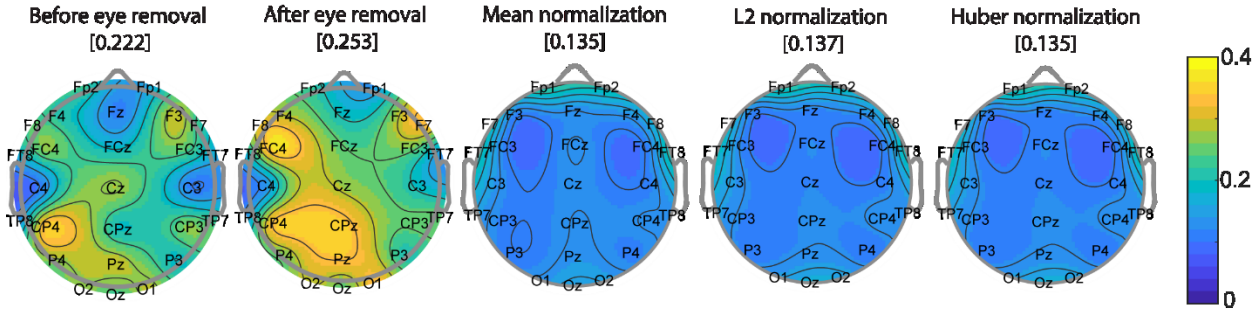

Figure 3. Effect of preprocessing and channel normalization on channel dispersion. Each scalp map shows the median over all recordings in all studies of the channel dispersion for each channel. The bracketed number over each scalp map is the overall median dispersion corresponding to the scalp map. Normalization always occurs after eye removal.

Interestingly, removal of eye artifacts changes the spatial distribution of dispersion, but does not reduce the amount of dispersion. Further, the maximum dispersion before eye artifacts are removed does not occur in the frontal channels. Note that dispersion is a dimensionless quantity and scaling a recording by multiplying by a constant does not change the dispersion. Normalization using any of the three methods (mean, L^2, or Huber) significantly reduces the channel dispersion in comparison with non-normalized recordings. Huber mean and mean are essentially the same for datasets without outliers, but Huber mean produces a slightly lower dispersion when there are outliers. The median and standard deviation methods of normalization (not shown) have slightly higher dispersion than the three methods shown in Figure 3.

Figure 4 shows the distribution of channel dispersions segregated by study before removal of eye artifacts (black box plots), after removal of eye artifacts (blue box plots), and after eye artifact removal followed by Huber normalization (green box plots). Huber normalization significantly reduces the median dispersion across studies in all cases. Normalization also reduces the number of channel dispersion outliers for most studies. Note that LKBase and LKCal were basic lane-keeping studies that included recordings acquired at three different experimental sites (HRED Army Research Laboratory in Aberdeen, MD, Teledyne Laboratories in Durham, NC, and SAIC Laboratories in Louisville, CO) using two different headset types. The channel dispersions across these tasks were well clustered and did not show a particular dependence on site location.

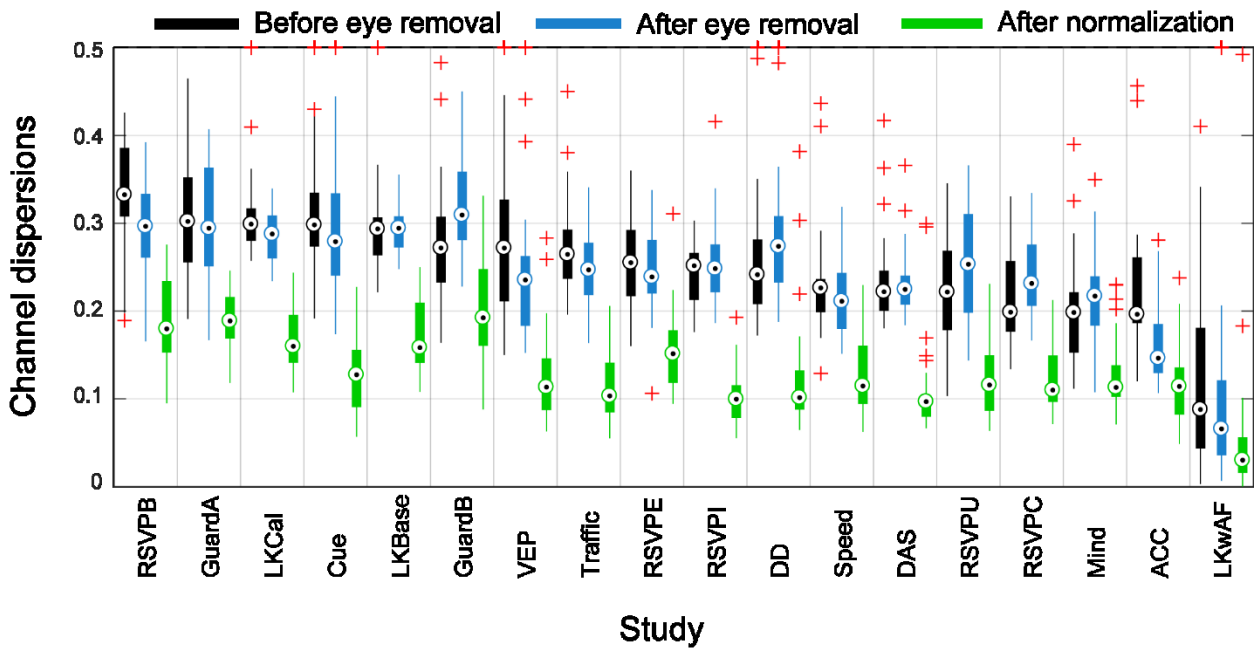

Figure 4. Channel dispersions segregated by study before eye artifact removal (black boxplots), after eye artifact removal (blue boxplots), and after Huber normalization (green boxplots). Studies are sorted by the median channel dispersion before eye removal.

3.2 Some channels consistently have a higher amplitude than others

Both removal of eye artifacts and scaling by a recording-specific constant significantly affect the relative distributions of channel amplitudes as illustrated in Figure 5. The first row of Figure 5 displays results computed on the composite channel amplitude matrix, corresponding to continuous EEG data after the 128 Hz down-sampling step in Figure 1, but before removal of eye artifacts. The middle row displays results computed on the composite channel amplitude matrix after removal of eye artifacts, corresponding to the continuous EEG without eye artifacts step in Figure 1. The bottom row of Figure 5 displays the results based on the composite channel amplitude matrix after eye artifacts have been removed and the result normalized using Huber mean.

The first column of Figure 5 shows the scalp maps of the row medians of the respective channel amplitude matrices: before eye artifact removal, after eye artifact removal, and after Huber normalization, respectively. The display uses the EEGLAB *topoplot()* function and the 26 common channels reported in the Methods section. The median robust channel amplitude is reduced after eye artifact removal (3.78 versus 4.77), and the spatial distribution of median robust channel amplitude goes from being concentrated in the frontal areas to a two-lobe form with amplitude more evenly distributed but concentrated in the frontal and occipital regions of the scalp. Notice that normalization does not markedly change the spatial distribution of the scalp map, since normalization just corresponds to a reweighting of the points contributing to the median. However, normalization does substantially reduce the overall median robust channel amplitude (to 1.02).

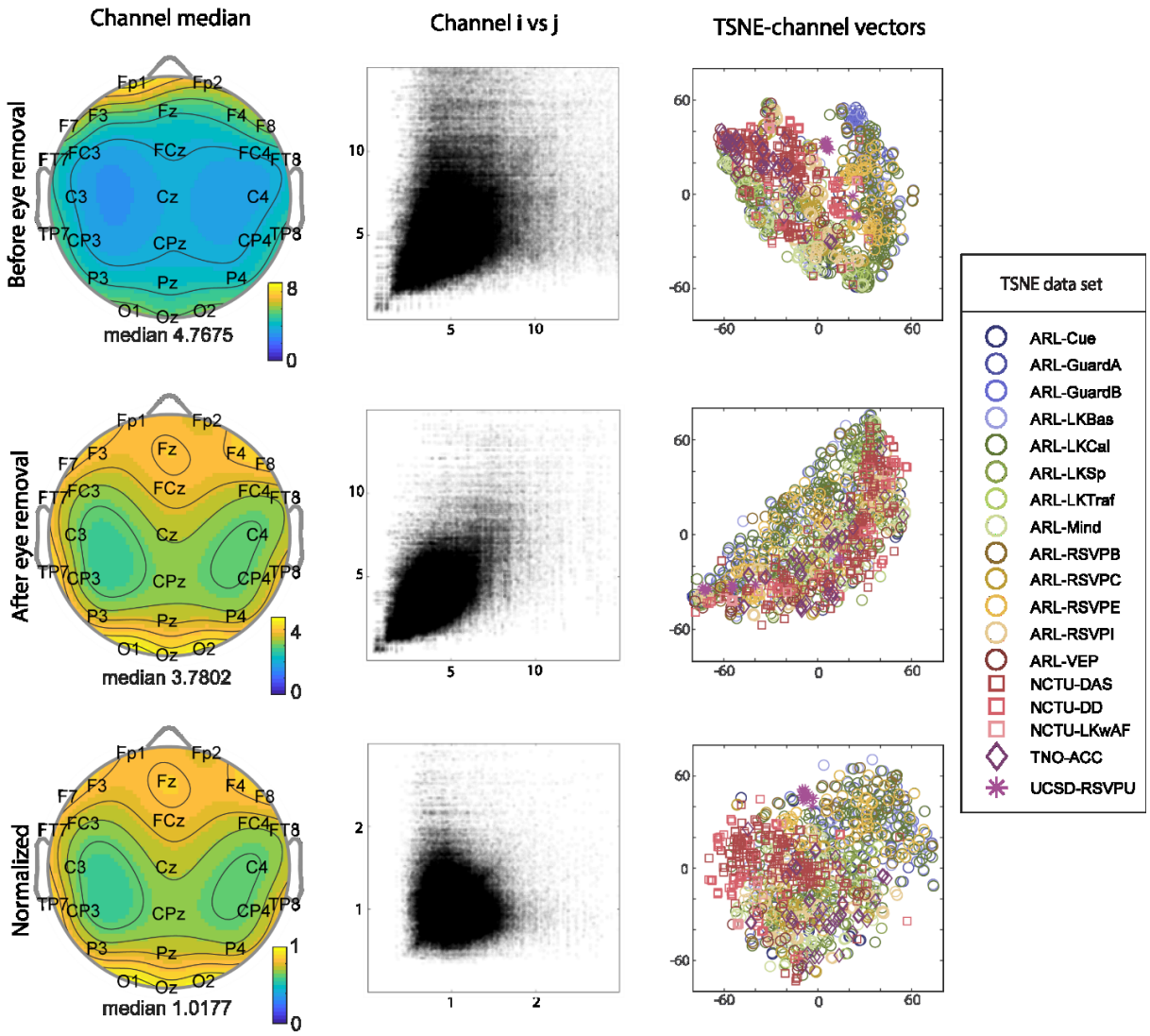

Figure 5. Composite statistics of robust channel amplitudes. Top row uses the composite channel amplitude matrix before eye artifact removal, the middle row uses the composite channel amplitude matrix after eye artifact removal, and the bottom row uses the composite channel amplitude matrix after eye artifact removal and normalization. The first column displays scalp maps of the row medians of the composite channel amplitude matrix at various stages in processing. The middle column plots the robust channel amplitude of channel i versus channel j ($j < i$) in each recording. The last column shows the two-dimensional *t-SNE* projection of recording channel amplitude vectors across all studies. The points for each study are depicted using unique colors, with point shape designating the institution (with the three ARL sites combined).

