

Introduction to EEG Decoding for Music Information Retrieval Research

DATA ACQUISITION AND PREPROCESSING

Electroencephalography (EEG)

- EEG is the measure of electrical activity of the brain
 - Sufficiently many neurons (hundreds to tens of thousands) firing synchronously will produce an electrical field strong enough to be measured noninvasively at the scalp
 - EEG is used here to refer to noninvasive measures
 - compare to electrocorticography (ECoG), in which electrodes are placed directly onto the cortical surface
-

General Acquisition Steps

- Amplification
 - EEG SNR thought to be around -20 dB
 - signals in microvolt (μ V) range
- Filtering
 - can be done at acquisition or offline
- Digitizing (sampling)
 - data are sampled at acquisition
- Storage
 - convention is one time series per channel
 - stored in electrodes-by-time matrix

EEG Equipment Overview

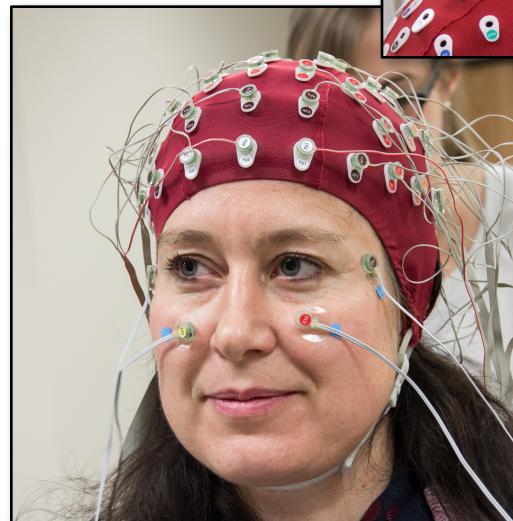
- electrodes
 - dry vs. wet (soak in solution / apply gel)
 - active vs. passive
 - net / cap / sticky (individual) / headset
 - amplifier
 - usually battery-powered (DC)
 - from analog to digital signals
 - recorder
 - connection: optical / wireless
-

EEG Electrodes

electrode net (EGI)



electrode cap
(Biosemi)



