Zhenxian Lu
Ian Nelson
Tamuz Hod

Expected and Pleasant vs. Unexpected and Unpleasant Visual Stimuli

LINK FOR THE APPENDIX (PROJECT.ipynb)

Abstract
Recent studies have shown that unexpected words and unpleasant pictures elicit significantly decreased electrophysiological activity. In the present study, we measured electroencephalography (EEG) to examine the effects of expected vs. unexpected visual stimuli. We expected to see a P300 ERP, but were not able to detect any. This may be the result of a number of factors including poor data collection procedures, confounds resulting from our choice of stimulus images, hardware limitations, or a false hypothesis. Consequently, we can neither support our hypothesis with our data nor rule it out. Despite this, we provide evidence that the Muse headset may be capable of recording EEG data that can answer this question with a better experimental setup, more extensive practice in its use, and a greater number of participants..

Descriptors: electroencephalography, EEG, expected stimuli, unexpected stimuli, pleasant stimuli, unpleasant stimuli, repetitive stimuli, repetition

Introduction

With the discovery of EEG in 1924, Hans Berger also revealed the voltage changes in the signal is correlated with external events, which is known as event-related potential (ERP). As time passes, many components of ERP were discovered, such as P1, N1, P2, N2, P3 and etc., where each of them corresponds to different types of stimuli. Recently, a number of researchers started to focus on studying unexpected and unpleasant stimulus effect on human brains. Marta Kutas and Steven A.Hillyard investigated incongruous words that occur unexpectedly would trigger event-related brain potential (ERP), and found out that a positive ERP component occurs 300 to 600 milliseconds (P300) after the stimuli were presented. However, a study led by Jason S. Moser and his associates (Greg Hakack, Emily Bukay, and Robert F. Simons) concluded that unpleasant visual evoked ERP which shows significantly decreased electrophysiological activity begins around 250 milliseconds post-stimulus and lasts several hundred milliseconds. This finding established our interests toward this research, which we were interested in looking for the results after combining both factors (unpleasant and unexpected stimuli) together.

In our research, we were aiming in analyzing the effects of expected and pleasant stimuli vs. unexpected and unpleasant stimuli. Additionally, each stimuli target was repeated consecutively two times to analyze the impact of repetitive stimuli. Based on the background research on this type of experiment, we expect to observe negative electrophysiological occurring 250ms~300ms after the stimuli were presented, and we expect to see a decrease of this affect when the repetitive stimuli are shown due to less of a surprise for the second time.

Iteration 1

Initial data collection and analysis

We created a program that recorded EEG from the Muse headset while displaying cat and snake pictures randomly to subjects, with snakes appearing more rarely than cats and never within the first 3 pictures, in order to create an expectation of cat pictures. Each of the first two subjects was shown around 20 pictures. When we analyzed the epoched data after averaging across trials and subjects, we found a very distinct response in the time-domain. However, after digging further we discovered that this was attributable to only the first trial for each subject, indicating that it may be a blink, eye movement, or muscle movement artifact as the subject reorients their head or gaze in response to the first displayed image. After re-analyzing the data with the first trial for each subject removed, there was no longer any obvious effect. This led us to alter both our analysis pipeline and experimental setup. We reduced the span of time each image was displayed, allowing us to display more images overall, thus increasing the number of trials per subject without increasing the recording session length. We also adjusted the rate at which snakes were displayed. We additionally began discarding the first trial in each recording session. Finally, we implemented a 0.5-15Hz butterworth bandpass filter, artifact rejection, and de-meaning in our data processing/analysis pipeline, as suggested here: https://www.krigolsonlab.com/muse-analysis-matlab.html?fbclid=IwAR0TpOqwCd8rsUXTIOv3h3n3KFJVuiIoAxMjcEVYyMFGpGgcKUDkXI6jTaU

Example of the data that showed a false effect:

