Statistical Modeling of EEG Brain Mapping Data: Challenges and Opportunities

Jay Magidson,

Statistical Innovations Inc.

www.statisticalinnovations.com

Presented at the M3 Conference on May 24, 2017

Introduction

Mixture latent Markov (MLM) modeling often provides new insights from longitudinal data where the goal is to identify latent (hidden) states and track changes in these states over time. Also known as the mixture *hidden* Markov model (MHM) when applied to data from only one case, among various potential new applications for this methodology is the analysis of electroencephalogram (EEG) data obtained from a single individual.

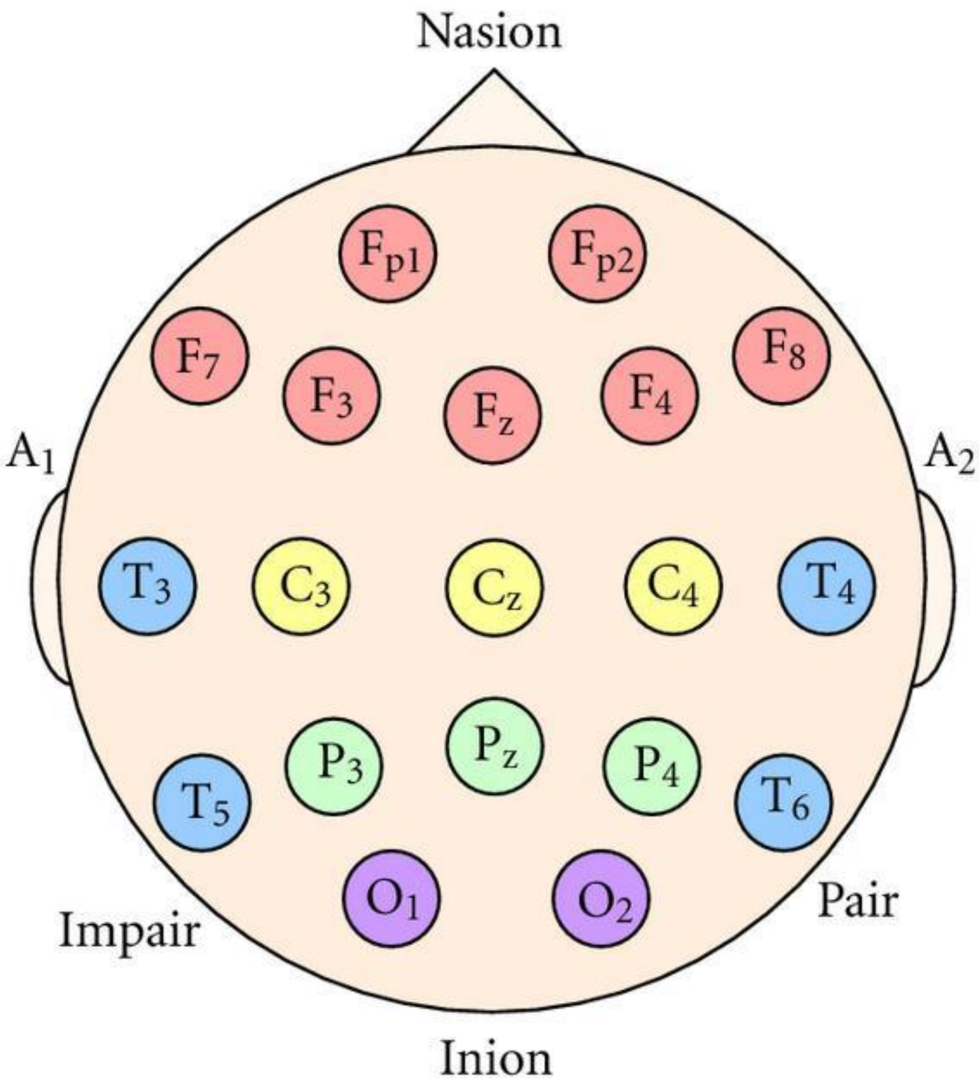
Background

- EEG data tracks brain wave patterns over time. A fast Fourier transform is commonly applied to separate the EEG data stream into multiple streams, each associated with a different frequency band, such as delta (1-5 Hz), theta (5-8 Hz), alpha (8-12 Hz) and Beta (12-20nHz).
- With 128 measurements per second collected over a two minute interval, a vast amount of data exists across each frequency band obtained from each of 19 brain regions.
- Interest in the field of neuro-feedback has been growing due to many success stories where ADD, anxiety disorders and other psychological problems have been greatly reduced or eliminated completely.
- In this paper we obtain data from brain maps pre and post neuro feedback, transform the band wave frequency data to be susceptible to MLM (MHM) modeling and perform exploratory analyses to attempt to find state changes that can explain the reported changes in behavior.