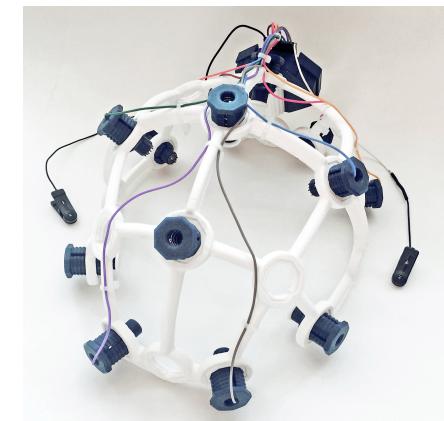
individual (sticky)
electrodes



headsets



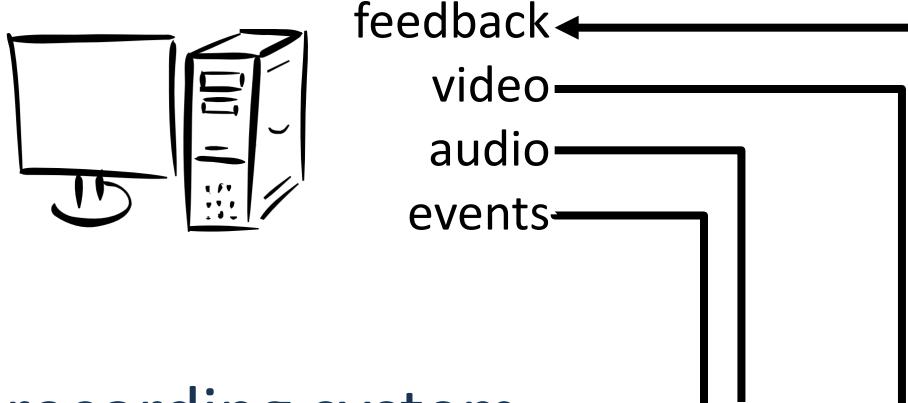
<http://emotiv.com>



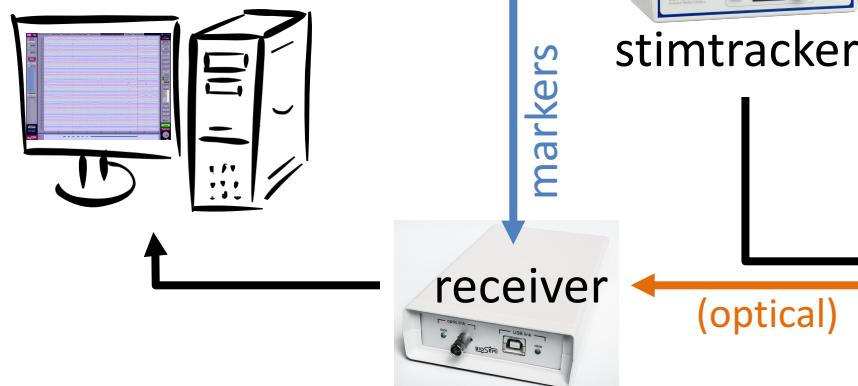
<http://openbci.com>

Typical Lab Recording Setup

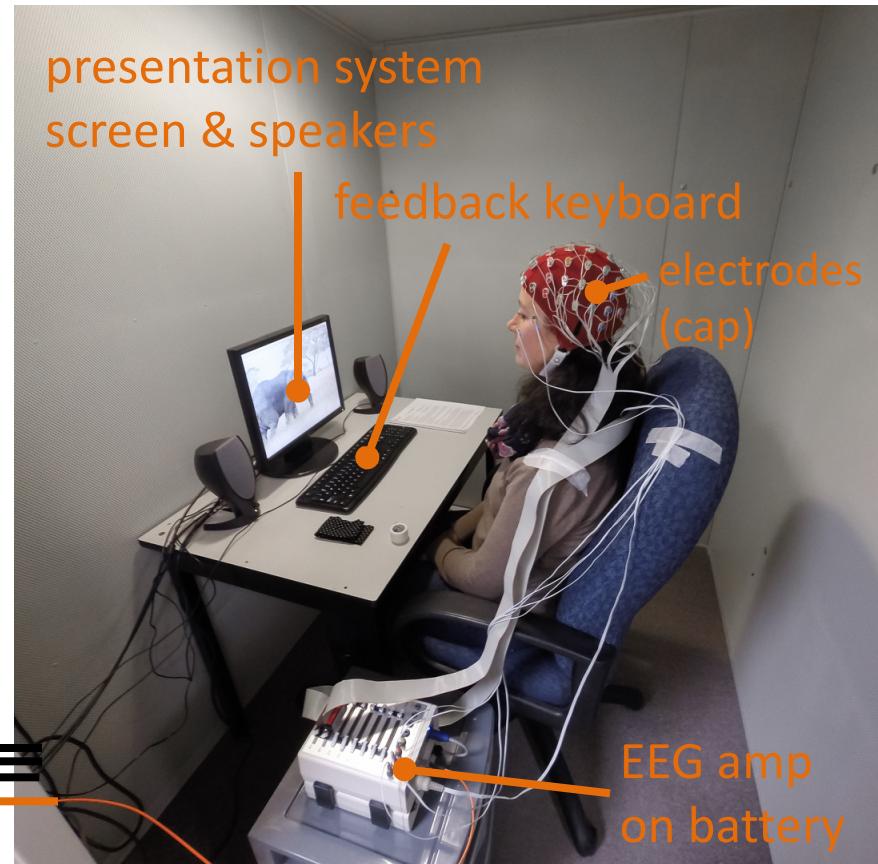
presentation system



recording system

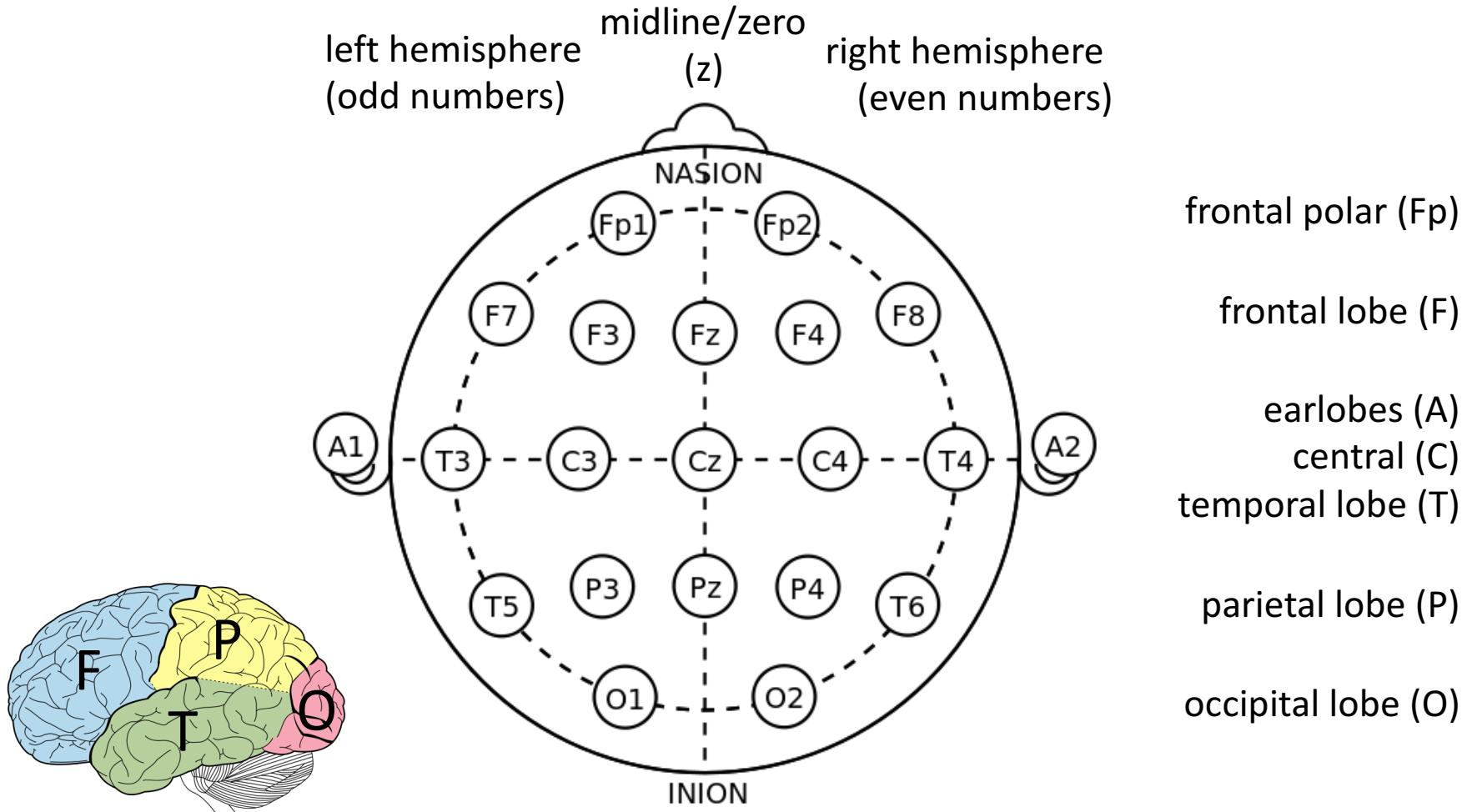


EEG room

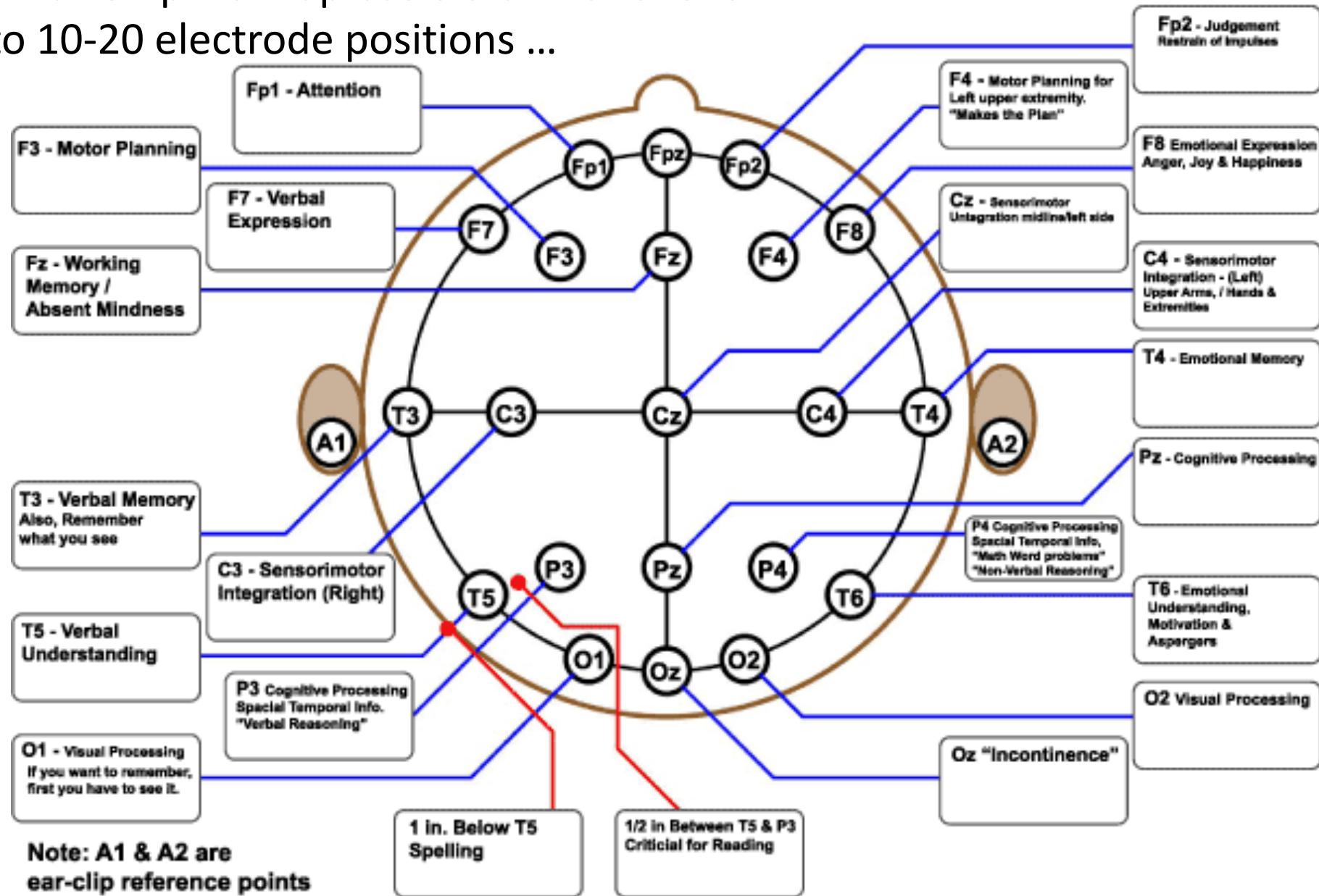


Electrode Naming Conventions