Statistical testing (Wilcoxon signed rank) of robust channel amplitudes for the 26 common channels found that, before eye artifact removal all channels had robust amplitudes that differed significantly ($p < 0.01$, *FDR* corrected) from the median of the robust channel amplitude in each

recording. After eye artifact removal and after normalization, only channel Fp2 was not significantly different from the recording median.

The middle column of Figure 5 plots robust amplitude of channel i versus channel j ($j < i$ to eliminate duplicates) in each recording before removal of eye artifacts (top row), after removal of eye artifacts (middle row), and after Huber normalization (bottom row). The top graph is somewhat asymmetric with respect to the 45° line, reflecting the amplitude dominance of the frontal channels. After eye artifacts are removed, the relationship becomes more symmetric, with a visible linear tendency, suggesting the existence of an overall recording-dependent scaling factor for robust channel amplitude. After dividing by the recording Huber mean normalization factor, this linear dependence on amplitude appears to be essentially eliminated. The median adjusted R-squared value, resulting from fitting a linear regression model to all channel amplitude pairs, is 0.413 before eye artifact removal, 0.463 after eye artifact removal, and 0.023 after normalization. In over 99% of the linear fits after eye artifact removal, the p value for the slope factor is significantly nonzero ($p < 0.01$, *FDR* corrected). This linear relationship, which explains about half of the variability in channel pair amplitudes, almost fully disappears after Huber mean normalization.

The rightmost column of Figure 5 shows the *t-SNE* projections of the recording amplitude vectors into a two-dimensional plane. Although the area occupied by the projections is more circular after normalization, the points are highly overlapping both before and after Huber normalization. The top row, which corresponds to the recordings before removal of eye artifacts shows more segregation by study, emphasizing the importance of consistent eye-artifact removal in cross-study analysis.

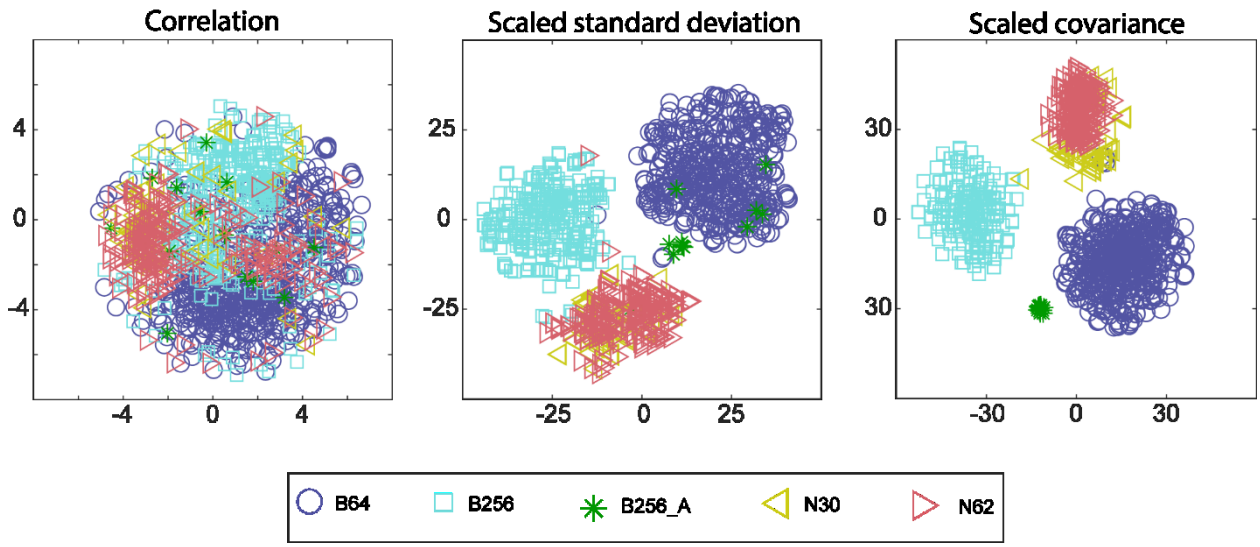

Figure 6. Two-dimensional *t-SNE* visualization of selected vectorized recording signal functions. Left plot shows recording channel correlation vectors using Riemannian distance. Center plot shows recording channel standard deviation vectors divided by Huber mean of the vector using Euclidean distance. Right plot shows recording channel covariance matrix divided by Huber mean of the variance using Riemannian distance. Points are colored by headset type: B64 designates a Biosemi 64-channel headset, B256 designates a Biosemi 256-channel headset, B256_A designates the UCSD Biosemi 256-channel headset with a custom cap, N30 designates a Neuroscan 30-channel headset, and N62 designates a Neuroscan 62-channel headset.

3.3 Channel covariance exhibits a strong dependence on EEG cap

To understand whether there are systematic variations in channel covariance across studies, we vectorized recording covariance matrices (using the 26 common channels defined in previous sections) and applied *t-SNE* to project these vectors into a two-dimensional plane.

Figure 6 shows the *t-SNE* projections plotted with headset type labeled using a unique symbol-color pairing. The N30 (gold) and N62 (red) clusters correspond to three studies (DAS, DD, LKwAF) performed at NCTU using Neuroscan headsets in two electrode configurations. These points correspond to different tasks and many different subjects. The B256_A cluster, which uses a custom cap and an older Biosemi 256-channel headset, corresponds a single study (RSVPU) of 14 recordings from 8 subjects performing an RSVP task. The B256 cluster corresponds to data from studies conducted at ARLS using a 256-channel headset. The points in this cluster represent a number of subjects and several studies (GuardA, GuardB, RSVPB and RSVPE), as well as a subset of LKBase and LKCal recordings. All of these studies were conducted at a single site.

The points in the B64 cluster correspond to data using a Biosemi 64-channel headset with a standard cap of various sizes acquired at three different institutions: TNO, ARLH, and ARLT. Some of the studies in this group were performed at solely at a single ARL site (Cue, Speed, Traffic, Mind, RSVPC, RSVPI, and VEP). ACC was only performed at TNO. Note that the LKBase and LKCal studies were conducted at three different sites using identical protocols and

equipment, but two-different headset types. The points corresponding to these studies fall into two different clusters, depending on the headset type used.

The plots show a very strong relationship between headset type and recording covariance (rightmost plot) scaled by a recording-specific constant. A similar relationship appears when the covariance is not scaled, but the clusters are slightly less well-defined. A headset relationship also appears in the *t-SNE* visualizations of the recording channel correlations (left plot) and the scaled channel standard deviation (center plot).

To quantify the strength and statistical significance of these relationships, we applied Representational Similarity Analysis (RSA) to several different transformations of the recording channel signals, including the vectorized channel correlation, channel covariance, and channel standard deviation. The RSA question we investigated was whether membership of two recordings in the same headset class implied that their signal behavior was more likely to be similar than would be expected if the recordings were assigned to a membership class at random.

Table 2 presents the results for different signal functions, different metrics, and different methods of normalization. All of the results of the table were significant with $p < 0.0005$. A higher RSA correlation means that a signal function has a stronger relationship with headset class. Huber mean normalization eliminates the effect of uniform scaling in channel amplitudes, focusing on relative changes in amplitude.

Table 2: RSA results for statistical significance of the relationship of several channel signal characteristics (correlation, covariance, and standard deviation) and the membership function of being of the same headset type. All results were statistically significant $p < 0.0005$.

Signal function	Base metric	Normalization	RSA correlation
Correlation	Riemannian		0.02
Covariance	Riemannian		0.02
Covariance	Riemannian	Huber mean covariance	0.65
Standard deviation	Euclidean		0.17
Standard deviation	Euclidean	Huber mean amplitude	0.34

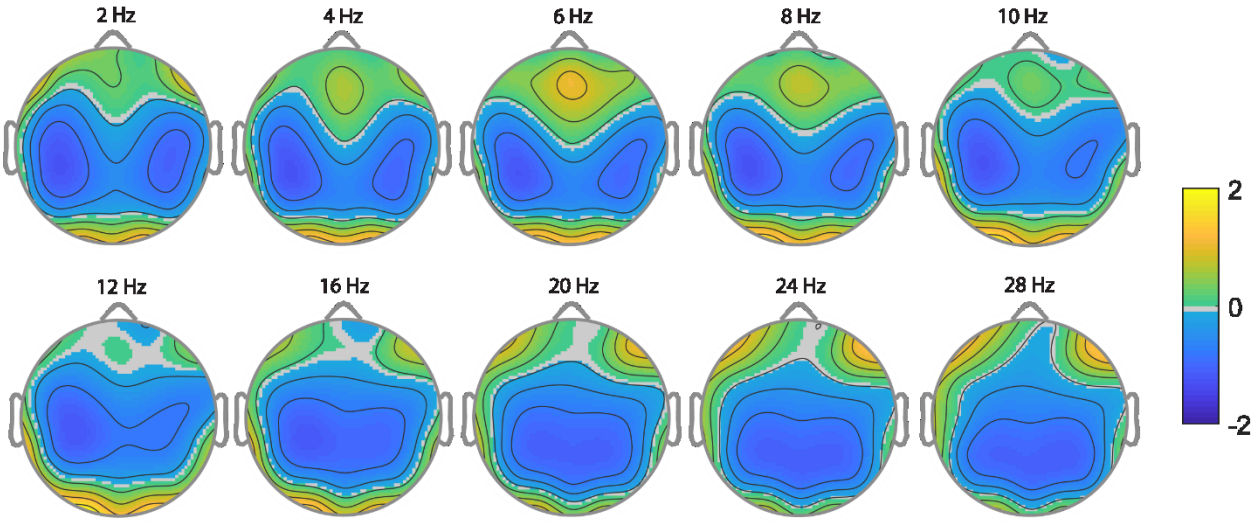

Figure 7. Median normalized channel amplitudes across recordings from all studies at different frequencies, masked by $p < 0.01$, *FDR* corrected after removal of eye artifacts and normalization by a recording-specific constant.

