A replication study on the relationship between single-trial N200 peak latencies and visual encoding time

Parineeta Ekhande
Supervisor: Dr. Micheal D. Nunez
Amsterdam Mathematical Psychology Lab
Faculty of Social and Behavioral Sciences

UNIVERSITY OF AMSTERDAM

Index

Abstract

Sequential sampling models provide underlying insights into human decision-making processes, especially in conditions of uncertainty. Fitting these models to participants' choices and reaction times provides insights into the steps humans undergo before, during, and after making a decision. Sequential sampling models rely on the idea that perceptual decisions are made possible by the process of evidence accumulation. Recent evidence from Nunez et al. (2019) suggests a positive correlation between the peak latency of an N200 event-related potential and Visual Encoding Time (VET; the period of time that precedes information processing), an increase in the former corresponding to an increase in non-decision time (NDT) – an independent estimate that is thought to include VET. A replication study by Pinier (2022) attempted to provide further evidence for this result by replicating the findings in three different experimental conditions (two-back continuous performance task, gabor patches embedded in visual noise, and varying spatial frequency and random dot motion task). In the present study, we extend on Nunez et al. (2019) and Pinier (2022)'s study by calculating single-trial estimates of N200 peak latencies for an n-back task (Swann et al. 2007). This was done by applying the method of Singular Value Decomposition (SVD) to every individual trial. We applied linear regression analysis using the subject's Reaction Time as an estimate of non-decision time (which is also the VET). The analysis provided no evidence for this relationship in the n-back task. However, various limitations underpinned this result, and thus, it is quite early to draw any definite conclusions about this relationship, necessitating future research.

Keywords: N200, event-related potentials, decision-making, single-trial ERP estimation, electroencephalography (EEG), visual encoding time, n-back task

Introduction

As humans, we are exposed to numerous situations wherein we have to make small decisions. These could be trivial, like which TV show to watch or what to have for dinner. Despite the triviality of the decisions, we make them after we think we have collected enough information to make the said decision. As intuitively as it comes to us, the process itself is complex. The quality of data accumulated determines the quality of the decision (Forstmann et al. 2016). So, how do we gather information from our surroundings? Ratcliff and McKoon (2008) propose that decisions are made based on sequential sampling of information. The theories devoted to understanding the sequential sampling process presume that the gathered information is often noisy and imprecise. We continue to gather imprecise bits of information till a certain threshold of evidence is reached, and such accumulation-to-threshold models are referred to as sequential sampling models (Forstmann et al. 2016)

One such sampling model is the Drift Diffusion Model (DDM). It relies on the notion that perceptual decisions are made based on the process of evidence accumulation (Pinier, 2022) and is also the model of interest for the current study. The DDM explains Reaction Time (RTs) in individual trials. Most research involving decision-making experiments involves stimulus presentation and participant responses. RT is believed to encompass cognitive functions beyond mere evidence accumulation, and it comprises of visual encoding time, response preparation period, perceptual information encoding, motor execution period, and motor execution (Nunez et al. 2019; Weindel, 2021; Pinier, 2022).

Visual Encoding Time (VET) is one potential cognitive function to help understand the process of information accumulation. It is the period that precedes information processing in visual decision-making tasks. VET typically spans around 150 ms to 200 ms following the presentation of a stimulus and is believed to represent cognitive processes such as figure-ground segregation (Nunez et al. 2019). However, the duration can vary based on factors like visual noise and levels of attention. Nunez et al. (2019) provide additional evidence indicating a potential correlation between the peak latency of the N200 component (a negative peak in EEG recording occurring between 150 and 275 ms after stimulus presentation) and VET.

The authors demonstrate that longer durations of the N200 component correspond to proportional increases in non-decision time (NDT), an independent estimation that encompasses VET and is derived from cognitive models of reaction times. Nevertheless, this was only shown in one type of task.

The relationship was further explored by Pinier (2022), wherein the author attempted to validate this relationship in three perceptual decision-making tasks: two-back continuous performance task, gabor patches embedded in visual noise, and varying spatial frequency and random dot motion tasks. The relationship did not reach statistical significance in either datasets. However, a weak negative relationship for the first dataset (n-back task) ($\beta = -0.14$) and a strong positive relationship for datasets two and three (gabor patch and random dot motion task) ($\beta = 0.852$ and $= 0.505$ resp.) was observed. In order to further test the theory, the current study attempts to replicate the results of Nunez et al. (2019) on a single-trial basis on the above-mentioned datasets. The evidence provided by Nunez et al. (2019) and Pinier (2022) provides the foundation for testing this relationship, and we aim to re-inspect this using a different methodology; however, our hypothesis remains the same: that single-trial peak latencies of N200 ERP will tract visual encoding time.

The N200 ERP and VET

An EEG recording can hold the summed-up data of neural activity elicited by an external trigger. The advantage of EEG is that it can capture neural activity on a millisecond timescale, making it suitable for analyzing rapid cognitive processes. This allows the extraction of single-trial data evoked in response to a specific event or stimuli in an individual experimental trial.

Responses to specific events or stimuli are known as Event-Related Potentials (ERPs). The N200 ERP is a deflection in the EEG that happens around 200 ms after the presentation of a stimulus. Loughnane et al. (2016) distinguished two sets of N200 ERP latencies after changes in visual stimuli. They demonstrated that these N200 latencies play a role in influencing the onset of evidence accumulation in the context of a task involving a random dot motion task. The N200 ERP has also been associated with decision-making processes. In tasks involving risky choices (such as gambling tasks), a negative N2 deflection can be seen during the occurrence of unexpected events in the task (Mushtaq et al. 2016).

These negative deflections are also characteristic of Feedback Related Negativity (FRN) as they both peak after stimulus presentations (Wang et al. 2021). The N200 is also shown to be strongly influenced by conflict (Enriquez-Geppert et al. 2010). Thus, we can conclude that N200 is involved in outcome evaluation and learning.

In the Nunez et al. (2019) study, the authors provide concrete evidence for the relationship between the N200 component and visual encoding time (VET). They defined VET as the time taken before the decision-making process begins, lasting 150-200ms based on noise, subject attention, etc. Their study showed N200 latencies to reflect the completion of visual encoding and the onset of evidence accumulation. The authors found that N200 deflection times were more pronounced in participants who made quicker decisions or were more confident in their decisions. They did not find a 1ms-to-1ms relationship between single-trial N200 latencies and VET. Their motivation for investigating the cognitive processing behind single-trial estimates of N200 latencies predates to their Nunez et al. (2017) study. The authors proposed that single-trial evidence accumulation rates can only be estimated using single-trial evoked potentials. Non-decision time was defined as the time frame within a subject's reaction time that is not associated with decision-making. A positive linear relationship was found between single-trial non-decision times and latency of N200 following stimulus onset (in low noise conditions). N200 amplitudes were also positively related to evidence accumulation rates per trial.

