Modulation of a Neural Marker of Distractibility through Neurofeedback - ECE 385J

Alexandra Mikhael, *am229763*, Arman Paydafar, *ajp4357*, and Blanca Álvarez del Río, *ba27793*

Abstract

Distractor suppression is critical for maintaining attentional control in complex environments. This study investigates the Distractor Positivity (Pd), a neural marker for suppressing irrelevant stimuli, using EEG and real-time neurofeedback. Three participants performed a visual attention task across one offline and two online sessions. EEG signals from parietal (P) and occipito-parietal (PO) electrodes contralateral to distractor locations were analyzed. Results revealed significant Pd components in a 200–500 ms window post-stimulus onset, with an additional late positivity emerging beyond 400 ms during online sessions. A Linear Discriminant Analysis (LDA) classifier was employed to decode distractor suppression signals. The offline-trained model achieved reliable performance ($\geq 60\%$) on first-day online data but showed a decline in the second session due to signal variability. Retraining the decoder with combined offline and online data improved accuracy, demonstrating the benefits of incorporating real-time modulation signals into model training.

I. Introduction

OUR visual field is constantly exposed to a multitude of stimuli, some relevant to our actions and others irrelevant, leading to distractions. Distractibility significantly impacts daily functioning and cognitive performance. In educational settings, increased distractibility can impair learning and academic achievement, while in professional settings, it can reduce productivity and increase error rates. Clinically, elevated distractibility is a key feature of attention-deficit/hyperactivity disorder (ADHD) and is also associated with age-related cognitive decline. Since our visual field contains more information than the visual system can process, selective attention is essential to extract relevant stimuli while ignoring attention-grabbing distractors [1], [2]. This process integrates top-down mechanisms, guided by our goals or expectations, and bottom-up mechanisms, driven by the most salient stimuli in a visual scene. However, studies have shown that salient distractions often capture our attention involuntarily, reducing our ability to perform tasks effectively [3]. In recent years, various mechanisms for distractor filtering have been studied, focusing on reducing the behavioral costs associated with distractors. [4], [5].

In neurophysiology, suppression refers to a reduced neural response to one stimulus compared to others. To quantify distractor suppression, neuroscientists usually calculate the difference between neural responses to distracting and target stimuli. [6] In this context, distractor suppression occurs when the neural response to a distractor is lower than the response to a target. The Distractor Positivity (Pd) is an event-related potential (ERP) component recorded over occipito-parietal electrodes, that serves as a neural marker for attentional control processes, specifically for the active suppression of distractors during visual attention tasks. The Pd component can occur in the time range between 100 and 500 msec after the onset of a distractor and is characterized by a positive deflection in the EEG signal [7]. Research has shown a direct relationship between Pd amplitude and the efficiency of distractor suppression. Greater Pd amplitudes are associated with better performance in tasks that require filtering of irrelevant stimuli, suggesting that Pd reflects a mechanism of selective attention [7]. Conversely, individuals with ADHD exhibit reduced Pd amplitudes compared to control groups, suggesting that diminished Pd activity may underlie distractor-related impairments and contribute to difficulties in suppressing irrelevant stimuli [8]. Targeting and modulating the Pd component could offer a means to enhance distractor suppression and improve attentional control in both healthy individuals and clinical populations.

One promising approach involves using neurofeedback to target and enhance the Pd component. EEG neurofeedback, which provides real-time feedback on neural activity, has gained recognition as an effective tool for cognitive modulation. Neurofeedback enables individuals to self-regulate specific brain patterns, offering the potential to enhance attentional control and decrease susceptibility to distractions [9]. However, the application of neurofeedback to the Pd component remains largely theoretical, as no studies have yet explored whether individuals can train themselves to actively increase or suppress Pd amplitudes in real time. This project aims to detect the Distractor Positivity (Pd) component in EEG signals and explore the feasibility of modulating Pd through neurofeedback, contributing to a deeper understanding of the neural mechanisms underlying distractor suppression.

The University of Texas at Austin.

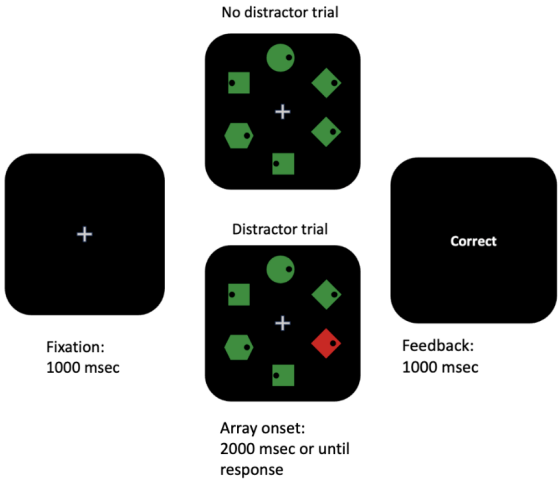

Fig. 1. Experimental Paradigm

II. Methods

This section describes the signal processing pipeline and machine learning implementation. Before addressing these components, we will first provide an overview of the experimental setup and design developed by Ana Melnichuk (Section II-A).

A. Experimental Setup

In this project, the experimental paradigm, shown in Figure 1, was designed by Ana Melnichuk to investigate neural markers of distractibility and evaluate whether subjects could modulate these markers to reduce the impact of distractors on attention.

