UNIVERSITY
OF TURKU

Integrating mobile-EEG with a naturalistic extended reality task

Human Neuroscience / Faculty of Medicine
Master's thesis

Author(s):
Aarni Seppälä

19.6.2024
Turku

Master's thesis

Subject: Human Neuroscience
Author(s): Aarni Seppälä
Title: Mobile-EEG in a Naturalistic Extended Reality Task
Supervisor(s): Docent, PhD. Juha Salmitaival, Assistant Professor, PhD. Henry Railo
Number of pages: 47 pages
Date: 19.6.2024

Naturalistic tasks can enhance the ecological validity of neuroscience research, particularly facilitate unravelling cognitive brain functions underlying developmental disorders where the symptoms are defined based on everyday life problems. Emergence of mobile electroencephalography (mEEG) allows more naturalistic tasks in EEG studies by enabling study paradigms that permit full body movement. However, movement causes notable artifacts into EEG data that require denoising. This thesis evaluated denoising methods and preprocessing pipelines for mEEG data in naturalistic extended-reality (XR) settings involving naturalistic whole body movement movement. The main objective of the present study was to develop a preprocessing pipeline to denoise mEEG data collected in a naturalistic paradigm. Twenty adults participated in the study, including 14 neurotypical (NT) and 6 diagnosed with attention-deficit/hyperactivity disorder (ADHD). Main methods for denoising the data included high-pass filtering at 2 Hz, inclusion of electromyography (EMG) sensors, and intercorrelating them with components derived by independent component analysis (ICA). Furthermore, other methods including filtering, EMG regression, Artifact Subspace Reconstruction (ASR) were compared. Results showed that high-pass filter at 2 Hz and inclusion of EMG-guided ICA improved denoising via making muscle-related ICs more distinguishable. Event-related potentials (ERPs) were calculated for distraction sounds and target-related responses for NT and ADHD adults to demonstrate the clinical potential of integrating mEEG with a naturalistic XR paradigm. Further studies with greater sample size should be conducted to confirm and expand the obtained findings.

Key words: Mobile-EEG, Naturalistic Neuroscience, Extended-Reality

Table of contents

1 Introduction

1.1 General background

For roughly one century, electroencephalography (EEG), has been used to study human cognitive processes and related neuronal functions by recording postsynaptic potentials of pyramidal cells detected by an EEG cap with electrodes placed on the scalp of the participant. Although EEG was already developed a century ago by Hans Berger (see St. Louis et al., 2016), the related methods have been improving, and the use cases where it can be employed are constantly extended. The present study relates to using EEG in naturalistic experimental setting. More specifically, it focuses on paradigms in which free whole body free movement of the participant is included in a complex naturalistic environment, utilizing mobile-EEG (mEEG) to measure brain activation.

EEG has gained attention from neuroscience researchers due to its high temporal resolution, non-invasiveness, accessibility, and experimental flexibility. In EEG methodology, the quality of data assessment, noise levels, physical fit of the cap, ease of measurement, and live feed have improved significantly over time. Further, the advances in EEG-related software and signal detection methods have advocated innovations for research designs (Gramann, 2014). Developments in the field have provided EEG caps that produce higher quality signal (Ladouce et al., 2019), and allowed developing transparent devices increasing the flexibility of the measurements (Xu et al., 2019). Moreover, efficient preprocessing pipelines for data cleaning are constantly evolving, contributing to the broader adoption of EEG in various experimental paradigms. Other notable advancements in field of EEG research include higher-density caps, enhanced visualization options (Delorme et al., 2012), and optimized algorithms for artifact cleaning (Zhang et al., 2023). These advancements in EEG allow more flexible real-time recording of brain signals, including EEG in tasks where participants move around (Lau-Zhu et al., 2019). Consequently, the freedom of movement for the participant provides opportunities to create novel EEG paradigms, for example, such that more accurately mimic real-life situations.

The use cases that could benefit from naturalistic paradigms include, for instance, the current assessment of cognitive functions associated with Attention-Deficit/Hyperactivity Disorder (ADHD, see Merzon et al. 2022, Seesjärvi et al. 2022) that could potentially be complemented by simultaneous mEEG measurement. Previous studies examining neural correlates of ADHD have revealed atypical brain activity in the frontal lobe as well as fronto-striatal and fronto-parietal networks (Castellanos & Tannock, 2022; Dickstein et al., 2006). Although the majority of human neuroscience studies in this field have used functional magnetic resonance imaging (fMRI, see Kim et al., 2023 for a meta-analysis), there are also EEG experiments studying neural correlates of ADHD, in which the participant remains stationary (Michelini et al., 2022; Rostami et al., 2019). Considering the generalizability of non-naturalistic tasks to real-life scenarios, context-sensitive behavioural studies should be studied in naturalistic settings (Sugden, Thomas, & Kiernan, 2021). However, for the moment, naturalistic paradigms allowing free movement and human-environment interaction in an open-ended situation have not been used to unravel neural correlates related to ADHD. Furthermore, the methods for denoising mEEG measurements in these contexts are inconclusive, and it remains unknown which ones are optimal.

1.2 Event-related potentials

Event-related potentials (ERPs) represent averaged EEG responses that are time-locked to specific events (Luck, 2014). ERPs provide a widely used option for EEG analysis that can also be utilized for mEEG data acquired in naturalistic setups (see Ladouce et al., 2019; Park et al., 2019; Robles et al., 2021). ERPs have been employed for over a decade to study various cognitive domains, including perception, cognition, emotion, and psychiatric disorders, such as ADHD (Luck & Kappenman, 2013). Metrics derived from ERPs typically include amplitude, estimated source location, and latency relative to the event. One of the most common ERPs is P3 (Luck, 2014), a positive deflection evoked by target stimuli that peaks at about 300 milliseconds after a stimulus and deviates from responses to standard stimuli. P3 response is elicited in tasks that require higher-order cognitive processes such as stimulus evaluation or context processing. The oddball paradigm, often used to study the P3 response, can be used to investigate cognitive processing, including change detection (Barry et al., 2020), working memory (Gutiérrez-Zamora Velasco et al., 2021), and executive functions (Zani & Zani, 2020). Thus, ERPs provide a versatile tool for exploring cognitive functions in both traditional and naturalistic settings.

1.3 Naturalistic tasks

Naturalistic tasks are crucial for understanding active exploration of the environment, a key aspect of real-world human behaviour (Gottlieb & Oudeyer, 2018). In social tasks, contextual information influences the cognitive system of the participant, which can be seen in both behavioural and neuronal levels (Wolf, 2018). Natural environments also offer complex, multimodal sensory cues (Sonkusare et al., 2019), which are often not present in laboratory studies. On the contrary, in laboratory studies, which often use narrow environmental contexts, behaviours and related neural activity that a naturalistic environment might evoke may be suppressed (Krakauer et al., 2017; Dennis et al., 2021). To address this limitation, virtual reality (VR) environments can be used to simulate naturalistic settings while recording brain activity (Thurley, 2022). Portable brain imaging methods, such as mEEG, are particularly beneficial for tasks developed to study neurodevelopmental disorders in naturalistic settings (Artoni et al., 2018).

1.3.1 Ecological validity

Ecological validity, which refers to the generalizability of the observations to other contexts, is essential for ensuring that study findings are applicable across different real-world contexts (Chang et al., 2022). While maximizing experimental control by minimizing contextual factors can reduce variance caused by extraneous variables, it may also decrease ecological validity. This trade-off is critical, especially when studying phenomena defined by everyday life situations, such as symptoms of developmental disorders. Improving ecological validity by simulating real-world situations is vital for capturing these phenomena accurately. Previous studies have captured the importance of ecological validity and naturalistic paradigms in human studies. For example, Railo and colleagues (2023) found different fear-related ERP responses in naturalistic and non-naturalistic paradigms. Using mEEG, gives possibility for greater ecological validity, as it allows the whole-body movement of the participant. One way to enhance ecological validity, is with extended reality (XR) related technologies.

1.3.2 XR

XR refers to an umbrella term of VR, augmented reality (AR), and mixed reality (MR) (De Paolis et al., 2024). VR refers to computer-generated environments that allow replication or simulation of real-world or imaginary environments that can be interactive (Abbas et al., 2023). In turn, AR refers to superimposed, often interactive digital imagery that is projected onto real-world view (Abbas et al., 2020). MR applies both domains of VR and AR. In MR, in addition to the virtual environment that the participant sees, there are real interactable objects. When the participant interacts with such an object, for instance, by grabbing it with hands, its movement is coupled with movement of the corresponding object in the virtual environment (Gerini et al., 2023). MR has already been used successfully in several clinical applications, such as neurosurgery simulations (Sabbagh et al., 2020). Thus, XR technologies hold significant potential for research in fields relevant to medical sciences.

1.4 ADHD

ADHD is a neurodevelopmental disorder with three main symptom domains, inattention, hyperactivity, and impulsivity (Kim et al., 2014; Seesjärvi et al., 2021; American Psychiatric Association, 2022). Inattention can manifest, for example, as difficulty of sustaining attention, distractibility, and forgetfulness. Hyperactivity and impulsivity may present as behaviours, including fidgeting, excessive talking, and in inability to remain still (American Psychiatric Association, 2022). Compared to neurotypical (NT) population, ADHD individuals have also impairments in executive functions and working memory (Kim et al., 2014; Seesjärvi et al., 2021). Impairment in executive functions can cause difficulties with inhibiting impulsive behaviours, planning, and organization (Grassi-Oliveira et al., 2017). Impairments in working memory affect ability to retain and manipulate information over short periods. This can be seen, for instance, as difficulty of following instructions (Arjona Valladares et al., 2020; Kim et al., 2014). Persons with ADHD may learn to deal with and camouflage some of the impairments, but it may involve a cost: if one is constantly trying to find another way to cope out of everyday struggles, this may cause heavy burden for the afflicted individual (Moen et al., 2011). Due to compensatory skills and ability to deal with these impairments, identifying ADHD in the adulthood may be challenging. At least traditional laboratory tests, questionnaires, and interviews may not show the impairment as clearly as in children's case (Kysow et al., 2016).

Various studies have attempted to identify neural markers for ADHD (Arjona Valladares et al., 2020), which could complement objective diagnostic evaluation. Previous studies have shown ADHD individuals to have stronger P3-responses to target stimuli as compared with NT participants (Arjona Valladares et al., 2020). This could be possibly explained by reduced inhibitory control to irrelevant stimuli: for a person with ADHD, it is harder to ignore irrelevant stimuli such as loud noises (Barry et al., 2009; Aboitiz et al, 2014). While the results regarding aberrant P3-responses in ADHD children are promising, not much research has been conducted on P3-responses in adults with ADHD (Szuromi et al., 2010). Moreover, most of the experimental tasks that have been used in EEG studies attempting to detect neural underpinnings of ADHD are highly restricted with respect to movement of the participant and naturality of the stimulus. For detecting ADHD persons with hyperactivity, movement related tasks in which the tasks can reflect real life settings for the participants could be especially effective. Contextual information may influence the symptoms of ADHD, and novel laboratory setting might elicit different reactions than a mundane context would. ADHD symptoms may not fully manifest in a non-naturalistic task, and they may not reflect to responses in self-ratings or questionnaires either as inattention and forgetfulness may lead to biased answers.

1.5 Naturalistic tasks with mEEG

Portable EEG recordings via mEEG have been combined with naturalistic paradigms in various settings. These include tasks where participants can move around in VR (Rohani et al., 2014), AR (Krugliak, & Clarke, 2022), or MR (Li et al., 2023). For instance, Wang and colleagues (2023) used a school class simulation where mEEG was used to study the effects of learning-related factors on attention. Although attention is one of the most widely studied constructs in laboratory-based EEG studies, it has not been studied much in real-world experimental setups (Ladouce et al., 2019).

A study by Rohani and colleagues (2014) used VR to create a virtual classroom to study the P3 component in ADHD population. Ladouce and colleagues (2019), in turn, used mEEG to successfully measure P3 in attention-related tasks in which NT participants walked with in a naturalistic VR setting while auditory oddball sounds were played varying the gait speeds of the participant. VR setting can be further modified towards more naturalistic settings, via usage of MR. Combining MR with mEEG presents a solid approach to research cognition in a naturalistic setting while still holding stringent experimental control. Participants can experience real-worldlike phenomenon instead of laboratory setting, yet the stimuli in the paradigm can be manipulated and controlled. Movement tasks with mEEG and MR offer the addition of brain recordings to tasks mimicking real-world behaviour. These often require at least some movement of the head, limbs, and even the lower part of body of the participant. Whereas movement-related artifacts influence EEG data quality, in mEEG the effect of movement tends to be larger due to more prominent movements. Thus, mEEG data may require additional denoising steps. Besides issues related to experimental design introduced above, there are technical challenges related to mEEG signal to noise ratio improvement that will be brought up in the following section.