More evidence for the N200-VET relationship has been explored using rapid serial visual presentation (RSVP) tasks, which demonstrated that the N200 component is modulated by the presentation rate of successive stimuli (Luck & Hillyard, 1994; Vogel et al. 1998). Thus, N200 ERP and VET reflect cognitive processes in the initial visual perception and stimulus evaluation stages. The evidence so far points towards a direct relationship between N200 latency and VET, and the former positively correlated with non-decision time in low and high visual noise conditions (Nunez et al. 2017; 2019), although limited to only the gabor patch task. The pertinence of studying the relationship between N200 and VET also arises from the need to understand the brain's visual information processing mechanism. Thus, the amount of support warrants verifying this relationship in different experimental conditions that the present study proposes to do.

Material and Methods

The data has been obtained from three openly available EEG datasets. The condition for selecting a dataset was a two-alternative forced-choice task in the visual modality with a precise stimulus onset time and recordings of RT (Pinier, 2022).

Dataset 1:

Data from a Reward Two Back study was analyzed to answer the N200-VET hypothesis. The dataset is publicly available online at https://headit.ucsd.edu/studies/ba6d6d12-a236-11e2-9420-0050563f2612. The Swartz Center for Computational Neuroscience collected it at the University of California San Diego and has been utilized in several published works, including studies by Mullen et al. (2010), Onton & Makeig (2006), Onton & Makeig (2008), Onton & Makeig (2009), Swann et al. (2007), Thompson et al. (2011), and Pinier (2022) The dataset consisted of 25 participants in good health (11 females and 14 males), whose ages ranged from 19 to 43 years (average age of 24.3, standard deviation of 4.7). Among them, twenty-two were right-handed, while three were left-handed. For the present study, five participants from this dataset were excluded due to a lack of stimulus-onset time.

This dataset focused on exploring the cognitive effects of positive and negative reinforcement. The task is a two-back continuous performance task with auditory feedback and reward/punishment. In this task, participants were required to indicate whether the current letter presented on the screen was the same as the letter presented two letters earlier by pressing a button.

If participants responded correctly within the given time frame, they received positive ('correct') feedback in the form of a tone, and 1 point was added to their total points (as hypothetical money added to their pocket). On the contrary, if they answered incorrectly within the given time, they received negative ('wrong') feedback in the form of a tone, and 1 point was subtracted from their total points (hypothetical money taken out of their pocket). In 20% of correct trials and 30% of incorrect trials, a bonus or penalty tone was given instead of the standard tone, signaling that the points added or subtracted would be two instead of one, thus representing a 'bonus' or 'penalty' consequence for that particular trial. A rare reinforcement neutral tone was delivered for 6% of the trials to create an oddball effect.

For EEG data acquisition, a BioSemi cap of 256 electrodes with a sampling rate of 256 Hz was used. The amplifier type was not specified in the experiment documentation.

Dataset 2:

The second dataset was of a perceptual decision-making task utilizing Gabor patches. The dataset was acquired from an unpublished study at the University of California, Irvine. The primary aim of this study was to investigate the temporal aspects of various mental processes involved in decision-making (Lui et al. 2018.). The dataset encompassed 49 healthy individuals (29 females and 20 males) aged 18 to 25 years (average age of 20.2, standard deviation of 1.5); all were right-handed.

In this dataset, the participants were exposed to Gabor patches set against a backdrop of visual noise. These patches displayed variations in both their orientations and spatial frequencies. The study comprised of two tasks, each featuring three distinct tiers of difficulty.

The first task, named the "Signal task," aimed to classify Gabor patches into low or high spatial frequency categories. These patches were shown with orientations of either 45 or 135 degrees. The difficulty increased across the three levels by reducing the difference in spatial frequency between the two categories:

For the "Easy" level, low-frequency patches were shown at a rate of 2.35 cycles per degree (cpd), while high-frequency patches were presented at 2.65 cpd (a difference of 0.3 cpd).

At the "Medium" level, low-frequency patches were showcased at 2.4 cpd, while high-frequency patches appeared at 2.6 cpd (a difference of 0.2 cpd).

For the "Hard" level, low spatial frequency patches were displayed at 2.45 cpd, and high spatial frequency patches were showcased at 2.55 cpd (a difference of 0.1 cpd).

The objective of the second task, known as the "Signal-Response (SR) Mapping" task, was contingent upon the block’s difficulty level. The spatial frequency of Gabor patches remained constant for the two categories across various block types. Both low and high-frequency patches were exhibited at 2.4 cpd and 2.6 cpd, respectively. The task's goal differed as follows:

For the "Easy" level, participants were instructed to press a button whenever they detected any Gabor patch.

At the "Medium" level, participants were tasked with distinguishing between low and high-frequency patches by pressing one of two buttons.

In the "Hard" level, participants needed to differentiate patches based on both spatial frequency and orientation by pressing one of two buttons. For instance, one button corresponded to patches with both high spatial frequency and a 45-degree orientation. In contrast, the other button corresponded to patches with low spatial frequency and an orientation of 135 degrees.
The recording for this dataset was done with Electrical Geodesic Inc.'s 128-channel Geodesic sensor net with a Net Amps series amplifier with a sampling rate of 1000 Hz.

Dataset 3:
Participants participated in a task involving random dot motion in this dataset. The third dataset was collected during a study at the University of Amsterdam, focusing on the contingent negative variation (CNV) and its correlation with response caution (Boehm et al. 2014). In the dataset, there were 25 healthy participants (17 females and 8 males) whose ages ranged from 17 to 25 years (average age of 21.3, standard deviation of 1.67). Among these, twenty-four were right-handed, while one was left-handed.
In each trial, the goal was to determine if an array of moving dots was shifting either to the left or the right on the screen. This arrangement combined two types of dots that combined to form a boundary-free circle. One type of dot moved independently in random directions, while the other type moved cohesively, either to the left or right.
The task had two distinct conditions: accuracy trials and speed trials. During accuracy trials, participants were expected to respond with high precision. Conversely, in the speed trials, they were instructed to respond as rapidly as possible in the speed trials. Each trial displayed the dot arrangement for 1.5 seconds, and feedback was given visually after every response. The subjects were shown 'AC/SP' to indicate whether the trial was either speed or accuracy trial. The subjects had to determine the motion of the randomly moving dots. For the AC trials, for a correct response, "correct" feedback was provided for correct responses, and "incorrect" feedback was provided for incorrect responses. For SP trials, the feedback "in time" was displayed if they responded within a determined time of stimulus onset, and the feedback "too slow" was displayed if they did not respond within the determined response time.