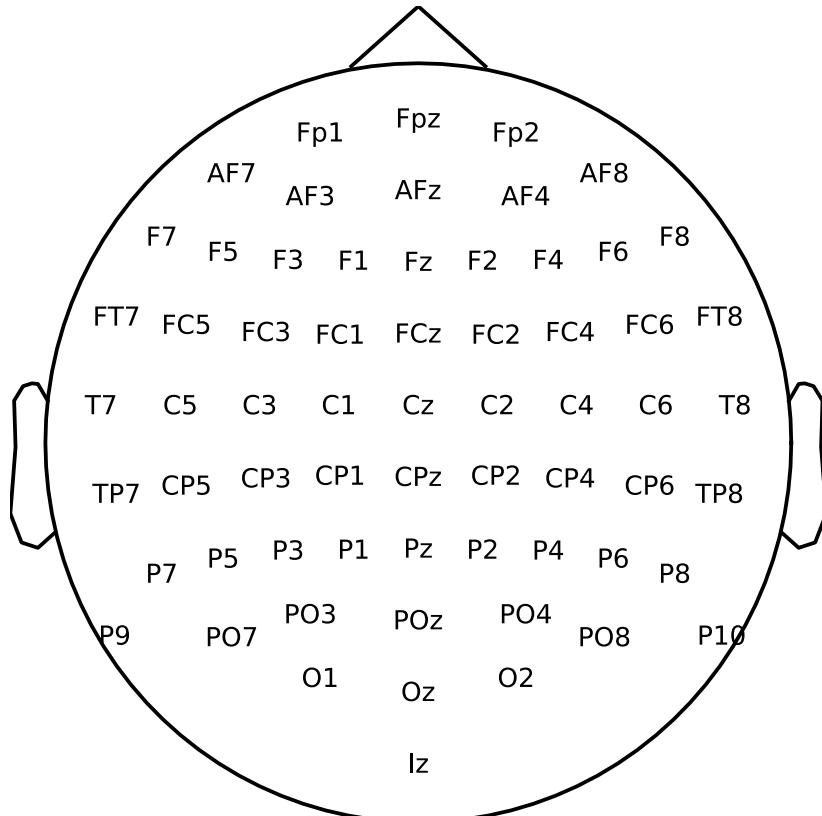
The International 10–20 System



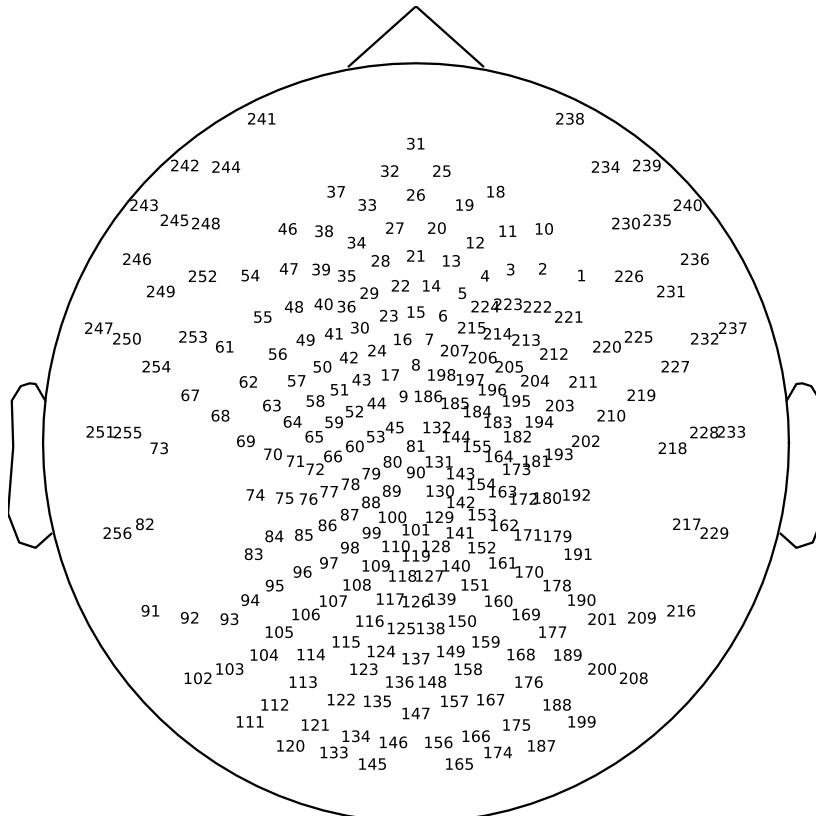
An attempt to map basic brain functions to 10-20 electrode positions ...



Some Electrode Montages



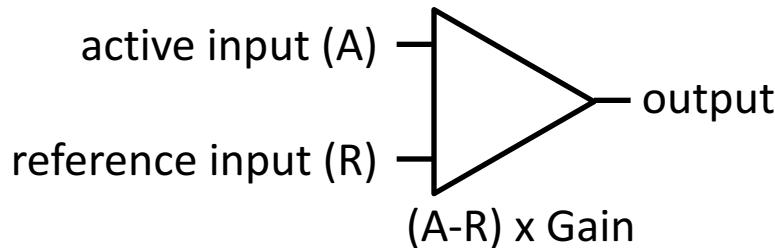
64 electrodes cap (Biosemi)



256 electrodes net (EGI)

Montage Types

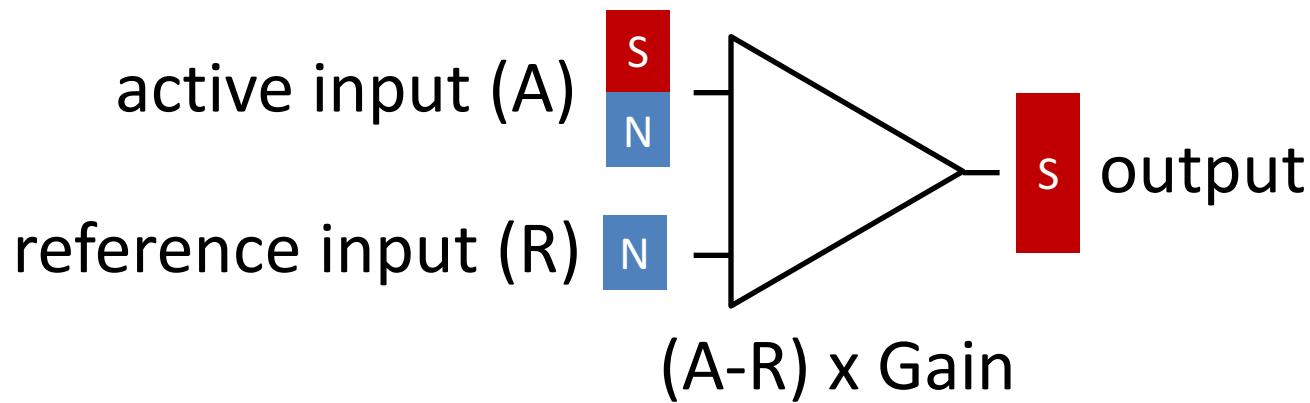
- electric potentials are only defined with respect to a **reference electrode** (“zero level”)