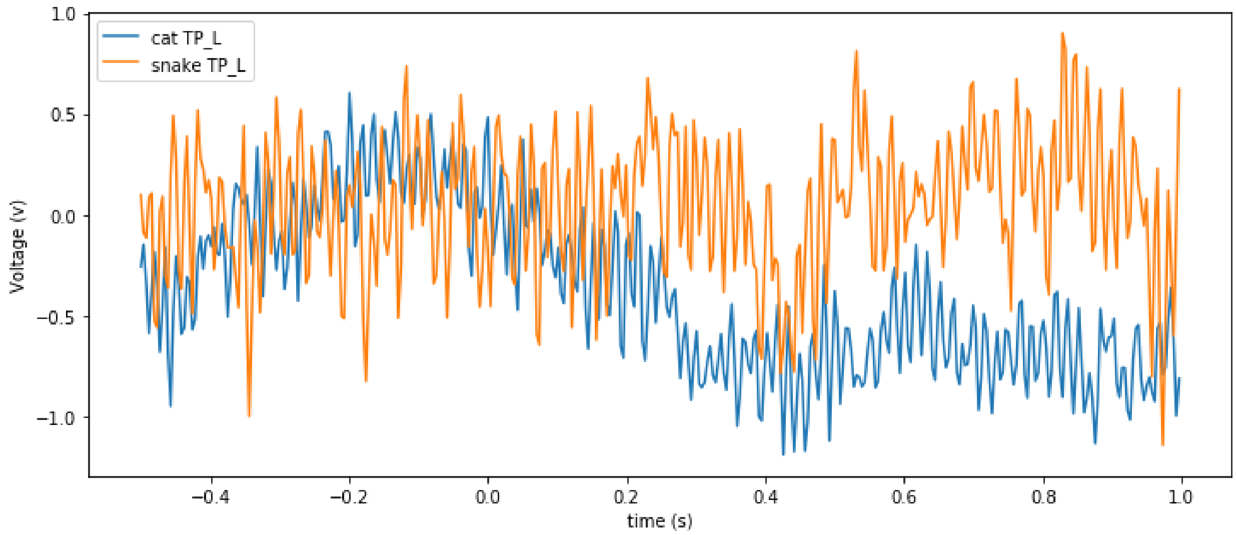

Fig 1

Iteration 2

Subjects

Six subjects (3 males and 3 females) were randomly chosen from the University of California, San Diego Geisel Library. One female data was dropped due to technical difficulties. Therefore, a total of five subjects' data were used for further analysis.

Stimuli and Procedure

Seventeen different pictures were selected from websites, of which twelves are pleasant pictures (cats), five are unpleasant pictures (snakes). Total thirty-eight pictures (trials) were randomly shown to the subjects. Besides, the same picture would be given twice in consecutive order, that is, if the first picture is "cat3", then the second picture would also be "cat3". It can be considered as a set of two same pictures. Same rule applies to snake pictures. Hence, total of fourteen sets of pictures were given in the experiment. However, the first four pictures were always cats in order to build up subjects' "expectations" toward cats. After that, there would be 5/14 chances to get snake picture every trial.

Before the experiment, subjects were only told that they would be shown several cats images, which becomes "expected" when the experiment starts. Then the first snake image would become "unexpected". During the experiment, subjects were asked to sit still on a comfortable chair and only looked at the center of the screen. Their eyes were 2ft away from the screen. As the experiment starts, there would be 3.75 seconds blank before the first picture pops up. Each picture has the same duration of 0.65625 second, and the time interval between two pictures is always 3.75 seconds. The screen kept blank when there was no pictures showing up.

Data Collection

Electrical activity was recorded using Muse-02 model that records raw EEG data using five electrodes: TP9 (Left ear), TP10 (Right ear), IF7 (Left forehead), IF8 (Right forehead) and AUXR (Right Auxiliary/baseline) with the range of 0 to 1682 μV .

The data we exported from Muse contains three files: the EEG file contains EEG data in each electrode and the corresponding index; the event file contains the picture numbers and the corresponding timestamps; and the accelerometer file which we didn't use in the present research.