The study utilized a TMS amplifier system and electrode cap (Electro-cap, Inc.) of 32 electrodes with a sampling rate of 500 Hz. EOG (Electroculography) activity was sampled via four electrodes, two placed on each eye's outer extremity (canthi) of each eye, and the other two were placed above and below the left eye.

A review of estimation of ERPs from single-trials

An EEG recording is a combination of muscle, biological factors, brain oscillations, and other artifacts (Britton et al. 2016). In order to extract valid signals and task-related activity, it needs to be cleaned of artifacts. This is achieved by the data undergoing several steps of pre-processing. Preprocessing allows to extract averages of EEG signals across experimental trials time-locked to specific events such as stimulus onset time (Luck et al. 2000) (e.g., Ghaderi et al. 2022). However, there are two main disadvantages of employing trial-averaged EEG signals. 1. Data loss is possible in each trial. 2. To test a hypothesis, a large amount of data is required (Zhao et al. 2005; Boudewyn et al. 2018). Thus, the alternative to overcome these disadvantages is to analyze single-trial EEG data (Bridwell et al. 2018).

Single-trial EEG data evoked in response to a specific event or stimuli can be extracted by mapping the data explicitly evoked in response to the said event. This response, i.e., the ERP synchronized to the onset of a stimulus, is typically computed by averaging EEG signals across trials aligned with the stimulus presentation. Even though ERPs offer a unique perspective into the cognitive processes underlying neural responses to specific events, their accurate estimation from single-trials is challenging. Subsequently, there have been many approaches to overcome these challenges (Zang et al. 2021; Ranjbar et al. 2018; Iyer et al. 2007; Li et al. 2007, 2009).

Averaging EEG signals time-locked to events has been a cornerstone in ERP estimation. It clearly represents the ERP component, enabling the analysis of trial-specific variability. However, this traditional approach hinges on acquiring a substantial number of trials for reliable outcomes. Averaging does not help account for data of a single trial’s sensitivity to noise. In response to this limitation, a paradigm shift towards leveraging the power of machine learning and Bayesian techniques to enhance ERP estimation accuracy from single-trials was adopted.

Zang et al. (2021) highlighted the potential for improved ERP detection accuracy, especially in low signal-to-noise ratio conditions. However, the approach still requires rigorous feature selection of ERP components and model tuning to ensure the accuracy of results.

Research has shown that the N100, P200, N200, and P300 components are commonly observed in event-related potentials in response to target stimuli. These components can be scored in terms of peak amplitude and peak latency using a previously published baseline-to-peak method (Luck et al. 2000). Estimating these ERP components' peak amplitude and latency on a single-trial basis has become a popular method for investigating inter-trial variability and dynamic cognitive processes. Subsequently, Ranjbar et al. (2018) presented an innovative method wherein they proposed a modified spatiotemporal filtering technique to increase the accuracy of ERP extraction. Their method allows the analysis of variability observed in different trials and participants, providing details of the intricate dynamics contributing to ERP responses. However, the technique still requires meticulous handling of correlations between the ERP components, and the accurate estimation of the polarity of peaks (P200, N200) still remains a challenge that necessitates continued refinement.

Another approach adopted by Iyer et al. (2007) investigated the efficacy of signal processing techniques. They studied *Iterative* Independent Component Analysis (*i*ICA) and wavelet transform (WT) denoising to enhance single-trial evoked potential estimation. Even though ICA can separate mixed EEG sources aiding in noise reduction, *i*ICA cannot extract evoked potential from actual EEG data (their methodology performed better on simulated data). Furthermore, even as WT denoising preserves temporal information while suppressing noise, the optimal selection of wavelet parameters is imperative for increasing denoising's effectiveness, which is a challenge.

Spatio-temporal filtering methods, proposed by Li et al. (2009a, 2009b), offer a promising avenue for enhancing single-trial ERP estimation. However, this approach has challenges, particularly its dependence on quality spatial filters and the condition of having uncorrelated ERP components for optimal feature modeling (Ranjbar et al. 2018).

Even though single-trial EEG data can provide relevant information for any particular trial, it is not entirely devoid of artifacts and noise, resulting in variability across trials (Ghaderi et al., 2022). To account for this variability, a promising technique, Single Value Decomposition, boosts SNR for single-trial ERP estimates. SVD, a matrix factorization method, shows potential in extracting robust ERP features from noisy EEG data (Nunez et al. 2017; 2019).

By decomposing the EEG data matrix into singular values and corresponding spatial patterns, SVD enables the separation of signal and noise components more effectively than traditional methods.

SVD for Single-Trial N200 ERP Estimation

ERPs evoked in response to the onset of a stimulus are measured typically by averaging EEG signals across trials aligned with stimulus presentation. Event-related components have been previously identified as indicators of Non-Decision Time (NDT; Nunez et al. 2019). The latencies and amplitude of the N200 are used as independent measures of NDT.

The limitation of ERPs is that they cannot be directly measured on a per-trial basis using individual electrodes due to limitations in signal-to-noise ratios (SNR). Although raw EEG signals could serve as single-trial measures, they often possess low SNR for task-specific brain responses. Nunez et al. (2017) provide a reliable explanation to address this issue to boost SNR. They did so by estimating single-trial latencies using Singular Value Decomposition (SVD), and a similar method is adopted for the present study.

Epochs denote discrete time segments extracted from a continuous data stream, often covering an experiment's particular event or significant time points. They provide time windows around stimulus presentation, responses, or any other event that occurs during the experiment. Epoching also involves partitioning continuous data, i.e., EEG in this case, into shorter intervals of fixed duration aligned within a period of interest. For the purpose of this study, the epoching interval is set at the default interval, which is from -200 ms before stimulus presentation to 500 ms post-stimulus presentation. This interval encompasses stimulus presentation, which allows us to set the time for observing the N200 ERP within 125 to 275 ms after stimulus presentation (this window is re-calculated for every dataset). Thus, every recorded epoch allowed us to track the data for the current trial, and every epoch holds the responses to visually evoked signals.

For every individual/unique epoch, singular value decomposition (SVD) was applied. It utilizes a component of the said epoch across all the channels to find an optimal weight to be used as a spatial filter to extract N200 latency for a single-trial. SVD also produces a topological map based on the selected component for every subject. In this study, the topographic map and N200 waveform templates were the same as those used in Nunez et al. (2019).

It is a method similar to Principal Component Analysis that allows the decomposition of the epochs into a time vs. component (U) and channel vs. component (V) matrixes. A visual representation of the procedure is provided in Figure 1.