1.6 Methods for denoising mEEG data

For a proper mEEG denoising and preprocessing, several different methods should be considered, as different tasks and paradigms benefit of specific signal analysis solutions. For example, tasks that cause quicker and more intense muscle movements may require other signal processing methods, than still-EEG tasks (Klug & Gramann 2020; Gwin et al., 2010). There are various views on what should be included in an mEEG preprocessing pipeline, particularly on paradigms that include large movements of the participant, for example, walking or running (Dimigen, 2020).

Removing certain electrodes or noisy parts of the data are often used strategies in preprocessing of EEG data. The first preprocessing option, however, should be to fix the signal by reducing the noise since removing channels may result to loosing important signal. Cutting parts of data may sometimes be appropriate for non-task-related parts. For example, in ERP analysis, cutting data around the epochs does not influence the number of epochs.

EEG data typically contains at least some types of artifacts. Most common artifacts include eyeblink, eye-movement, cardiac, powerline, and muscle artifacts (Luck, 2014). Whereas eye-related artifacts, cardiac artifacts, and power artifacts are traditionally easy to clean, muscle artifacts can be difficult to avoid due to their presence across large range of frequencies (Muthukumaraswamy, 2013). With high quality EEG-caps and more developed paradigms, noise related to head movement can be largely avoided (Lau-Zhu et al., 2019), and the remaining noise can be further polished from the mEEG signal with elaborated signal analysis pipelines (Klug & Gramann, 2020). Even more difficult artifacts, such as gait artifacts or head movement-related artifacts can be handled either by small adjustments to research setting, or different methods for preprocessing the data. In practice, denoising of the mEEG artifacts described above is commonly conducted with filtering methods that discard the noise signals of chosen frequencies and independent component analysis (ICA) -related methods that extract noise components in a data-driven manner. Moreover, some more nuanced methods have been developed to tackle other issues that filtering, and ICA are not able to rule out.

1.6.1 Filtering methods

One of the most effective preprocessing tools for EEG is filtering. As EEG data is essentially a complex time-varying signal, it can be represented in form oscillating voltages at various frequencies (Luck, 2014). Most widely used filters with EEG are high-pass and low-pass filters, which will let only certain frequencies through and attenuate the rest. High-pass filter set at any point allows only frequencies above that point through, cutting anything below it. Similarly, low-pass filter only allows frequencies below that set point go through. In still-EEG experiments, both are advised, at least high-pass filtering at frequencies 0.5 Hz or higher (Luck, 2014). In some cases, filtering data for mEEG is not much different from the one that is routinely conducted for still-EEG data. Depending on the amount and frequencies of noise, the filter set points can be adjusted. Muscle noise tends to be spread across a large range of frequencies, especially to frequencies above 20 Hz (Muthukumaraswamy, 2013). Previous studies have used filters such as bandpass filters for high-pass filter at 1 Hz, low-pass filter at 20-30 Hz (Krugliak & Clarke, 2022; 30 Hz Ladouce et al., 2019). Klug and Gramann (2020) found high-pass filter at 2 Hz to be most effective for mEEG when the preprocessing pipeline also contained ICA.

1.6.2 Independent component analysis (ICA)

ICA is a commonly used tool for EEG and mEEG to remove artifacts for eyeblink, eye movement, power artifacts and muscle-related noise. In ICA, the complex EEG signal is decomposed and separated into orthogonal subcomponents. From those subcomponents, researcher aims to detect those that are likely to relate to artifacts or measurement noise either via visual inspection or automatized methods, remove them, and interpolate the data back. ICA is an especially powerful tool for detecting, removing, and interpolating both eyeblink and eye movement artifacts, and it is widely used for detecting and removing muscle artifacts as well (Gorjan et al., 2022). Often ICA alone is enough to detect all major artifacts (Gorjan et al., 2022).

Distinctiveness of the typical artifact independent components (ICs) in ICA can be improved with additional sensors. With an addition of an EMG sensor near the source of the noise, artifacts can be recognized with more confidence. For example, electrooculogram (EOG) for eye-related artifacts, electrocardiogram (EKG) for heart related artifacts, and electromyography (EMG) for other muscle-related artifacts, such as gait artifact. Especially with a very noisy signal, one should consider using these sensors. More common artifacts, such as eyeblink, eye movement, and cardiac artifacts can be isolated and removed very effectively using ICA on either the raw EEG data or epoched EEG data (Ladouce et al., 2019). However, sometimes ICA is not enough to tackle the more difficult artifacts, such as muscle artifacts. ICA can manage denoising slower gait changes, but artifacts created by running or fast-paced walking cannot not be cleaned with ICA only (Gwin et al., 2010).

ICs detected by ICA can be classified by using automatized categorization tool, for example, ICLabel (Pion-Tonachini et al., 2019). The number of ICs detected by ICLabel can further be used as an index of the EEG data quality (Klug & Gramann, 2020). The type of ICA algorithm used affects the observed ICs, and their quality. There are five common ICA approaches, namely FastICA, InfoMax, extended InfoMax, Picard, and AMICA. The results between the new and fast Picard model, and slower extended InfoMax and AMICA tend to be of high quality (Hsu et al., 2018). One of the most frequently used ICA methods is InfoMax, as it is a default option in the EEGLAB software (Delorme et al., 2012)

1.6.3 Other denoising methods

Canonical correlation analysis (CCA) is a method for finding linear correlations between two multivariate datasets (Hotelling, 1936). CCA can be applied to identify muscle artifact contamination in mEEG data (Sweeney, et al., 2013). CCA is typically used together with EMG, by examining the canonical correlation components between the two signals. Those components that have high intercorrelation between mEEG and EMG signal can be denoted as muscle related artifacts. More specifically, if the EMG-sensor is placed on feet, the artifact component can likely be gait-related, and if the EMG-sensor is placed on the neck, the artifact component is likely to be neck-muscle (i.e., head rotation). After artifact identification, the data can be denoised via removal and interpolation of the artifacts (Janani et al., 2018). Reports on using CCA seem to be positive for movement tasks. For instance, Beckwith and Beckwith (2022) found it to be even more effective than ICA.

Another linear transformation-based denoising method of mEEG data is via MNE's function mne_eogregression (Gramfort et al. 2013), which computes a linear regression model based on the given EMG sources and regresses the EMG noise out of the mEEG signal. Usage of the method is applicable for eye-related artifacts (Croft & Barry, 2000).

More automatic method for cleaning muscle artifacts is artifact subspace reconstruction (ASR) (Mullen et al., 2013; Jacobsen et al., 2020). It can be used either on intense (i.e., high amplitude variations) or transient artifacts (Chang et al., 2020). ASR has been successfully used to remove muscle-related noise, and it is especially effective on high amplitude artifacts, including eye and muscle artifacts, but depending on the task it maybe also suited for reducing smaller artifacts (Chang et al., 2020).

Combined use of multiple preprocessing methods in the same pipeline can be used to take advantage of their complementary strengths. For instance, ICA can be applied to more complex movement artifacts, including neck muscle and gait artifacts. Gorjan and colleagues (2022) investigated combining ICA with other methods such ASR and CCA. Especially by combining ASR with ICA, large-scale muscle artifacts were successfully removed. Combination of CCA with other methods, such as ICA was effective for muscle artifact removal as well (Gorjan et al., 2022). Chang and colleagues (2022) concluded that ASR is a powerful tool for improving the quality of ICA decomposition and produced better overall data quality by removing muscle artifacts and retaining brain signal. It still remains unclear, which combinations of preprocessing methods are optimal for denoising mEEG data from naturalistic tasks that include full body movement of participants.

1.7 Research questions and hypotheses

The first aim of this study is to examine the denoising possibilities of mEEG signal by using filters, ASR, and ICA in a naturalistic paradigm involving full body movement. Preprocessing pipelines are typically optimized already for non-movement EEG measurement but for movement tasks, novel preprocessing solutions are needed in accordance with study design used. To examine methods of denoising mEEG data, different preprocessing methods are analysed in denoising complex mEEG data, in which the participants perform everyday chore like tasks that require whole body movements like gait, limb movements, and head movements in an MR game. Previous studies have demonstrated the effectiveness of ASR, ICA, and EMG regressions alone, but to our knowledge different combinations them have not been studied in naturalistic MR paradigms. Based on preprocessing methods used by previous studies (Artoni et al., 2018; Butkeviciute et al., 2019; Chang et al., 2020; Dimigen et al., 2020; Gorjan et al., 2022; Gwin et al., 2010; Klug et al., 2020; Wang et al., 2023), five different preprocessing pipelines were built and compared. The denoising methods for preprocessing pipelines included ASR, ICA, EMG regression, and filtering methods for the new pipelines. To evaluate denoising capabilities of the preprocessing pipelines used, the number of epochs removed are compared, Power spectral density (PSD) plots of mEEG were visually compared to still-EEG PSD-plots.

The secondary aim of this study is demonstrating the potential of the new preprocessing pipelines built to measure neural correlates of cognitive phenomena in a naturalistic mEEG XR paradigm that involves full body movement. For this, data of NT and ADHD participants was epoched based on target related feedback audio cues that they received when completing a task. Data was also epoched based on task irrelevant distraction audios, and P3 ERPs were calculated for these distraction audios. Whereas feedback response components and P3 for ADHD adults have already been studied, it has not been studied in naturalistic tasks in XR paradigms using mEEG. This study aims to demonstrate the possibility of studying these ERPs in a naturalistic setting that includes full body movement.

2 Method

2.1 Participants

The present study comprises of two measurements: pilot and the main measurements. The pilot involved 14 participants, who were young NT adults. These participants did not receive any compensation for their involvement. The main measurement included 20 participants, among whom 14 were neurotypical adults, and 6 were adults with ADHD. The latter group was required to have been officially diagnosed, and only those without any comorbid psychiatric disorders were accepted. Participants in the main measurement received 50-euro compensation for their involvement. Participation was voluntary for all participants, and the participants were allowed to withdraw from the experiment at any time. All participants in the main measurement provided their consent for participation and data usage.

2.2 Apparatus

In both the pilot experiment and the main measurement, participants engaged in an MR game that simulated everyday life environments, wherein players performed routine tasks (such as turning on a lamp, picking up a chair, sitting down on a sofa), as instructed by the game, and completed these tasks to the best of their ability. This MR-game, referred to as XR-EPELI involves movement of the player, and is a follow-up game for the previous VR-version, EPELI. XR-EPELI was designed by Juha Salmitaival, Gautam Vishwanath, Sofia Tauriainen, and implemented with Unity (v. 2022.3.5f1; Unity Technologies, 2023). XR-EPELI was played using Varjo XR-3 XR-goggles (Varjo, Finland, 2023). XR-goggles were high-resolution (2880 x 2720 pixels) with a field of view of 115 degrees. The refresh rate for the XR-goggles was 90 Hz, and it was colour calibrated to 99% sRGB and 93% DCI-P3. The cabling of the XR goggles was arranged to hang from the ceiling, as this design provided more cable length as participants moved away from the starting point, ensuring unhindered movement, fewer collisions, and enhanced immersion for the player. Several items and some furniture in the virtual room of the game were digital copies of actual physical objects. Physical objects to be moved were equipped with object trackers (HTC Vive tracker, USA). As the participants interacted with and moved physical objects in real life to perform tasks, the objects mirrored their movements on the head-mounted display. To precisely locate the trackers, sensor base stations (HTC Vive Base Station 2.0, USA) were utilized.

Participants were equipped with a high-density mEEG using a 32-channel mEEG cap (ANT-Neuro eego™sports, The Netherlands), a fully portable mEEG-setup with the battery of 500 grams, housed in a lightweight backpack worn by the participant. The sampling rate of the mEEG cap was 2 kHz, and it was monitored with a windows tablet compatible with the mEEG cap. For the mEEG, reference was originally on the CPz channel. The mEEG cap had an active shielded waveguard™ (ANT-Neuro, The Netherlands) that reduced movement related to neck muscle artifacts due to minimal movement of the electrode cables.