The sample data

- Brain map EEG data was obtained pre-and post neuro-feedback from 19 scalp locations on a person reporting excessive anxiety.
- Comparing the baseline (pre) data with a normal reference group discovered that the magnitude of alpha waves was relatively low and that the frontal brain regions showed relatively high levels of dissociation.
- A protocol of neuro-feedback was designed to positively reinforce the occurrence of increased levels of alpha when they were found to simultaneously occur in four selected regions.

19 Scalp Locations (regions, sites)



4 Channels at a time (frontal brain regions) Reward alpha (8-12 Hz) above a threshold in all 4 selected regions:

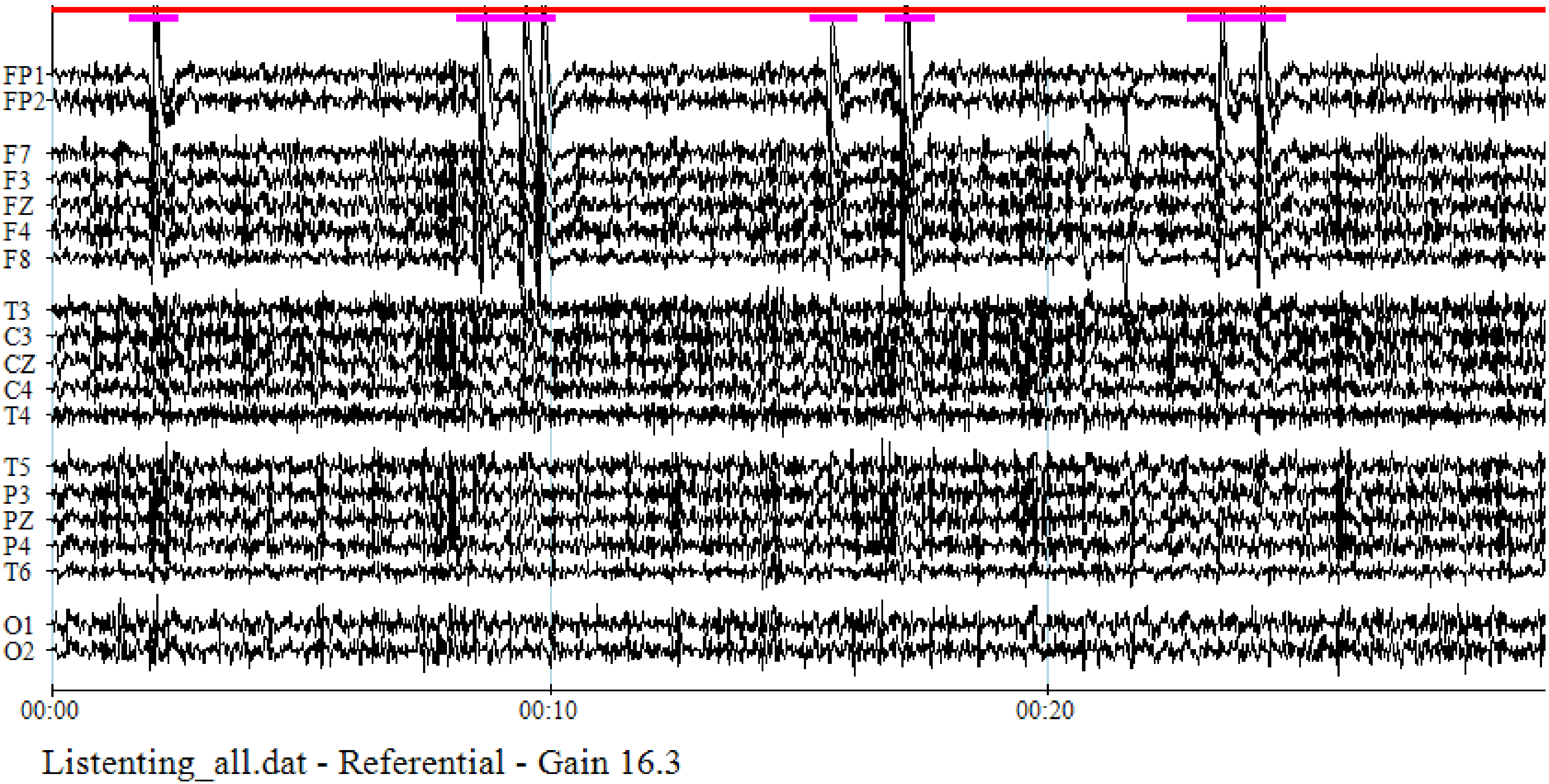
F7, Fz, F3, F4
Fp1, Fz, F3, F4
Cz, F4, C3, C4