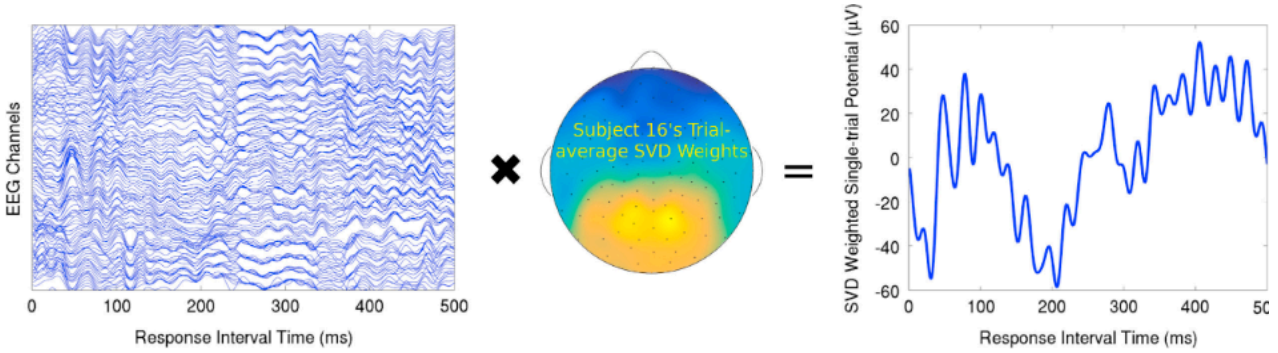


Fig. 1. Portrays a graphical SVD representation for deriving single-trial evoked response estimations within EEG data. The EEG signals shown are synchronized with the onset of the signal during the response interval, effectively capturing the single-trial-evoked data in response to a signal's initiation. The EEG data on the top left can be envisioned as a matrix structured with time along rows and channels along the columns (*T x C*). The SVD weights V (*C x Ch*) are obtained from the trial data and represented on the schematic scalp depiction with intermediate interpolated values (middle image). Thus, the ERP of a specific trial (top right) is derived by multiplying the temporal data of each channel from that trial with the corresponding weight in the vector V and subsequently summing across all weighted channels.

(Reprinted from How attention influences perceptual decision making: Single-trial EEG correlates of drift-diffusion model parameters (Nunez et al., 2017), Copyright 2017 by Elsevier Inc. Reprinted with permission.)

Applying SVD to epochs generated in every trial allowed the extraction of the N200 component, peak latency, peak negative amplitude, and associated topographic maps for the first 10 components, from which the component that explained the most variability in the original data was chosen.

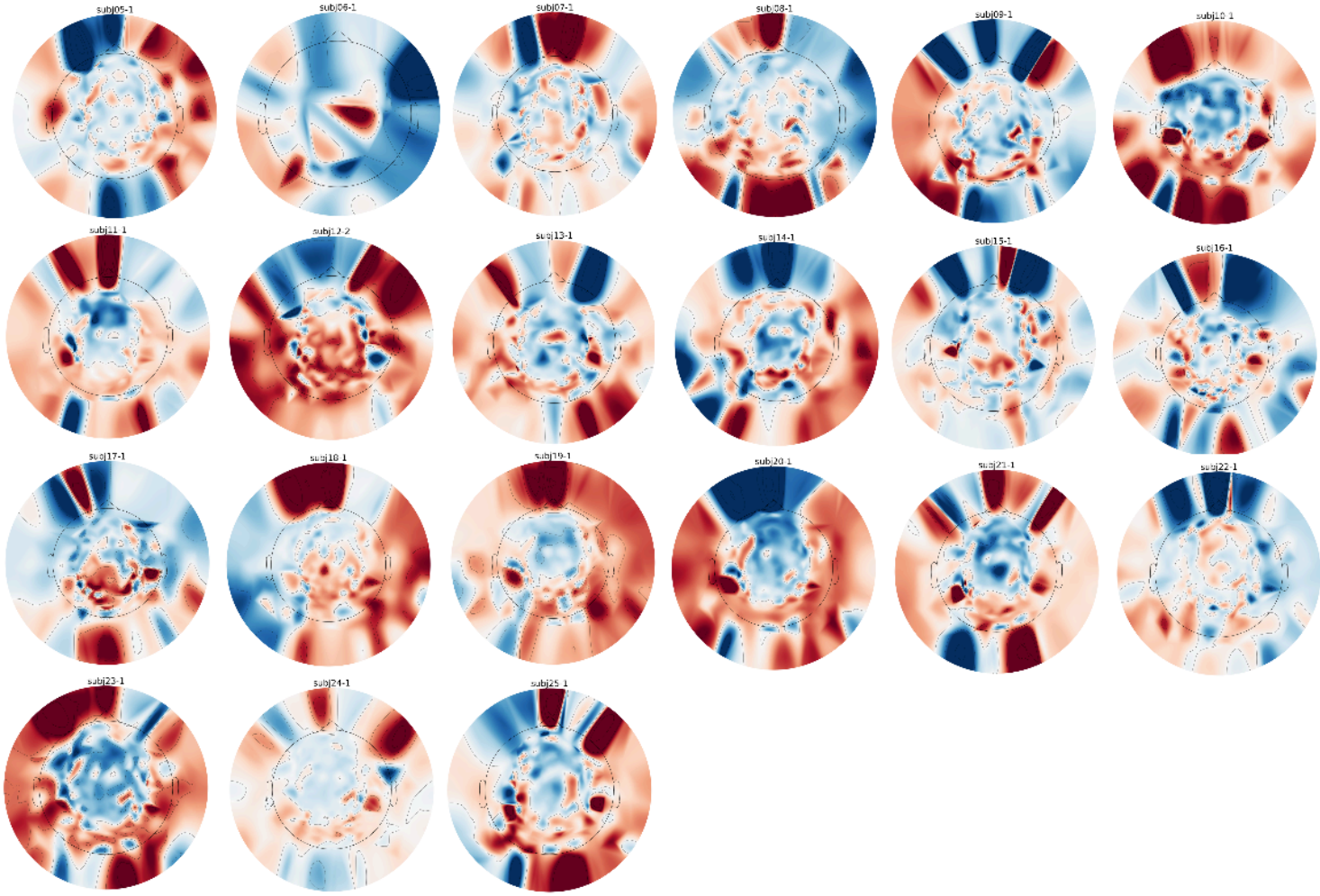


Fig. 2. Topographic maps for selected components of all the subjects in experiment 1 for a random trial -150.

Pre-Processing and Statistical Analysis

The EEG data was preprocessed using a Python package called MNE. The data for datasets 1 and 3 was downloaded as raw data, whereas the data for dataset 2 was already preprocessed.