The autonomous nervous system sensors were collected with Noraxon Ultium EMG device (Noraxon, USA). Wireless sensors were employed to allow participants free movement. These 24-bit EMG sensors had 2 kHz sampling rate and had less than 1µv baseline noise. Eight EMG sensors were placed on the participants: calf muscles, arms (flexor carpi radialis muscle), corners of the eyes (EOG), chest (EKG), and neck (trapezius muscle). EMG data were subsequently utilized to remove muscle-, and movement-related artifacts from the data in preprocessing.

Data from mEEG measurements were visually compared to data from still-EEG measurements. For this, EEG was measured using MEG-EEG device (MEGIN TRIUX, MEGIN Oy, Helsinki) with 32 EEG-Channels. Tasks for still-EEG measurements were matching to those of mEEG, but without movement.

2.3 Procedure

Data collection for the first experiment was conducted during last quarter of 2022 to the first quarter of 2024, and the data collection for the second experiment was conducted during 2024. All measurements were done at Aalto University. Each measurement took approximately 30 minutes, after setting up the hardware, including mEEG, XR-game, and EMGs, which took another 30 minutes.

Before the measurement, mEEG cap was fitted on the participant, their scalp was prepared, and gel was applied. The mEEG system was placed in a backpack worn by the shoulders of the participant. The participant's skin was cleaned in the areas EMG sensors were attached. EMG sensors were placed on the participants' legs, arms, chest, and neck. XR-headset was adjusted by the participant, ensuring it was comfortable but secure, and the headset was calibrated with its built-in calibration procedure.

Participants then began the game, following the tasks given by the virtual game guide. During the gameplay, research assistant monitored over the participant, to ensure sure that they would not bump into objects, in case of possible calibration issues of the in-game objects. After each block of tasks, the next block would begin only after the participant pressed the continue button.

Before beginning the actual game, a short training session was provided to the participants. The training session lasted approximately three minutes and it included showing the controls, calibrating the system, and orienting the participants to the gameplay by allowing them to practice moving within the play area.

The XR-EPELI game session comprised of 8 blocks, each consisting of every day routine tasks to perform. At the beginning of each block participants were instructed to conduct seven tasks by a virtual guide, a humanlike character. All the tasks for one block were instructed at the same time, and the participant had to remember the tasks to be able to complete them. There was no predesignated order in which the tasks had to be completed within a block. Time to complete a block of tasks was three minutes, but it was possible to conduct the blocks faster, if the participant managed to do so. The tasks were performed in an environment that was designed to look like a typical homelike environment with a kitchen, bedroom, living room, and bath room (Figure 1).

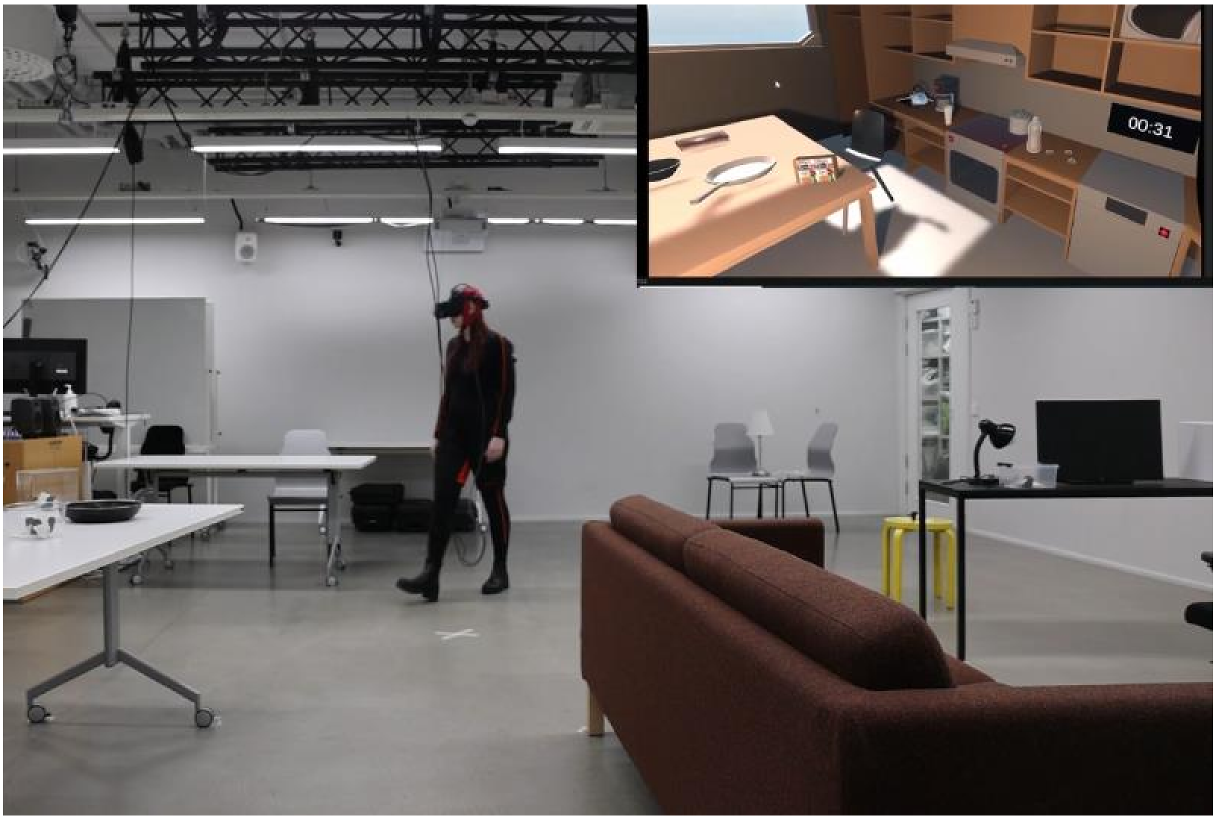


Figure 1. Setting of the XR-EPELI measurements. The participant's view is seen on the right top corner. Some of the objects have trackers, that when moved physically, also move virtually in the game.

Task completion sound was a bell sound with length of 2 seconds and was triggered after any task was completed. Task completion sound had two notes (B6, and D7) closely together, with pitches of 1973.53 Hz and 2093.00 Hz, respectively. Within the gameplay, with a random time interval, a distraction audio was played. Distraction audio used was task irrelevant and would occur throughout the gameplay. Distraction audios chosen for the task were designed to resemble ambient sounds that occur in everyday life environments, and they were apple bite sound, eating sound, doorbell buzz sound, closing drawer sound, drinking sound, fridge door closing sound, and swallowing sound.

2.4 EEG Data Processing

Most of the data processing for the mEEG was done with MNE-Python software (Gramfort et al., 2014), using Python version 3.10.10 (Python Software Foundation, 2023). If a plugin was not available for MNE-Python, Matlab version R2023a (The MathWorks Inc., 2022) along with its EEG data-processing toolbox EEGLAB (Delorme & Makeig, 2004) was used.

Due to data synchronization issues and missing event data, large portion of our data had to be excluded from the present study. After exclusion of these datasets, 10 datasets of NT adults were remaining. For the ADHD group, only 2 datasets were available after the exclusion.

The present study used several different EEG preprocessing pipelines to clean the data and identify the best preprocessing pipeline for this naturalistic task. These preprocessing pipelines (Figure 3) were compared to control pipelines (Figure 2), that were created for this study and resembled preprocessing pipelines typically used for denoising EEG signal. All datasets were denoised using both control preprocessing pipelines and all three alternative preprocessing pipelines.

For the first control pipeline, data were first imported, and the data format was converted to work appropriately with MNE-Python, by renaming channels to match the MNE-Python's channel montage. For the first control pipeline, data was first bandpass filtered at the frequencies 1 Hz to 40 Hz. Bad channels were visually removed by hand, and the mEEG signal was re-referenced to the mean of all channels that were not removed initially. Removed channels were interpolated, and extended infomax ICA was used to find most common artifacts: eye movement, eyeblink, and cardiac artifact. MNE_ICALabel was used to label artifacts, and ICs that were not recognized in prior visual investigation that had above 90% confidence and category "muscle" or "eye-artifact" were removed. Artifacts were removed and interpolated with ICA. Data was then epoched. Epoch rejection was done using with autorejection via MNE-Python's function 'epochs.drop_bad' with a criteria of 100μV. This removed all epochs with +-100μV amplitudes at any time point.

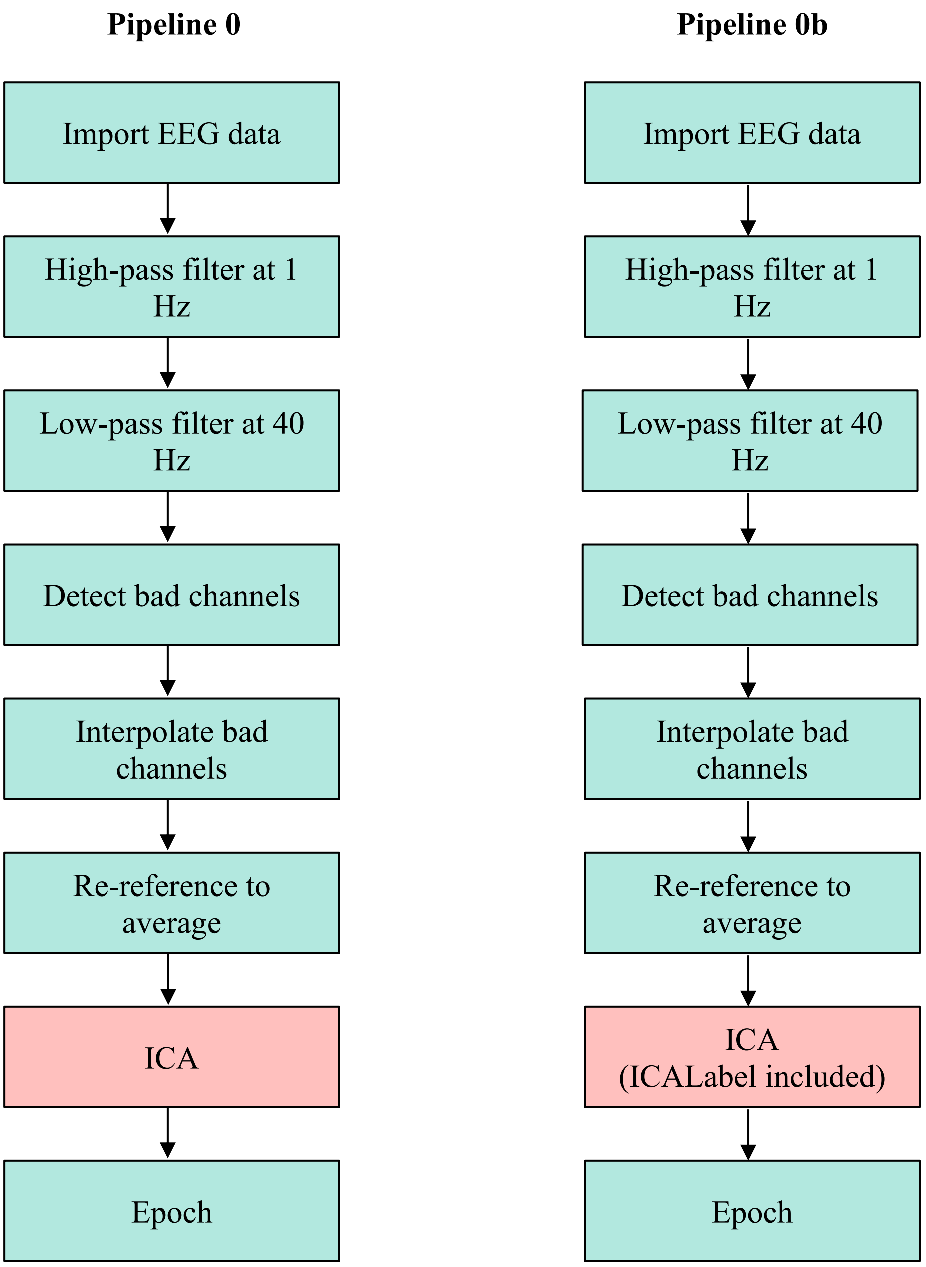
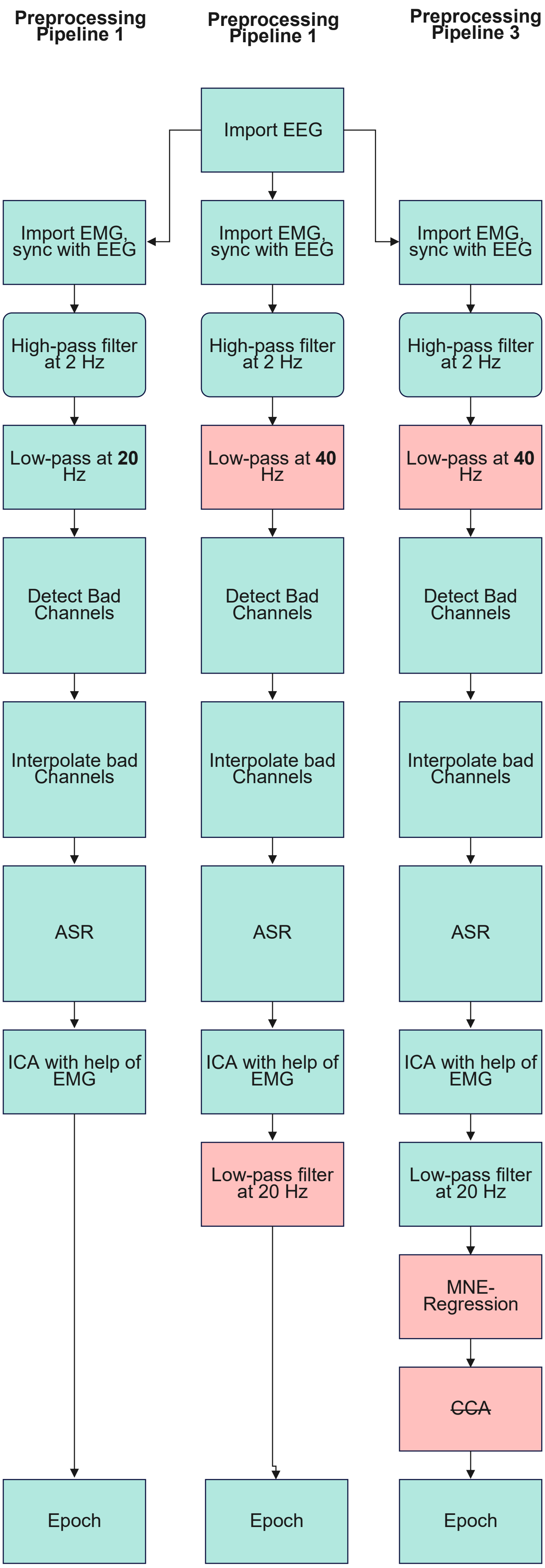