Three subjects participated in this experiment. At the start of each trial, subjects, equipped with a 64-channel EEG cap, were seated in front of a computer screen. The horizontal electrooculogram (HEOG) was recorded using a pair of bipolar electrodes placed lateral to the external canthus of each eye, and the vertical electrooculogram (VEOG) was recorded using a pair of bipolar electrodes placed above and below the left eye. The EEG was recorded with a reference to right mastoid, electrode impedance were kept below 5 kOhms.

First, subjects were presented with a central cross on the screen and instructed to focus on it during a 1-second fixation period. Then, the screen displayed various shapes for 2 seconds or until a response was detected. In distractor trials, a red shape (distractor) appeared among green shapes, whereas in no-distractor trials, only green shapes were displayed. subjects were required to maintain their gaze on the central cross and identify the position of a dot on the green circular shape (target) by pressing either the left or right arrow key on the keyboard. They were also instructed to try and ignore the distractor.

After each trial, subjects received visual feedback for 1 second, indicating whether the neural marker for distractor suppression was detected. The word ”Correct” was displayed if Pd was detected during distractor trials or not detected during no-distractor trials. Conversely, the word ”Incorrect” appeared if Pd was detected during no-distractor trials or not detected during distractor trials. Each subjects completed one offline session and two online sessions over two days. On the first day, the offline session included 8 training runs, followed by 8 decoding runs as part of online session 1. On the second day, participants completed 8 additional decoding runs for online session 2. Each run consisted of 60 trials, evenly divided between distractor and no-distractor trials. During distractor trials, the red shape appeared on the left side of the screen in 15 trials and on the right side in the other 15.

B. Signal Processing

The data was acquired using the LOOP software developed by the Clinical Neuroprosthesis and Brain Interfaces Lab, with an antNeuro64 amplifier [10] operating at a sampling rate of 512 Hz. In the loaded data, the first 64 channels were EEG data, with channel 69 used as the trigger channel. Channels M1, M2, and EOG were excluded, leaving a total of 61 EEG channels of raw data for further analysis.

1) Temporal Filtering: The first step in our processing pipeline involved temporal filtering of the raw EEG data. To determine the optimal filter, we evaluated several types, including Butterworth, Chebyshev Type I, Bessel, and Elliptic filters. Among these, the 4th-order Butterworth filter was found to introduce the least distortion to the signal. As a result, we selected an FIR 4th-order Butterworth filter with a pass band of 1–30 Hz. Filtering was performed using two MATLAB functions: *filtfilt* for

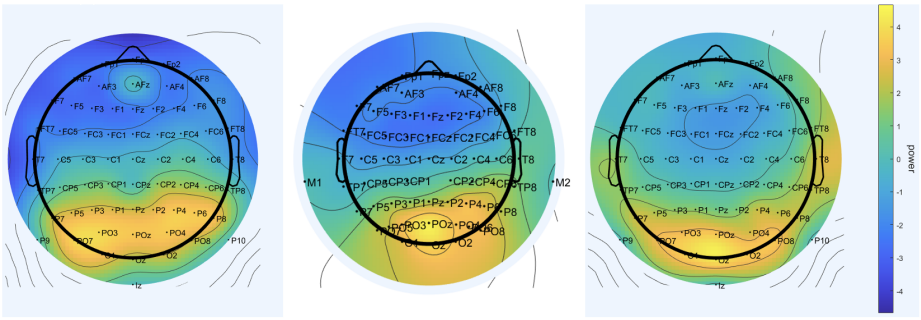

Fig. 2. Topoplot of brain activity during distractor trials (around 350 ms) of offline session for Subject 1 (left), Subject 2 (middle) and Subject 3 (right).

non-causal filtering and *filter* for causal filtering. The *filtfilt* function is inherently non-causal, applying zero-phase filtering by running the filter both forward and backward, thereby eliminating phase distortions. This non-causal filter was used to analyze offline data and accurately characterize the Pd component. In contrast, the causal filter was employed for simulating real-time analysis and developing the decoder.

2) Spatial Filtering: The second step in the processing pipeline involved spatial filtering, implemented using Common Average Referencing (CAR). This method reduces common noise and artifacts present across all electrodes by subtracting the average signal of all electrodes from each individual electrode's signal. In our context, CAR is the most suitable spatial filtering method. In fact, since the signal of interest is localized in posterior (occipito-parietal, PO, and parietal, P) electrodes, using Laplacian filtering would be inappropriate, as it could remove the signal under study. The advantage of CAR is thus its ability to reduce noise and artifact components in the EEG signals without affecting the signal of interest, given that this signal is not present across all electrodes.

One limitation of CAR is that a single channel with excessive noise or a large artifact can affect the signal across all other channels. To address this, we visually inspected the signals from all 61 channels and excluded trials exhibiting significant artifacts or noise before applying spatial filtering.