Average referencing was applied to all electrodes. Independent Component Analysis was used to eliminate artifacts in the dataset. In both datasets 1 and 3, ICA analysis addressed components related to eye movements. Since EOG electrode data was not available for dataset 1, it was simulated by creating a bipolar reference using two frontal EEG electrodes ("1EX4" as anode and "1EX5" as cathode). Finally, another bandpass filter with frequencies ranging from 1 to 10 Hz was implemented. For more details on preprocessing, refer to Pinier's (2022) thesis.

After pre-processing, the processed data held information on all the events that occurred within the trials. The processed data was segmented into epochs to identify N200 peak latencies for every trial. Every epoch contained information on all the events within the trial (e.g., stimulus presentation, feedback response, RT). These epochs were segmented temporally, initiating at -200ms before stimulus presentation and 500ms post-stimulus onset. The segmented epochs were further sliced into the N200 window to observe the N200 ERP, i.e., from 125ms to 275ms (for dataset 1). This was applied to all the trials. SVD was then applied to decompose it into respective *u, s,* and *v* matrixes. The N200 peak latencies were calculated by correlating the first ten components of the *u* and *v* matrixes (Pearson's correlation) with the waveform and topographic template derived from the Nunez et al. (2019) study, respectively (shown in Fig. 3). The correlations were multiplied, and thus, the component with the highest product value was chosen to represent the subject's N200 latency for that trial. This process was applied to every trial the subject underwent.

For the NDT estimate, the subject's RT was used. Linear regression was performed using Python library *statsmodels,* and Bayesian regression was performed using JASP. The analysis presented below is only for dataset 1. It is important to note that due to a lack of computing power, the analysis was performed in batches of 150 trials, with ICA being run individually for every batch of trials.

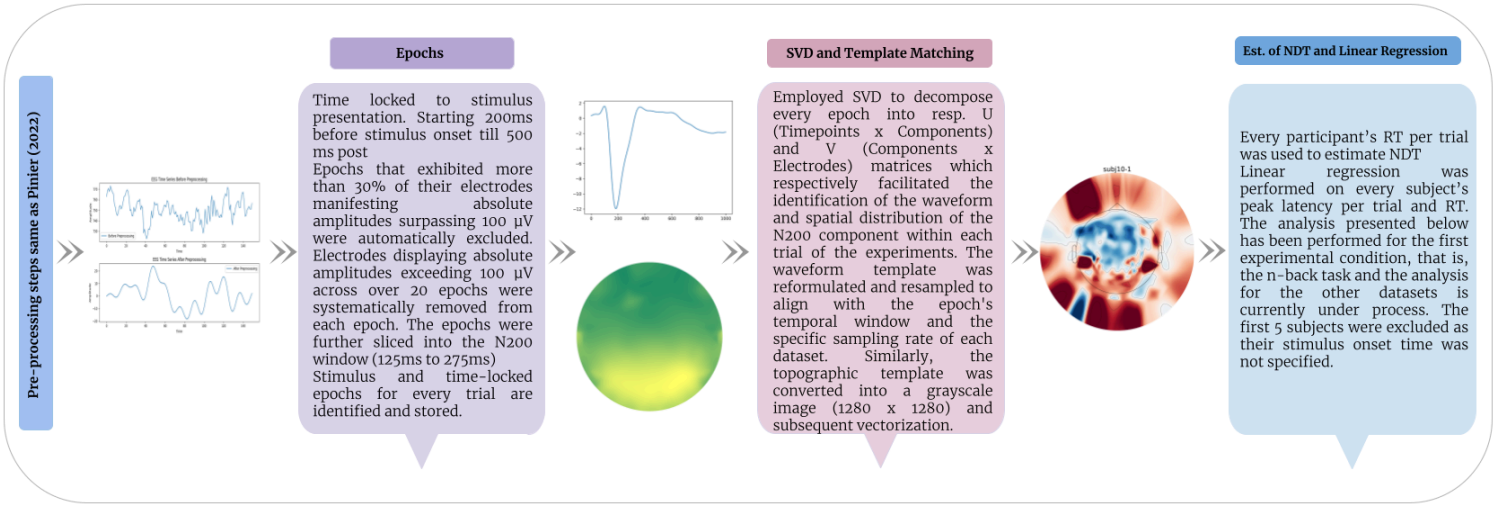


Fig.3. An illustration of the extraction process. The waveform and topographic template retrieved from the Nunez et al. (2019) study are shown in between the "Epochs" block and "SVD and Template Matching" block, respectively.

Results

N200 peak latencies, peak values, and selected components were identified and calculated for every trial the subject went through in one experiment session.

It is important to note that the analysis was performed only on dataset 1. Single-trial estimates of N200 latencies were compared to response times across all data points in dataset 1. The current methodology of running the analysis on batches of 150 trials, along with individual ICA for every batch, would not be adopted for the other two datasets. This is because the amount of noise and other cognitive variance per trial is very high in the estimated N200 latencies and RT. Thus, the following results are reported only for dataset 1.

We calculated two regressions by using both N200 latencies and RT as independent variables. Following this, we also ran Bayesian regression for both analyses. A significant relationship expected to be observed in single-trials was not found. The regression coefficient in analysis 1, wherein RT is considered as IV, was $R^2 = 0.000$ (R^2_{adj}= -0.000) with a 95% confidence interval (-0.007, 0.004) (p=0.676). Bayes factor, in comparison to the null model, was calculated. For analysis 1, $BF_{(10)}$ is 0.017, and inverted $BF_{(01)}$ was 58.023 (bayesian $R^2 = 0.000$). A slight marginal effect of RT on peak latency was observed, indicating a weak negative relationship between RT and N200 peak latency, a result similarly reported in Pinier (2022). For analysis 2, wherein the N200 peak latency is considered as IV, the regression coefficient was $R^2 = 0.000$ (R^2_{adj}= -0.000) with a 95% confidence interval (-0.041, 0.027) (p=0.676). Bayes factor in comparison to the null model was $BF_{(10)}$ is 0.017, and inverted $BF_{(01)}$ was 58.023 (bayesian $R^2 = 0.000$). No marginal effect of peak latency on RT was observed.

Overall, the evidence does not support the effect of N200 latency on visual encoding time in dataset 1. There could be several potential reasons, such as the batching of the trials, that could have led to the lack of the relationship and other reasons discussed in the following section. It is again important to note that these findings are specific to dataset 1, and this methodology of batching would not be applied to the other two datasets.