Figure 2. Control pipelines (pipeline 0 and pipeline 0b). These pipelines were constructed to mimic typical preprocessing pipelines used in EEG studies.



Preprocessing Pipeline 1
Preprocessing Pipeline 1
Preprocessing Pipeline 3
Import EEG
Import EMG, sync with EEG
Import EMG, sync with EEG
Import EMG, sync with EEG
High-pass filter at 2 Hz
High-pass filter at 2 Hz
High-pass filter at 2 Hz
Low-pass at **20** Hz
Low-pass at **40** Hz
Low-pass at **40** Hz
Detect Bad Channels
Detect Bad Channels
Detect Bad Channels
Interpolate bad Channels
Interpolate bad Channels
Interpolate bad Channels
ASR
ASR
ASR
ICA with help of EMG
ICA with help of EMG
ICA with help of EMG
Low-pass filter at 20 Hz
Low-pass filter at 20 Hz
MNE-Regression
~~CCA~~
Epoch
Epoch
Epoch

Figure 3. Three main preprocessing pipelines used to clean raw mEEG data. ASR = Artifact Subspace Reconstruction algorithm, ICA = Independent component analysis, CCA = Canonical Correlation Analysis.

Alternative preprocessing pipelines were built to fit the study paradigm, utilizing methods that were expected to be most suitable. In the following, the methods used in alternative pipelines are described first, and then the differences between pipelines.

Firstly, the EEG data was imported, and the data format was changed to work appropriately with MNE-Python (i.e., channel names were changed to represent conventional naming system). EMG data was imported to MNE-Python using functions in the Scipy toolbox (Jones et al., 2001). All the EMG sensor data and EKG data was resampled and time locked to match the EEG data.

EEG, EKG, and EMG data were imported, and noisy channels were removed with visual inspection. Data was then filtered with bandpass filter with lower cutoff frequency of 2 Hz and higher cutoff frequency of 20 Hz. Data was re-referenced from the single channel 'CPz' to mean as mastoids were noisy and thus the average of mastoids might have not reflected the signal properly. To ensure logical and orderly arrangement of the EEG channels, the EEG montage was set to 'standard_1005'.

ICA was performed to find eye-related noise components from the data. Both eye movement, and eyeblink artifacts were considered. These artifacts were first visually investigated by finding representative peaks from ICs for eyeblink and blocky movement that represents eye movement in the component activation. Eye-related artifacts were confirmed with both topographical plots. Eye movement was seen as activation on the frontal electrodes – negative for one eye, positive for one – and eyeblink as in positive activation on the frontal electrodes. Further, MNE-ICALabel toolbox (Li et al., 2022) was used to aid eye-artifact detection: those ICs that were classified as eye related with more than 90% of the confidence factor in MNE-ICALabel, were considered as eye-related artifacts, and they were cleaned with ICA.

A copy of the ICs was saved, and their intercorrelations were compared to the EMG data. ICs that had over 10% intercorrelation with an EMG sensor across all time points in the dataset were removed. Similarly to eye-related artifact removal, MNE-ICAlabel was used to determine muscle-related noise. If ICLabel classified an IC as a muscle-related artifact with the confidence above 90% the respective IC was removed, and the data interpolated with ICA. CCA was planned to be used to find intercorrelations between EMG and EEG signals within the full data.

CCA was calculated using scikit-learn package in Python (Pedregosa et al, 2016). However, due to technical issues with CCA, the data was riddled with power artifacts, and thus rather made the data even more noisy. CCA was not used as planned for this reason.

ASR was used to clean motion artifacts from the data. To use it with MNE, the ASRrpy package (Gütlin, 2021) was used. ASRpy's intensity factor was set to 10, which cuts down the more intense amplitude artifacts. This was set to fit with the intensity of the motion artifacts in the data. For rest of the options in ASR default settings were used. ASR was then used to interpolate the noisy segments with high intensity movement artifacts.

Linear regression method was used to regress EMG-related noise (neck muscle artifacts, and gait artifacts, cardiac artifacts) from the EEG signal. This was done via MNE-Python's EOGRegression method (Gramfort et al. 2013), which built a regression model based on all the sensors (EEG, EMG, EKG), and regressed out noise captured by EMG sensors.

The data was then epoched for two conditions: distraction audio (unattended deviant stimuli, task irrelevant) and task completion sound (attended deviant stimuli, task relevant). Epochs were captured using a time window from -0.2 seconds to +0.7 seconds from the stimulus onset. Noisy epochs were removed first via visual investigation, and then by removing all the epochs based on rejection criteria of +-100µV. Based on epochs, ERPs were plotted for both types of epochs, distraction audio for P3, and task completion sound for feedback related event potential. ERPs were plotted using electrodes Fz, Cz, Pz, C3, C4. Plotting was done with help of matplotlib.pyplot Python package (Hunter et al., 2012).

From the different aforementioned preprocessing methods, and combinations of these methods, three different main preprocessing pipelines were created (Figure 3). The main difference for pipeline 3 was to start with stronger filter, low-pass filter at 20 Hz, whereas other two preprocessing pipelines used only 40 Hz first, and further on filtered with 20 Hz. For the third pipeline, MNE's EOGRegression, and CCA for MNE-Python were used. Pipelines one and two did not use EOGRegression or CCA.

2.5 Statistics

Statistical analyses were calculated with Python and IBM SPSS Statistics (Version 27).

To measure denoising of the mEEG-signal, the three preprocessing pipelines were compared. To compare noise left in the data using different preprocessing pipelines, the number of epochs removed were compared between the preprocessing pipelines. To find an objective way to remove epochs, autorejection with a limit of +-100µV was used. The proportions of removed epochs were compared using Chi-Square test of independence, and as a post hoc test for the groups, multiple comparisons with Bonferroni corrections were applied using methods similar to Garcia-Perez and colleagues (2003).

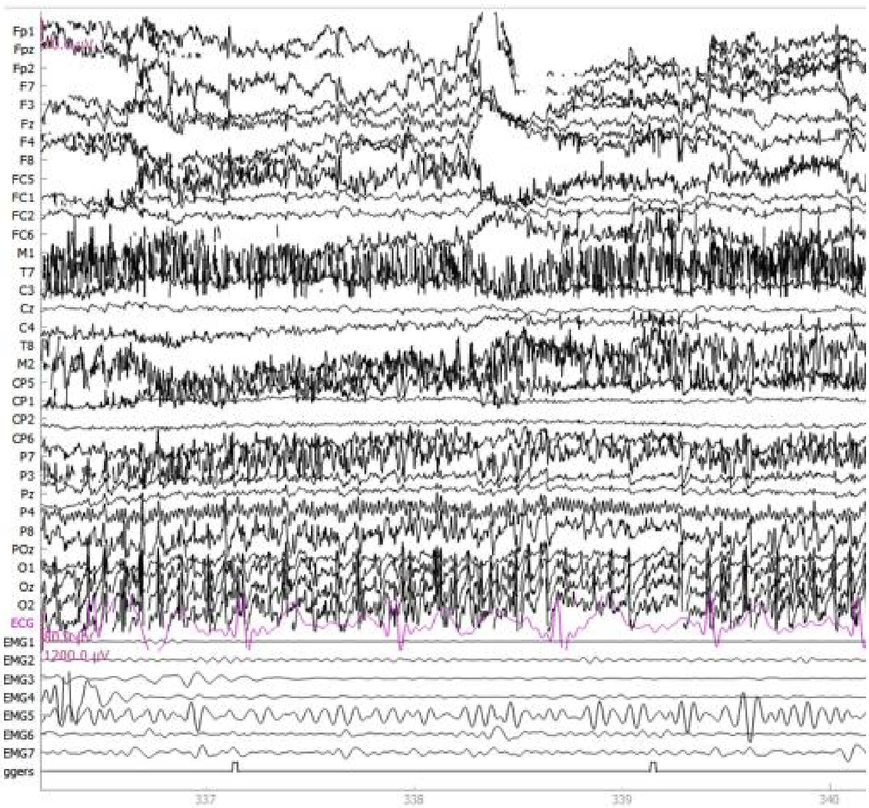
For the second hypothesis, regarding inhibition of feedback-related cue response of ADHD population, the data was epoched, and the feedback related cue response potential was plotted for both groups, ADHD and neurotypical. As the sample size was not sufficient for statistical testing, data was visually inspected.

3 Results

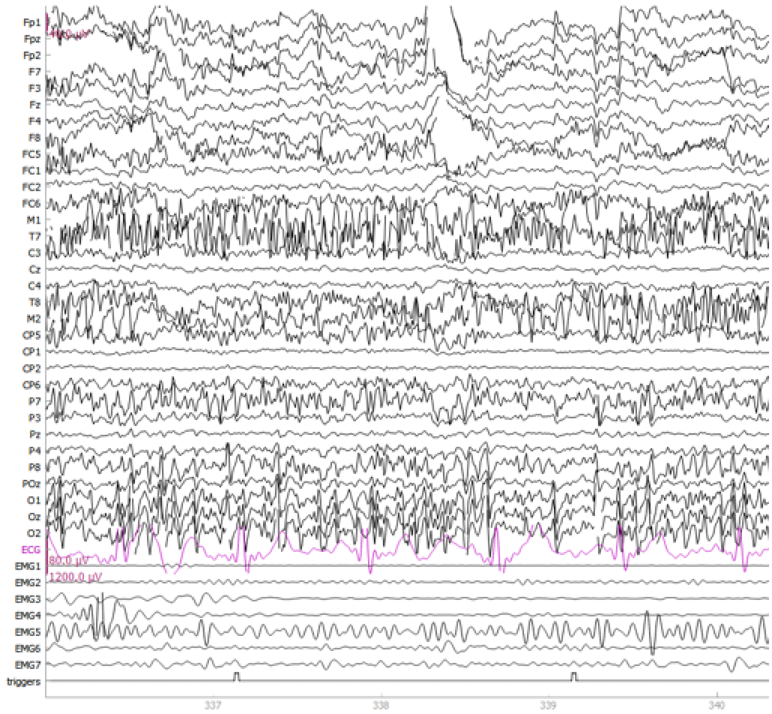
3.1 Data quality of mEEG

Filtering improved the signal by cleaning portion of the noise. This prominent effect was seen most, when comparing the control pipelines 0 and 0b that had 1 Hz high-pass filter to all the other pipelines in the study. The effects of the filter can be seen in Figure 4. The effects of each part of the methods used in preprocessing pipelines can be seen step by step in Figures 4-7 on the same data sample snippet.