F3, F4, Fz, F8
F7, F4, Fz, F8
Fp1, F4, Fz, C4

Better to do pre-post EEG brain maps simultaneously with pre-post fMRIs

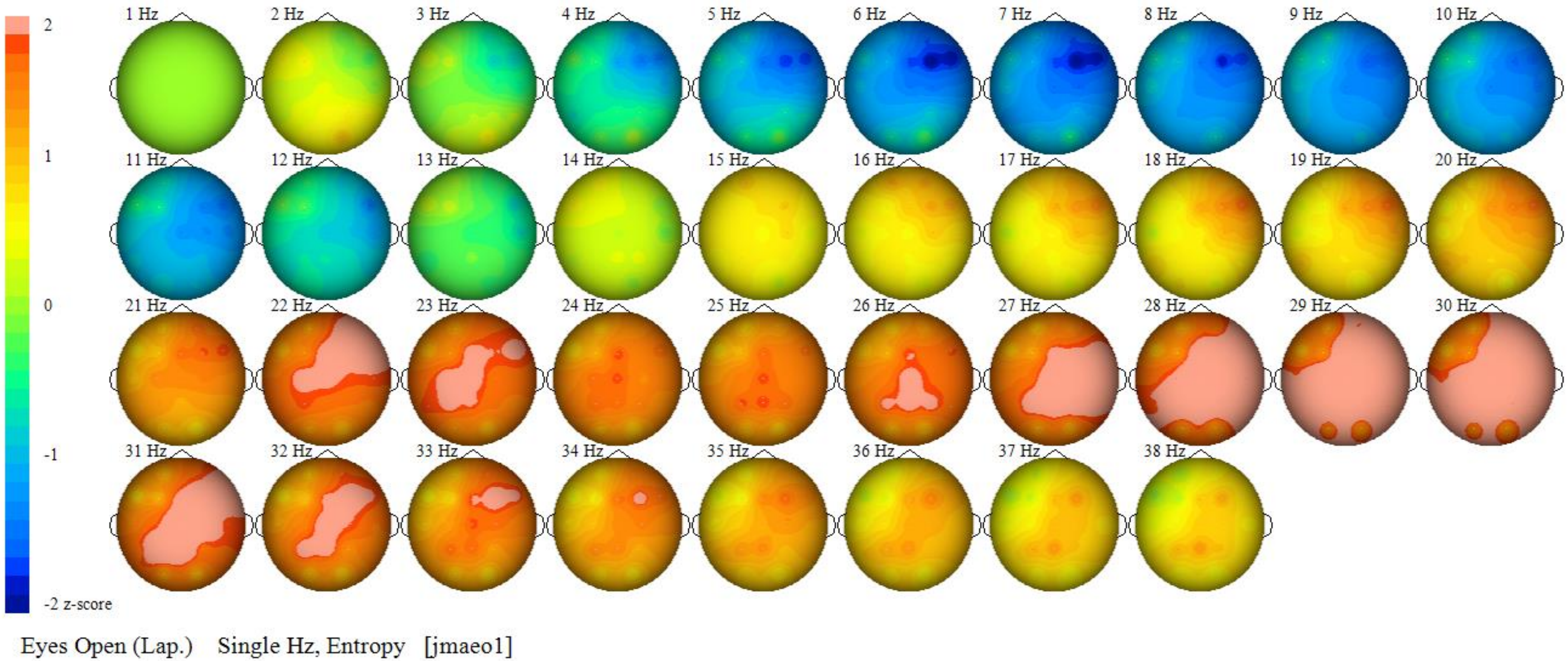
Original Raw Time Series Data from 19 sites