3) Data Parsing: The third step in the processing pipeline was data parsing, which involved segmenting the data into individual trials for each session. As outlined in Section II-A, each run consisted of 60 trials: 30 distractor trials, evenly split between those where the distractor appeared on the right side of the screen and those where it appeared on the left, and 30 no-distractor trials. Since visual information is encoded contra-laterally in the brain (with the left visual field processed by the right hemisphere and the right visual field by the left hemisphere), we analyzed distractor trials using electrodes contralateral to the distractor's position. For no-distractor trials, left-side electrodes were analyzed as outlined in Section II-C. Trial start times and associated information were extracted from the trigger channel. For trial parsing, a time window was defined from 200 ms before trial onset to 1 second after onset, resulting in a 1.2-second window per trial. All trials were concatenated and stored in a 3-D array.

4) Baseline Correction: The final step in the pre-processing pipeline involved baseline correction, which normalizes the EEG signal relative to a predefined baseline period. In this case, the baseline was set as the 200 ms window prior to trigger onset, corresponding to the fixation period. This step is crucial for mitigating offsets or low-frequency drifts in the EEG signal caused by noise, skin potentials, or non-stationarity. By aligning the signal to a common reference, baseline correction ensures that variations reflect neural activity rather than external factors and improves the comparability of trials with differing baseline offsets. The grand average of the trials was subsequently calculated and plotted for characterization.

C. Cortical Activation

As noted in previous research [11], [12], the physiologically relevant channels for attention and distractibility suppression are predominantly found in the parietal and occipito-parietal regions. We generated topoplots of cortical activation during distractor trials around the time of occurrence of the Pd component (around 350 ms) for offline sessions, to support our decision to focus exclusively on the specific PO and P electrodes. These topoplots also allowed us to examine the cortical activation patterns for each subject during these trials. Topoplots of brain activity during distractor trials for the three subjects are shown in Figure 2. The results clearly demonstrate significant activation in the parietal and occipito-parietal areas, aligning with prior findings.

Based on this analysis, we decided to focus on a subset of channels for signal characterization and as input features for our classifier. Specifically, we selected the P and PO channels. Additionally, because visual information is processed contralaterally in the brain, the selected channels were categorized as follows:

- Electrodes positioned over the right hemisphere were included for distractor trials where the distractor appeared on the left side of the screen. Specifically, P2, P4, P6, P8, PO4, and PO6.
- Electrodes positioned over the left hemisphere were included for distractor trials where the distractor appeared on the right side of the screen. Specifically, P1, P3, P5, P7, PO3, and PO5.
- Any of the two hemispheres can be used as reference during no-distractor trials [11], but since Ana chose to use the left hemisphere, we wanted to be consistent with that to be able to compare our decoder performance to hers during online session. Specifically, P1, P3, P5, P7, PO3, and PO5.

D. Machine Learning

The final phase of this project focused on implementing a machine learning pipeline to assess whether the Distractor Positivity (Pd) component could be decoded. The data used to design and test the model was pre-processed as discussed in Section II-B using a causal filter.

1) Training and Test Sets: For this step, the data was grouped into distractor and no-distractor trials, regardless of the distractor's position on the screen. The 3-D arrays (time × channels × trials) were flattened into 2-D arrays ((time × channels) × trials) to serve as inputs for the decoder. Label vectors were then created, assigning a label of 1 to distractor trials and 0 to no-distractor trials. With 8 runs per subject during the training session, each consisting of 60 trials (30 distractor trials — 15 with the distractor on the left side of the screen and 15 on the right —and 30 no-distractor trials), we allocated 7 runs (420 trials) for training and 1 run (60 trials) for testing on offline session data alone. After the decoder was built and assessed on offline data, the model was retrained on all offline data and tested on online session 1 and 2 data separately.

Since distractor positivity components can occur at any time during the trial and be narrow in time, down-sampling was not used. In fact, this could lead to some components being ignored/removed from the signal during the down-sampling process.

2) Decoder Model: We first implemented and compared different decoder models: Linear Discriminant Analysis (LDA), Linear Regression (LR) and Gradient Boost (GB). No significant differences were obtained among the three models performances when trained on offline data and tested on online session 1 data. The LDA decoder was finally chosen because of its simplicity, effectiveness and ease of real-time implementation.

The first step was data normalization using Z-score, which ensured that all features contributed equally to classification by standardizing them to have a mean of zero and a standard deviation of one. This step minimized biases caused by differing feature scales. LDA was then applied to separate the two classes (distractor and no-distractor) by finding a linear decision boundary. We used the *fitcdiscr* MATLAB function to implement our LDA using offline data as follows:

```
m = fitcdiscr(cepoch, trainlabels','Prior','uniform','DiscrimType','pseudolinear');
w = m.Coeffs(2,1).Linear; mu = m.Coeffs(2,1).Const;
distance = cepoch*w + mu; p1 = 0.25; p2 = 1-p1;
bcoeff1 = -log((1-p1)/p1)/prctile(distance,100*p1);
bcoeff2 = -log((1-p2)/p2)/prctile(distance,100*p2);
b = (bcoeff1+bcoeff2)/2;
```

We then computed the posterior probabilities for each epoch in the test set by applying the sigmoid function (activation function), as shown below:

$$s(x) = 1 \div (1 + e^{-b(x \times w + \mu)}) \quad (1)$$

Where $\boldsymbol{b}$ is the scaling factor of the sigmoid function, $\boldsymbol{w}$ is the vector that separates the two classes in the feature space, and μ is a constant for the discriminant function distinguishing Class 1 from Class 2. With the posterior probabilities calculated, we needed to set a threshold to classify each epoch as either Class 1 (Pd detected) or Class 2 (Pd not detected). For this, we chose a threshold of 0.5. The threshold value was established through an iterative process of trial and error, with the chosen value yielding the best results.