A



B



C

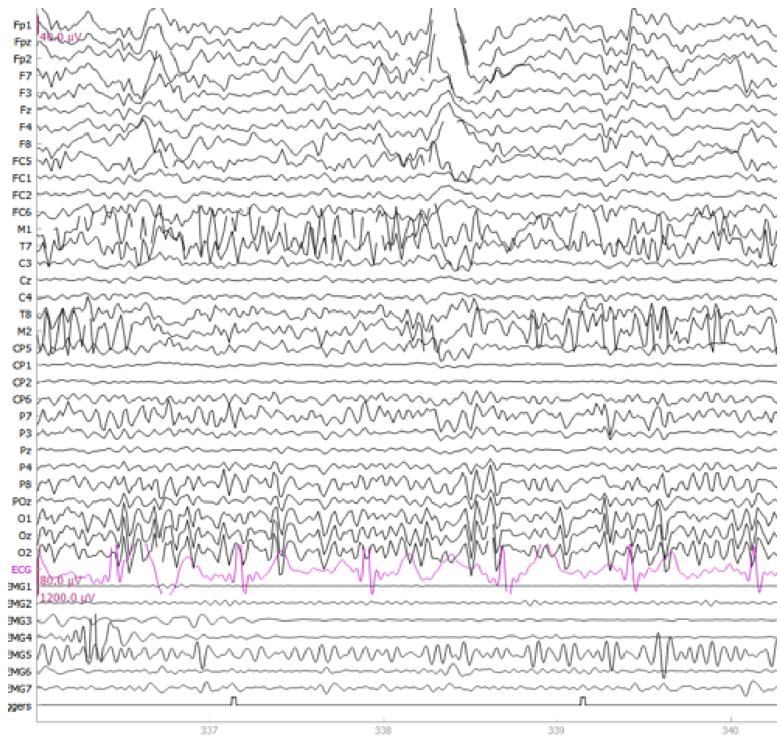


Figure 4. Effect of filters on example snippet of data. **A** Raw data with no filters with EMGs included. In **B** filter for control pipelines (filters at 1 Hz and 40 Hz) were used. In **C**, filters for preprocessing pipelines 1-3 (filters at 2Hz and at 20 Hz) were used. All electrodes and EMGs are plotted. Eye-related artifacts are still seen in frontal electrodes, at 337, 338.5, and 339.5 seconds.

High intensity artifacts were denoised using ASR, among other aforementioned methods. Data quality improvement was prevalent in very noisy snippets of mEEG data. ASR improved eye-related high intensity artifacts as well. The effects of ASR can be seen in Figure 5, in which high amplitude peaks in frontal electrodes that were seen in Figure 4, were removed.

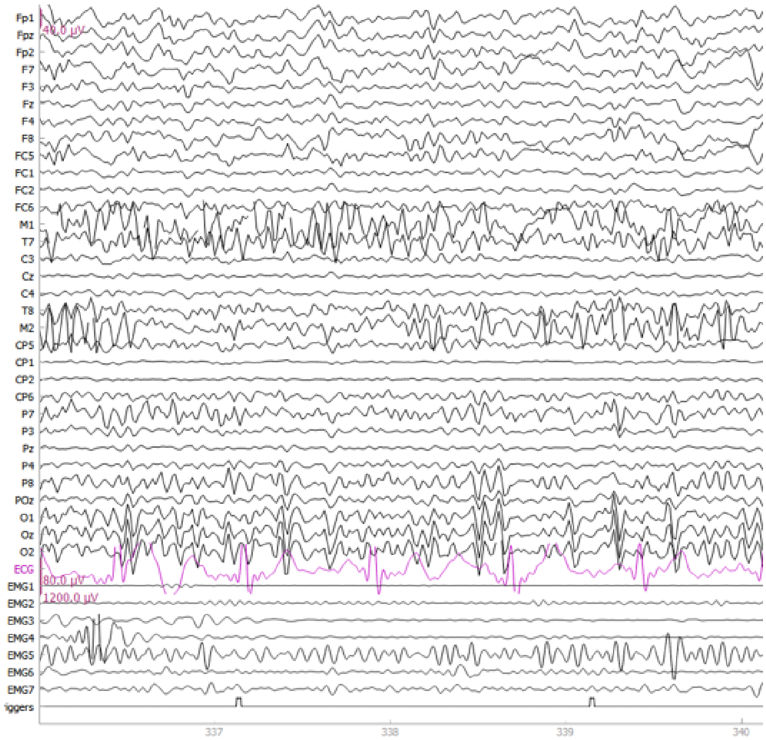


Figure 5. Sample snippet of the data with filtering and ASR. ASR further removed a lot of high intensity artifacts (see Figure 4 for comparison of same snippet without ASR), as well as eye-related artifacts. All electrodes and EMGs are plotted.

ICA was very effective in removing eye-related artifacts. This was true for both eye-movement and eyeblink artifacts. Eye-related artifacts were cleaned heavily already by ASR, but removing eye-related artifacts with ICA improved the data even further, especially for frontal electrodes. (Figure 6).

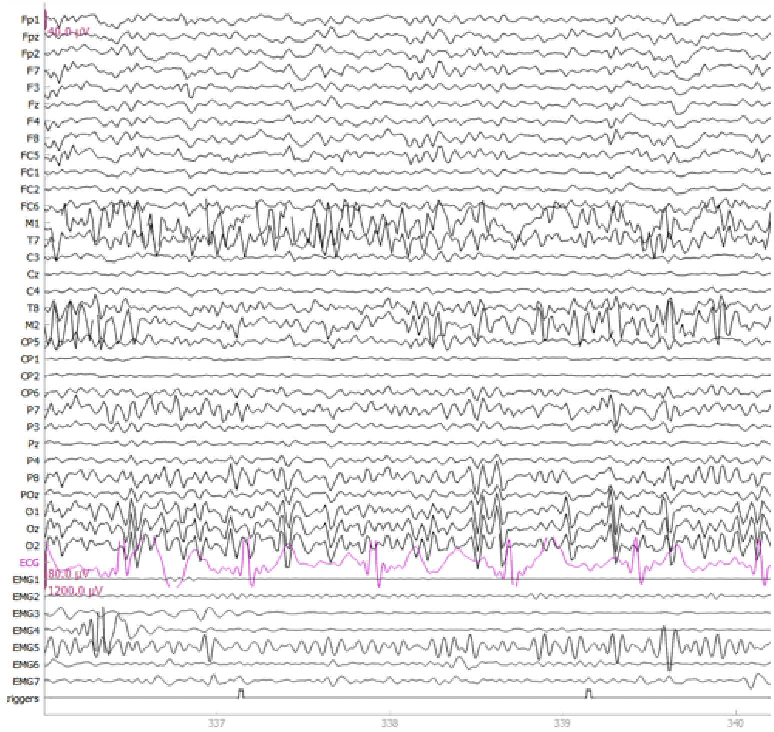


Figure 6. The effect of removing eye-related ICs from the data. Most prominent effects can be seen in frontal electrodes. All electrodes and EMGs are plotted. Muscle related noise still remains, especially in M1, M2, and occipital channels.

The detection of muscle-related ICs in the mEEG data improved with the inclusion of EMG sensors. The intercorrelations calculated for ICs and EMG sensors helped to confirm some of the muscle-related ICs, and the removal helped to clean mEEG signal in some of the electrodes considerably. Effect of denoising on data was especially large on the electrodes further from the centre, such as occipital and mastoid electrodes (Figure 7). Moreover, the inclusion of EMG sensors helped to confirm that some of the suspected muscle activity -related ICs were not linked to brain signal and could be cleaned.

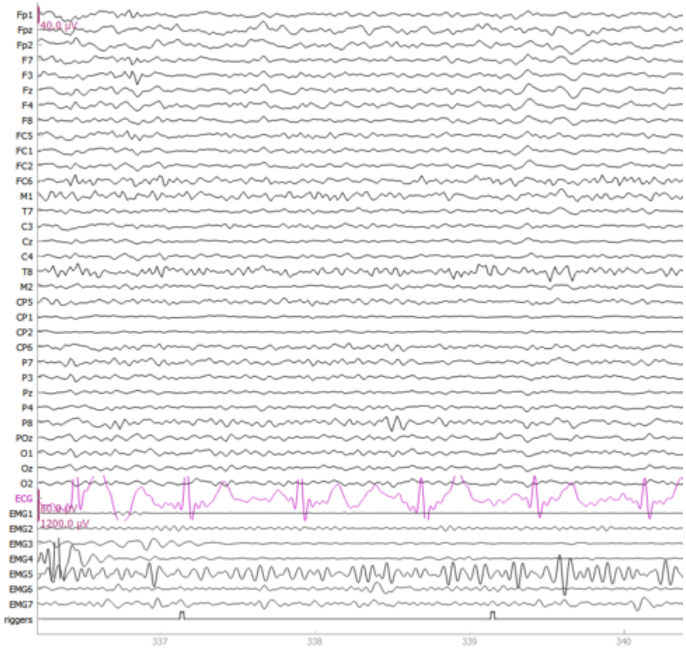


Figure 7. The effect of removal of neck muscle artifacts with ICA, with the help of EMG sensors on neck. Noise in data has decreased drastically, and the largest effects are seen in mastoid and occipital channels. All electrodes and EMGs are plotted.

Comparison of noise in the mEEG data to still-EEG data can be seen in Figure 8. Both datasets are denoised successfully, and mEEG data is cleaned to match the quality of still-EEG data. No visible eye-related artifacts or muscle artifacts can be seen in either data snippet. Movement-related artifacts have been successfully denoised from the mEEG data, to have match the quality of still-EEG data.

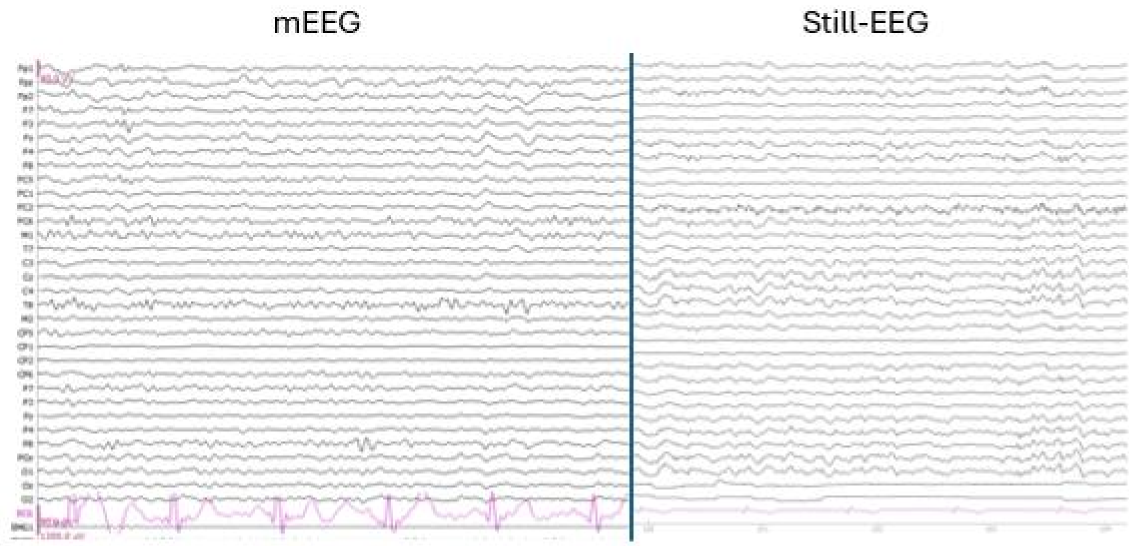


Figure 8. 5 second data snippets of same participant performing cognitively similar tasks in mEEG and still-EEG settings. Both data snippets are scaled with 40 µV scaling. Best performing pipeline, preprocessing pipeline 1 was used for mEEG, and still-EEG was preprocessed using typical preprocessing pipeline for still-EEG data, here done with preprocessing pipeline 0b.