with Artifacts (Blinking, etc.) flagged

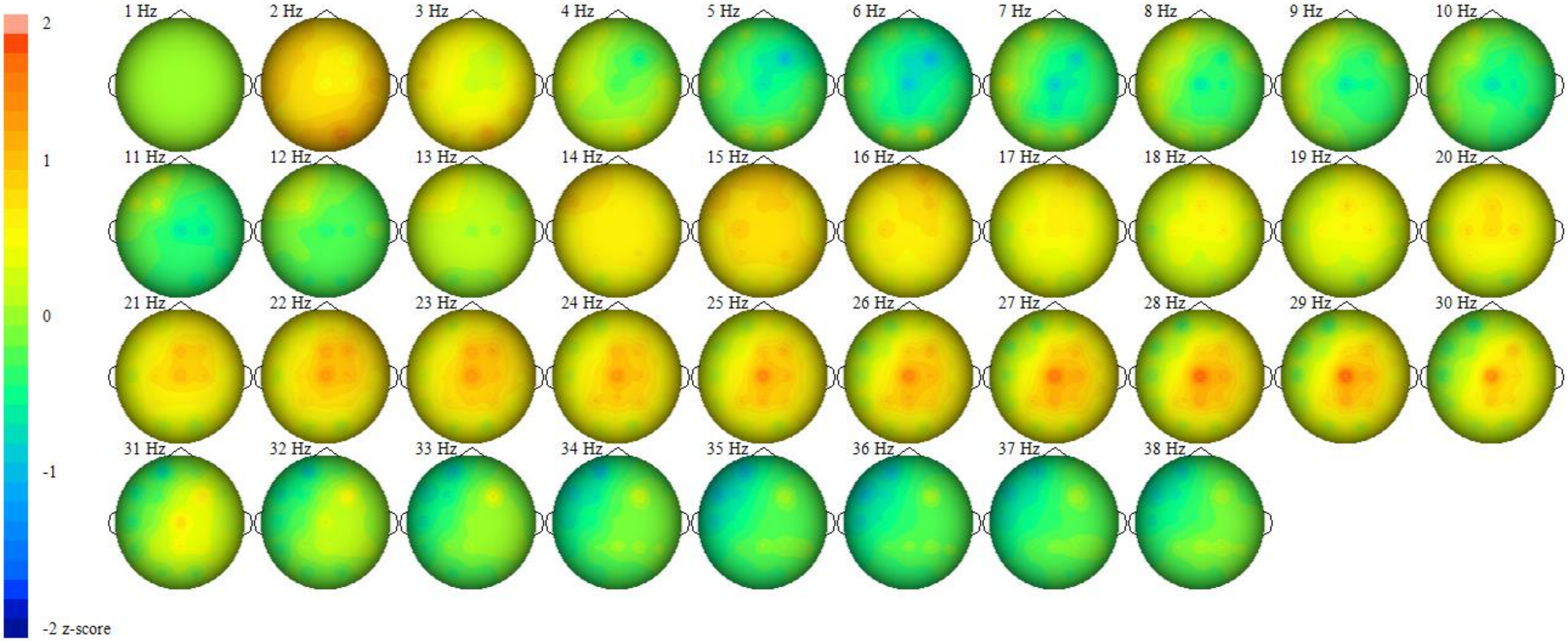


Pre-brain mapping (eyes open) Traditional Analysis

Low magnitude of alpha waves (8-12 Hz)