Data Pre-processing

Prior to analysis, EEG data was demeaned and both a 60Hz notch filter and butterworth bandpass filter (passband: .5 to 15Hz) were applied. For two subjects it was necessary to manually apply a time offset to the timestamps in their stimulus timing data, as the values recorded differed by thousands of seconds from their EEG recording timestamps. Subjects recorded immediately before and after these two subjects had consistent first EEG timestamp to first stimulus timestamp offsets (~4.4 seconds), and both differed from this value by approximately the same amount in opposite directions, so this was taken to be the 'correct' offset and the timestamps were adjusted accordingly.

Method 1 & Results

Epoching and pre-stimulus vs post stimulus power spectra

Data was epoched and epochs containing artifacts were rejected (rejection criteria: variance values in the EEG data within the window in excess of 200). For each channel, epochs corresponding to cat and snake images were averaged separately, allowing the pre- and post-stimulus power spectra for cat images and snake images to be computed. Initial window length parameter selection of -0.5 seconds pre-stimulus and 1.5 seconds post-stimulus showed a small increase in power post-stimulus compared to pre-stimulus in most of the 0.5 to 15Hz range. A pre-stimulus window length of 0.5 seconds and post stimulus window length of 1 second was selected to minimize the inclusion of artifacts and thus minimize the amount of noise by increasing the number of trials averaged across. These lengths resulted in poor frequency resolution, however.

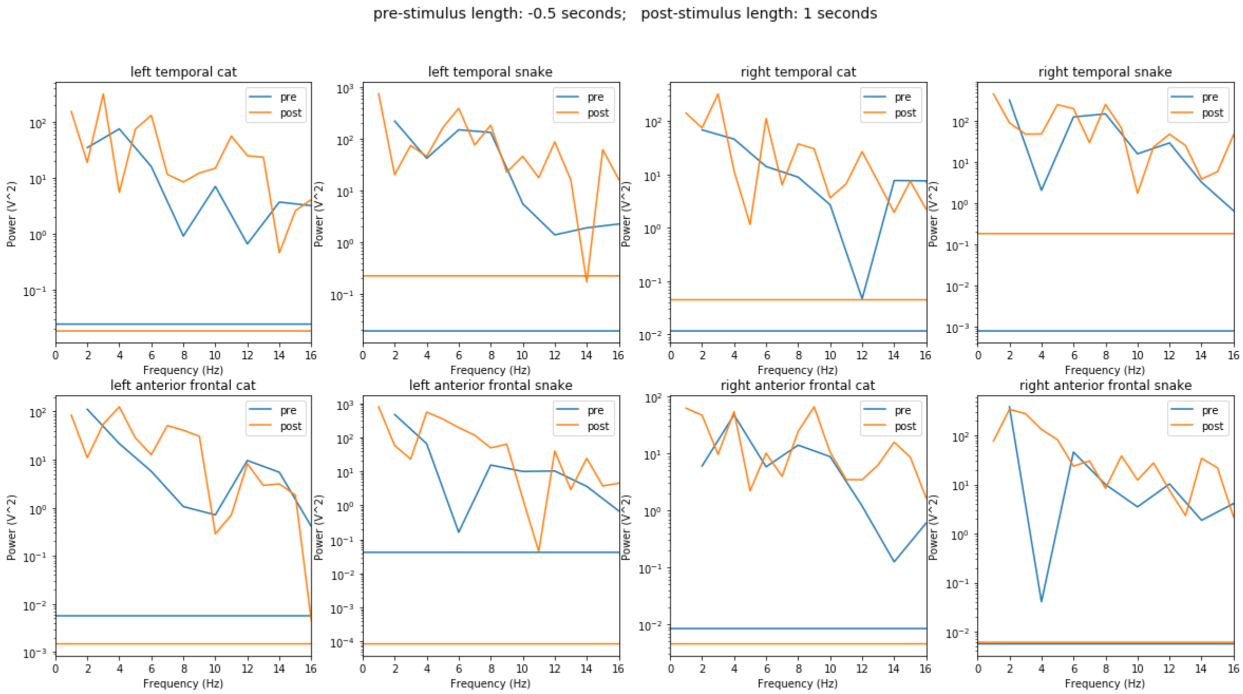

Fig 2: Figure 2 shows the power spectra obtained from each non-baseline electrode, separated by stimulus. Each graph shows the pre- and post-stimulus power spectrum for that electrode and stimulus, calculated from the waveform obtained by averaging across epochs for each electrode and stimulus.