To evaluate performance, we performed 8-fold cross-validation, dividing the data into eight subsets for iterative training and testing. In each iteration, one run was used as the test set, while the remaining seven were used to train the decoder. This approach preserved the temporal structure of the EEG data, ensuring that the model generalized well and avoided overfitting. Additionally, we conducted a chance-level analysis by shuffling class labels and recalculating accuracy, establishing a baseline for comparison to validate the decoder's performance.

Finally, the full offline dataset was used to train the final model, which was saved along with the normalization parameters for future use. This approach ensures consistent application of the decoder to new data while maintaining accuracy.

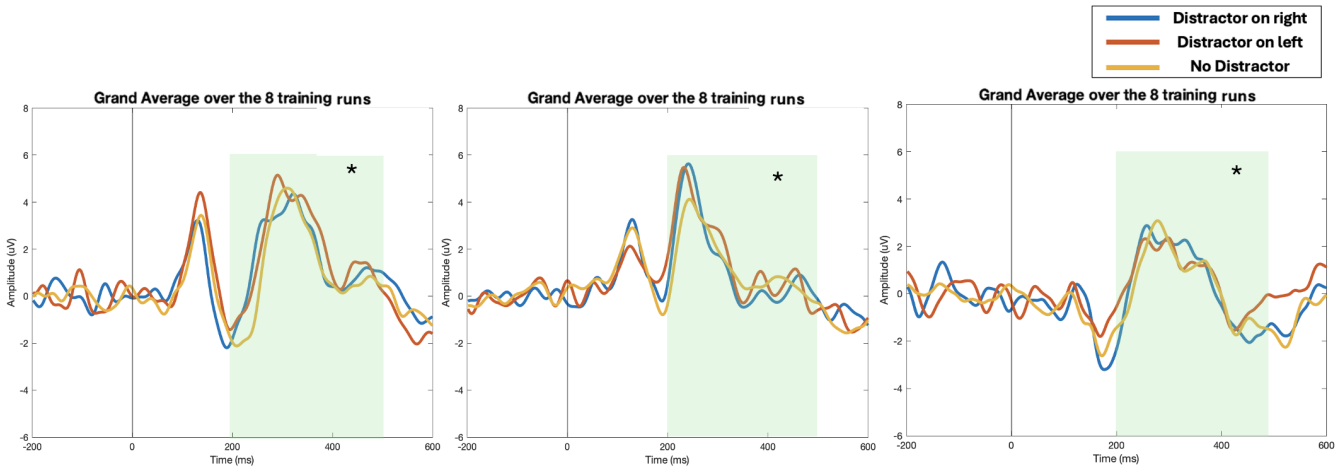

Fig. 3. Time-domain grand average signals for the offline session: no-distractor trials (yellow), distractor trials on the left side of the screen (red), and distractor trials on the right side (blue) for Subject 1 (left), Subject 2 (middle) and Subject 3 (right). t=0 indicates trial onset. The highlighted regions show the final window used for analysis (200-500ms). * $p = 0.02 \pm 0.01$

III. Results

A. Signal Characterization

The raw data was filtered using temporal and spatial filtering techniques, following the pipeline outlined in Section II-B but using a non-causal filter.

1) Signal Visualization: For signal visualization, shown in Figure 3, we applied non-causal filters to the data from the offline session. This session was chosen for visualization because participants were not instructed to modulate brain activity, allowing us first to see if we can detect the distractor positivity component (Pd) in the neural data, and second, to compare it with the online sessions.

After visually inspecting the offline sessions for all three participants (Figure 3), we determined the regions of interest. This refers to regions where the Pd component was identified, appearing the third positive deflection. A positive deflection is defined as any peak observed during distractor trials (positive or negative in amplitude) that is more positive compared to no-distractor trials. An initial time window from 300 ms to 350 ms after trial onset (time 0) was analyzed, solely to cover the area where the Pd component was observed. A two-sample t-test conducted between the distractor and no-distractor classes yielded statistically significant results, with $p = 0.008^{**} \pm 0.001$.

The 300-350 ms window was determined to be too narrow to account for other components that may change during online sessions while still reflecting distractor suppression mechanisms. Additionally, subject-to-subject and run-to-run variability reduced performance when the bounds were restricted to this range. As previously noted, the Pd component can appear within a range of 100–500 ms [7]. Thus, a window from 200 ms to 500 ms after trial onset was used for the remaining of the analysis. A two-sample t-test was also performed for this second window to further support our choice, yielding $p = 0.02^{*} \pm 0.01$.

To ensure a valid real-time analysis, a causal filter was applied during the subsequent analysis. The non-causal filter used for visualization purposes was compared to a causal filter (Butterworth 4^{th} order) to evaluate whether the 200–500 ms window remained appropriate, considering potential phase shifts in the signal. The results indicate that this time window may be preserved when using causal filters. To simulate real-time analysis, causal filtering was applied to both the training and testing data to design and evaluate the model. The training data for the machine learning model was taken from the offline session, while the model was tested using data from online Sessions 1 and 2.