Post-brain mapping (eyes open)
Traditional Analysis

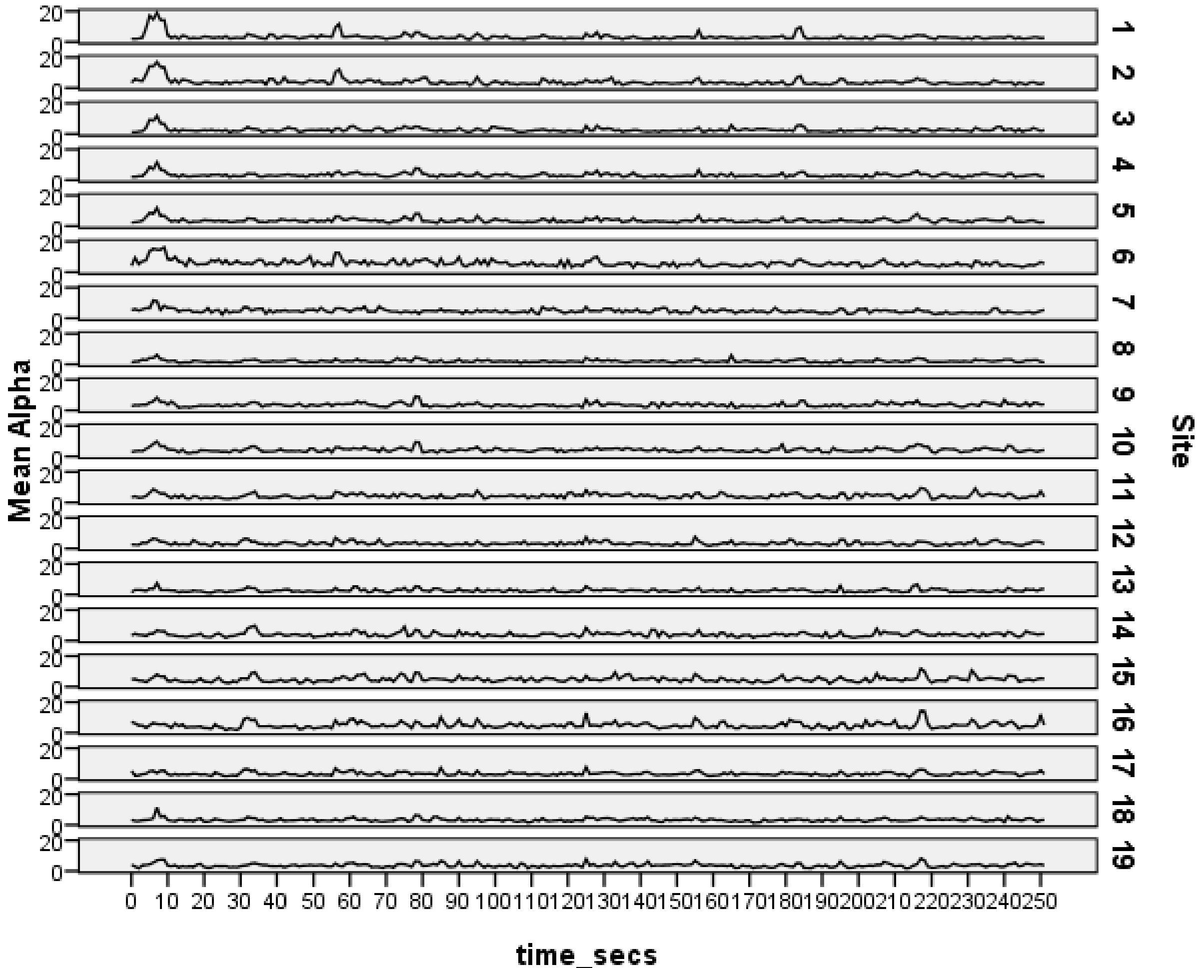


Eyes Open (Lap.) Single Hz, Entropy [j2meo1]