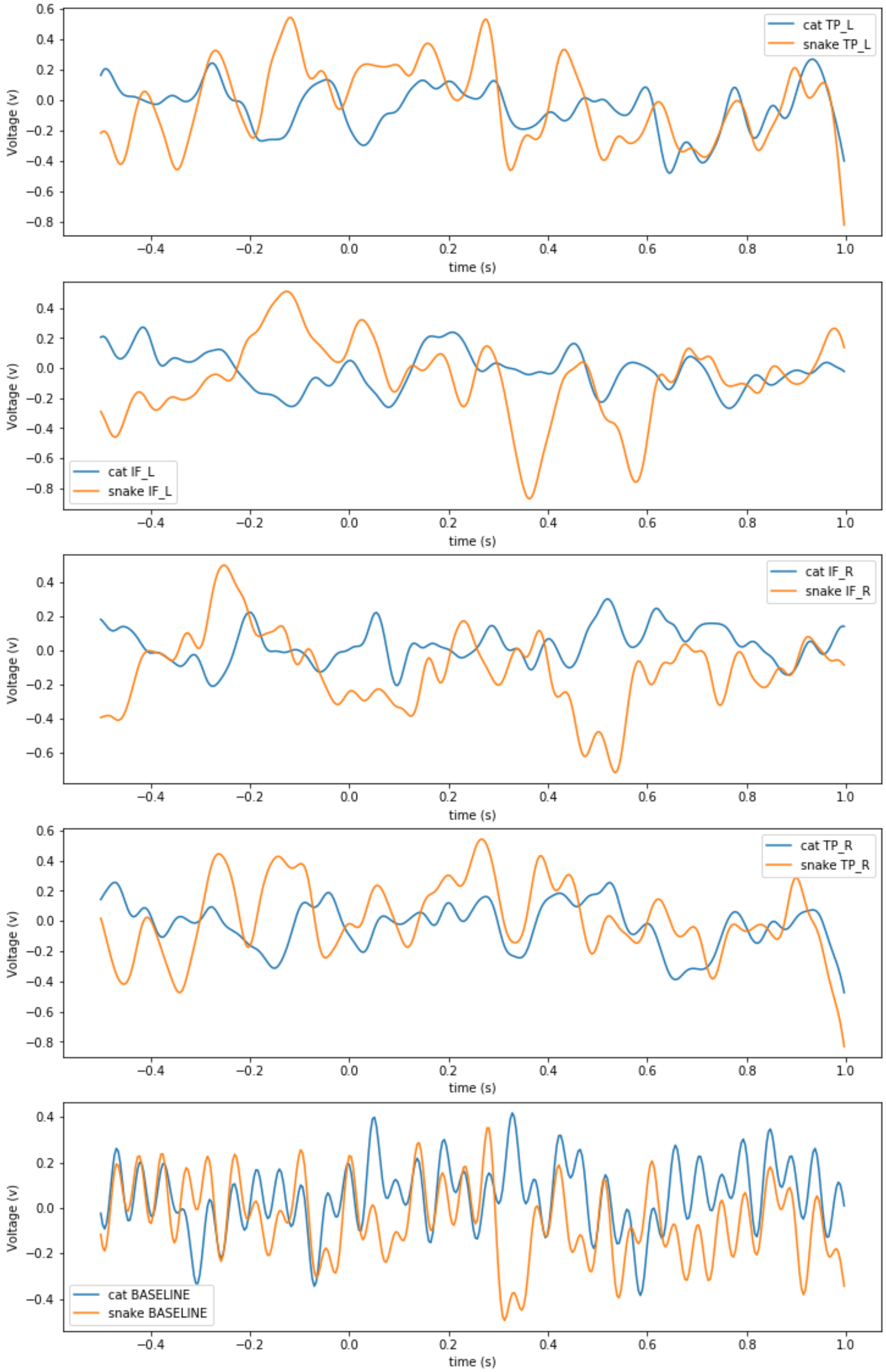

Fig 3: Figure 3 shows the averaged epoch data for each electrode and stimulus

Method 2 & Results

Pearson Correlation

We performed Pearson correlation coefficient method to test the correlation between cats and snakes in each electrode (comparison 1) and the correlation between cats/snakes across electrodes (comparison 2). We did so using SciPy's statistical function to calculate the pearson correlation coefficient. This function returns two parameters: coefficient that measures the linear relationship between two datasets, and p-value that determines the significance of the results. In comparison 1, we got (0.298, $p<0.00001$) in electrode 1 (TP Left), (0.039, $p=0.285$) in electrode 2 (IF Left), (-0.026, $p=0.468$) in electrode 3 (IF Right), (0.195, $p<0.00001$) in electrode 4 (TP Right), and (0.228, $p<0.00001$) in electrode 5 (Baseline). In comparison 2, we calculated all possible correlations across different electrodes (10 for cats, 10 for snakes), and here are some of the results: (0.673, $p<0.00001$) between electrode 1 and 4 in cats, (-0.304, $p<0.00001$) between electrode 2 and 3 in cats, (0.117, $p<0.001$) between electrode 3 and 4 in cats, (0.613, $p<0.00001$) between electrode 1 and 4 in snakes, (-0.152, $p<0.00001$) between electrode 2 and 3 in snakes, (0.138, $p<0.0001$) between electrode 3 and 4 in snakes.

Figure 4 reveals the correlation between cats and snakes in each electrode. TP Left and Baseline electrode have weak positive linear relationship between 0.2 to 0.29, but TP Right, IF Left and IF Right have coefficient between 0.01 to 0.19 that represents no linear relationship. Hence, we conclude that there is no correlation between cats and snakes in any electrode.