3.3 Some channels consistently have higher amplitude at certain frequencies

Statistical testing (Wilcoxon signed rank) also shows that channel amplitudes consistently vary at different frequencies. Figure 7 shows normalized channel amplitudes at different frequencies, each masked by $p < 0.01$ (*FDR* corrected). The amplitude patterns between 4 and 8 Hz resemble the median channel amplitude plots in Figure 5 after eye activity removal, suggesting that overall channel amplitude patterns are highly influenced by amplitude differences in the 4 to 8 Hz range. Note that the channel amplitude vector before eye removal more closely resembles the spatial distribution observed in the 15 to 30 Hz region of Figure 7.

3.4 Some brain areas consistently have higher power in certain frequencies

To investigate whether there are systematic variations in amplitude associated with particular frequency bands and brain areas, we applied measure projection analysis (MPA) to the spectrogram deviations calculated from the IC in-brain dipoles across the corpus and mapped the results to brain areas using the AAL atlas. The left column of Figure 8 shows three-dimensional views of brain areas with high or low local spectral amplitude deviation averages in different frequency ranges: delta ([2, 4] Hz), theta ([4, 8] Hz), alpha ([8, 14] Hz) and beta ([14, 31] Hz).. Brain areas where the average local amplitude deviation, obtained via 12 mm Gaussian smoothing, within the specified frequency band was at least 5% are shown in red. Regions where the amplitude deviation was at most −5% are coded in blue.

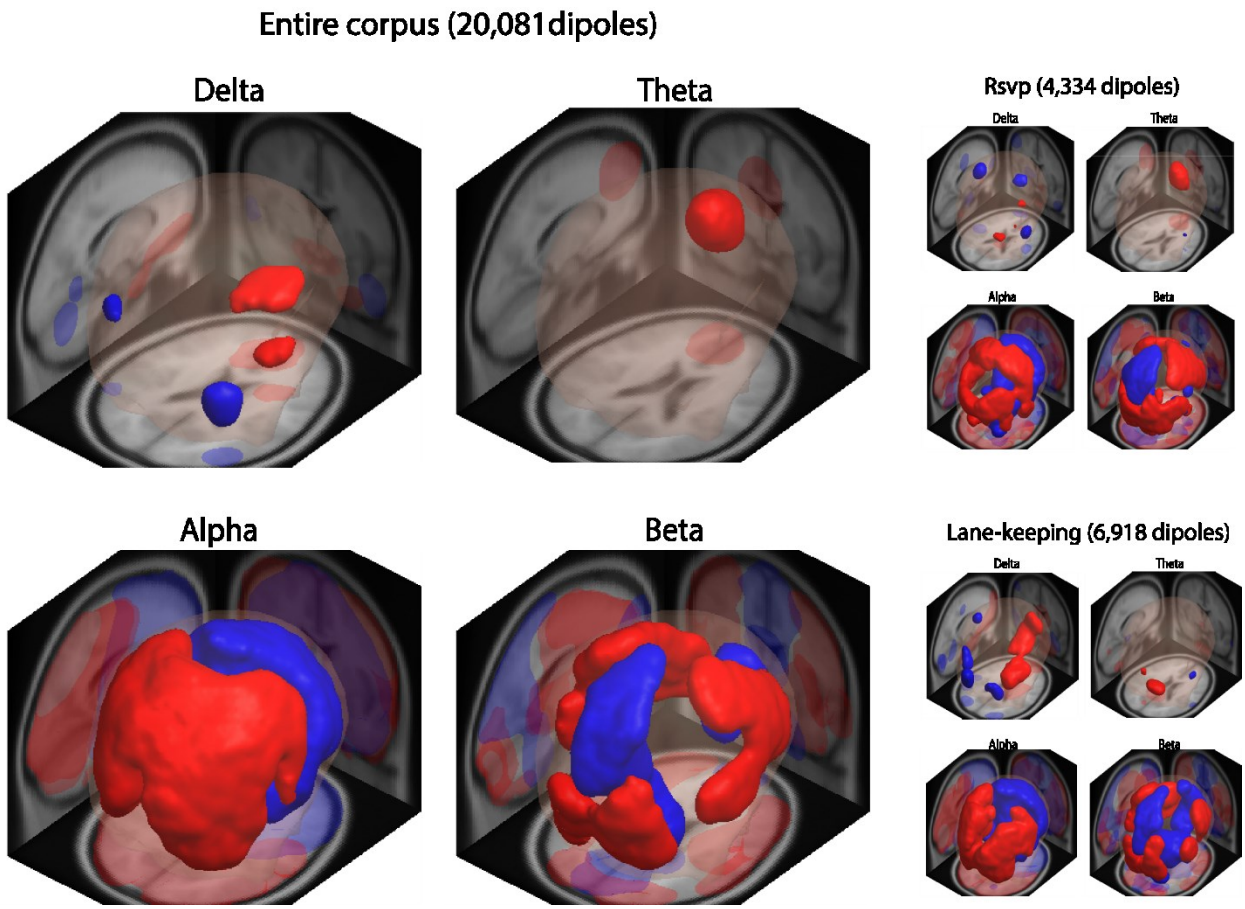

Figure 8. Brain regions with local average (12 mm spatial Gaussian smoothing) spectral amplitude deviation over 5% (red) and less than −5% (blue) for common frequency bands. The left column shows results from the entire corpus, while the right column splits out the results for two distinct groups within the corpus – one group (*Rsvp* group) consists of the five Rsvp studies, and the other (*Lane-keeping* group) consists of data from LKBase and LKCal.

These 3D views show a clear anterior/posterior dichotomy in alpha amplitudes: posterior brain areas (in red) generally have more alpha amplitude than anterior areas (in blue). Beta activity

appears to be increased in sensorimotor, middle frontal, and temporal regions. We also note increased activity in inferior occipital and cerebellar regions, which may be consistent with neck muscle activity (Gramann et al., 2010). Beta activity is reduced along central midline structures, including anterior cingulate, superior frontal, superior parietal, posterior cingulate, and medial occipital regions. Increased frontal theta activity in anterior cingulate and superior frontal gyrus is also prominent.

We investigated whether these differences are statistically significant, as shown in Figure 9. The top row of the figure shows transverse brain slides overlaid with dipole density images. The dipole density was computed using Gaussian spatial smoothing with a standard deviation of 12 mm. While dorsal and midline brain regions have a higher dipole density, dipoles are distributed throughout the brain volume.

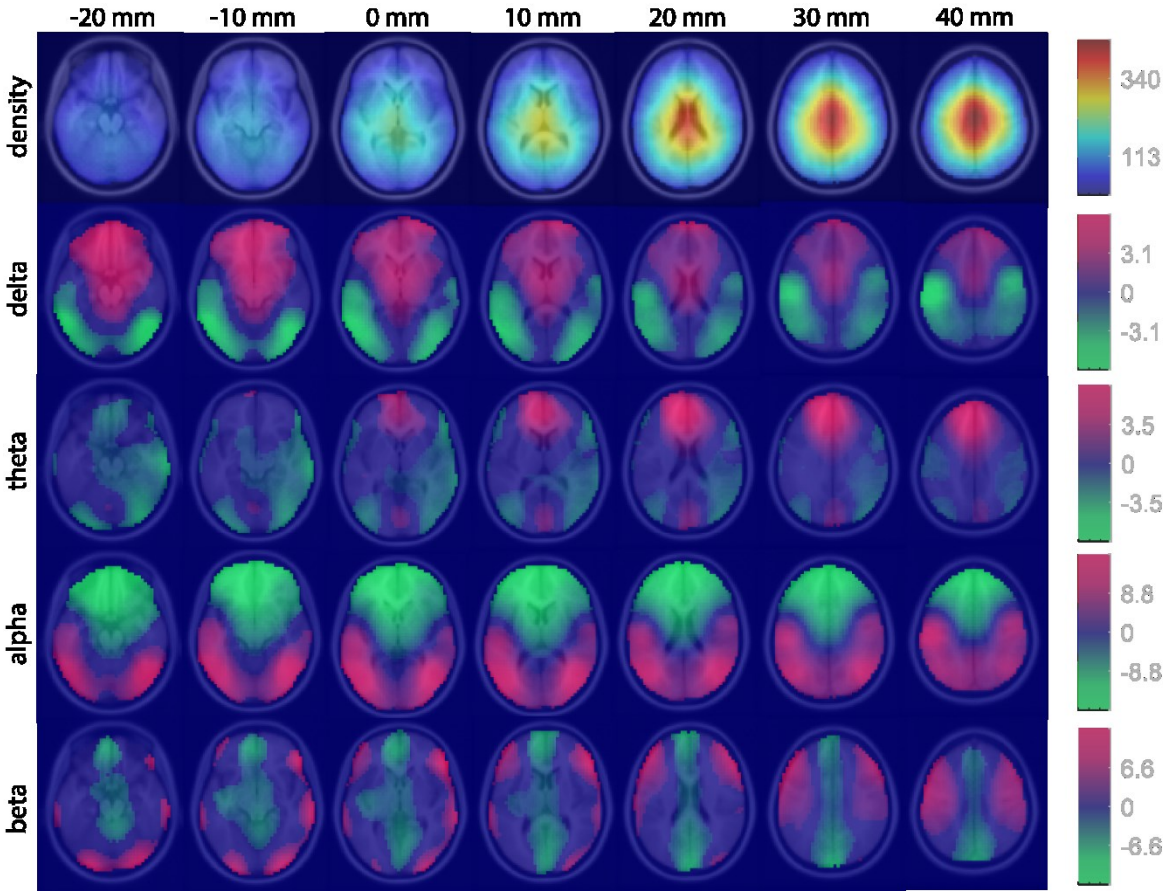

Figure 9. Statistically significant variations of spectral amplitude in common frequency bands. Dipole density plot using 12 mm spatial Gaussian smoothing (top row). Spatially smoothed spectral amplitude deviation percentage for different frequency bands (rows 2-5). Magenta is associated with higher-than-average baseline amplitude at that frequency and green is associated with lower than average baseline amplitude. The images are masked for significance ($p < 0.01$, *FDR*-corrected).