B. Online Sessions Modulations

To investigate whether subjects had been modulating Pd during online sessions, a comparison between sessions was conducted. For that, we calculated the difference potential between contralateral and ipsilateral electrodes to the distractor position. This difference waveform is used in literature to detect and quantify Pd [13]. Subject 1 results are shown in Figure 4, where an increase in Pd amplitude during both online sessions can be seen. These results are also observed for the other subjects.

We also wanted to determine whether the modulation was specific to Pd or if it also affected other distractor suppression components. As shown in Figure 5, the results show changes in amplitude for components after 400 ms for subjects 1 and 2.

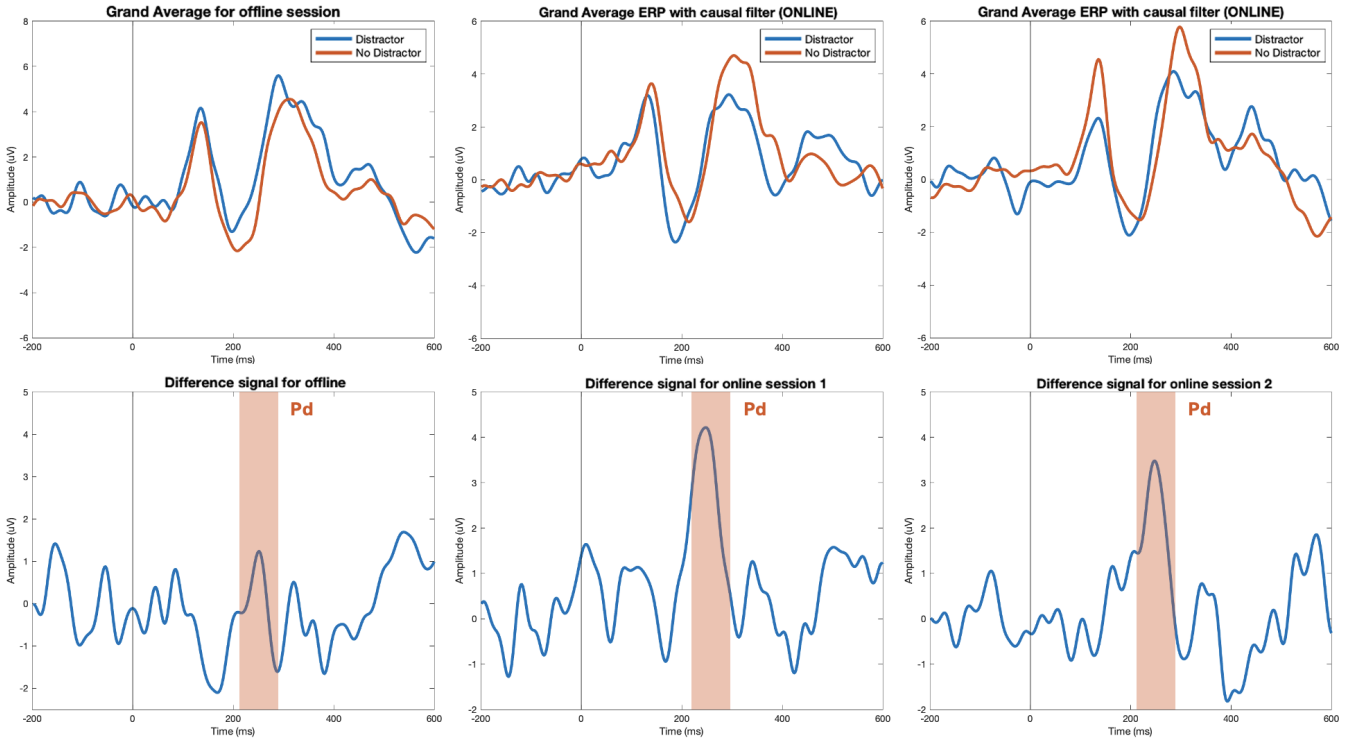

Fig. 4. Grand Average signals for distractor (blue) and no-distractor (red) trials (top) and difference potentials (bottom) for offline (left), online Session 1 (middle) and online Session 2 (right) sessions for Subject 1.