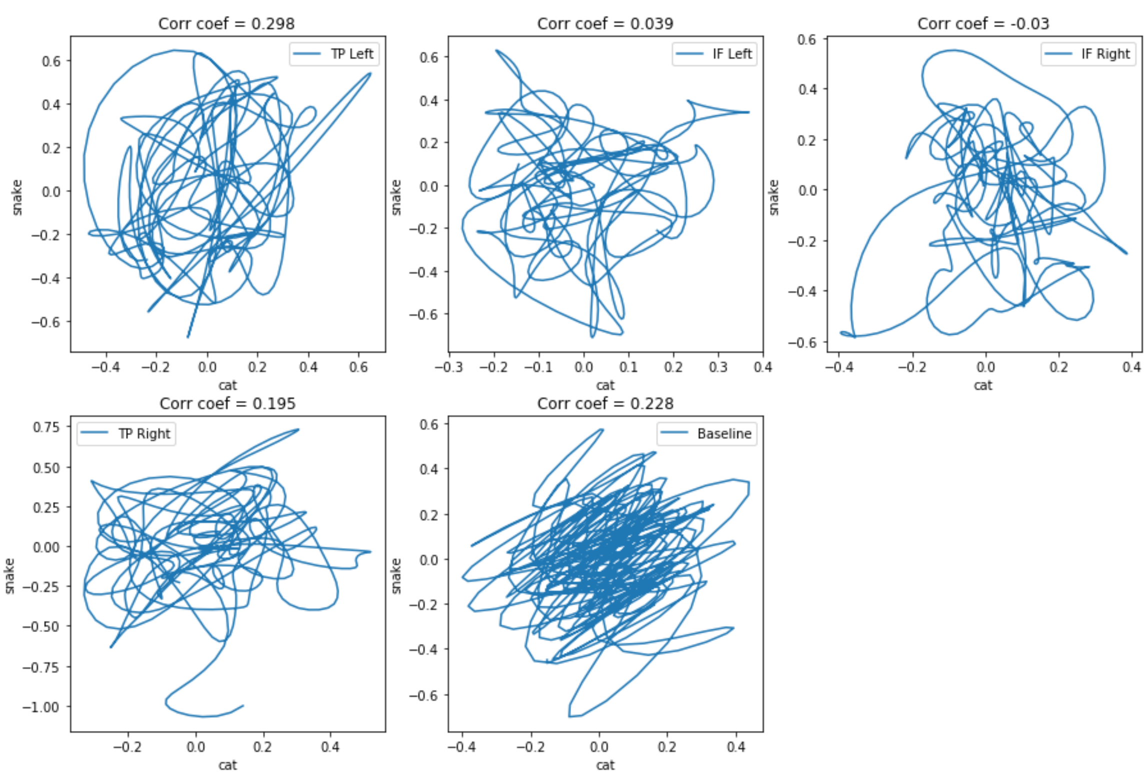

Fig 4: the above figure contains 5 subplots, which each indicates the correlation between cats and snakes in TP Left/IF Left/IF Right/TP Right/Baseline electrode. The x-axis is cats and the y-axis is snakes. The plot titles tell the correlation coefficient for each plot. As you can see, neither of the plot shows significant correlation.

Figure 5 plots the correlation between different electrodes in cats/snakes. We noticed that plot 1 and plot 4 show strong positive correlation and plot 2 shows moderate negative correlation. We think this might happen because TP/IF Left and TP/IF Right measure the same part of the brain but in different side. When we compared the electrodes that selected randomly, the correlation coefficient suddenly decreased to ~0.01 (plot 3 and plot 6). We tested 20 pairs of electrodes in total, 6 were shown in the figure, and the rest of them perform the same as plot 3 and plot 6 (no correlation). In this case, we conclude that there is no correlation across all five electrodes except TP Left/IF Left and TP Right/IF Right.