The bottom four rows of Figure 9 show slice displays of the spatially smoothed (12 mm Gaussian kernel) amplitude deviation distribution in standard frequency bands. These values are masked by significance ($p < 0.01$, *FDR* corrected), computed by permuting spectral amplitude deviations in each frequency band across dipoles (400 permutations, allowing substitution). The slice views reflect the distinctive patterns shown in the 3D visualizations of Figure 8.

As in Figure 8, increased beta activity is observed in temporal, lateral frontal, and sensorimotor regions as well as inferior occipital and cerebellar regions, while activity is decreased along the central midline and cingulate structures. Alpha is increased in posterior regions and decreased in anterior regions, while theta is increased in superior frontal and anterior cingulate regions. Modulations in frontal theta and posterior alpha have been well studied and associated with a number of cognitive processes (Ward, 2003). Specifically, decreases in posterior alpha power have been linked to increasing demands of attention, alertness, and task loading. In contrast, frontal theta power tends to increase during the encoding phase of memory tasks (Klimesch, 1999). Conversely to alpha, we note increased anterior and decreased posterior activity in the delta band.

The statistical tests represented in Figure 9 examine whether, for each frequency range, the local amplitude deviation average at a given voxel has a value that is significantly higher or lower than zero. We also performed the same analysis for each study individually and noticed that although the basic patterns were common to most studies, there appeared to be some variations between the studies focusing on visual target detection and those focusing on driving activities.

To examine this variation more carefully, we analyzed two sub-groups that together comprised a large subset of the overall corpus, an *RSVP* group and *Lane-keeping* group. The *RSVP* group (RSVPB, RSVPC, RSVPE, RSVPI, and RSVPU) consisted of studies with different types of visual stimulus presentation within the RSVP paradigm. The recordings in this subset consisted of periods of rapid visual stimulus presentation interspersed with frequent resting periods to alleviate subject fatigue. The *Lane-keeping* group (LKBase and LKCal studies) contained relatively few interspersed events. The right column of Figure 8 shows the spectral amplitude deviations across different frequency bands for these two groups. Although some of the general characteristics of the two groups are similar, there appear to be some differences in detail, particularly for alpha in occipital regions, and for beta in the temporal regions. To test whether these spatial differences were statistically significant, we performed permutation tests to assess whether, at each voxel, the local average spectral amplitude deviation for the *RSVP* group minus that of the *Lane-keeping* group was significantly different from zero. The differences were not significant after controlling for multiple comparisons using *FDR*, indicating a very weak effect, if any.

3.5 Some brain areas consistently have higher or lower aperiodic slope

To investigate whether there are systematic variations in aperiodic slope, we applied measure projection analysis (MPA) to the aperiodic slope deviations calculated from the in-brain dipolar ICs that met the log spectral linear fit criterion (15,055 out of 20,081 dipoles). The average aperiodic slope was 1.1237 +/− 0.0022, a highly significant deviation ($p < 10^{-10}$) from pink noise ($1/f^{\alpha}$, $\alpha \approx 1$).

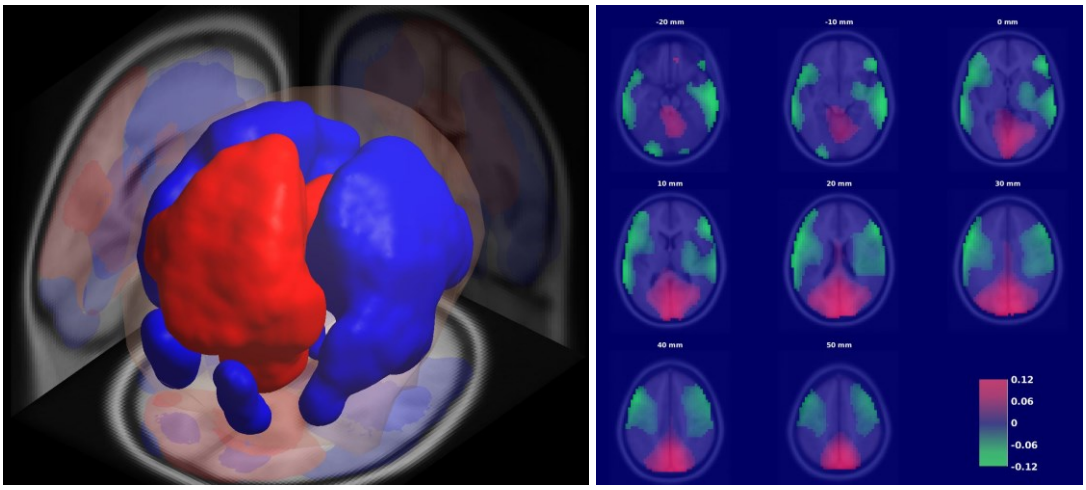

Figure 10. Statistically significant variations of local aperiodic slope. Left shows three-dimensional view of brain areas with significantly ($p < 0.05$, *FDR*-corrected) high (red) or low (blue) local aperiodic slope deviations. Right shows local average (using 12 mm spatial Gaussian smoothing) aperiodic slope deviations masked by significance.

Figure 10 shows areas with statistically significant variations of local aperiodic slope. The left plot shows a three-dimensional view of brain areas with significantly ($p < 0.05$, *FDR*-corrected) high (red) or low (blue) local aperiodic slope deviations. The right plot shows local average aperiodic slope deviations masked by significance. Both left and right temporal areas, plus some occipital areas have lower than average aperiodic slope while posterior midline areas have higher than average slope. The average slope (1.1237 +/− 0.0022) reported here for dipolar sources is consistent with the slope for EEG channel data (1.33 +/− 0.19) reported by Dehghani et al. (Dehghani et al., 2010). These authors also report lower slopes in parieto-temporal regions, overlapping with the areas of reduced slope shown in Figure 10.

4 Discussion and conclusions

The results of this paper were based on 1,173 recordings representing 633 hours of EEG collected in 18 studies from 6 different institutions on 3 continents using a number of different paradigms. We demonstrated the viability of mega-analysis for EEG and presented a variety of methods for EEG mega-analysis. We used a particular automated pipeline to preprocess the corpus in a standardized way, but the focus of the work is not to establish one preprocessing approach as preferable over another. Rather, the goal of this paper was to show that given a reasonable preprocessing strategy, EEG mega-analysis was possible and informative. We believe that several automated preprocessing packages such as MARA (Winkler et al., 2011b), AutoReject (Jas et al., 2017), HAPPE (Gabard-Durnam et al., 2018), or BEAPP (Levin et al., 2018) might be effectively employed in this regard. However, the results of this paper reinforce the importance of effective removal of eye-artifacts before results from multiple recordings are combined.

We have introduced several simple, robust measures for ascertaining whether a given recording is an outlier, whether the study as a whole has expected behavior, and whether the process for removing eye artifacts has been effective. A scalp map plot of a recording's amplitude vector (left column of Figure 5) provides a simple test of whether a recording is an outlier and whether the eye-artifact removal process has been effective. A simple view of the study as a whole can be achieved by plotting a scalp map of the row median of the study channel amplitude matrix at different stages of processing (left column of Figure 5). The scalp map of the study channel dispersion vector (Figure 3) plotted at various stages of processing can identify data issues and bad channels across a study. Box plots of study channel dispersion (Figure 4) can also provide useful information about outlier channels and unexpected behavior.

We also showed that simply normalizing each recording by a recording-specific constant reduced (almost halved) channel amplitude variability across studies. Our results illustrate the importance of normalization when attempting cross-study analysis for neuroscientific inquiry or Brain-Computer Interface (BCI) applications. Further, the results emphasized the importance of aggressively removing eye activity from the signal before making comparisons across recordings.

This paper also demonstrated consistent differences in power spectral amplitude within several frequency bands for particular channels and brain regions. Theta increased in superior frontal and anterior cingulate regions. Alpha increased in posterior (occipital, temporal, parietal) regions and decreased in frontal cortex, while delta showed the opposite pattern. Beta increased in temporal, middle frontal and sensorimotor regions and decreased in central midline structures, including superior frontal, superior parietal, posterior cingulate, and medial occipital regions. Although many of these differences were already known qualitatively (reported in a number of separate studies, each using a small sample), our quantitative analysis produces a more comprehensive image of these differences and can inform models of whole-brain oscillatory dynamics. Importantly, since we did not subtract evoked brain activity, e.g. ERPs, from the data before computing the spectrum, the spatial distributions of amplitude deviations at different frequency bands, both for channels and dipolar ICs, may not only reflect ongoing activity but also the prevalence of event-related evoked activity in different brain areas and its spectral characteristics.

But the fact that we did not find statistically significant differences in spectral deviations between *Rsvp* and *Lane-keeping* corpora suggests that task-related contributions to recording-average IC spectral amplitudes are relatively small in our data, compared to other sources of variability, e.g. subject differences. The lane-keeping tasks are monotonous driving tasks during which prominent resting state network activity has been observed (Lin et al., 2016). Although the RSVP tasks had periods of rapid visual stimulus presentation, these periods were also frequently interspersed with periods of rest. Also RSVP presentation rates differed greatly across studies (0.5, 2, 5, 10 or 12 Hz), significantly reducing the frequency-domain effect of visually evoked potentials when spectral averages are computed over all RSVP studies.