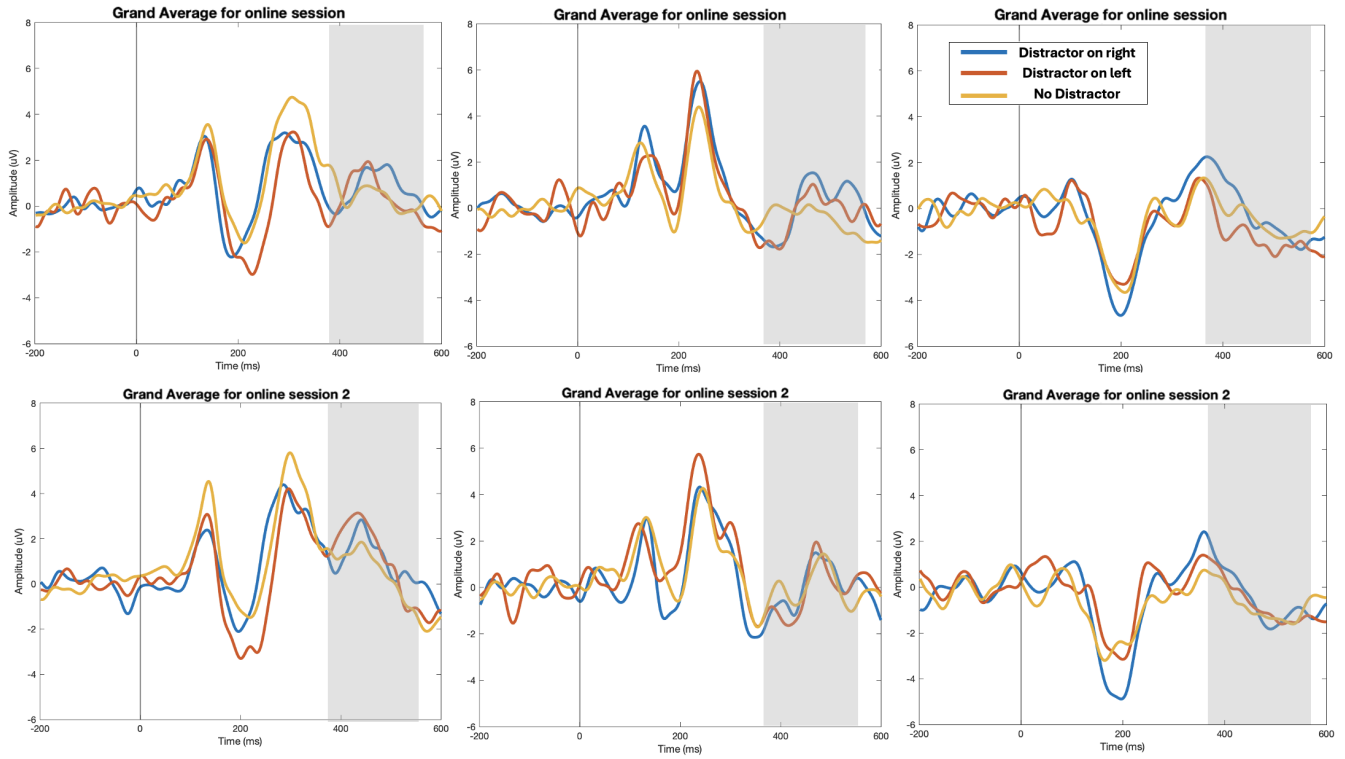

Fig. 5. Grand average signals for online Session 1 (top) and online Session 2 (bottom) for Subject 1 (left), Subject 2 (center) and Subject 3 (right). The gray areas show modulation of additional distractor positivity components.

This finding indicates that distractor suppression involves neural processes beyond the Pd region and that other neural markers may also be modulated.

TABLE I
Test accuracies of different models.

	LDA (%)	LR (%)	GB (%)
Subject 1	57.39	56.4	58.2
Subject 2	59.53	54.1	53.8
Subject 3	57.91	53.26	51.94

TABLE II
Test accuracies of the LDA model trained on offline data.

Subjects	k-fold (%)	Online 1 (%)	Online 2 (%)	Chance Level (%)	Ana's model (%)
1	65.43	57.39	46.78	49	47.92
2	70.25	59.53	42.1	48	42.67
3	64.1	57.91	43.622	49	47.50

TABLE III
Test accuracies of the LDA model trained on offline and online Session 1.

Subjects	Online 2 (%)
1	56.36
2	59.78
3	57.27

C. Decoder Performance

First, we compared different decoder architectures: LDA, Logistic Regression (LR) and Gradient Boosting (GB). Using the offline data to train the model and testing on online data of Session 1, after Z-score normalization, the results can be seen in Table I.

Seeing that the three decoders have similar performances when used on online data, we decided to choose LDA for its simplicity. Thus, an LDA binary classifier was utilized using k-fold cross validation. Data was kept sequentially consistent during the trial due to non-stationarity of the signal. The input to our decoder was all time points from 200 ms to 500 ms after trial onset.

The k-fold cross validation scheme was then compared with a chance-level analysis using randomized label shuffling to understand whether the performance of the classifier was better or equivalent to the chance level for the model.

Table II summarizes the results for all 3 subjects for our decoder when trained only on the offline session data, as well as the performance of the decoder used during the actual online sessions for comparison (denoted as Ana's model).

Now, when using both the offline and first online sessions to train the decoder, we get new performances when tested on data from online Session 2 as shown in Table III. We observe significant improvements in performance during online session 2.

IV. Discussion

We investigated detection and modulation of the Distractor Positivity (Pd) component through neurofeedback.

A distractor-elicited voltage difference was observed between 200-300 ms after trial onset, denoted as the Pd component. This positivity was detected in both offline and online sessions, confirming the extractability of Pd, as shown in Figure 4. We also saw an increase in Pd amplitude during online sessions 1 and 2, further affirming its role in distractor suppression. It shows the brain's ability to dynamically shift between enhancing relevant targets (e.g., N2pc) and suppressing irrelevant stimuli (Pd), which is essential for efficient attentional control. This also shows the ability of subjects to modulate this neural marker during online sessions. However, we had concerns regarding subject-to-subject and run-to-run variability, as well as not including all distractor suppression components when only looking at Pd. Gaspar and McDonald [14] demonstrated that the detection window for the Pd component could be extended, while Gaspelin, N [7] further noted it can appear up to 500 ms after stimulus onset.