To improve ICA, in pipelines 2 and 3, low-pass filter was set first at 40 Hz, so the muscle-related noise would be more prevalent to ICA. In these preprocessing pipelines, 20 Hz filter was later set. The expected improvement in the data quality did not appear, even if ICA found more muscle-related components, the data ended up noisier than in the pipeline 1, in which low-pass filter was set to 20 Hz. This was especially true for epoch wise data quality: preprocessing pipelines 2-3 autorejected more epochs than preprocessing pipeline 1.

In the preprocessing pipeline 3, MNE-EOGRegression was used to build a regression model based on the EMG sensors, and to regress the EMG-related noise out of mEEG signal. Improvements were very modest at best.

Out of all epochs (n = 886), in respect to epochs removed, preprocessing pipeline 1 performed the best, rejecting only 98 epochs (11.1%), as seen in Table 1. The best results for mEEG datasets, in respect to epochs retained, was with preprocessing pipeline 1 (see Figure 9). Preprocessing pipeline 0 rejected a lot more epochs. Using preprocessing pipeline 0 without ICLabel, 773 epochs (83.6%) were autorejected. Along with the help of ICLabel, the preprocessing pipeline performed poorly, autorejecting 741 epochs (83.6%).

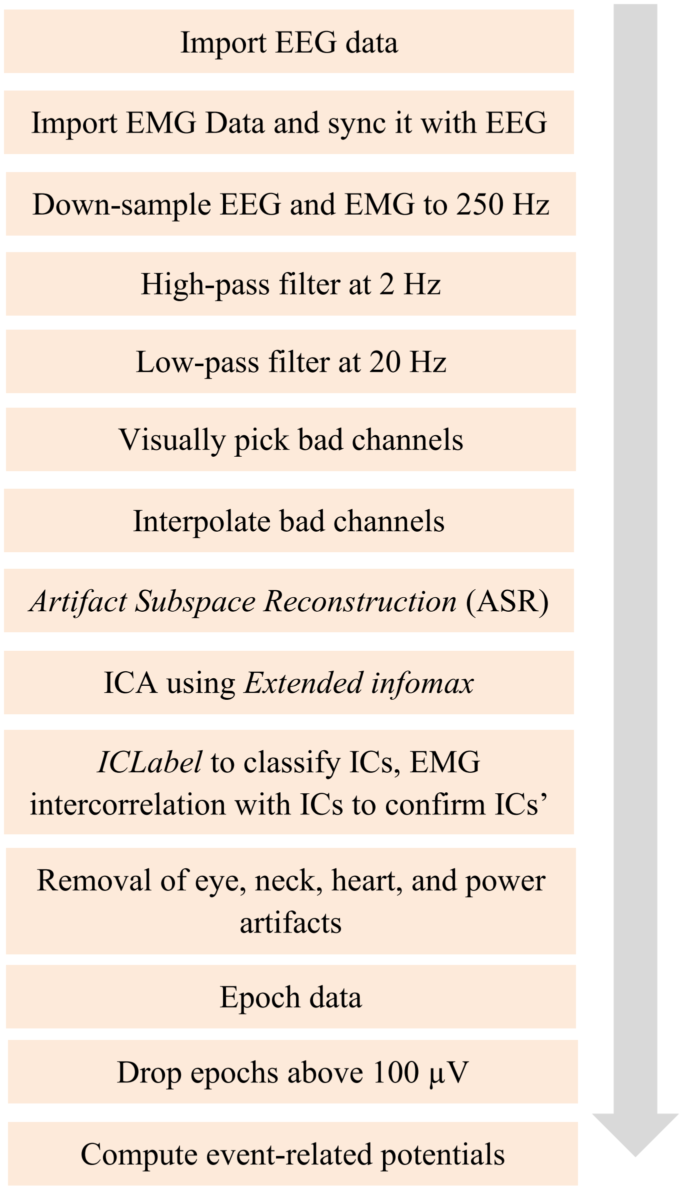


Figure 9. The best performing pipeline which was used in the ERP analysis.

Table 1

Number of epochs removed in each preprocessing pipeline by autorejection with 100 μV as a rejection criterion.

	Pipeline 0	Pipeline 0b	Pipeline 1	Pipeline 2	Pipeline 3
N Epochs Total	886	886	886	886	886
N Epochs Rejected	773	741	98	129	127
% Epochs Rejected	87.2	83.6	11,1	14.6	14.3

Preprocessing pipelines 1-3 had considerably less epochs removed than the first two pipelines (Table 1). Hence, as expected, the Chi-Square test revealed a significant difference in number of removed epochs with autorejection of 100μV between the five preprocessing pipelines (X^2 (4, 4430) = 2272.9, $p < .001$).

Post-hoc analyses using planned comparisons for Chi-Square tests show a significant difference was found for pipeline 1, X^2 = (5.87), $p = .046$, while no significant differences were found for preprocessing pipelines 2, X^2 = (1.77), $p = .55$, or preprocessing pipeline 3, X^2 = (1.19), $p = .83$.

PSD plots were created for mEEG and still-EEG datasets of single participant (Figure 10). Although much of the variance caused by noise in mEEG has been cleaned, the mEEG PSD-plot still shows more within-electrode variance compared to the still-EEG PSD-plot.

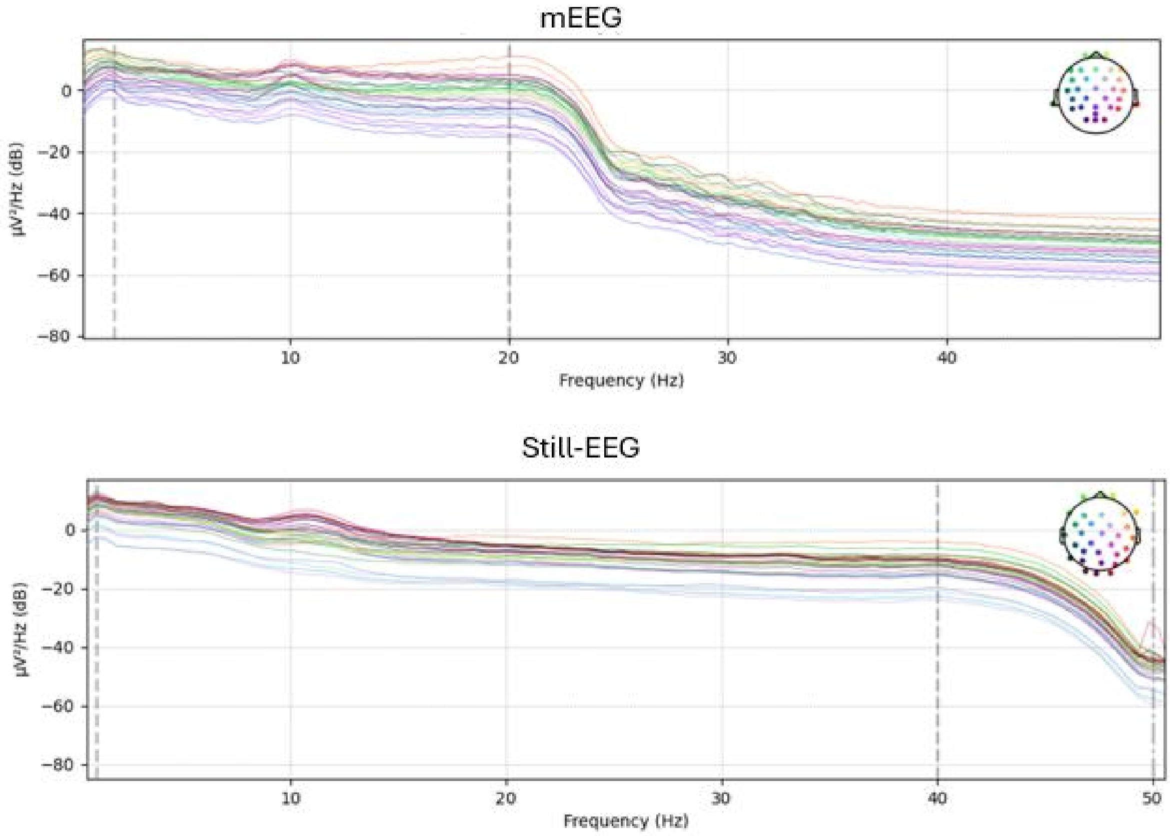
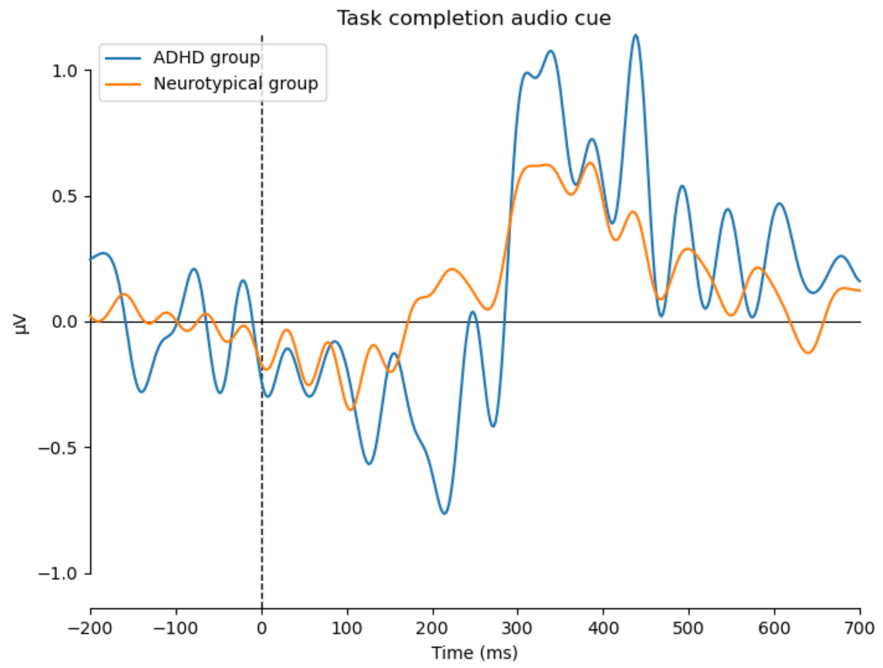


Figure 10. Spectral plots from mEEG and still-EEG for the same participant were created. In y-axis, power of the signal of the electrode in decibels, in X-axis frequency in hertz. Each individual line represents an electrode. Dashed lines represent filters.

3.2 Hypothesis 2: Reduction is ERP amplitude in ADHD group.

ERPs for the distraction audio stimuli and task completed sound cue stimuli were calculated for both the ADHD-group and the control group. For the ERP-analysis, preprocessing pipeline 1 was used (Figure 11), as it had the highest performance of the pipelines used in the present study.

A



B

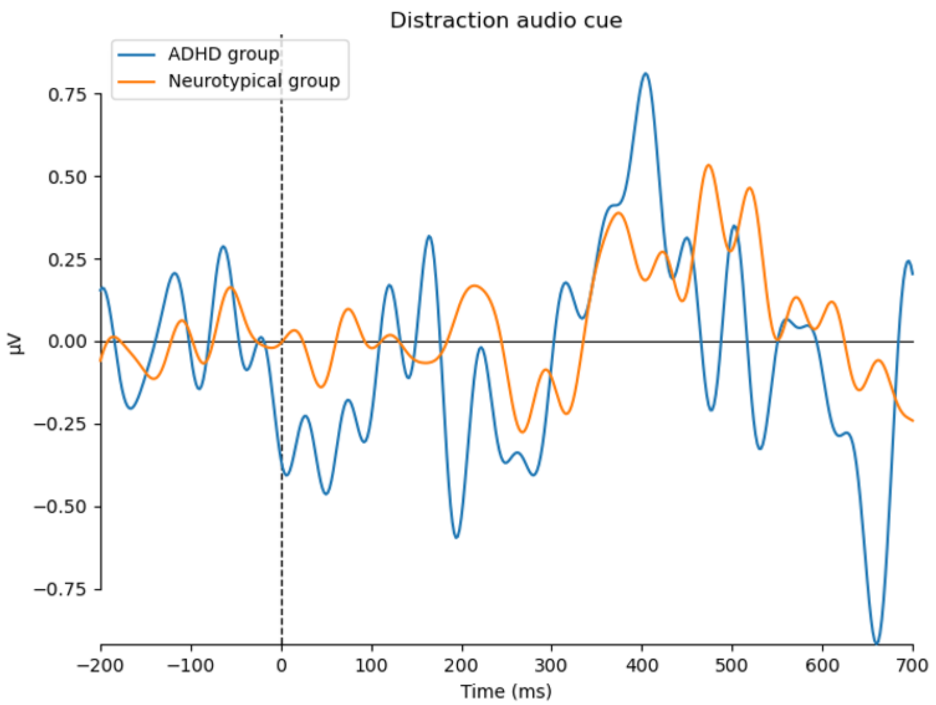


Figure 11. ERP following audio cue for task completion (A) and distraction audio (B) for ADHD and neurotypical groups. These ERPs were computed using all epochs for both groups and events. On x-axis is time in milliseconds, and on Y-axis is amplitude in microvolts.