By exploiting a large number of dipoles across studies we reduced the effect of random errors in source localization, but systemic errors cannot be reduced by simply including more data. We modelled brain sources with ICA-derived single equivalent dipoles and used standard EEG sensor locations and brain templates to localize them. The accuracy of such modelling approaches is a subject of ongoing discussion and investigation (Delorme et al., 2012) (Akalin Acar et al., 2016). We are currently exploring the use of distributed source analysis methods such as LCMV Beamforming (Van Veen et al., 1997), variants of LORETA (Low Resolution Electromagnetic Tomography) (Pascual-Marqui, 2002), and Sparse Bayesian Learning (Wipf and Nagarajan, 2009) in EEG mega-analysis to address issues with single equivalent dipole modeling assumptions. Additionally, source localization accuracy could be improved by utilizing individual head models derived from MRI along with digitizing EEG sensor locations in each session (Akalin Acar and Makeig, 2013).

Improved accuracy of electrode location information may also reduce differences in channel covariance statistics across headsets, e.g. by interpolating EEG data into more accurate common anatomical positions. Meta-analysis of fMRI has relied on consistent mapping of responses to a common brain atlas. Large-scale studies of the properties of neural sources inferred by various techniques might pave the way for similar analyses in EEG. This could be done, e.g. by discovering different sources of variability and methods to “standardize” EEG features before aggregation by mitigating variability factors irrelevant to the hypothesis being investigated, a subject of our ongoing and future work.

Many EEG methods for classification, source, and connectivity estimation make assumptions about or directly make use of the channel covariance matrix. Our investigation of covariance variability points to systematic differences in covariance caused by EEG headsets. This highlights the importance of data normalization for BCI transfer learning applications and large-scale analysis. A straightforward way to deal with covariance differences is to whiten the data in each recording, as a preprocessing step, using its own covariance matrix. This has been shown to improve the similarity of regressed ERPs across study recordings and improve cross-subject BCI performance (Bigdely-Shamlo et al., Accepted)

Also, Table 2 shows that normalizing channel data by channel Huber amplitude increases the strength of the relationship between channel-derived features (covariance and amplitude) and EEG headset factor, as indicated by RSA correlation. This suggests that differences in EEG headsets primarily affect relative channel amplitudes and are not the main contributing factor to uniform scaling of channel amplitudes. Future work in this area could explore this hypothesis and potential underlying electromechanical reasons for this relationship, e.g,. a particular EEG headset design may apply more mechanical force on a subset of EEG sensors, resulting in a better electrical coupling with scalp.

The study of aperiodic slope in EEG is relatively new. The aperiodic spectrum has been shown to change with task (Podvalny et al., 2015) and aging (Voytek et al., 2015) (Dave et al., 2018). It is also hypothesized to be associated with tonic differences in excitation/inhibition balance (Voytek and Knight, 2015) (Gao et al., 2017). In a four-subject comparison of EEG and MEG, Dehghani et al. (Dehghani et al., 2010) found spatial variations of the EEG aperiodic spectrum at the channel level that are consistent with the variations found in this study. Brain areas with smaller slope exponent, α, seem to have a large overlap with areas exhibiting significantly higher beta, and vice versa. This relationship suggests the possibility of a common underlying factor inversely relating beta power to aperiodic slope and warrants further investigation.

EEG mega-analysis is a nascent subfield and this work is an initial demonstration of the power and potential for this approach. The automated nature of these analyses and the assembly of a large corpus of data will permit more systematic investigation of brain dynamics associated with cognitive phenomena, along with an assessment of generalizability of results obtained from single studies or paradigms. In a companion paper, we apply mega-analysis to assess commonalities in event-related EEG features across studies, while ongoing and future work aims to extend these analyses to investigate distributed cortical source activity and large-scale brain connectivity.

Acknowledgments

The authors would like to thank Scott Makeig for inspiring this work. We would like to acknowledge Tony Johnson and Michael Dunkel of DCS Corporation for their careful assembly and curation of the ARL data. We would also like to acknowledge Ching-Teng Lin and Jung-Tai King of NCTU and the other experimenters who contributed their data to this effort. This work received computational support from UTSA's HPC cluster Shamu, operated by the Office of Information Technology. This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-10-2-0022 (CAST 076910227001). The views and the conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S Government. The U.S Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

Akalin Acar, Z., Acar, C.E., Makeig, S., 2016. Simultaneous head tissue conductivity and EEG source location estimation. NeuroImage 124, 168–180. https://doi.org/10.1016/j.neuroimage.2015.08.032

Akalin Acar, Z., Makeig, S., 2013. Effects of forward model errors on EEG source localization. Brain Topogr. 26, 378–396. https://doi.org/10.1007/s10548-012-0274-6

Barquero, L.A., Davis, N., Cutting, L.E., 2014. Neuroimaging of reading intervention: A systematic review and activation likelihood estimate meta-analysis. PLOS ONE 9, e83668. https://doi.org/10.1371/journal.pone.0083668

Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. Neural Comput. 7, 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129

Bigdely-Shamlo, N., Cockfield, J., Makeig, S., Rognon, T., La Valle, C., Miyakoshi, M., Robbins, K.A., 2016a. Hierarchical Event Descriptors (HED): Semi-structured tagging for real-world events in large-scale EEG. Front. Neuroinformatics 10. https://doi.org/10.3389/fninf.2016.00042

Bigdely-Shamlo, N., Ibagon, G., Kothe, C., Mullen, T., Accepted. Finding the optimal cross-subject EEG data alignment method for analysis and BCI. Presented at the 2018 IEEE SMC, Miyazaki, Japan.

Bigdely-Shamlo, N., Kreutz-Delgado, K., Kothe, C., Makeig, S., 2013a. EyeCatch: Data-mining over half a million EEG independent components to construct a fully-automated eye-component detector, in: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Presented at the 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5845–5848. https://doi.org/10.1109/EMBC.2013.6610881

Bigdely-Shamlo, N., Makeig, S., Robbins, K.A., 2016b. Preparing laboratory and Real-World EEG data for large-scale analysis: A Containerized approach. Front. Neuroinformatics 10. https://doi.org/10.3389/fninf.2016.00007

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., Robbins, K.A., 2015. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. Front. Neuroinformatics 9. https://doi.org/10.3389/fninf.2015.00016

Bigdely-Shamlo, N., Mullen, T., Kreutz-Delgado, K., Makeig, S., 2013b. Measure projection analysis: A probabilistic approach to EEG source comparison and multi-subject inference. NeuroImage 72, 287–303. https://doi.org/10.1016/j.neuroimage.2013.01.040

Bigdely-Shamlo, N., Vankov, A., Ramirez, R.R., Makeig, S., 2008. Brain activity-based image classification from rapid serial visual presentation. IEEE Trans. Neural Syst. Rehabil. Eng. 16, 432–441. https://doi.org/10.1109/TNSRE.2008.2003381

BigEEG Workflow, 2018.

Brooks, J., Kerick, S., 2015. Event-related alpha perturbations related to the scaling of steering wheel corrections. Physiol. Behav. 149, 287–293. https://doi.org/10.1016/j.physbeh.2015.05.026

Brooks, J.R., Kerick, S.E., McDowell, K., 2015. Novel measure of driver and vehicle interaction demonstrates transient changes related to alerting., Novel Measure of Driver and Vehicle

Interaction Demonstrates Transient Changes Related to Alerting. J. Mot. Behav. J. Mot. Behav. 47, 47, 106, 106–116. https://doi.org/10.1080/00222895.2014.959887, 10.1080/00222895.2014.959887

Brouwer, A.-M., Snelting, A., Jaswa, M., Flascher, O., Krol, L., Zander, T., 2017. Physiological effects of adaptive cruise control behaviour in real driving, in: Proceedings of the 2017 ACM Workshop on An Application-Oriented Approach to BCI out of the Laboratory, BCIforReal '17. ACM, New York, NY, USA, pp. 15–19. https://doi.org/10.1145/3038439.3038441

Costafreda, S.G., 2009. Pooling fMRI data: meta-analysis, mega-analysis and multi-center studies. Front. Neuroinformatics 3. https://doi.org/10.3389/neuro.11.033.2009

Costafreda, S.G., Brammer, M.J., David, A.S., Fu, C.H.Y., 2008. Predictors of amygdala activation during the processing of emotional stimuli: A meta-analysis of 385 PET and fMRI studies. Brain Res. Rev. 58, 57–70. https://doi.org/10.1016/j.brainresrev.2007.10.012

Dave, S., Brothers, T.A., Swaab, T.Y., 2018. 1/f neural noise and electrophysiological indices of contextual prediction in aging. Brain Res. 1691, 34–43. https://doi.org/10.1016/j.brainres.2018.04.007

Dehghani, N., Bédard, C., Cash, S.S., Halgren, E., Destexhe, A., 2010. Comparative power spectral analysis of simultaneous elecroencephalographic and magnetoencephalographic recordings in humans suggests non-resistive extracellular media. J. Comput. Neurosci. 29, 405–421. https://doi.org/10.1007/s10827-010-0263-2

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Mullen, T., Kothe, C., Akalin Acar, Z., Bigdely-Shamlo, N., Vankov, A., Makeig, S., 2011. EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing [WWW Document]. Comput. Intell. Neurosci. https://doi.org/10.1155/2011/130714

Delorme, A., Palmer, J., Onton, J., Oostenveld, R., Makeig, S., 2012. Independent EEG sources are dipolar. PLOS ONE 7, e30135. https://doi.org/10.1371/journal.pone.0030135

fooof: Parameterizing neural power spectra - fitting oscillatory peaks and 1/f background, https://github.com/voytekresearch/fooof., 2018. . Not a Polish bear.