Distractor positivity refers to any positivity in neural signal during distractor trials compared to no-distractor trials. From Figure 5, there is an emergence of a new positivity component (shaded gray area) that was consistent across the 2 online sessions. We hypothesize that this is related to distractor suppression modulation since cognitive processes often occur at a later time after trial onset. It could be easier for subjects to modulate this later peak since they need time to achieve cognitive control. With all that said, we decided to widen our decoding window up to 500 ms after trial onset. With this new window,

the results of the t-test yielded a p-value of $p = 0.02^{*} \pm 0.01$, reinforcing our choice for the decoder input and the significance of other peaks in the distractor suppression model.

The decoder's performance offers valuable insights into the feasibility of using machine learning for real-time neurofeedback. In offline sessions, the LDA decoder consistently performed above chance ($\geq 60\%$), validating the reliability of the distractor positivity components as features for classification. During the first online session, the decoder maintained a performance above chance ($\geq 55\%$) for all subjects, indicating that the offline-trained model could generalize effectively to next day data. However, a significant drop in performance ($\geq 47\%$), below chance level, was observed during the second online session for all subjects. Despite this decline, our results align with those recorded during the actual online session, as indicated by "Ana's decoder" in Table II. The drop in performance could be attributed to various factors, such as variability in participants' engagement and non-stationarity in neural signals across different days.

To address this issue, we trained the decoder using a combination of data from the offline and online session 1. This approach resulted in improved performance, when tested on online Session 2. Our hypothesis for this increase in performance is that, since we saw modulation in the signal, specifically in later peaks not present during offline sessions, adding this information to the decoder's input would help it predict online session 2 modulations. Thus, incorporating data from online sessions during BCI decoding enhances performance because it includes data collected while the subject is actively modulating the target brain signals in real time, rather than relying solely on task execution data. This allows the decoding model to better capture the specific brain activity patterns associated with intentional modulation, improving accuracy and system performance.

Another approach to address day to day drifts in signals would be to use a Reimannian classifier to project data from different days into a unified geometric space before passing it to the decoder. This method could help minimize the effects of non-stationarity and improve the decoder's consistency across sessions [15].

Signal processing choices were also critical to the outcomes of this study. Temporal filtering using both causal and non-causal filters allowed for a detailed characterization of the signals while maintaining real-time applicability. Spatial filtering via common average referencing (CAR) effectively reduced noise and artifacts without compromising the Pd signal localized in posterior electrodes.Our analysis also confirmed the importance of using contralateral channels to the distractor's position on the screen, consistent with the physiology of visual information processing.

Regarding behavioral analysis, we could look at two parameters : Reaction Time (RT) and Behavioral Response. Reaction time refers to the time interval between the presentation of the stimulus and the subject's corresponding response, here the pressing of the key. It is often used as a behavioral measure to evaluate cognitive processing speed, task performance, and attentional engagement during BCI experiments. Behavioral Response refers to the observable and measurable actions or outputs produced by participants in response to a task or stimulus. Here, it would be if we pressed the button correctly as well as modulating brain activity.

In order to quantify the first parameter, we would have compared RT between offline and online sessions. We hypothesize that we should see a decrease in reaction time during distractor trials until they are close to the RT recorded during no-distractor trials. This would show that we are suppressing the distractor and its presence is not affecting behavior anymore. AS for behavioral response, we could look at the trial output, specifically what key was pressed and if the answer was correct (e.g pressing the right arrow key when the dot is on the right in the target shape) or incorrect. We hypothesize that the fraction correct responses should increase as the sessions progress.

However, we did not have access to RT information in the data, only trigger and classification outputs. Also, during the online sessions, we were asked to focus on modulating our brain activity and disregard what key we were pressing and take our time to try and modulate. So some subjects just pressed keys at random and later in the trial while focusing on Pd modulation and classifier outputs. So the trial output as well as RT would not reflect behavioral changes.

V. Conclusion and Future Work

This study successfully demonstrated the detection of the Distractor Positivity (Pd) component, offering valuable insights into its role in distractor suppression. Pd was consistently observed across participants in offline and online sessions, with significant differences validating its relevance as a neural marker.

This study revealed a significant positive peak around 450 ms. This suggests that distractor positivity components are not confined to a single temporal region or marker but may manifest at different times, reflecting broader neural processes involved in distractor suppression. Future work should explore the functional relevance of these additional regions of positivity and investigate their relationship to attentional control mechanisms.

The LDA decoder performed reliably above chance in offline and initial online sessions, confirming the utility of Pd for classification. A performance drop in the second online session highlighted challenges such as signal variability across days. However, training the decoder with combined session data improved generalization. This paves the way for future research into the role of distractor positivity components in selective attention, with potential applications in enhancing neural decoding techniques and developing clinical interventions for attention-related conditions.

A key direction for future research is to look at all positivity components and their roles in distractor suppression mechanisms. By analyzing the neural dynamics of these regions, we can determine whether they represent distinct components or extensions of the Pd response. Furthermore, integrating data-driven approaches, such as clustering or decomposition methods, may help identify temporal regions of positivity without predefining strict windows.