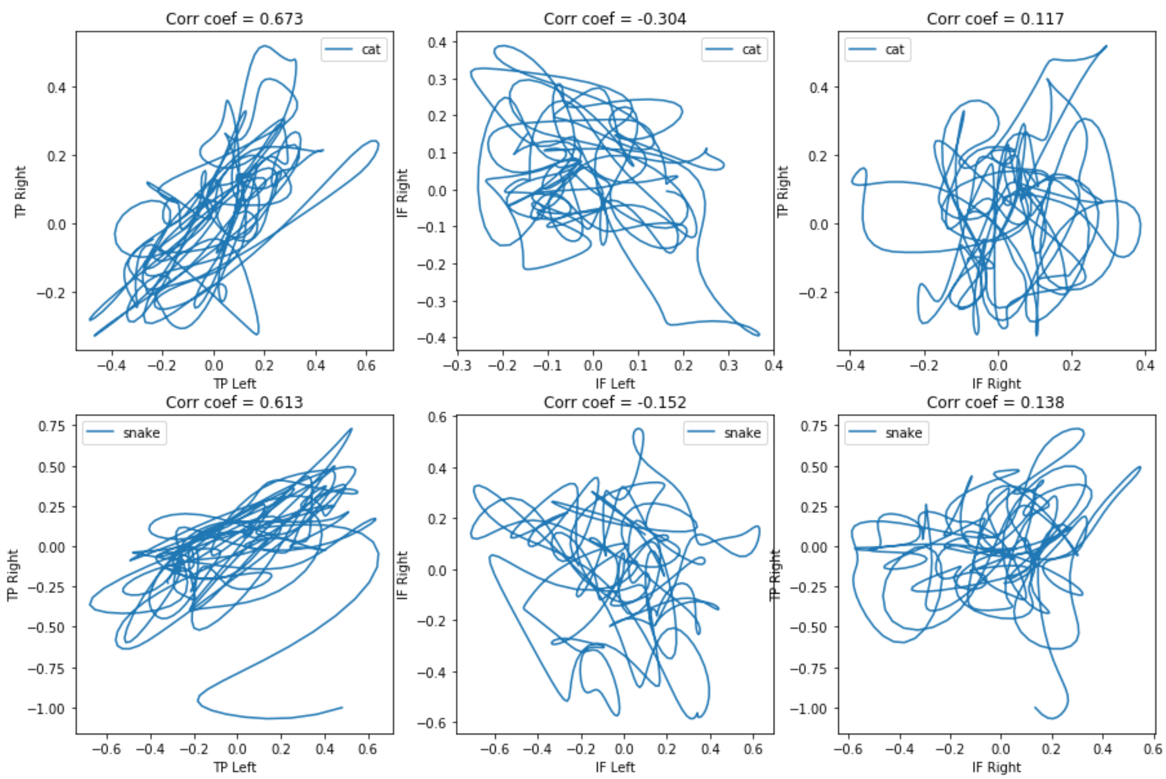

Fig 5: the figure has 6 subplots, of which 4 are chosen due to the high correlation, and 2 are chosen randomly from the 20-pairs correlation set. Each plot compares between electrodes. The above 3 plots are cats, the below 3 plots are snakes. The plot titles tell the correlation coefficient for each plot. The high correlations indicate that legismite signal is being picked up by Muse in the same brain area, but not in the random areas (electrodes).

Method 3 & Results

Coherence

A final analysis was done on the coherence of the signal pre and post stimuli. The coherence value represents how in sync are the oscillations of a specific frequency at a specific time. We used a short-time Fourier transformation with a window size of 2s and an overlap of 1s. This was chosen so that the pre and post stimuli times will always fall into consecutive windows (stimuli were shown for approximately 0.5 seconds) while a frequency resolution of 0.5Hz was deemed acceptable for the frequency range we were looking at 0.5-15Hz. The coherence was calculated for each frequency on a per channel and per subject basis. The different subjects were then averaged together to produce the graphs in Fig x. As the power spectrum density was relatively flat for all 5 channels we did not focus on any particular frequency range.

No significant difference in coherence either pre or post stimuli were found between the cat and snake pictures. Both cat and snake trials exhibited significant elevation in coherence for the post-stimulus window for all relevant frequencies across all channels. The mean elevation coherence is as follows:

mean elevation in coherence

channel	cat	snake
TP_L :	0.143	0.1
IF_L :	0.108	0.123
IF_R :	0.141	0.161
TP_R :	0.106	0.117
BASELINE :	0.11	0.094