Transform the Data to remove Cycle

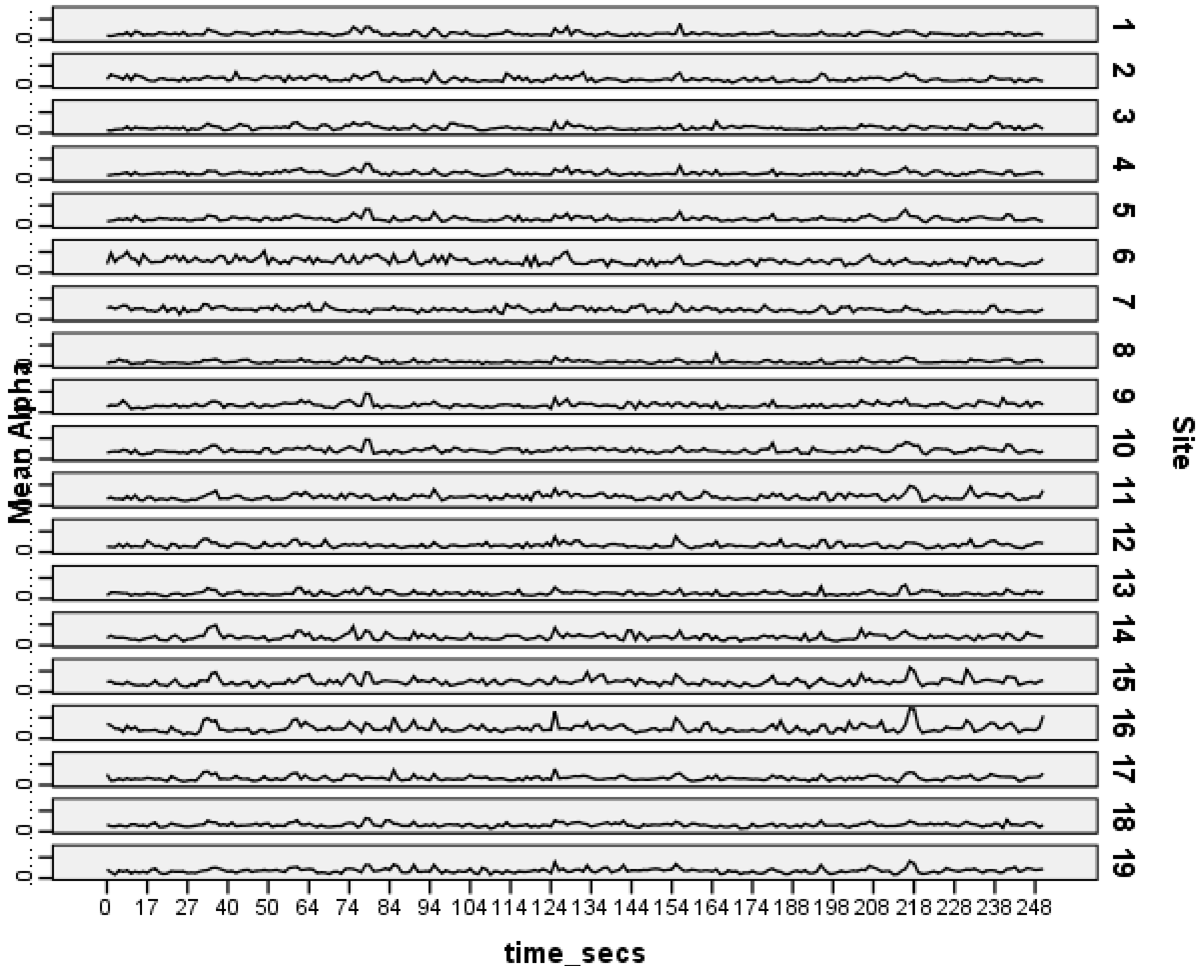
Compute peak magnitude among the 128 measurements in each second for each frequency range. Approximately 1 cycle per second (depends on frequency).

Listening Time Series: Peak Alpha by Site (Site represents channel -- scalp location)



Post data – Pre and Post data analyzed separately

Listening Time Series: Peak Alpha by Site
(Artifacts identified and recoded as missing)



After removing seconds showing artifacts (sites 1 and 2 used to identify eye blinking)

Descriptive Results:
Mean Alpha Magnitude by Site

Site	Post Alpha	Pre-Post Change	Included in Protocol?	Right brain
FP1	3.0	0.1	X	
FP2	3.6	0.8		X
F7	2.6	-0.8	X	
F3	3.2	-0.2	X	
Fz	3.6	0.2	X	
F4	5.6	2.4	X	X
F8	4.7	2.2	X	X
T3	2.1	-1.6		
C3	3.6	-0.4	X	
Cz	4.2	0.3	X	X
C4	4.3	0.7	X	X
T4	3.4	0.8		
T5	2.6	-1.5		
P3	4.0	-0.6		
Pz	4.9	0.3		
P4	5.0	0.5		X
T6	3.3	0.6		X
O1	3.3	-1.1		
O2	3.6	-0.3		X
Total	3.7	0.1		

Overall, there is only a slight pre-post gain in Alpha magnitude (0.1).