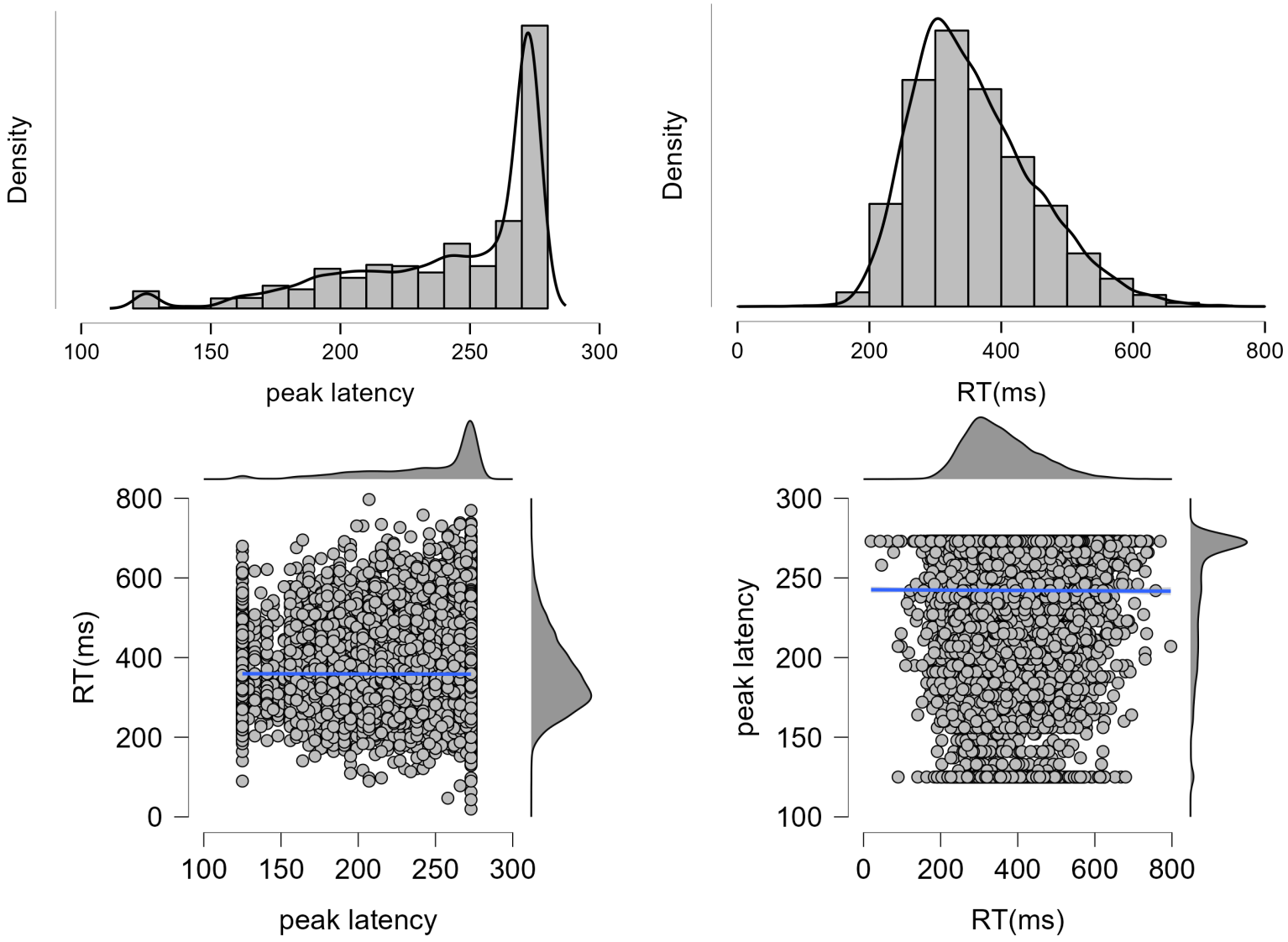
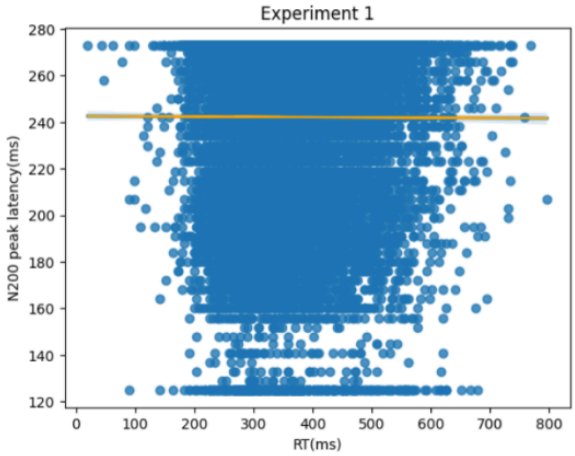


Fig. 4. Distribution plots for single-trial N200 latencies and Reaction Time. The mean for single trial N200 latencies is 242.18 ms, and for Reaction Time is 358.6 ms. Standard Deviation of 37.6 for peak latencies and 93.03 for Reaction Time.

Coefficient	-0.0012
Std Error	0.003
95% CI	[-0.007 , 0.004]
R^2	0.000
R^2_{adj}	-0.000
BF_(10)	0.017
BF_(01)	58.023
Bayesian R^2	0.000



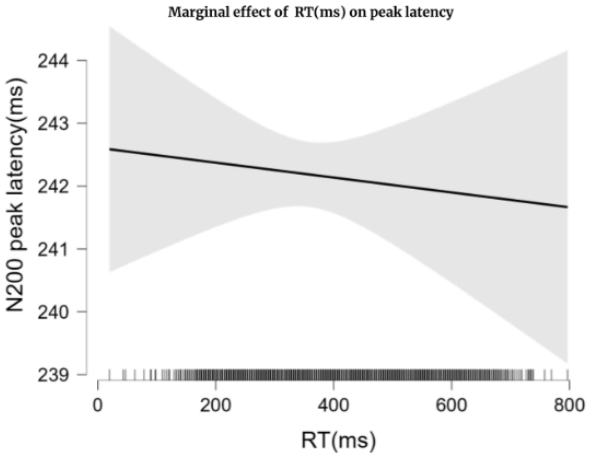
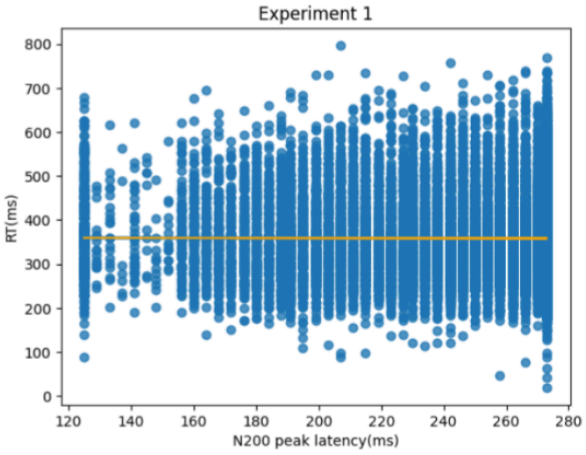


Fig. 5(a)