- monopolar
 - potential difference between reference electrode and electrode of interest
- bipolar
 - 2 electrodes per 1 channel
 - 1 individual reference electrode for each channel
- referential
 - 1 **common reference electrode** for all channels

Referencing

- goal: choose reference to capture the noise



Popular Common References

- one mastoid (e.g., TP10)
 - linked mastoids
 - vertex electrode (Cz)
 - single or linked earlobes
 - nose tip
-
- reference-free recording with **active electrode systems** (e.g., Biosemi Active Two)
 - reference chosen during data import
-

Re-Referencing

- change reference during data import
 - = linear transformation
- **average reference** often applied for high-density EEG recordings
 - assumption: if sufficiently dense and evenly distributed, sum of electric field values recorded at all scalp electrodes is always 0 by Ohm's law
 - > any differences would be noise

Sampling

- similar to recording / analyzing audio
 - part of analog-to-digital conversion
 - too low -> aliasing [Nyquist–Shannon Theorem]
 - low-pass filter at Nyquist-frequency
 - too high -> unnecessarily large data files
 - better over-sample than under-sample
 - down-sampling as dimensionality reduction
-

Real-Time EEG Analysis

- typical application: Brain-Computer Interfaces
 - not all systems allow real-time data access
 - usually through streaming (e.g., TCP/IP)
 - might require extra SDK license
 - how many channels?
 - bandwidth limit (especially wireless)
 - processing load limit (especially mobile devices)
 - special method requirements (e.g., buffering)
-

Spatial filters for EEG

- EEG data are noisy and also record redundant information from adjacent electrodes.
 - One way to improve SNR and concentrate relevant information into a lower-dimensional subspace is to compute a **spatial filter** over the data.
 - Spatial filters typically comprise a linear weighting of the electrodes (a vector of length nChannels).
 - Multiple filters may be computed at once and represented in an “un-mixing” matrix.
-

Spatial filters for EEG

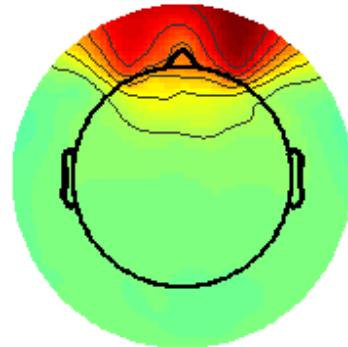
- Spatial filters are typically computed subject to the optimization of some criterion.
- Example spatial filters
 - Principal Components Analysis (PCA): Maximize variance explained
 - Independent Components Analysis (ICA): Derive temporally independent activations^{1,2}
- Techniques developed specifically for EEG
 - Reliable Components Analysis (RCA): Maximize covariance of two data frames³
 - Common Spatial Pattern (CSP): Minimize variance for one stimulus condition, maximize for other⁴

[1] Bell et al. (1995). Neural Computation. [2] Jung et al. (1998). Adv Neur Inf Proc Sys. [3] Dmochowski et al. (2012). Frontiers in Human Neuroscience. [4] Blankertz et al. (2008). IEEE Signal Processing Magazine.

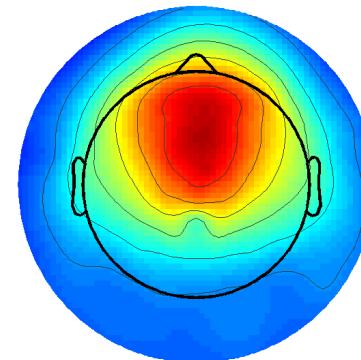
Spatial filters for EEG

- It's often helpful to visualize the spatial filter on a topoplot of the scalp.
- Important! Do not plot the weight vector directly. Instead, plot the **forward-model projection** of the weight vector.^{1,2}
 - Consider spatial filter matrix W , data covariance matrix R
 - Forward model $A = RW(W^T RW)^{-1}$

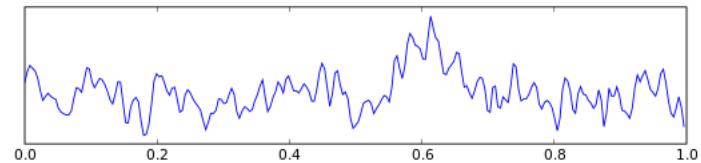
Example Independent Component
related to eye movements



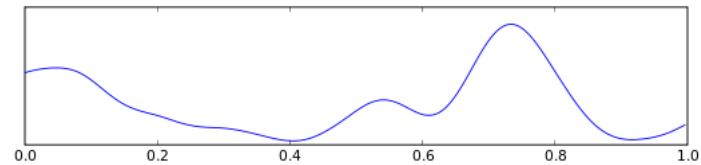
Example Reliable Component
related to naturalistic music listening