Finally, conducting additional practice sessions or obtaining more data could have strengthened the conclusions drawn in this study. Future research should also focus on evaluating the generalization of these findings in larger and more diverse participant groups. Examining how these additional positive peaks manifest in populations such as individuals with ADHD, other attentional disorders, or older adults could provide valuable insights and pave the way for clinical applications of Pd-based neurofeedback.

ACKNOWLEDGMENT

The authors would like to thank Deland Liu and Prof. Jose del.R Millan for their feedback and availability during the semester. Also, the authors would like to thank Ana Melnichuk for her continuous feedback, time, answering our questions and inquiries about the project as well as conducting the experiments.

VI. CONTRIBUTIONS

- **Alex:** I did my recording session and assisted my teammate session. Me and Arman worked on the MATLAB code equally, developing our own implementation of all the pipeline steps and comparing results before choosing the best way to go. I worked on the difference waveforms and trying the different time windows. I implemented in addition to the pipeline the CAR and t-test analysis. I also implemented the LDA classifier, data reshaping and activation function for testing. Arman also created the topoplots and tested different ML models. We also helped each other with debugging when something was not working and had a great duo experience. Me and Blanca wrote the different parts of the report, adjusting the sections where the other left comments. I also wrote the whole report in the IEEE format on overleaf and adjusted all the figures and formats.

- **Arman:** I conducted my recording session and helped with my teammate's session. I worked on the code in parallel with Alex so that we may best understand the full BCI pipeline. I generated topo-plots for the offline analysis, generated different ML models for analysis, and conducted the chance-level analysis. I also implemented CCA, however, the results were notably worse than CAR so we decided to move forward with CAR. In addition to developing the full pipeline in parallel, Alex was able to achieve better results, so we ended up going with her results for the purposes of the report. She also focused on the t-test, comparison of the contra/ipsilateral signals in distractor/no distractor conditions. Alex and Blanca contributed more heavily to the report. I drafted several sections pertaining to the results. I also provided feedback and edits for the other sections.

- **Blanca:** I recorded my own sessions for the experiment and assisted in one of my colleague's sessions. Regarding the code, I didn't contribute to the programming. I spent time understanding the code and the decisions my colleagues made so that I would be prepared for the presentation. Arman shared his code and made himself available for any questions or help I might have needed. As for the report, Alex and I worked on most of the writing. I wrote several sections from scratch, gave feedback for possible improvement on Alex's parts, and continuously reviewed the entire document for language and style consistency.

REFERENCES

[1] Forster, Sophie, and Nilli Lavie. "Failures to ignore entirely irrelevant distractors: the role of load." Journal of Experimental Psychology: Applied 14.1 (2008): 73.

[2] Marini, Francesco, Leonardo Chelazzi, and Angelo Maravita. "The costly filtering of potential distraction: evidence for a supramodal mechanism." Journal of Experimental Psychology: General 142.3 (2013): 906.

[3] Theeuwes, Jan. "Irrelevant singletons capture attention." Neurobiology of attention. Academic Press, 2005. 418-424.

[4] Chelazzi, Leonardo, et al. "Neural basis of visual selective attention." Wiley Interdisciplinary Reviews: Cognitive Science 2.4 (2011): 392-407.

[5] Chelazzi, Leonardo, et al. "Getting rid of visual distractors: The why, when, how, and where." Current opinion in psychology 29 (2019): 135-147

[6] Moran, Jeffrey, and Robert Desimone. "Selective attention gates visual processing in the extrastriate cortex." Science 229.4715 (1985): 782-784.

[7] Gaspelin, Nicholas, et al. "The distractor positivity component and the inhibition of distracting stimuli." Journal of cognitive neuroscience 35.11 (2023): 1693-1715

[8] Rodríguez-Martínez, Elena I., et al. "Neurophysiological differences between ADHD and control children and adolescents during the recognition phase of a working memory task." Neuroscience Research 164 (2021): 46-54

[9] Enriquez-Geppert, S., R. J. Huster, and C. S. Herrmann. "EEG-neurofeedback as a tool to modulate cognition and behavior: a review tutorial. Front Hum Neurosci. 2017; 11: 51." This review gives an overview of the key aspects relevant to the development of neurofeedback protocols and discusses the relevant aspects for the practical application of neurofeedback (2017).

[10] ant, "Copyright © ant-neuro.com 2024", "ant neuro","https://www.ant-neuro.com/products/eego-mylab"

[11] van Moorselaar, Dirk, et al. "Neural mechanisms underlying expectation-dependent inhibition of distracting information." elife 9 (2020): e61048.

[12] Noonan, MaryAnn P., et al. "Selective inhibition of distracting input." Behavioural brain research 355 (2018): 36-47.

[13] Gaspelin, Nicholas, et al. "The distractor positivity component and the inhibition of distracting stimuli." Journal of cognitive neuroscience 35.11 (2023): 1693-1715.

[14] Gaspar, John M., and John J. McDonald. "Suppression of salient objects prevents distraction in visual search." Journal of neuroscience 34.16 (2014): 5658-5666.

[15] Tibermacine, Imad Eddine, et al. "Riemannian Geometry-Based EEG Approaches: A Literature Review." arXiv preprint arXiv:2407.20250 (2024).