Both groups displayed a positive peak in amplitude 300 to 400 milliseconds after task completion sound was played, reaching peak of 1.5 μV for the ADHD group, and 0.8 μV for the neurotypical group. Distraction sound cue produced positive raise in amplitude after 300 to 400 milliseconds for both groups.

4 Discussion

The present study aimed to create preprocessing pipeline for naturalistic mEEG tasks that include whole body movement of the participant and compare denoising pipelines for EEG and mEEG data. A successful method for denoising the data was based on optimizing the mEEG data for ICA to find as many muscle-noise related components as possible. Using a high-pass filter at 2 Hz instead of traditional 1 Hz made ICA results easier to interpret regarding muscle sourced noise (i.e., due to traditional source localization of the components, and general shape of the components). Compared to traditional methods for denoising EEG data, we found that using additional EMGs on the neck muscles, and identifying ICs based on these EMG sensors helped the process of denoising mEEG signal. Using EMG with mEEG increased confidence in determining the source of the IC, and thus resolving whether it was due to muscle -related noise or brain signal.

Additionally, pipeline 3 using a regression-based method to directly remove muscle-noise from the full signal was found to be not effective. Other than EMG -related methods, data quality was improved by methods such as ASR. ASR was used exclusively to clean high intensity artifacts, which are typically caused by quick and large movements, such as rapid head turning. Whereas ASR cleaned these artifacts very well, it remains unclear how much brain signal was retained. ASR captured high intensity artifacts, for example proportion of the muscle artifacts, removed them, and interpolated the data. Whereas the variance removed led to low amount of noise, judging by the figures, some of the signal was removed and interpolated as artificial data.

MNE-EOGRegression was applied in the present study, but the effects were minimal. In fact, for most datasets no difference was seen in visual inspection. Regression tools like MNE-EOGRegression could be useful in cleaning eye-related artifacts, but they were inefficient for more complex movement artifacts, such as neck movement artifacts and gait artifacts.

In our experiment, there were large differences in data quality among participants. Within the same paradigm, some datasets required more intensive preprocessing than others. Individuals who had faster gait and swift movements, exhibited less noise in the data, requiring methods such as ASR. On the other hand, slower gait and slower movements led to less noise in the data. In less noisy datasets, methods like ASR may not be as necessary as in noisier datasets. Future studies should compare preprocessing pipelines with and without ASR, while evaluating both signal and noise.

Naturalistic tasks reflect different movement styles and patterns of the participants. More careful and slow-paced walking and head movement does not cause as extensive noise due to decrease in quick head movements. VR related tasks, including XR used in our study, seemingly made the participants walk slower and with more care, with especial care on head movements. This could be, for instance, due to muscle tension caused from the weight of the XR-device or the participants' careful behaviour with expensive technology. Different tasks cause different types of movement, and this is rather important to consider when building preprocessing pipelines for mEEG. With naturalistic tasks, movements can vary significantly depending on the task, and the behaviour of the participant. Further studies could explore tailoring different preprocessing pipelines for different movement types. This could be achieved by categorizing participants' movement, for example, according to the intensity of movements. However, this may not be plausible, as tailoring different preprocessing methods for participants within a study could cause data to be skewed.

Although head movements cause substantial artifacts, this can be tackled with help of EMGs, as was done in the present study. Further studies could look into additional EMGs, especially different muscles on the neck area. Previous studies have used EMGs in the legs of the participants, but intercorrelations with ICs and leg EMGs were too small to be considered for removal of the related ICs. Previous studies, such as the one by Jacobsen and colleagues (2020), found that EMGs were useful for denoising gait-related noise, but in our study this was not the case. It could be due to the quality of our EEG cap or the task structure that did not cause quick leg movements. Different types of experimental tasks the participants perform cause different types of artifacts, thus design of the paradigm is extremely important for the quality of mEEG data.

Some of the preprocessing methods that modified the data greatly, for example ASR, appeared to decrease the peak amplitude of the ERP. While the data itself was cleaned more and there were less artifacts, the absolute response amplitudes were reduced. With exploration of the datasets here, larger amplitudes could be seen with worse performing denoising pipelines. This reflects the notion that peak amplitude alone is perhaps not the most suitable ERP metrics (Clayson et al., 2013)

Both neurotypical and ADHD groups exhibited positive peak amplitudes for both types of audio stimuli played. While the sounds used to produce these ERPs were not standard but naturalistic, they show some resemblance to P3 component. The application of mEEG in naturalistic paradigms offers a promising way to study cognition, especially the neural correlates for ADHD.

Further, different denoising methods may also negatively impact the signal-to-noise ratio if not closely inspected. As there is no ground truth on how the data should look like, one can is not sure whether what part of the removed data is artifact, and how much signal is accidentally removed with the denoising methods.

4.1 Limitations of the study

One major limitation in this study was small sample size, which was partially due to technical difficulties causing datasets to be unusable, especially for the ADHD group. Further studies could aim to replicate the study setting and study neural correlates of ADHD with mEEG.

Another limitation was that due to technical difficulties, CCA and MNE-Regression did not perform as expected. Further studies could investigate using these methods in different paradigms and datasets, along with different variations of preprocessing methods.

Technological complexity of this study caused additional work, as for data synchronization for the different formats of data had to match perfectly in order to be useful in denoising. A lot of data were not applicable due to event data being lost either from EMG or EEG data. As the recordings for EEG and EMG, and event data came all from different sources, even minor latency issues could have caused the EMG-data and EEG-data to be off sync leading to smaller intercorrelations for EMG and EEG data, and cause issues with denoising the EEG and EMG data. Further studies should also consider using already synchronized EMG to avoid latency and synchronization related issues.

The methods for measuring denoising capabilities of the preprocessing pipeline were suboptimal. Comparing the number of epochs removed with high absolute value amplitudes only assesses the noise left in the data, without assessing the amount of signal retained. Although some of the signal was detected via ERPs, it remains unclear how much of the signal was lost due to denoising methods. Future studies should aim to replicate common tasks with mEEG and compare the results of the same participants performing similar tasks with still-EEG.

Concerns about reliability can arise with naturalistic tasks. In the present study, ERPs of the participants were calculated based on sounds played. However, one cannot be sure how the stimuli was perceived, as there was a lot of different stimuli showing and playing at the same time. Especially in paradigms as in the present study, large amount of information available for the participant may influence their cognition. For instance, the inclusion of XR might have been stimulating for the participants, as XR is a novel technology that most had not experienced prior to the study. Further, the inclusion of XR was not necessary for the present study, as the sounds could have been played by speakers or headphones, and only real objects could have been used instead.

4.2 Future directions

The application of mEEG in naturalistic XR paradigms has a significant promise for investigating the neural underpinnings of developmental disorders such as ADHD in ecologically valid contexts. However, it is important to keep in mind concerns regarding generalizability and reliability of these findings, as it is necessary to have larger sample sizes. Future studies should consider using other denoising methods, such as Unfolding toolbox (Dimigen et al., 2019), which allows the denoising of the (m)EEG data within one package. Novel methods for denoising mEEG data are actively developed. With the breakthroughs in artificial intelligence, it could be interesting to see novel denoising methods that apply related methods for denoising movement artifacts. For instance, more efficient preprocessing methods that not only denoise mEEG data but also effectively retain signal could enhance the validity of mEEG as a tool for measuring brain activation during tasks with full-body movement.

4.3 Conclusion

In this thesis, preprocessing pipelines using combination of methods of ICA with muscle EMGs, filtering, ASR, and regression that successfully denoised mEEG data from a naturalistic task were built. Whereas most studies have not applied neck and limb EMGs, they were found to be important in denoising of the data. Using the developed pipeline opens a possibility to measure brain signal with mEEG from naturalistic paradigms that include movement, and it could be used to investigate the neural underpinnings of developmental disorders. Initial piloting for studying ADHD using naturalistic XR-paradigm and mEEG was conducted. The present results advocate conducting further studies using naturalistic settings and mEEG.

4.4 Acknowledgements

I would like to thank my supervisors Juha Salmitaival and Henry Railo for helping me through the thesis, and Translational Cognitive Neuroscience Lab (Salmi Lab) for the opportunity to using their study design, materials, and data for the study. Further, I would like to thank Sofia Tauriainen and Aquilino Beatrice for data collection, Marijn van Vliet for assistance with MNE-python, and Benedikt Ehinger for suggestions on building preprocessing pipelines.

References

Abbas, J. R., Kenth, J. J., & Bruce, I. A. (2020). The role of virtual reality in the changing landscape of surgical training. The Journal of Laryngology & *Otology,* 134(10), 863–866. https://doi.org/10.1017/s0022215120002078

Aboitiz, F., Ossandon, T., Zamorano, F., Palma, B., & Carrasco, X. (2014). Irrelevant stimulus processing in ADHD: Catecholamine dynamics and Attentional Networks. *Frontiers in Psychology*, 5. https://doi.org/10.3389/fpsyg.2014.00183

American Psychiatric Association. (2022). Diagnostic and statistical manual of mental disorders (5th ed). https://doi.org/10.1176/appi.books.9780890425787

American Psychiatric Publishing. (2015). Attention-deficit hyperactivity disorder -- Diagnosis, Autism spectrum disorders -- Diagnosis. In *Neurodevelopmental Disorders: DSM 5® Selections* (1st ed., pp. 78–87).

Arjona Valladares, A., Gómez, C. M., Rodríguez-Martínez, E. I., Barriga-Paulino, C. I., Gómez-González, J., & Diaz-Sánchez, J. A. (2020). Attention-deficit/hyperactivity disorder in children and adolescents: An event-related potential study of working memory. *The European Journal of Neuroscience*, *52*(10), 4356–4369. https://doi.org/10.1111/ejn.14767

Artoni, F., Barsotti, A., Guanziroli, E., Micera, S., Landi, A., & Molteni, F. (2018). Effective synchronization of EEG and EMG for Mobile Brain/Body Imaging in clinical settings. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00652

Beckwith, J., Beckwith, C. (2022). Automatic EEG artifact removal using blind-source separation methods. *Clinical Investigation*, DOI: 10.37532/2041-6792.2022.12(12).253-258

Barry, R. J., Clarke, A. R., McCarthy, R., Selikowitz, M., Brown, C. R., & Heaven, P. C. L. (2009). Event-related potentials in adults with attention-deficit/hyperactivity disorder: An investigation using an inter-modal auditory/visual oddball task. International *Journal of Psychophysiology*, 71(2), 124–131. https://doi.org/10.1016/j.ijpsycho.2008.09.009

Barry, R. J., Steiner, G. Z., De Blasio, F. M., Fogarty, J. S., Karamacoska, D., & MacDonald, B. (2020). Components in the P300: Don't forget the Novelty P3. *Psychophysiology*, *57*(7), e13371-n/a. https://doi.org/10.1111/psyp.13371

Beasley, T. M., & Schumacker, R. E. (1995). Multiple regression approach to analyzing contingency tables: Post hoc and planned comparison procedures. The Journal of Experimental Education, 64(1), 79-93.

Blum, S., Jacobsen, N. S. J., Bleichner, M. G., & Debener, S. (2019). A Riemannian Modification of Artifact Subspace Reconstruction for EEG Artifact *Handling. Frontiers in Human Neuroscience*, 13. https://doi.org/10.3389/fnhum.2019.00141

Butkeviciute, E., Bikulciene, L., Sidekerskiene, T., Blazauskas, T., Maskeliunas, R., Damasevicius, R., & Wei, W. (2019). Removal of movement artefact for mobileEEG analysis in sports exercises. *IEEE Access,* 7, 7206–7217. https://doi.org/10.1109/access.2018.2890335

Castellanos, F.X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews Neuroscience*, 3, 617–628.