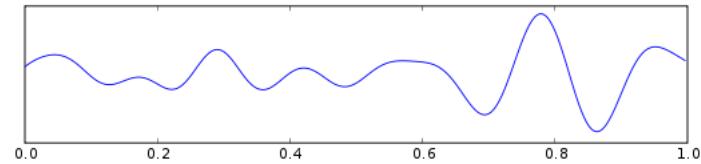
EEG Oscillations



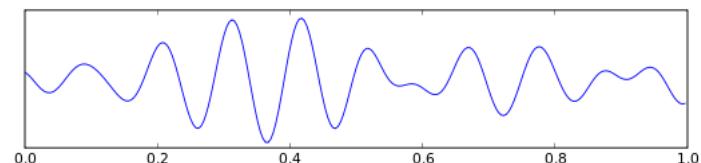
delta $< 4 \text{ Hz}$



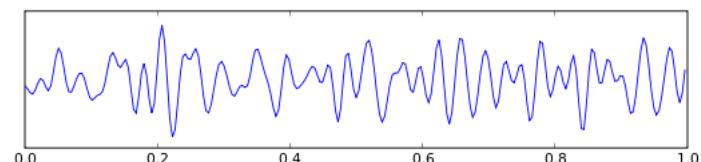
theta $4 - 8 \text{ Hz}$



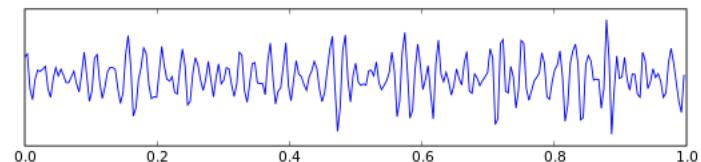
alpha $8 - 15 \text{ Hz}$



beta $15 - 30 \text{ Hz}$



gamma $30 - 100 \text{ Hz}$

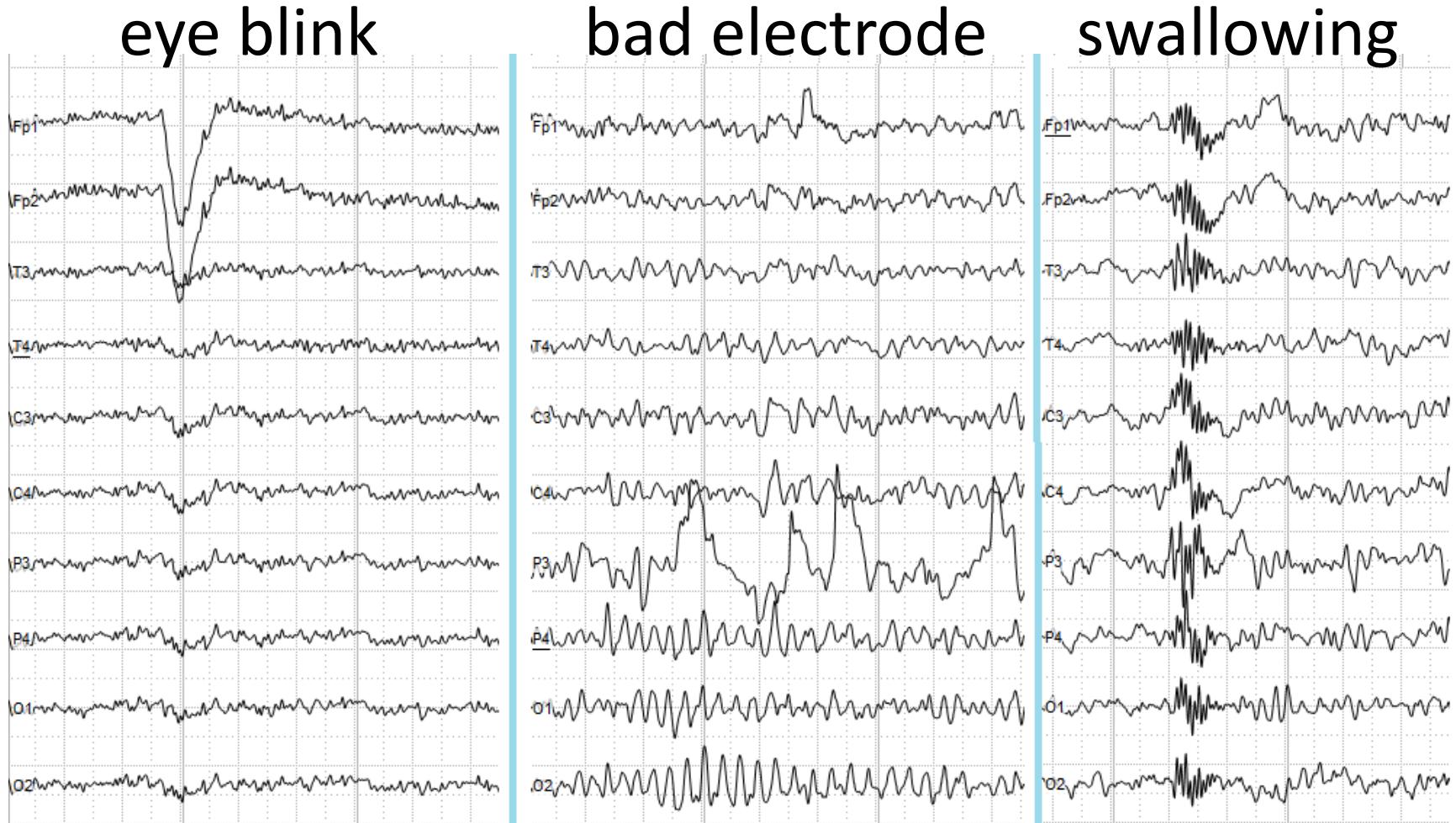


EEG Artifacts

- a) physiologic – from the subject
(sources other than the brain)
 - muscles (eyes, jaws, heart, ...)
 - sweat (impedances change!)

 - b) extra-physiologic – from outside the body
 - 50/60 Hz AC noise
 - bridged electrodes (electrolyte/salt bridges)
 - bad/dead electrodes (high impedance)
-

Common Artifacts



Dealing with EEG artifacts

- **Exclude** the data – possible when working with a large number of trials and/or electrodes
 - **Remove** the artifact by filtering (e.g., high-frequency muscle artifacts), regression (e.g., eyeblinks), identifying and removing artifact components (e.g., eyeblinks)
 - **Impute** affected data – e.g., replace a bad electrode with the spatial average of its neighbors
-

Removing eye artifacts with Extended ICA

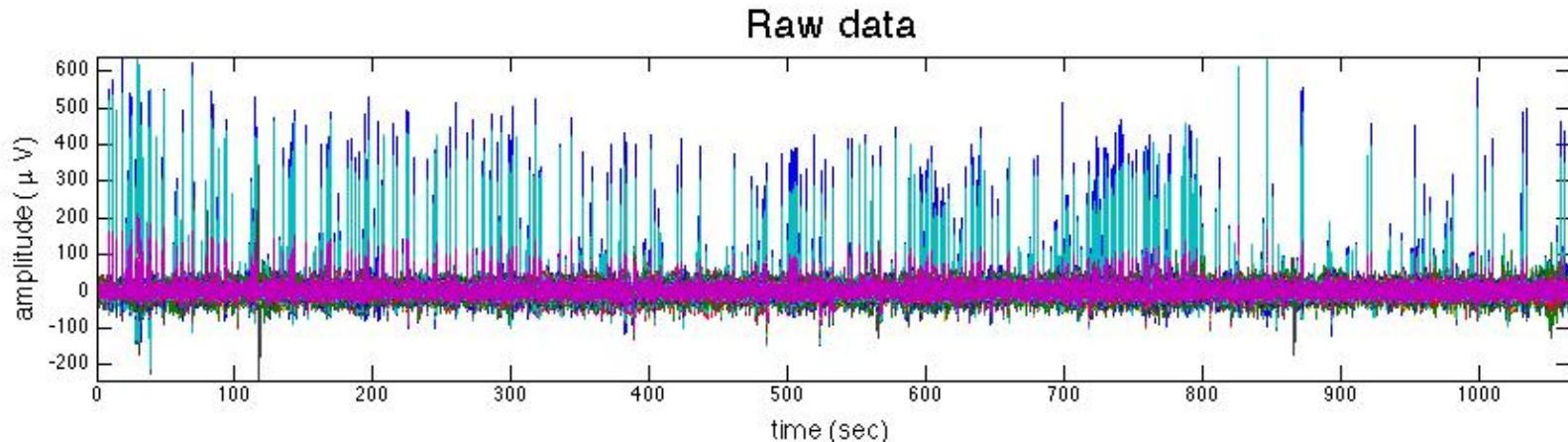
ICA is a validated method for removing eye artifacts from EEG data, and is available in the widely used toolboxes^{1,2}

- Convert data from **electrode** space to **component** space.
- Eye activity should concentrate into a few components.
- Set those component activations to zero and convert back to “clean” electrode space.

[1] Bell et al. (1995). Neural Computation. [2] Jung et al. (1998). Adv Neur Inf Proc Sys.²³

Preprocessing example: ICA

- 18 minutes of music listening (4 4.5-min trials)¹
- EEG recorded with 128 electrodes, 124 used for analysis; data bandpass filtered (0.3-50 Hz) and downsampled to $fs=125$.
- No bad electrodes; “good” participant (sits still)
- Raw data: We have overplotted all 124 electrodes to get sense of activity over time.