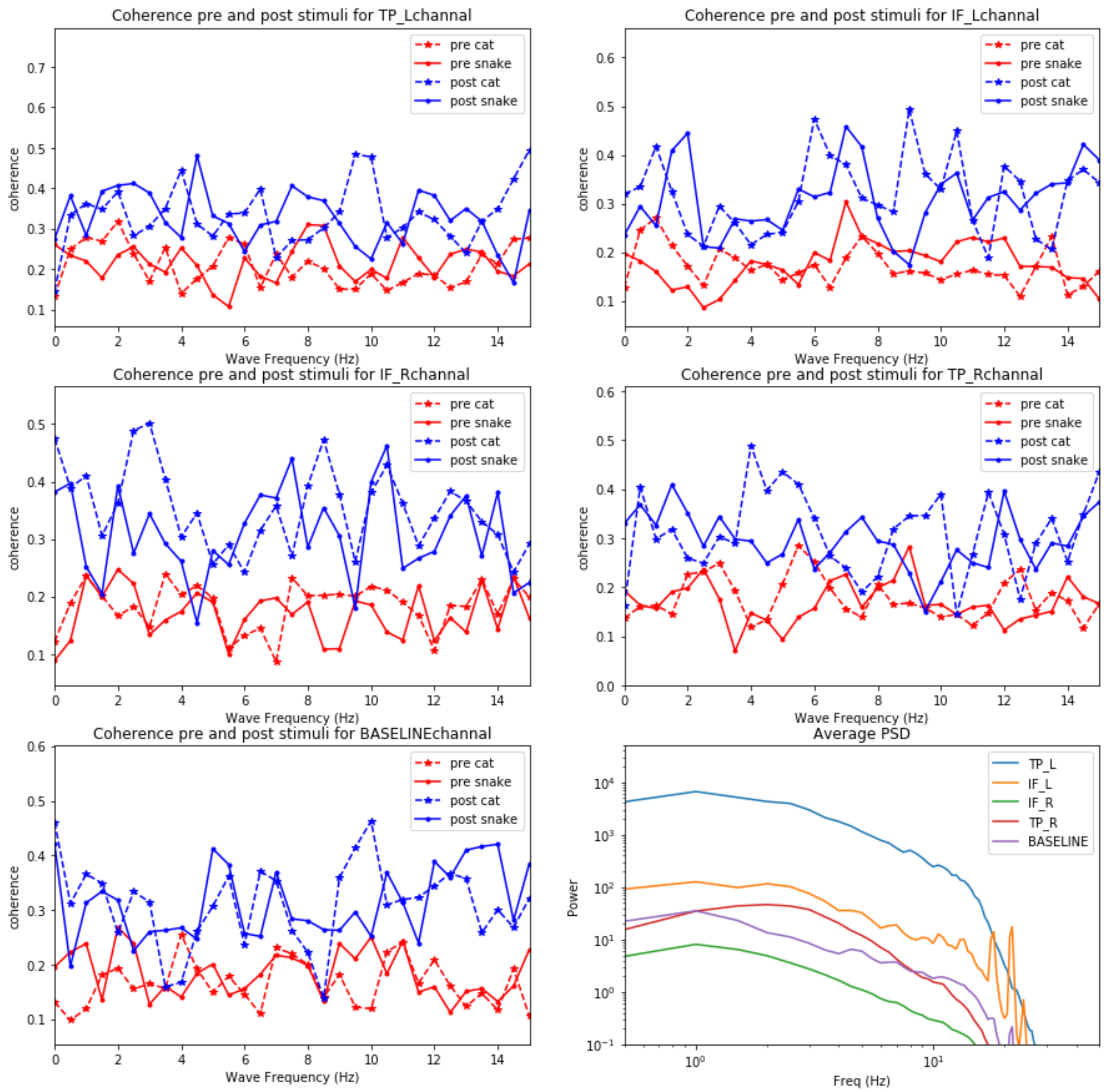

Fig 6: the first 5 plots are the pre-post stimuli from each electrode. The post stimuli increases coherence in each electrode. The last plot is the power spectrum density. We can see the power for each electrode is flat, no significant frequency that can be chosen.

Conclusion and Limitations

Our results suggest that while the Muse 2 headset is poorly suited to studies with small sample sizes/trial numbers and poorly controlled environments, it is capable of extracting at least some data under these conditions. Results from the pre- and post-stimulus power spectra are inconclusive, as they are too erratic to read meaningfully or compare across channels/stimuli or even pre- and post-stimulus. On the other hand, although comparison between stimuli or channel of the phase coherence data does not bear meaningful results, there is a clearly visible increase in post-stimulus coherence for both stimuli. Additionally, while there was little correlation between randomly paired channels, there was a significant correlation between the left and right temporal channels when both cats and snakes were presented.

While the inherent limitations of the Muse headset cannot be controlled, factors that could be improved which imposed limitations on the quality of data were: number of participants, number of trials per participant, recording environment, and stimulus presentation methods.