Förstner, W., Moonen, B., 2003. A metric for covariance matrices, in: Geodesy-The Challenge of the 3rd Millennium. Springer, Berlin, Heidelberg, pp. 299–309. https://doi.org/10.1007/978-3-662-05296-9_31

Fox, P.T., Lancaster, J.L., 2002. Mapping context and content: the BrainMap model. Nat. Rev. Neurosci. 3, 319–321. https://doi.org/10.1038/nrn789

Fox, P.T., Lancaster, J.L., Laird, A.R., Eickhoff, S.B., 2014. Meta-analysis in human neuroimaging: Computational modeling of large-scale databases. Annu. Rev. Neurosci. 37, 409–434. https://doi.org/10.1146/annurev-neuro-062012-170320

Gabard-Durnam, L.J., Leal, M., S, A., Wilkinson, C.L., Levin, A.R., 2018. The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data. Front. Neurosci. 12. https://doi.org/10.3389/fnins.2018.00097

Gao, R., Peterson, E.J., Voytek, B., 2017. Inferring synaptic excitation/inhibition balance from field potentials. NeuroImage 158, 70–78. https://doi.org/10.1016/j.neuroimage.2017.06.078

Garcia, J.O., Brooks, J., Kerick, S., Johnson, T., Mullen, T.R., Vettel, J.M., 2017. Estimating direction in brain-behavior interactions: Proactive and reactive brain states in driving. NeuroImage 150, 239–249. https://doi.org/10.1016/j.neuroimage.2017.02.057

Gorgolewski, K.J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S.S., Maumet, C., Sochat, V.V., Nichols, T.E., Poldrack, R.A., Poline, J.-B., Yarkoni, T., Margulies, D.S., 2015. NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. Front. Neuroinformatics 9. https://doi.org/10.3389/fninf.2015.00008

Gramann, K., Gwin, J.T., Bigdely-Shamlo, N., Ferris, D.P., Makeig, S., 2010. Visual evoked responses during standing and walking. Front. Hum. Neurosci. 4. https://doi.org/10.3389/fnhum.2010.00202

Hairston, W.D., Whitaker, K.W., Ries, A.J., Vettel, J.M., Bradford, J.C., Kerick, S.E., McDowell, K., 2014. Usability of four commercially-oriented EEG systems. J. Neural Eng. 11, 046018. https://doi.org/10.1088/1741-2560/11/4/046018

Haller, M., Donoghue, T., Peterson, E., Varma, P., Sebastian, P., Gao, R., Noto, T., Knight, R.T., Shestyuk, A., Voytek, B., 2018. Parameterizing neural power spectra. bioRxiv 299859. https://doi.org/10.1101/299859

Hart, H., Radua, J., Nakao, T., Mataix-Cols, D., Rubia, K., 2013. Meta-analysis of functional magnetic Resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. JAMA Psychiatry 70, 185–198. https://doi.org/10.1001/jamapsychiatry.2013.277

Huang, K.-C., Huang, T.-Y., Chuang, C.-H., King, J.-T., Wang, Y.-K., Lin, C.-T., Jung, T.-P., 2016. An EEG-based fatigue detection and mitigation system. Int. J. Neural Syst. 26, 1650018. https://doi.org/10.1142/S0129065716500180

Huber, P.J., 1964. Robust estimation of a location parameter. Ann. Math. Stat. 35, 73–101. https://doi.org/10.2307/2238020

Jas, M., Engemann, D.A., Bekhti, Y., Raimondo, F., Gramfort, A., 2017. Autoreject: Automated artifact rejection for MEG and EEG data. NeuroImage 159, 417–429. https://doi.org/10.1016/j.neuroimage.2017.06.030

Jurcak, V., Tsuzuki, D., Dan, I., 2007. 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. NeuroImage 34, 1600–1611. https://doi.org/10.1016/j.neuroimage.2006.09.024

Kleifges, K., Bigdely-Shamlo, N., Kerick, S.E., Robbins, K.A., 2017. BLINKER: Automated extraction of ocular indices from EEG enabling large-scale analysis. Front. Neurosci. 11. https://doi.org/10.3389/fnins.2017.00012

Klem, G.H., Lüders, H.O., Jasper, H.H., Elger, C., 1999. The ten-twenty electrode system of the International Federation. The International Federation of Clinical Neurophysiology. Electroencephalogr. Clin. Neurophysiol. Suppl. 52, 3–6.

Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res. Rev. 29, 169–195. https://doi.org/10.1016/S0165-0173(98)00056-3

Kriegeskorte, N., Mur, M., Bandettini, P., 2008. Representational similarity analysis – connecting the branches of systems neuroscience. Front. Syst. Neurosci. 2. https://doi.org/10.3389/neuro.06.004.2008

Kristensen, E., Guerin-Dugué, A., Rivet, B., 2017. Regularization and a general linear model for event-related potential estimation. Behav. Res. Methods 1–20. https://doi.org/10.3758/s13428-017-0856-z

Levin, A.R., Leal, M., S, A., Gabard-Durnam, L.J., O'Leary, H.M., 2018. BEAPP: The Batch Electroencephalography Automated Processing Platform. Front. Neurosci. 12. https://doi.org/10.3389/fnins.2018.00513

Lilly, J.M., Olhede, S.C., 2012. Generalized Morse wavelets as a superfamily of analytic wavelets. IEEE Trans. Signal Process. 60, 6036–6041.

Lin, C.-T., Chuang, C.-H., Kerick, S., Mullen, T., Jung, T.-P., Ko, L.-W., Chen, S.-A., King, J.-T., McDowell, K., 2016. Mind-wandering tends to occur under low perceptual demands during driving. Sci. Rep. 6, 21353. https://doi.org/10.1038/srep21353

Lin, C.-T., Huang, K.-C., Chao, C.-F., Chen, J.-A., Chiu, T.-W., Ko, L.-W., Jung, T.-P., 2010. Tonic and phasic EEG and behavioral changes induced by arousing feedback. NeuroImage 52, 633–642. https://doi.org/10.1016/j.neuroimage.2010.04.250

Mallet, S., 2008. A Wavelet Tour of Signal Processing, 3rd ed. Academic Press.

Marathe, A.R., Ries, A.J., Lawhern, V.J., Lance, B.J., Touryan, J., McDowell, K., Cecotti, H., 2015. The effect of target and non-target similarity on neural classification performance: a boost from confidence. Front. Neurosci. 9. https://doi.org/10.3389/fnins.2015.00270

Marathe, A.R., Ries, A.J., McDowell, K., 2014. Sliding HDCA: Single-trial EEG cassification to overcome and quantify temporal variability. IEEE Trans. Neural Syst. Rehabil. Eng. 22, 201–211. https://doi.org/10.1109/TNSRE.2014.2304884

Maumet, C., Auer, T., Bowring, A., Chen, G., Das, S., Flandin, G., Ghosh, S., Glatard, T., Gorgolewski, K.J., Helmer, K.G., Jenkinson, M., Keator, D.B., Nichols, B.N., Poline, J.-B., Reynolds, R., Sochat, V., Turner, J., Nichols, T.E., 2016. Sharing brain mapping statistical results with the neuroimaging data model. Sci. Data 3. https://doi.org/10.1038/sdata.2016.102

Melnik, A., Legkov, P., Izdebski, K., Kärcher, S.M., Hairston, W.D., Ferris, D.P., König, P., 2017. Systems, subjects, sessions: To what extent do these factors influence EEG data? Front. Hum. Neurosci. 11. https://doi.org/10.3389/fnhum.2017.00150

Miller, K.J., Sorensen, L.B., Ojemann, J.G., Nijs, M. den, 2009. Power-law scaling in the brain surface electric potential. PLOS Comput. Biol. 5, e1000609. https://doi.org/10.1371/journal.pcbi.1000609

Mullen, T.R., Kothe, C.A.E., Chi, Y.M., Ojeda, A., Kerth, T., Makeig, S., Jung, T.P., Cauwenberghs, G., 2015. Real-time neuroimaging and cognitive monitoring using wearable dry EEG. IEEE Trans. Biomed. Eng. 62, 2553–2567. https://doi.org/10.1109/TBME.2015.2481482

Onton, J., Makeig, S., 2006. Information-based modeling of event-related brain dynamics, in: Neuper, C., Klimesch, W. (Eds.), Progress in Brain Research, Event-Related Dynamics of Brain Oscillations. Elsevier, pp. 99–120. https://doi.org/10.1016/S0079-6123(06)59007-7

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data [WWW Document]. Comput. Intell. Neurosci. https://doi.org/10.1155/2011/156869

Pascual-Marqui, R.D., 2002. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. Methods Find. Exp. Clin. Pharmacol. 24 Suppl D, 5–12.

Podvalny, E., Noy, N., Harel, M., Bickel, S., Chechik, G., Schroeder, C.E., Mehta, A.D., Tsodyks, M., Malach, R., 2015. A unifying principle underlying the extracellular field potential spectral responses in the human cortex. J. Neurophysiol. 114, 505–519. https://doi.org/10.1152/jn.00943.2014

Poldrack, R.A., Baker, C.I., Durnez, J., Gorgolewski, K.J., Matthews, P.M., Munafò, M.R., Nichols, T.E., Poline, J.-B., Vul, E., Yarkoni, T., 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. Nat. Rev. Neurosci. 18, 115–126. https://doi.org/10.1038/nrn.2016.167

Poldrack, R.A., Barch, D.M., Mitchell, J., Wager, T., Wagner, A.D., Devlin, J.T., Cumba, C., Koyejo, O., Milham, M., 2013. Toward open sharing of task-based fMRI data: the OpenfMRI project. Front. Neuroinformatics 7. https://doi.org/10.3389/fninf.2013.00012

Reid, A.T., Bzdok, D., Genon, S., Langner, R., Müller, V.I., Eickhoff, C.R., Hoffstaedter, F., Cieslik, E.-C., Fox, P.T., Laird, A.R., Amunts, K., Caspers, S., Eickhoff, S.B., 2016. ANIMA: A data-sharing initiative for neuroimaging meta-analyses. NeuroImage, Sharing the wealth: Brain Imaging Repositories in 2015 124, 1245–1253. https://doi.org/10.1016/j.neuroimage.2015.07.060

Robbins, K., Su, K., Hairston, W.D., 2018. An 18-subject EEG data collection using a visual-oddball task, designed for benchmarking algorithms and headset performance comparisons. Data Brief 16, 227–230. https://doi.org/10.1016/j.dib.2017.11.032

Sajda, P., Gerson, A., Parra, L., 2003. High-throughput image search via single-trial event detection in a rapid serial visual presentation task, in: First International IEEE EMBS Conference on Neural Engineering, 2003. Conference Proceedings. Presented at the First International IEEE EMBS Conference on Neural Engineering, 2003. Conference Proceedings., pp. 7–10. https://doi.org/10.1109/CNE.2003.1196297

Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D., Nichols, T.E., 2009. Meta-analysis of neuroimaging data: A comparison of image-based and coordinate-based pooling of studies. NeuroImage 45, 810–823. https://doi.org/10.1016/j.neuroimage.2008.12.039

Touryan, J., Apker, G., Lance, B.J., Kerick, S.E., Ries, A.J., McDowell, K., 2014. Estimating endogenous changes in task performance from EEG. Front. Neurosci. 8. https://doi.org/10.3389/fnins.2014.00155

Touryan, J., Lance, B.J., Kerick, S.E., Ries, A.J., McDowell, K., 2016. Common EEG features for behavioral estimation in disparate, real-world tasks. Biol. Psychol. 114, 93–107. https://doi.org/10.1016/j.biopsycho.2015.12.009

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15, 273–289. https://doi.org/10.1006/nimg.2001.0978

Van Der Maaten, L., 2014. Accelerating t-SNE Using Tree-based Algorithms. J Mach Learn Res 15, 3221–3245.