Preprocessing example: ICA

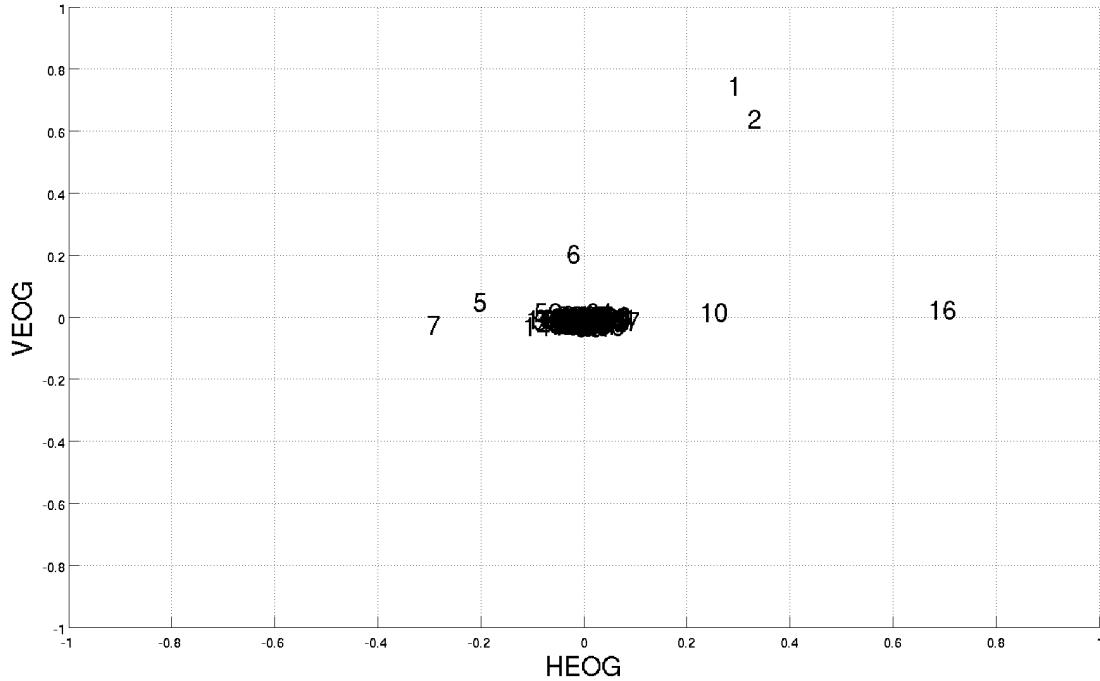
- We have already computed HEOG and VEOG channels using combinations of electrodes to the sides and above/below the eyes, respectively.
 - We then compute ICA and use the **spatial filter W** to convert the data from electrode space to component space.
 - We correlate the timecourse of each ICA component with each of the EOG channels to identify correlated components.
 - We also visualize the **forward-model projection** of the components (weight vectors) on a topographical map.
 - We will use the correlations, along with visual inspection of the activation and topography, to identify EOG components.
 - Automatically reject any component for which $|\rho| \geq 0.3$
 - By inspection, reject components for which $0.2 \leq |\rho| < 0.3$
-

Preprocessing example: ICA

Components 1, 2, 16 are over 0.3 threshold – automatic reject.

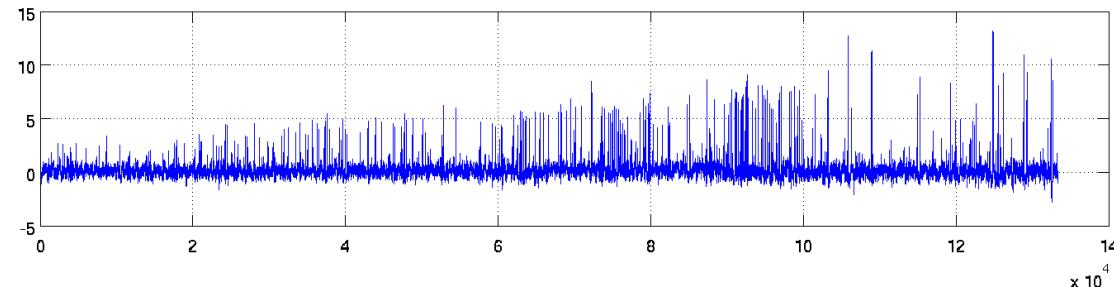
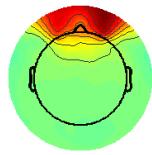
Components 6, 7, 10 are marginal – will check component characteristics.

Component 5 is outside of cloud but not high enough to trigger manual reject.

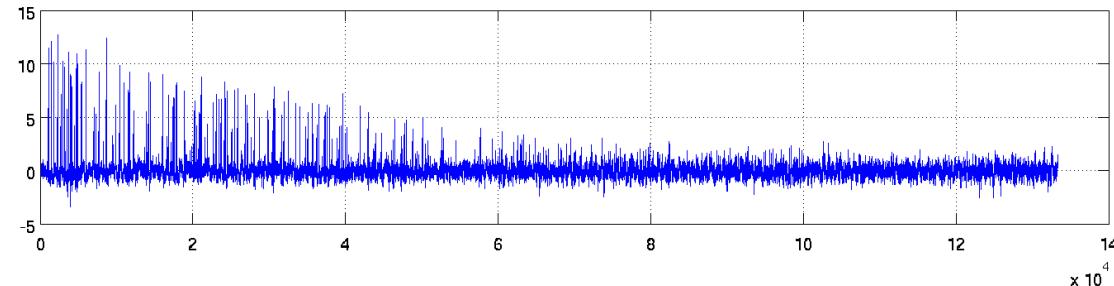
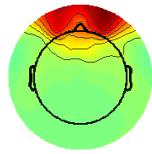


Preprocessing example: ICA

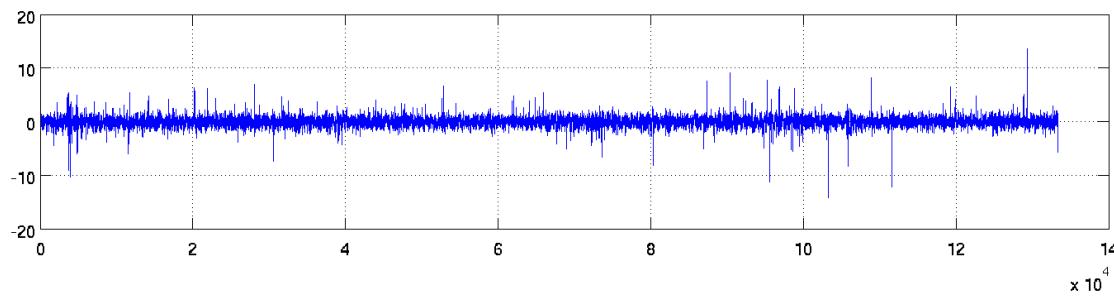
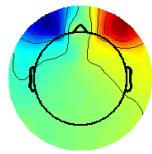
Source 1
corrV = 0.7562
corrH = 0.2935



Source 2
corrV = 0.6502
corrH = 0.3332



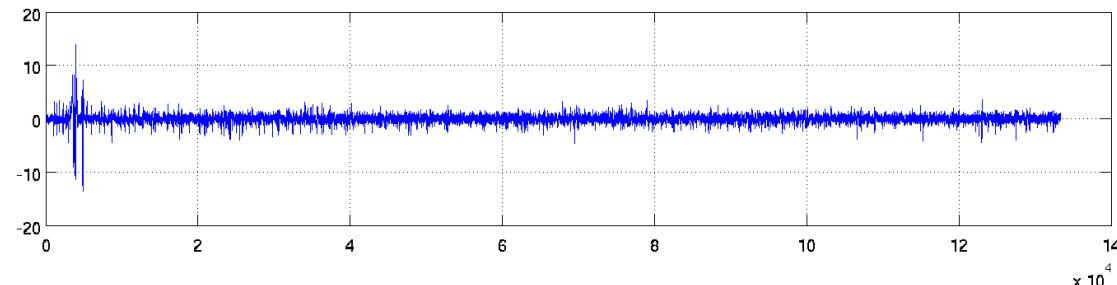
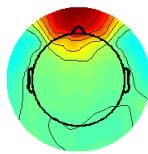
Source 16
corrV = 0.0343
corrH = 0.6975