The variables in this experiment may also cause some issues. Our selection of stimuli, cat and snake pictures, may also be viewed as pleasant and unpleasant, respectively. While our hypothesis was that the unexpected snake images would produce a P300 effect, we found one research paper (Moser et. al.) which found that both pleasant and unpleasant pictures make significant changes in electrophysiological activity in the form of an N250 ERP but not in neural imagery (eg. fMRI). This finding may play a part in the failure of our experiment. In future research, neural imaging may also be utilized to further explore this.

References:

Krigolson, O. E., Williams, C. C., Norton, A., Hassall, C. D., & Colino, F. L. (2017). Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research. *Frontiers in neuroscience*, *11*, 109.

Kutas, M., & Hillyard, S. A. (1980b). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207(4427), 203-205.

Moser J. S., Hajcak G., Bukay E., Simons R. F. (2006). Intentional modulation of emotional responding to unpleasant pictures: an ERP study. Psychophysiology 43, 292–296. 10.1111/j.1469-8986.2006.00402.x

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. Psychophysiology, 37(2), 257– 261. https://doi.org/10.1111/1469-8986.3720257

Contributions

Tamuz Hod: data collection, develop the experimental and recording setup , data cleaning, Coherence.

https://github.com/TamuzHod/Muse_EEg_Exeriment

Ian Nelson: data collection, epoching, data cleaning, background searching, writing paper.

Zhenxian Lu: data collection, pearson correlation, data cleaning, background searching, writing paper.