Van Der Maaten, L., Hinton, G., 2008. Visualizing data using t-SNE. J Mach Learn Res 9, 2579–2605.

Van Veen, B.D., van Drongelen, W., Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Trans. Biomed. Eng. 44, 867–880. https://doi.org/10.1109/10.623056

Voytek, B., Knight, R.T., 2015. Dynamic network communication as a unifying neural basis for cognition, development, aging, and disease. Biol. Psychiatry, Cortical Oscillations for Cognitive/Circuit Dysfunction in Psychiatric Disorders 77, 1089–1097. https://doi.org/10.1016/j.biopsych.2015.04.016

Voytek, B., Kramer, M.A., Case, J., Lepage, K.Q., Tempesta, Z.R., Knight, R.T., Gazzaley, A., 2015. Age-related changes in 1/f neural electrophysiological noise. J. Neurosci. 35, 13257–13265. https://doi.org/10.1523/JNEUROSCI.2332-14.2015

Wang, Y.K., Jung, T.P., Lin, C.T., 2015. EEG-based attention tracking during distracted driving. IEEE Trans. Neural Syst. Rehabil. Eng. 23, 1085–1094. https://doi.org/10.1109/TNSRE.2015.2415520

Ward, L.M., 2003. Synchronous neural oscillations and cognitive processes. Trends Cogn. Sci. 7, 553–559. https://doi.org/10.1016/j.tics.2003.10.012

Winkler, I., Brandl, S., Horn, F., Waldburger, E., Allefeld, C., Tangermann, M., 2014. Robust artifactual independent component classification for BCI practitioners. J. Neural Eng. 11, 035013. https://doi.org/10.1088/1741-2560/11/3/035013

Winkler, I., Haufe, S., Tangermann, M., 2011a. Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals. Behav. Brain Funct. 7, 30. https://doi.org/10.1186/1744-9081-7-30

Winkler, I., Haufe, S., Tangermann, M., 2011b. Automatic classification of artifactual ICA-components for artifact Removal in EEG signals. Behav. Brain Funct. 7, 30. https://doi.org/10.1186/1744-9081-7-30

Wipf, D., Nagarajan, S., 2009. A unified Bayesian framework for MEG/EEG source imaging. NeuroImage 44, 947–966. https://doi.org/10.1016/j.neuroimage.2008.02.059

Appendix A

This appendix provides additional details about the 18 studies used in this paper with references and IRB information. Table A.1 provides some study summary statistics.

Table A.1: Additional summary statistics for the 18 studies used in this paper.

Dataset	Site	Set	Rec	Subjs	Events	Codes	Mins	Dipoles	
								Clean	Brain
Cue	ARLT	B-64	33	17	297142	33	1525	618	446
GuardA	ARLS	B-256	29	27	211378	37	1461	706	580
GuardB	ARLS	B-256	22	21	169845	31	1173	368	314
LKBase	ARLH	B-64	21	21	1061315	35	7919	2703	2167
	ARLS	B-256	81	59					
	ARLT	B-64	28	28					
LKCal	ARLH	B-64	21	21	518199	27	3921	6044	4751
	ARLS	B-256	127	59					
	ARLT	B-64	97	75					
Speed	ARLT	B-64	61	32	442085	26	2697	1185	881
Traffic	ARLT	B-64	30	29	177937	44	1389	605	478
Mind	ARLT	B-64	59	20	291979	50	1813	906	594
RSVPB	ARLS	B-256	31	27	683108	38	1970	725	611
RSVPC	ARLH	B-64	51	17	82366	64	392	1338	1059
RSVPE	ARLS	B-256	59	10	1292141	38	3652	1228	1027
RSVPI	ARLH	B-64	72	18	218098	45	1080	1927	1601
VEP	ARLH	B-64	18	18	31987	14	169	557	435
DAS	NCTU	N-62	201	103	1219933	67	5551	4256	3671
DD	NCTU	N-30	60	15	203240	26	1006	924	605
LKwAF	NCTU	N-30	11	10	95363	14	880	183	143
ACC	TNO	B-64	47	15	322528	57	1269	933	682
RSVPU	UCSD	B-256	14	8	484194	16	980	138	36
Total			**1173**		**7802838**	**662**	**38849**	**25344**	**20081**

A.1 ARL Auditory Cueing Task (ARL-Cue)

The ARL Auditory Cueing Task involved continuous driving using a driving simulator with brake and foot-pedals but no motion platform. Subjects were instructed to maintain the constant, posted speed and a position in the center of the lane. The vehicle was subjected to perturbations to the left or right at unpredictable times (lane-keeping). In task condition A, a tone was played at random times, unrelated to perturbation events. In task condition B, the tone was played within a 2-second window prior to the perturbation. The two task conditions were counter-balanced among participants and each lasted for 45 minutes. Directly prior to the auditory cueing task, each participant performed 15 minutes of the Calibration Driving Task (A.5). This study was conducted by Teledyne Laboratories in Durham, NC using a BioSemi 64-channel headset (Garcia et al., 2017). The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 12-040) before the study began.

A.2 ARL Advanced Guard Duty Task (ARL-GuardA)

The ARL Advanced Guard Duty Task (Touryan et al., 2016) involved the visual verification of IDs based on eight information fields on an ID card and a corresponding photograph. Participants were asked to determine if the individual in the image, paired with the corresponding ID card, should have access to a specified restricted area. Some of the ID cards were valid and some were not (e.g., expiration date passed, incorrect access area, or photos did not match). Participants were instructed to press either an "allow" or "deny" button for each image-ID pairing. The two-alternative forced-choice response was self-paced with a maximum time limit of 20 seconds. If the participants chose to deny access, they were subsequently asked to provide a reason. Reasons for denied access were selected from a numerical list of five options: incorrect access, expired ID, suspicious DOB, face mismatch, no watermark. If the participant did not respond within the allotted time, the computer forced a "deny" decision. The task was divided into ten blocks of five minutes each. The rate of ID presentation as well as the criteria for validating IDs varied among blocks. Data included 27 sessions from 27 unique subjects. This study was conducted by SAIC (Louisville, CO) using a BioSemi 256-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 12-041) before the study began.

A.3 ARL Baseline Guard Duty Task (ARL-GuardB)

The ARL Baseline Guard Duty Task was similar to the ARL Advanced Guard Duty Task. Rate of ID presentation, but not the criteria for verification varied among blocks. Data included 21 sessions from 21 unique subjects. This study was conducted by SAIC (Louisville, CO) using a BioSemi 256-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 12-041) before the study began.

A.4 ARL Baseline Driving Task (ARL-LKBase)

The ARL Baseline Driving Task was conducted at three different sites using an identical experimental apparatus that included a driving simulator with steering wheel and brake/foot pedals but no motion platform (Touryan et al., 2014) (Brooks et al., 2015) (Brooks and Kerick, 2015) (Garcia et al., 2017). The Baseline Driving Task included either 45 minutes of continuous driving or 60 minutes of driving organized into 10-minute blocks. The subject was instructed to maintain a center lane position and speed while the vehicle underwent lateral perturbations at random intervals. Experiments conducted at ARL HRED (Aberdeen MD) and at Teledyne (Durham, NC) used a BioSemi 64-channel headset. Experiments conducted at SAIC (Louisville, CO) used a BioSemi 256-channel headset. The site protocols were approved by the Army Human Research Protections Office (Protocol ARL-20098-10051, ARL 12-040, and ARL 12-041) before the studies began.

A.5 ARL Calibration Driving Task (ARL-LKCal)

The ARL Calibration Driving Task consisted of a 15-minute period of continuous driving in which the vehicle was subjected to random lateral perturbations. This task was typically performed at the beginning of experiments that included other, subsequent tasks. Like the Baseline Driving Task, these experiments were conducted at three different sites using an identical experimental apparatus that included a driving simulator with steering wheel and brake/foot pedals but no motion platform (Touryan et al., 2014) (Brooks et al., 2015) (Brooks and Kerick, 2015) (Garcia et al., 2017). The ARL HRED site (Aberdeen MD) and the Teledyne site (Durham, NC) used BioSemi 64-channel headsets. The SAIC site (Louisville, CO) used a BioSemi 256-channel headset. The sites protocols were approved by the Army Human Research Protections Office (Protocol ARL-20098-10051, ARL 12-040, and ARL 12-041) before the studies began.

A.6 ARL Lane-keeping with Speed Control Task (ARL-Speed)

The ARL Lane-keeping with Speed Control Task, which is similar to the ARL Baseline Driving Task, consisted of 45 minutes of continuous driving in a driving simulator while maintaining the posted speed limit. The driving simulator included a steering wheel and brake/foot pedals but no motion platform (Touryan et al., 2014) (Brooks et al., 2015) (Brooks and Kerick, 2015) (Garcia et al., 2017). Subjects were instructed to maintain a constant specified speed and stay in the center of the lane while the vehicle underwent random lateral perturbations. In condition A, speed was controlled by the simulator (i.e., cruise control) while in condition B, speed was controlled by the subject. The Speed Control Task was always counter-balanced with a 45-minute Baseline Driving Task following a 15-minute Calibration Driving Task. The study was conducted at Teledyne Corporation (Durham, NC) using a BioSemi 64-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 12-040) before the study began.