Coefficient	-0.0072
Std Error	0.017
95% CI	[-0.041 , 0.027]
R2	0.000
R^2_{adj}	-0.000
BF_(10)	0.017
BF_(01)	58.023
Bayesian R^2	0.000



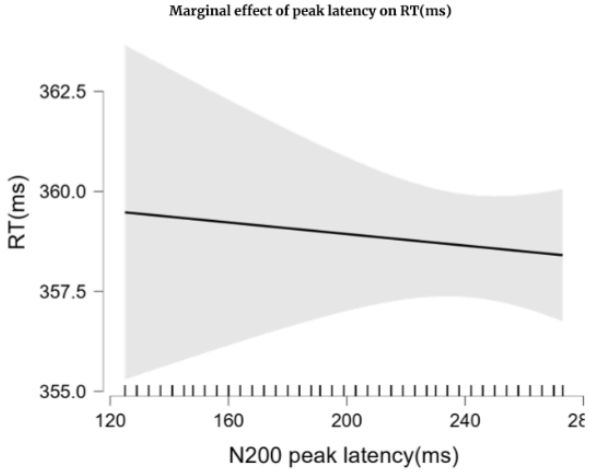


Fig. 5(b)

Fig.5(a). Linear regression plot for condition 1 (RT as IV). N200 peak latencies were calculated for every trial against the respective reaction time. The marginal effect plot shows a weak negative correlation but is not significant.

Fig. 5(b). Linear regression plot for condition 2 (peak latency as IV). N200 peak latencies were calculated for every trial against the respective reaction time. The marginal effect plot shows no relationship between the variables.

Discussion

In the present study, we sought to extend the findings of Nunez et al. (2019), who posited a relationship between N200 peak latencies and Visual Encoding Time (VET) within the context of a gabor patch task, and of Pinier (2022) who replicated the findings in the n-back task, gabor patch task and random dot motion task by extracting trial averaged N200 latencies. We extracted single-trial N200 latencies in the n-back task dataset and expected to see a 1ms-to-1ms relationship on single-trials. However, the said relationship was not observed, except for a weak marginal effect of reaction time on N200 peak latencies. There could be several reasons for this, which are discussed as follows.

Primarily, the reason why the present study is not a direct replication of Nunez et al. (2019) is due to the minor change in methodology as a test of theory. The Nunez et al. (2019) study extracted an averaged weight map using SVD for all the trials. However, we treated every trial's data as a unique EEG time series. We extracted N200 latencies by applying SVD to all the trials, leading to every trial having its unique weight map. We theorized that every trial must hold the data for all the events that occur in the experiment, including evidence accumulation, thus, we could treat every trial as an individual time series. We expected every trial to have some trace of the evidence accumulation, visualized through N200 ERP and extracted through ICA and SVD. However, the current results do not substantiate the relationship between N200 and VET or the theory. Both analyses (of treating both the variables are independent variables) for the regressions proved inconclusive for the hypothesis that N200 latencies would track visual encoding time (VET) in the n-back task. We only observed a slight marginal effect indicating a weak negative correlation for reaction time tracking N200 latencies. This is consistent with the weak negative correlation found in Pinier's (2022) study for the n-back task. The absence of the relationship could be attributed to the batching of trials for analysis, which may have potentially introduced biases in the extraction process. Analyzing the trials in batches also resulted in ICA running individually for every batch, which can account for the more significant variance observed in the extracted topographic maps of the subjects (refer to Fig.2.).

It is also essential to consider that the dataset itself might explain the lack of the relationship. The task used by Nunez et al. (2019) was a perceptual decision-making task that resulted in prominent N200 latencies with significant deflection times of the N200. In comparison, the n-back task is a working memory task and not a perceptual decision-making task. It is likely that the cognitive noise elicited due to the nature of the task could interfere with detecting the signals we are looking for. Furthermore, it is plausible that this particular dataset contains noisy signals and does not have quality N200 representations. This can also be observed by examining the extracted latency distributions (Fig.4. top-left corner). Thus, the task and the dataset itself may not be suitable for validating the N200-VET relationship. Additionally, integrating the n-back task into a drift-diffusion framework is challenging; however, a modified version of the n-back suggested by Boag et al. (2021) could provide further insights if applied in this context.

One possible way to increase the likelihood of observing quality N200s is by designing experiments with a constrained task goal. We speculate that complex tasks can affect the occurrence of quality N200 signals independent of evidence accumulation. For example, in the n-back task, participants were required to remember and recall presented letters, while in a garbor patch experiment or random dot motion task, participants simply needed to indicate whether they saw a patch or report the direction of moving dots. In order to effectively extract the N200 signal, we should focus on managing behavioral noise by imposing constraints on the task goal during experimentation.

As a future work possibility, we can adopt an alternate modeling technique proposed by Ranjbar et al. (2018). Their spatiotemporal filtering technique offers promise in examining single-trial variance and peak latency estimation, aligning with the Nunez et al. (2019) theory. We could also explore within-subject variability in order to test the robustness of the N200 signal during evidence accumulation in different perceptual decision-making tasks.

Despite the preliminary findings, our methodology of treating the trial data as an individual EEG time series presents a valuable test of the N200-VET relationship theory, offering insights beyond a direct replication of the Nunez et al. (2019) study. Although our current results do not decisively support the relationship, the ongoing extraction of data from the remaining datasets may contribute further evidence, potentially substantiating the N200-VET link demonstrated by Nunez et al. (2019) in the context of different perceptual decision-making tasks.

The limitations of our study, including data processing constraints leading to analyses in batches of trials, potential code adaptation issues, and deviations from the original Nunez et al. (2019) paper's methodology, necessitate a cautious interpretation of our findings. Collecting a more diverse sample, controlling for noise through behavioral adaptations seen through constraints in experimental conditions, and adhering to more stringent replication methodologies are crucial aspects of this continued journey.

Further research directives should involve a holistic analysis of N200 extraction, exploring Ranjbar et al. (2018) methodology, and adherence to the exact methodology proposed by Nunez et al. (2019) for direct replication across all and future datasets. Despite current setbacks, our investigation remains ongoing for the other two experiments, albeit without batching, maintaining optimism for potential relationship verification through our replication attempts.

References

Boag, R. J., Stevenson, N., van Dooren, R., Trutti, A. C., Sjoerds, Z., & Forstmann, B. U. (2021). Cognitive Control of Working Memory: A Model-Based Approach. Brain Sciences, 11(6), 721. https://doi.org/10.3390/brainsci11060721

Boehm, U., van Maanen, L., Forstmann, B., & van Rijn, H. (2014). Trial-by-trial fluctuations in CNV amplitude reflect anticipatory adjustment of response caution. NeuroImage, 96, 95–105. https://doi.org/10.1016/j.neuroimage.2014.03.063

Boudewyn, M. A., Luck, S. J., Farrens, J. L., & Kappenman, E. S. (2018). How many trials does it take to get a significant ERP effect? It depends. Psychophysiology, 2018 Jun;55(6):e13049. doi: 10.1111/psyp.13049. Epub 2017 Dec 20. PMID: 29266241; PMCID: PMC5940511.