However, in sites targeted in neurofeedback protocol:

All 4 right brain sites had gains, including the 2 largest gains in the frontal sites F4 and F8 (+2.4 and + 2.2)

All but 1 of the 8 right-brain sites showed gains--Occipital site O2 showed reduction of 0.3

Descriptive Results: Reduced Correlations

		Alpha_F4	Alpha_F8	Alpha_PZ	Alpha_P4	Alpha_CZ	Alpha_C4
Alpha_F4	Pearson Correlation	1	.677	.482	.481	.650	.559
	Sig. (2-tailed)		.000	.000	.000	.000	.000
	N	185	185	185	185	185	185
Alpha_F8	Pearson Correlation	.677	1	.479	.491	.579	.513
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	185	185	185	185	185	185
Alpha_PZ	Pearson Correlation	.482	.479	1	.825	.701	.677
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	185	185	185	185	185	185
Alpha_P4	Pearson Correlation	.481	.491	.825	1	.637	.823
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	185	185	185	185	185	185
Alpha_CZ	Pearson Correlation	.650	.579	.701	.637	1	.699
	Sig. (2-tailed)	.000	.000	.000	.000		.000
	N	185	185	185	185	185	185

PRE: 185 time points

Correlations

		Alpha_F4	Alpha_F8	Alpha_PZ	Alpha_P4	Alpha_CZ	Alpha_C4
Alpha_F4	Pearson Correlation	1	.300	.062	.134	.163	.190
	Sig. (2-tailed)		.000	.347	.041	.013	.004
	N	233	233	233	233	233	233
Alpha_F8	Pearson Correlation	.300	1	.166	.305	.185	.240
	Sig. (2-tailed)	.000		.011	.000	.005	.000
	N	233	233	233	233	233	233
Alpha_PZ	Pearson Correlation	.062	.166	1	.716	.636	.597
	Sig. (2-tailed)	.347	.011		.000	.000	.000
	N	233	233	233	233	233	233
Alpha_P4	Pearson Correlation	.134	.305	.716	1	.486	.762
	Sig. (2-tailed)	.041	.000	.000		.000	.000
	N	233	233	233	233	233	233
Alpha_CZ	Pearson Correlation	.163	.185	.636	.486	1	.607
	Sig. (2-tailed)	.013	.005	.000	.000		.000
	N	233	233	233	233	233	233

POST: 233 time points

Post shows reduced correlations between the frontal sites (F4, F8) and the other sites (e.g., CZ, C4, PZ, P4)

Theory – improved executive functioning to suppress amygdala

Classification of Latent Class Models for Longitudinal Research

Latent Markov (LM) models are cluster models for longitudinal data where units (e.g., persons) can switch between clusters. In the corresponding *growth* models, persons stay in the same cluster. Clusters in LM models are called *latent states*, while in the growth model clusters are called *latent classes*. In LM models, *transition probability* parameters spell out how switching between states occurs from time t to t+1.

Model name	**Transition structure**	**Unobserved heterogeneity**	**Measurement error**
Mixture latent Markov (MLM)	yes	yes	yes
Mixture Markov	yes	yes	no
Latent (Hidden) Markov (LM)	yes	no	yes
Mixture latent growth	no	yes	yes
Mixture growth	no	yes	no
Standard latent class	no	no	yes

Vermunt, Tran and Magidson (2008) "Latent class models in longitudinal research", chapter 23 in Handbook of Longitudinal Research, S. Menard Editor, Academic Press.