A.7 ARL Lane-keeping with Traffic Complexity Task (ARL-Traffic)

The ARL Lane-keeping with Traffic Complexity Task, which is similar to the ARL Baseline Driving Task, consisted of 45 minutes of continuous driving in a driving simulator with steering wheel and brake/foot pedals but no motion platform (Touryan et al., 2014) (Brooks et al., 2015) (Brooks and Kerick, 2015) (Garcia et al., 2017). The visual environment consisted of a long, straight highway with visual complexity derived from vehicle traffic, including oncoming traffic and traffic in the direction of travel (in the passing lane). The experiment also included pedestrians on either side of the road, but not crossing the road. Subjects were instructed to maintain a constant specified speed and stay in the center of the lane while the vehicle underwent random lateral perturbations. The 45-minute Traffic Complexity Task was always counter-balanced with a 45-minute Baseline Driving Task following a 15-minute Calibration Driving Task. The study was conducted by Teledyne Corporation (Durham, NC) using a BioSemi 64-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 12-040) before the study began.

A.8 ARL Lane-keeping with Auditory Mind-wandering Distraction Task (ARL-Mind)

The ARL Lane-keeping with Auditory Mind-wandering Distraction Task was similar to the ARL Baseline Driving Task with additional features (Garcia et al., 2017). The experiment was performed using the same driving simulator that included a steering wheel and brake/foot pedals but no motion platform (Touryan et al., 2014) (Brooks et al., 2015) (Brooks and Kerick, 2015). The subject was instructed to maintain lane position and a specified constant speed as the vehicle was subjected to random lateral perturbations. The environment included a mix of regular traffic and police vehicles. The subject was instructed to press a button when spotting a police vehicle. The experiment consisted of three counter-balanced 30-minute blocks corresponding to conditions A, B, and C, respectively. During condition A, the subject listened to an audio podcast on traffic safety; during condition B, the subject listened to a sports-related audio podcast; and in condition C the subject listened to an audio podcast on mediation. The study was conducted by Teledyne Corporation (Durham, NC) using a BioSemi 64-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 12-040) before the study began.

A.9 ARL RSVP Baseline Task (ARL-RSVPB)

The ARL RSVP Baseline Task (Touryan et al., 2014) was a rapid serial visual presentation (RSVP) task utilizing natural photographs containing common objects (chairs, containers, doors, posters, and stairs). A different object was the target for each 10 minute block with a 2-minute break between blocks (6 blocks in total). Images were presented at approximately 5 Hz, and the subject was instructed to press a button each time a target image was perceived. This study was conducted by SAIC (Louisville, CO) using a BioSemi 256-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 12-041) before the study began.

A.10 ARL RSVP Cognitive Technology Threat Warning System Task (ARL-RSVPC)

The ARL RSVP Cognitive Technology Threat Warning System Task (Marathe et al., 2014), which was designed to detect neural responses during multitasking, incorporated three simultaneous tasks on different displays (RSVP, target detection, and formation deviation). The center display had a primary RSVP task that presented video clips consisting of five consecutive images, each 100 ms in duration. There was no interval between videos so that the first frame was presented immediately after the last frame of the prior video. Targets included vehicles, animals, and humans, both moving and stationary with varied backgrounds. After receiving a target detection cue, the subject initiates the target detection task on the left display by pressing the “Start” button. The subject sees a set of 20 images selected from the previous RSVP for further evaluation and performs a more detailed assessment by labeling each image as “Target” or “No Target”. After receiving a formation deviation cue, the subject evaluates (vehicle) formation deviation information and indicates the appropriate corrective action to be taken. The cue was presented on the central display (condition 1) or one of the secondary displays (condition 2). The experiment was conducted by ARL HRED (Aberdeen, MD) using a BioSemi 64-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 20098-10025) before the study began.

A.11 ARL RSVP Expertise Task (ARL-RSVPE)

ARL RSVP Expertise Task (Touryan et al., 2014) used a paradigm similar to the ARL RSVP Baseline Task with images of common objects (chairs, containers, doors, posters, and stairs). A different object was the target for each session. The subject was instructed to press a button each time a target image was perceived. Each subject competed 5 sessions, run on subsequent days. Images were presented at approximately 5 Hz. The experiment was conducted by SAIC (Louisville, CO) using a Biosemi 256-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 12-041) before the study began.

A.12 ARL RSVP Insurgent-Civilian Task (ARL-RSVPI)

The ARL RSVP Insurgent-Civilian Task (Marathe et al., 2015) was an RSVP study with visually similar targets and non-targets presented at a rate of 2 Hz. Target images contained one or more persons holding guns (target), while non-target images contained scenes without people or with people who did not have guns. The experiment was conducted under 4 different RSVP conditions: 1) presentation of only targets with subject counting, 2) presentation of only targets with subject counting and pressing a button for each, 3) presentation of targets and non-targets with subjects counting targets, and 4) presentation of targets and non-targets with subjects both counting and pressing a button for targets. Subjects also performed a baseline condition in which they listed to tones, blinked their eyes in time to audio tones, fixated on a cross, and closed their eyes, each for one minute. The experiment was conducted by ARL HRED (Aberdeen, MD) using a BioSemi 64-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 20098-09021) before the study began.

A.13 ARL Visually Evoked Potential Task (ARL-VEP)

The ARL Visually Evoked Potential Task (Hairston et al., 2014) was part of a larger study comparing neural responses for the same subjects on four different commercial headsets and several distinct paradigms. In the VEP task, 18 subjects performed a visual oddball task where target (enemy combatants) and non-target (US soldiers) images were presented at a rate of 0.5 Hz with roughly 1 out of 7 images being targets. The experiment conducted by ARL HRED (Aberdeen, MD) using a BioSemi 64-channel headset. This dataset has been used in a number of papers as summarized in a recent Data In Brief (Robbins et al., 2018). The dataset has also been made publicly available on NITRC (https://www.nitrc.org/projects/vep_eeg_raw/). The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 14-042) before the study began.

A.14 NCTU Dynamic Attention Shifting Task (NCTU-DAS)

The NCTU Dynamic Attention Shifting Task was a lateral perturbation lane-keeping task that was performed on a motion platform and included auditory or visual distractions. On some perturbations, subjects received an attend cue (auditory or visual) along with a direction (left or right). Subjects then received a stimuli (spoken or written word). If the word was a target (animal) subjects pressed the left or right button (corresponding to stimulus position). Subjects heard a "ding" for correctly identified targets and a "buzz" in all other cases. Data was acquired using a Neuroscan 62-channel headset. The study was approved (VFHIRB no.: 2013-01-029BC) by the Taipei Veterans General Hospital.

A.15 NCTU Distracted Driving Task (NCTU-DD)

The NCTU Distracted Driving Task (Wang et al., 2015) was a lateral perturbation lane-keeping task performed on a motion platform combined with visual math problems intermittently presented on the dashboard to mimic distracted driving. Subjects pressed a button on the right side of dashboard if the presented equation was correct, and a button on the left side of the dashboard if the presented equation was incorrect. Stimulus onset synchrony was randomly selected to be -400, 0, or 400 ms. Subjects performed four 15-minute blocks separated by 10 minute rest periods. Five conditions included left perturbation (with or without math), right perturbation (with or without math), and math alone. Data was acquired using a Neuroscan 30-channel headset. The study was approved (VFHIRB no.: 2013-01-029BC) by the Taipei Veterans General Hospital.

A.16 NCTU Lane-keeping with Arousing Feedback Task (NCTU-LKwAF)

The NCTU Lane-keeping with Arousing Feedback Task (Lin et al., 2010) (Huang et al., 2016) was a lateral perturbation lane-keeping task conducted in VR-based driving simulator with perturbations presented at random intervals of between 8 and 12 seconds. Trials that had reaction time 3x longer than trials at the beginning of the session were designated as "drowsy trials". The alert reaction time was computed as the mean reaction time in the first 5 minutes of the task. In 50% of the "drowsy" trials, the system triggered a 1,750 Hz tone-burst. Data was acquired using a Neuroscan 30-channel headset. The study was approved (VFHIRB no.: 2012-01-019BCY) by the Taipei Veterans General Hospital.

A.17 TNO Adaptive Cruise Control Task (TNO-ACC)

The TNO Adaptive Cruise Control Task was conducted at the Netherlands Organization for Applied Scientific Research (Brouwer et al., 2017). Subjects drove a Toyota Prius on a closed test track with cruise control set at 35 km/h. An automated human voice indicated whether strong or weak deceleration was desired. The subject then engaged ACC deceleration by pressing a button. Before deceleration a computer voice announced whether the ACC would decelerate strongly or weakly. The ACC decelerated to 25 km/h between 0.5 and 3.5 s after the computer announcement following a steep or a shallow velocity profile (strong or weak deceleration). In 20% of the trials, the deceleration profile did not match the stated desire. The automated human voice then asked the driver to accelerate back to 35km/h. Data was acquired using a BioSemi 64-channel headset. The study was reviewed and approved by the U.S. Army Human Research Protections Office (Protocol ARL 16-038) before the study began.

A.18 UCSD RSVP Target Detection Task (UCSD-RSVPU)

The UCSD RSVP Target Detection Task (Bigdely-Shamlo et al., 2008) used a standard RSVP paradigm with 4.1 second bursts of images presented at 12-Hz. Images consisted of satellite photograph vignettes of London with small airplane targets overlaid. Subjects pressed a button at the end of each 49-image block (left = saw plane, right = didn't see plane) to indicate whether or not the target was detected. Half of the datasets (train) had "correct/incorrect" visual feedback after button press. The other half of the datasets (test) did not provide feedback. Data was acquired using a BioSemi 256-channel headset using a custom cap design. The study was approved (#071254) by the Institutional Review Board (IRB) of the University of California, San Diego.