Bridwell, D. A., Cavanagh, J. F., Collins, A. G., Nunez, M. D., Srinivasan, R., Stober, S., et al. (2018). Moving beyond ERP components: a selective review of approaches to integrate EEG and behavior. Frontiers in Human Neuroscience, 12, 106.

Britton, J., Frey, L., & Hopp, J. (2016). The normal EEG - electroencephalography (EEG): An introductory text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants. Britton JW, Frey LC, Hopp JLet al., authors; St. Louis EK, Frey LC. https://www.ncbi.nlm.nih.gov/books/NBK390343/

Enriquez-Geppert, S., Konrad, C., Pantev, C., & Huster, R. (2010). Conflict and inhibition differentially affect the N200/p300 complex in a combined go/Nogo and stop-signal task. NeuroImage, 51(2). https://doi.org/10.1016/j.neuroimage.2010.02.043

Ghaderi-Kangavari, A., Rad, J. A., Parand, K., & Nunez, M. D. (2022). Neuro-Cognitive Models of Single-Trial EEG Measures Describe Latent Effects of Spatial Attention during Perceptual Decision Making. https://doi.org/10.1101/2022.04.07.487571

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. Annual Review of Psychology, 67(1), 641–666. https://doi.org/10.1146/annurev-psych-122414-033645

Iyer, D., & Zouridakis, G. (2007). Single-trial evoked potential estimation: Comparison between independent component analysis and wavelet denoising. Clinical Neurophysiology, 118(3), 495–504. https://doi.org/10.1016/j.clinph.2006.10.024

Li, R., Principe, J. C., Bradley, M., & Ferrari, V. (2009a). A spatiotemporal filtering methodology for single-trial ERP component estimation. IEEE Transactions on Biomedical Engineering, 56(1), 83–92.

Li, R., Keil, A., & Principe, J. C. (2009b). Single-trial P300 estimation with a spatiotemporal filtering method. Journal of Neuroscience Methods, 177, 488–496.

Loughnane, G. M., Newman, D. P., Bellgrove, M. A., Lalor, E. C., Kelly, S. P., & O'Connell, R. G. (2016). Target selection signals influence perceptual decisions by modulating the onset and rate of evidence accumulation. Current Biology, 26(4), 496–502.

Lui, K., Nunez M. D., Vandekerckhove, J., & Srinivasan, R. (2018). Mental chronometrics of decision making. https://osf.io/ufdk6

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. Psychophysiology, 31(3), 291–308. https://doi.org/10.1111/j.1469-8986.1994.tb02218.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. Trends in Cognitive Sciences, 4, 432–440

Mullen T., Onton J., Delorme A., Makeig S. (2010). Analysis and visualization of theta-band information flow dynamics in an ERN-producing task Human Brain Mapping [Poster session]. Barcelona, Spain.

Mushtaq, F., Wilkie, R. M., Mon-Williams, M. A., & Schaefer, A. (2016). Randomised prior feedback modulates neural signals of outcome monitoring. NeuroImage, 125, 868–879. https://doi.org/10.1016/j.neuroimage.2015.10.046

Nunez, M. D., Vandekerckhove, J., & Srinivasan, R. (2017). How attention influences Perceptual Decision Making: Single-trial EEG correlates of drift-diffusion model parameters. Journal of Mathematical Psychology, 76, 117–130. https://doi.org/10.1016/j.jmp.2016.03.003

Nunez, M. D., Gosai, A., Vandekerckhove, J., & Srinivasan, R. (2019). The latency of a visual evoked potential tracks the onset of decision-making. NeuroImage, 197, 93–108. https://doi.org/10.1016/j.neuroimage.2019.04.052

Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. In C. Neuper & W. Klimesch (Eds.), Progress in Brain Research (Vol. 159, pp. 99–120). Elsevier. https://doi.org/10.1016/S0079-6123(06)59007-7

Onton J., Makeig S. (2008). Relations between sensory and behavioral context and EEG spectra in a working memory Task [Poster session]. Society for Neuroscience, Washington DC, United States.

Onton J., Makeig S. (2009). Independent modulators of regional EEG alpha sub-band power during a working memory task [Poster session]. Cognitive Neuroscience Society, San Francisco, CA, United States.

Pinier, C. (2021). Exploration of a Potential Relationship between N200 Peak-Latency and Visual Encoding Time: a Replication Study (thesis).

Ranjbar, M., Mikaeili, M., & Khorrami Banaraki, A. (2017). Single-trial estimation of peak latency and amplitude of multiple correlated ERP components. Journal of Medical and Biological Engineering, 38(2). https://doi.org/10.1007/s40846-017-0309-2

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. Neural computation, 20(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Swann N., Onton J., Makeig S. (2007, November). Single-Trial EEG changes associated with specific behavioral contexts in a Two-Back Task [Poster session]. Society for Neuroscience, San Diego, CA, United States.

Thompson W., Mullen T., Onton J., Makeig S. (2011) A Bayesian spatiotemporal model for multi-subject EEG source dynamics and effective connectivity [Poster session]. ICON XII, Palma, Mallorca.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. Journal of Experimental Psychology: Human Perception and Performance, 24(6), 1656-1674. Doi: 10.1037//0096-1523.24.6.1656

Wang, G., Li, J., Zhu, C., Wang, S., & Jiang, S. (2021). How do reference points influence the representation of the N200 for consumer preference? Frontiers in Psychology, 12. https://doi.org/10.3389/fpsyg.2021.645775

Weindel, G. (2021). On the measurement and estimation of cognitive processes with electrophysiological recordings and reaction time modeling. https://doi.org/10.31237/osf.io/342zp

Zang, S., Zhou, C., & Chao, F. (2021). Estimation of event-related potentials from single-trial EEG. Advances in Intelligent Systems and Computing. https://doi.org/10.1007/978-3-030-87094-2_37

Zhao, L., Cao, B., Hoya, T., & Cichocki, A. (2005). Analysis of single-trial EEG data using a combined robust pre-whitening technique and ICA approach. Proceedings of 2005 IEEE International Workshop on VLSI Design and Video Technology, 2005. https://doi.org/10.1109/iwvdvt.2005.1504595

Appendix

	N200 Peak Latency (ms)	Reaction Time (ms)
Mean	242.185	358.628
Standard Deviation	37.659	93.045
Minimum	125.000	19.531
Maximum	273.000	796.875