Latent GOLD® 5.1 syntax*: Mixture Latent Markov

```
variables
   caseid Site;
   replicationweight wgt ;                 // to exclude the artifacts from analysis
   dependent Alpha continuous;
   latent
      Class nominal 3,                     // classes consist of sites
      State dynamic nominal 3;
equations
   Class <- 1;
   State[=0] <- 1 |Class;                  // initial state distribution depends on class
   State <- (b1~tra) 1 | State[-1] Class;
   Alpha <- 1 + (+)State;
   Alpha | State;
```


- Vermunt and Magidson (2016) "Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax", Belmont MA: Statistical Innovations Inc.

http://www.statisticalinnovations.com/wp-content/uploads/LGtecnical.pdf

Preliminary Results

	Class					Class		
	1	2	3			1	2	3
	0.42	0.32	0.27			0.37	0.22	0.42
PRE					POST			
	State[=0]					State[=0]		
Class	1	2	3		Class	1	2	3
1	0.15	0.77	0.08		1	0.02	0.44	0.55
2	0.15	0.77	0.08		2	0.95	0.03	0.03
3	0.15	0.77	0.08		3	0.35	0.62	0.04
Pre	Alpha				Post	Alpha		
State	Mean				State	Mean		
1	2.4				1	2.2		
2	3.6				2	3.5		
3	5.5				3	5.3		

PRE		State			POST		State		
Class	State[-1]	1	2	3	Class	State[-1]	1	2	3
1	1	0.87	0.10	0.03	1	1	0.58	0.35	0.07
1	2	0.56	0.44	0.00	1	2	0.00	0.84	0.16
1	3	0.11	0.54	0.35	1	3	0.00	0.10	0.90
2	1	0.73	0.22	0.05	2	1	0.88	0.10	0.02
2	2	0.14	0.77	0.09	2	2	0.33	0.66	0.01
2	3	0.15	0.46	0.40	2	3	0.29	0.32	0.39
3	1	0.79	0.21	0.00	3	1	0.52	0.48	0.00
3	2	0.01	0.89	0.11	3	2	0.12	0.79	0.09
3	3	0.00	0.23	0.77	3	3	0.00	0.51	0.49

(mean alpha | state) similar for PRE and POST

PRE initial state distribution turns out to be class independent

BIC selects 3 classes for PRE for 3-state analysis

For POST, class 1 (sites) provided the highest mean.

Sites assigned to class 1 were the 7 sites highlighted previously (right-brain sites included in the protocol).

POST: F4, F8, CZ, C4, PZ and P4

(for 4 or 5 classes) class 1 includes only F4, F8, PZ and P4

Further Preliminary Result

The magnitude in the post-neurofeedback data was found to be higher not only for alpha but also for the other frequencies (Theta, Delta, and Beta), which may be indicative of higher amounts of overall activity in the right brain, especially at sites F4, F8, CZ, C4, PZ and P4.

Summary and Conclusion

For each of the 19 brain regions measured, transformation of the raw EEG data for the alpha frequency range from 128 measurements per second to 233 peak alpha magnitudes (one peak for each of the 233-second time period measured) successfully removed the cyclical nature of the data.

In addition, flagging artifacts associated with blinking as missing, resulted in data that was assumed to meet the latent Markov model assumptions.

Despite the fact that there were approximately 200 time points, the Latent GOLD Markov module had no problems estimating a mixture latent Markov model -- the elapsed time was only 8 seconds to estimate a 3-state 3-class model.

The results for the analysis of the **post**-neurofeedback data, as compared to the **pre** model results suggested that there were significant increases with respect to the magnitude of alpha brain waves – that is, increased occurrence of the high magnitude latent state, state #3. Such increases were predominantly found to exist in brain regions grouped into latent Class 1, which included the frontal right-brain sites F4 and F8 targeted by the neurofeedback as well as the more posterior right brain sites Cz, C4, Pz, and P4.

Possible Explanation

Reduced correlations for the post-neurofeedback data between the peak alpha associated with the right frontal regions and those associated with the more posterior right-brain regions is consistent with the activity in these post regions being more independent of the activity occurring in the frontal regions. This result is consistent with what might be expected if the right frontal regions were better performing their function of suppressing the right amygdala, thus providing an explanation for the reduced level of anxiety reported.

Further understanding of how the brain operates and further analyses are needed to confirm that this conclusion is reasonable.