Preprocessing example: ICA

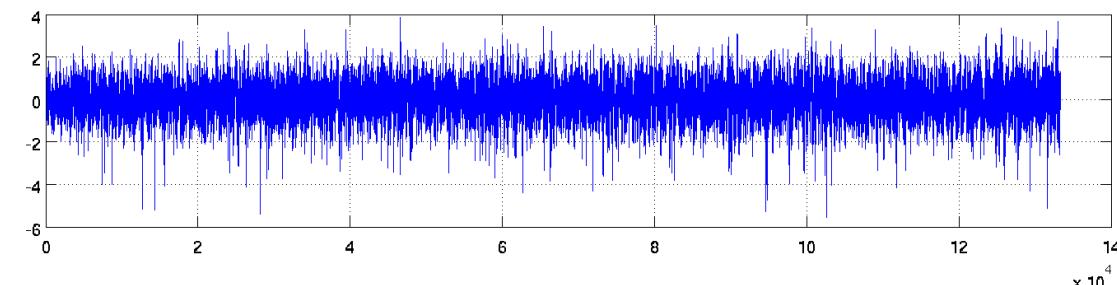
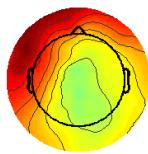
Source 6

corrV = 0.2131
corrH = -0.0174



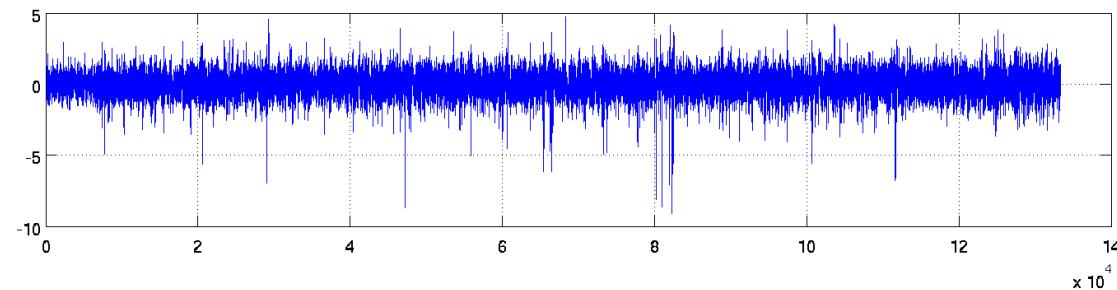
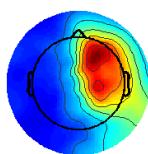
Source 7

corrV = -0.0157
corrH = -0.2899



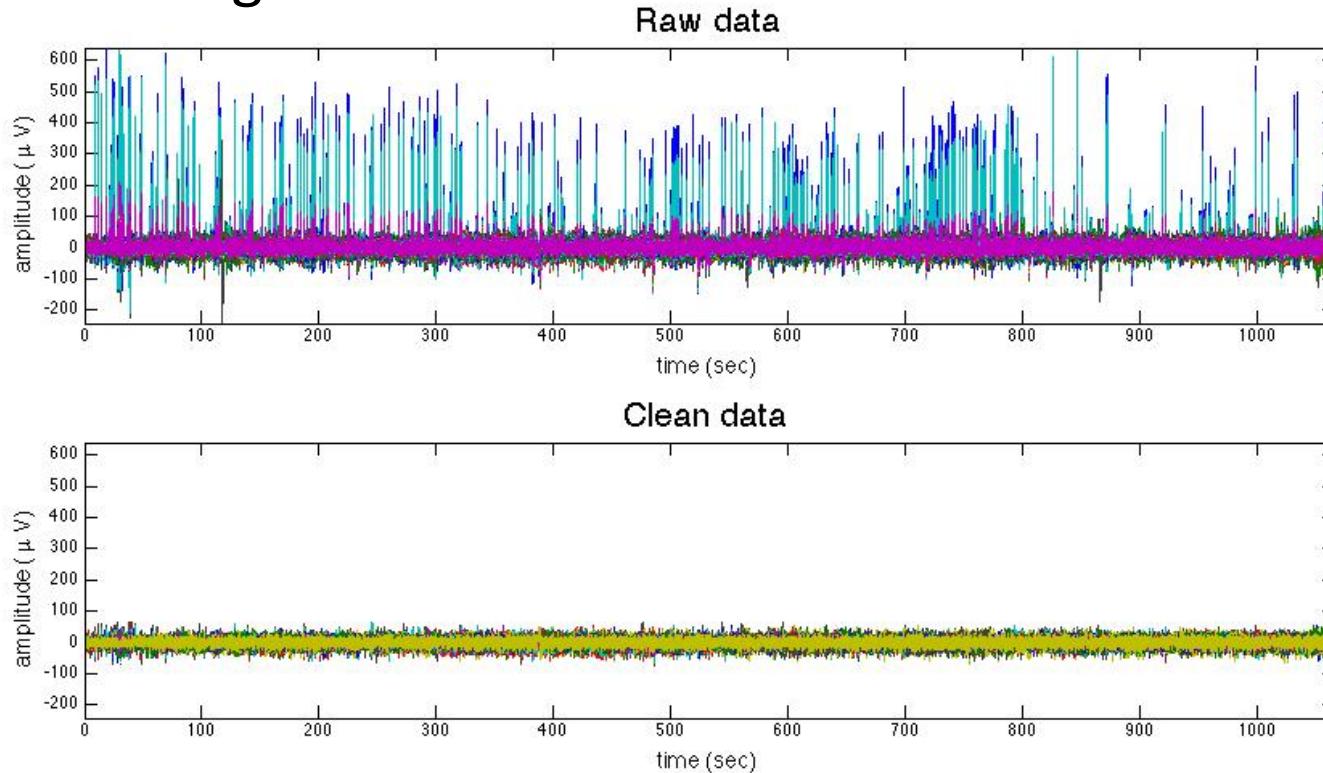
Source 10

corrV = 0.0250
corrH = 0.2538



Preprocessing example: ICA

- We set activations for high-reject components to zero.
- Convert back to “clean” channel space, remove noisy transients, convert to average reference.



Event-Related Potentials (ERPs)

Definition:

“Scalp-recorded neural activity that is generated in a given neuroanatomical module when a specific computational operation is performed.”¹

Not the focus of today’s tutorial, but makes up the majority of the research literature on music and EEG.

[1] Luck (2005). *An Introduction to the Event-Related Potential Technique*.

Event-Related Potentials (ERPs)

Idea:

- averaging over many trials (epochs)
averages out random brain activity and
keeps only the relevant waveform
- compare waveform components



Event-Related Potentials (ERPs)

Interpretation:

observe:

- time course ...
- amplitude ...
- distribution across scalp ...

differences in ERP

infer:

- timing ...
- degree of engagement ...
- functional equivalence ...

of underlying cognitive process

Event-Related Potentials (ERPs)

- example:
musical stimuli (dotted) vs. spoken syllables (solid)

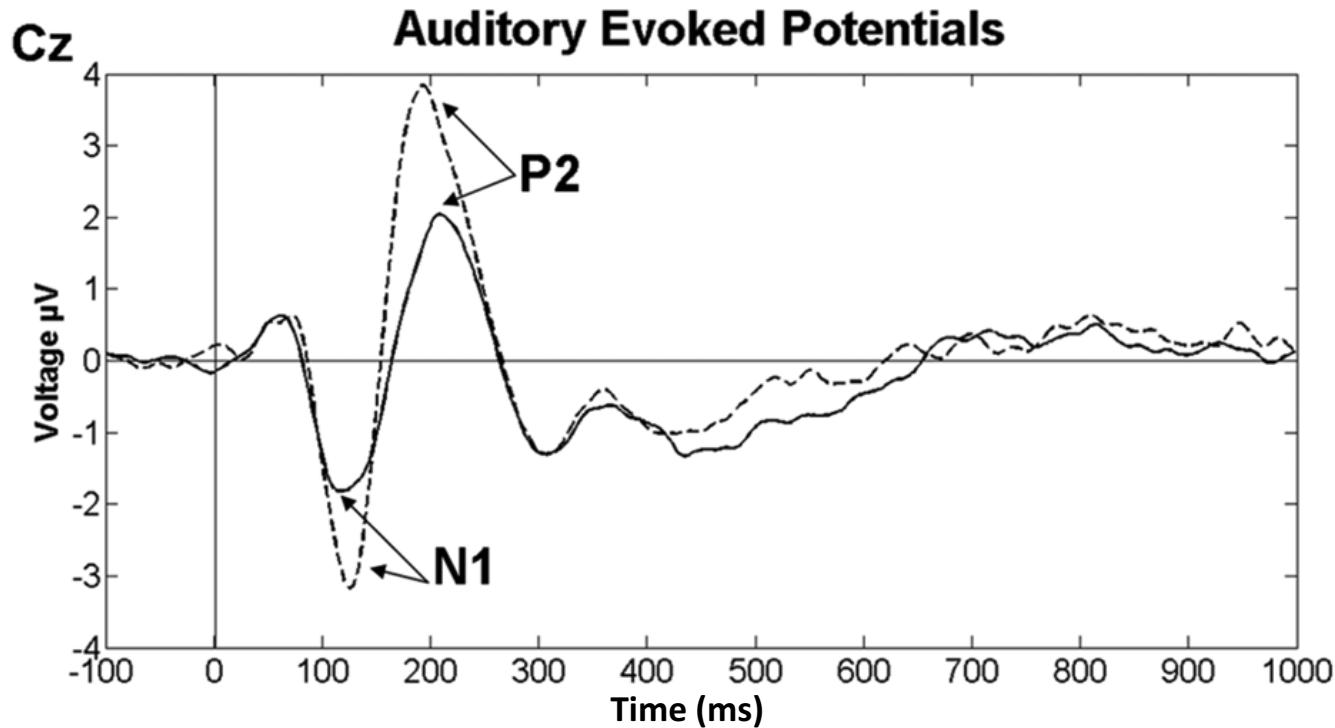


figure from Meyer et al., "Short-term plasticity in the auditory system: differential neural responses to perception and imagery of speech and music" in Restorative neurology and neuroscience 25(3-4):411-31, February 2007.

Experiment Design: General

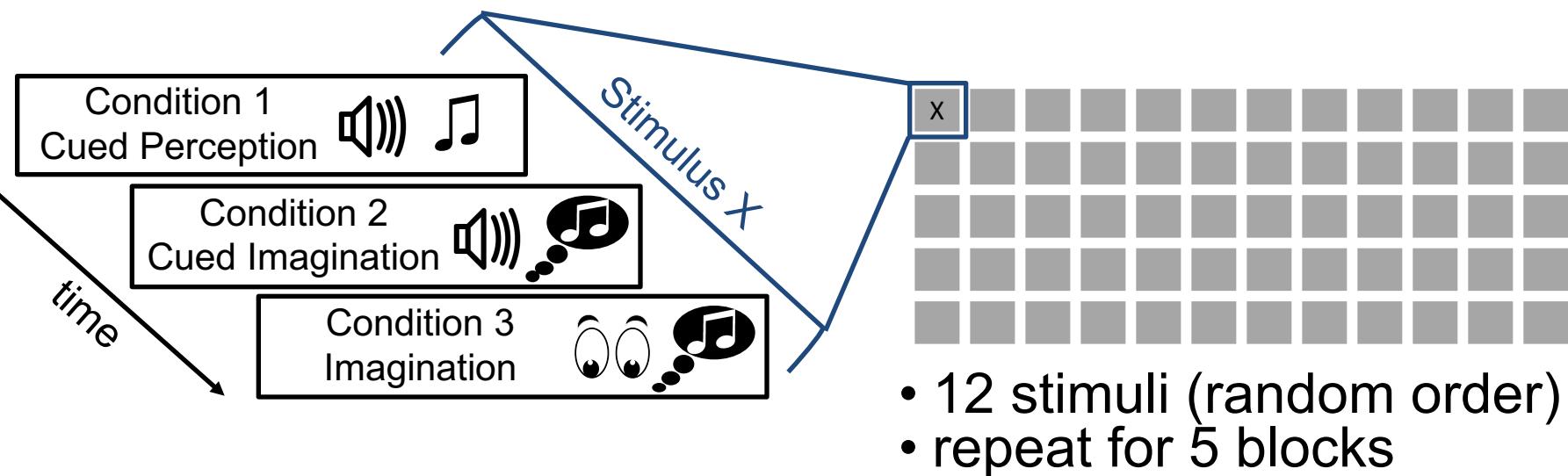
- general design / analysis approach
 - epoched (ERP analysis), i.e., average over **many** short time windows (epochs) to reduce noise
 - single-trial
 - active vs. passive task
 - focus on a specific effect,
exclude secondary effects
 - how many trials?
-

Experiment Design: Timing

- separate events in time (avoid overlaps!)
- self-paced vs. controlled timing
- constant or variable inter-trial timing
 - anticipation vs. surprise effect
- event marker / trigger recording
 - for time-locking events / stimuli
 - confirm values and timing accuracy!!!

Experiment Design: Blocks

- vary conditions/stimuli within blocks rather than between blocks



Experiment Design: Artifacts

- avoid / reduce distractions (phone!)
 - avoid / reduce movement (good chair!)
 - avoid fatigue, allow pauses
-
- electrically (and acoustically) shielded room
 - remove unused equipment
 - test different electric circuits
 - control light & temperature in recording room
-

Experiment Design: Technical

- amplifier and filter settings
- sampling frequency
 - use more than just 2x target frequency range (!)
- number, type, location of electrodes
 - impact on preparation time & data volume
- reference electrode(s)

Experiment Design: Technical

- electrode localization equipment?
 - better spatial precision than general montage data
- feedback & trigger device(s)?
 - intuitive use
 - timing
- additional physiological measures?
 - eye-tracking , EOG (eye muscles), ECG (heart) etc.
 - helps cleaning the data

Experiment Design

- Be aware of biases!
 - participant selection
 - participant group assignment
 - selection of stimuli
- most importantly:

Test, test again and pilot!!!
(Be prepared for things to go wrong!)

Live Demo – Part 1

- importing / pre-processing of a single (1h) EEG recording session
 - from OpenMIIR:
<https://github.com/sstober/openmiir>
 - using MNE-Python:
<https://github.com/mne-tools/mne-python>
- jupyter notebook available at:
<https://github.com/sstober/ismir2016eeg-tutorial>