Chang, M., Büchel, D., Reinecke, K., Lehmann, T., & Baumeister, J. (2022). Ecological validity in Exercise Neuroscience Research: *A systematic investigation. European Journal of Neuroscience,* 55(2), 487–509. https://doi.org/10.1111/ejn.15595

Chang, C.-Y., Hsu, S.-H., Pion-Tonachini, L., & Jung, T.-P. (2020). Evaluation of Artifact Subspace Reconstruction for automatic artifact components removal in multi-channel EEG Recordings. *IEEE Transactions on Biomedical Engineering*, *67*(4), 1114–1121. https://doi.org/10.1109/tbme.2019.2930186

Clayson, P. E., Baldwin, S. A., & Larson, M. J. (2013). How does noise affect amplitude and latency measurement of event-related potentials (ERPs)? A methodological critique and simulation study. *Psychophysiology*, 50, 174-186.

Croft, R. J., & Barry, R. J. (2000). Removal of ocular artifact from the EEG: A Review. *Neurophysiologie Clinique/Clinical Neurophysiology*, *30*(1), 5–19. https://doi.org/10.1016/s0987-7053(00)00055-1

Delorme A & Makeig S (2004) EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics, *Journal of Neuroscience Methods* 134:9-21.

Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PLoS ONE*, *7*(2). https://doi.org/10.1371/journal.pone.0030135

Dennis, E. J., El Hady, A., Michaiel, A., Clemens, A., Tervo, D. R. G., Voigts, J., and Datta, S. R. (2021). Systems neuroscience of natural behaviors in rodents. *J. Neurosci*. 41, 911–919. doi: 10.1523/JNEUROSCI.1877-20.2020

Dickstein, S. G., Bannon, K., Xavier Castellanos, F., & Milham, M. P. (2006). The neural correlates of attention deficit hyperactivity disorder: An ale meta-analysis. *Journal of Child Psychology and Psychiatry*, *47*(10), 1051–1062. https://doi.org/10.1111/j.1469-7610.2006.01671.x

Dimigen, O. (2020). Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments. *NeuroImage*, *207*, 116117. https://doi.org/10.1016/j.neuroimage.2019.116117

Dimigen, O., & Ehinger, B. V. (2019). Analyzing Combined Eye-Tracking/EEG Experiments with (Non)Linear Deconvolution Models. https://doi.org/10.1101/735530

Ehinger, B.V., & Dimigen, O., (2019). Unfold: An integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis, *peerJ*, https://peerj.com/articles/7838/

Garcia-Perez, M. A., & Nunez-Anton, V. (2003). Cellwise residual analysis in two-way contingency tables. *Educational and psychological measurement*, *63*(5), 825-839.

Gramann, K. (2014). Towards a new cognitive neuroscience: Modeling Natural Brain Dynamics. *Frontiers Research Topics*. https://doi.org/10.3389/978-2-88919-271-7

Gerini, L., Solari, F., & Chessa, M. (2023). Passive haptic feedback for more realistic and efficient grasping movements in virtual environments. *Extended Reality*, 3–22. https://doi.org/10.1007/978-3-031-43401-3_1

Gottlieb, J., & Oudeyer, P.-Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, 19(12), 758–770. https://doi.org/10.1038/s41583-018-0078-0

Gramfort, A., Luessi, M., Larson, E., Engeman, D., Sthrohmeier, D., Brodbeck, C., Goj. R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2014). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience,* 7(267):1–13, 2013. doi:10.3389/fnins.2013.00267.

Gramfort, A., Luessi, M., Larson, E., Engeman, D., Sthrohmeier, D., Brodbeck, C., Goj. R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2014). MNE software for processing MEG and EEG data. *NeuroImage,* 86:446–460. doi:10.1016/j.neuroimage.2013.10.027.

Rodrigo Grassi-Oliveira, Leandro Fernandes Malloy-Diniz, & Debora Marques de Miranda. (2017). *Executive Functions in Psychiatric Disorders*. Frontiers Media SA.

Gorjan, D., Gramann, K., De Pauw, K., & Marusic, U. (2022). Removal of movement induced EEG artifacts: Current state of art and guidelines. *Journal of Neural*

Engineering. https://doi.org/10.1088/1741-2552/ac52d2

Gütlin, D (2021) Artifact Subspace Reconstruction for Python (ASRpy) [Python3 package for MNE-Python]

Gutiérrez-Zamora Velasco, G., Fernández, T., Silva-Pereyra, J., Reynoso-Alcántara, V., & Castro-Chavira, S. A. (2021). Higher cognitive reserve is associated with better working memory performance and working-memory-related p300 modulation. *Brain Sciences*, *11*(3), 1–10. https://doi.org/10.3390/brainsci11030308

Gwin, J. T., Gramann, K., Makeig, S., & Ferris, D. P. (2010). Removal of movement artifact from high-density EEG recorded during walking and running. *Journal of Neurophysiology,* 103(6), 3526–3534. https://doi.org/10.1152/jn.00105.2010

Hasler, R., Perroud, N., Meziane, H. B., Herrmann, F., Prada, P., Giannakopoulos, P., & Deiber, M.-P. (2016). Attention-related EEG markers in adult ADHD. *Neuropsychologia,* 87, 120–13 https://doi.org/10.1016/j.neuropsychologia.2016.05.008

Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, *28*(3/4), 321. https://doi.org/10.2307/2333955

Hsu, S.-H., Pion-Tonachini, L., Palmer, J., Miyakoshi, M., Makeig, S., & Jung, T.-P. (2018). Modeling brain dynamic state changes with adaptive mixture independent component analysis. *NeuroImage* (Orlando, Fla.), 183, 47–61. https://doi.org/10.1016/j.neuroimage.2018.08.001

Hunter, J.D. "Matplotlib: A 2D Graphics Environment", *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90-95, 2007.

Hunter, J., Dale, D., Firing, E., Droettboom, M., and the Matplotlib development team; 2012 - 2022. Matplotlib package for Python

IBM Corp. (2020). IBM SPSS Statistics for Windows (Version 27.0) [Computer software]. IBM Corp.

Jacobsen, N. S., Blum, S., Witt, K., & Debener, S. (2020). A walk in the park? characterizing gait-related artifacts in mobile EEG Recordings. European *Journal of Neuroscience,* 54(12), 8421–8440. https://doi.org/10.1111/ejn.14965

Janani, A. S., Grummett, T. S., Lewis, T. W., Fitzgibbon, S. P., Whitham, E. M., DelosAngeles, D., Bakhshayesh, H., Willoughby, J. O., & Pope, K. J. (2018). Improved artefact removal from EEG using canonical correlation analysis and spectral slope. *Journal of Neuroscience Methods*, *298*, 1–15. https://doi.org/10.1016/j.jneumeth.2018.01.004

Jylkkä, J. (2023). Assessment of Goal-Directed Behavior with the 3D Videogame Epeli: Psychometric Features in a Web-Based Adult Sample. *PLoS One* https://doi.org/10.31234/osf.io/7bquv

Kim, S., Liu, Z., Glizer, D., Tannock, R., & Woltering, S. (2014). Adult ADHD and working memory: Neural evidence of impaired encoding. *Clinical Neurophysiology*, *125*(8), 1596–1603. https://doi.org/10.1016/j.clinph.2013.12.094

Kim, S. J., Tanglay, O., Chong, E. H., Young, I. M., Fonseka, R. D., Taylor, H., Nicholas, P., Doyen, S., & Sughrue, M. E. (2023). Functional connectivity in ADHD children doing go/no-GO TASKS: An fmri systematic review and meta-analysis. *Translational Neuroscience*, *14*(1). https://doi.org/10.1515/tnsci-2022-0299

Klug M, Gramann K. Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments. *Eur J Neurosci*. 2021 Dec;54(12):8406-8420. doi: 10.1111/ejn.14992. Epub 2020 Oct 15. PMID: 33012055.

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., and Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. *Neuron* 93, 480–490. doi: 10.1016/j.neuron.2016.12.041

Krugliak, A., & Clarke, A. (2022). Towards real-world neuroscience using mobile EEG and augmented reality. *Scientific Reports,* 12(1). https://doi.org/10.1038/s41598-022-06296-3

Käthner, I., Wriessnegger, S. C., Müller-Putz, G. R., Kübler, A., & Halder, S. (2014). Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. *Biological Psychology*, *102*(1), 118–129. https://doi.org/10.1016/j.biopsycho.2014.07.014

Kysow, K., Park, J., & Johnston, C. (2016). The use of compensatory strategies in adults with ADHD symptoms. *ADHD Attention Deficit and Hyperactivity Disorders*, *9*(2), 73–88. https://doi.org/10.1007/s12402-016-0205-6

Ladouce, S., Donaldson, D. I., Dudchenko, P. A., & Ietswaart, M. (2019). Mobile EEG identifies the re-allocation of attention during real-world activity. *Scientific Reports,* 9(1). https://doi.org/10.1038/s41598-019-51996-y10

Lau-Zhu, A., Lau, M. P. H., & McLoughlin, G. (2019). Mobile EEG in research on neurodevelopmental disorders: Opportunities and challenges. *Developmental Cognitive Neuroscience*, 36, 100635. https://doi.org/10.1016/j.dcn.2019.100635

Li, A., Höchenberger, R., Donoghue, T., Scheltienne, M., Feitelberg, J., Saini, Anand. (2022) MNE-ICAlabel: Automatically annotating ICA components with ICLabel in Python. *Journal of Open Source Software*, doi: 10.21105/joss.04484

Li, Q., Sun, M., Song, Y., Zhao, D., Zhang, T., Zhang, Z., & Wu, J. (2023). Mixed reality-based brain computer interface system using an adaptive bandpass filter: Application to remote control of mobile manipulator. *Biomedical Signal Processing and Control*, *83*, 104646-. https://doi.org/10.1016/j.bspc.2023.104646

Luck, S.J. and Kappenman, E.S. (2013) *The Oxford Handbook of Event-related potential components*. Oxford: Oxford University Press.

Luck, S.J. *An Introduction to the Event-Related Potential Technique*. (2014). 2nd ed. Cambridge, US: Mit Press, Print.

MATLAB. (2010). version 9.14.0.2306882 (R2023a). Natick, Massachusetts: The MathWorks Inc.

Michelini, G., Salmastyan, G., Vera, J. D., & Lenartowicz, A. (2022). Event-Related Brain Oscillations in Attention-Deficit/Hyperactivity Disorder (ADHD): A Systematic Review and Meta-Analysis. https://doi.org/10.31234/osf.io/b73dy

Moen, Ø. L., Hall-Lord, M. L., & Hedelin, B. (2011). Contending and Adapting Every Day: Norwegian Parents' Lived Experience of Having a Child With ADHD. *Journal of Family Nursing*, *17*(4), 441–462. https://doi.org/10.1177/1074840711423924

Mullen, T., Kothe, C., Chi, Y. M., Ojeda, A., Kerth, T., Makeig, S., Cauwenberghs, G., & Tzyy-Ping Jung. (2013). Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. 2013 35th *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBC). https://doi.org/10.1109/embc.2013.6609968

Muthukumaraswamy, S. D. (2013). High-frequency brain activity and muscle artifacts in MEG/EEG: A review and recommendations. *Frontiers in Human Neuroscience*, 7. https://doi.org/10.3389/fnhum.2013.00138

Park, J. L., & Donaldson, D. I. (2019). Detecting the neural correlates of episodic memory with mobile EEG: Recollecting objects in the real world. *NeuroImage*, 193, 1–9. https://doi.org/10.1016/j.neuroimage.2019.03.013

Parsons, T. D., Carlew, A. R., Magtoto, J., & Stonecipher, K. (2015). The potential of function-led virtual environments for ecologically valid measures of executive function in experimental and clinical neuropsychology. *Neuropsychological Rehabilitation,* 27(5), 777–807. https://doi.org/10.1080/09602011.2015.1109524

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLABEL: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Railo, H., Kraufvelin, N., Santalahti, J., Laine, T. (2023) 'Rapid withdrawal from a threatening animal is movement-specific and mediated by reflex-like neural processing', *NeuroImage*, 283, p. 120441. doi:10.1016/j.neuroimage.2023.12044.

Robles, R., Kuziek, J., Wlasitz, N., Bartlett, N., Hurd, P., Mathewson, K. (2021). EEG in motion: Using an oddball task to explore motor interference in active skateboarding. *European Journal of Neuroscience.* 54(12), 8196-8213. https://doi.org/10.1111/ejn.15163

Rohani, D. A., Sorensen, H. B., & Puthusserypady, S. (2014). Brain-computer interface u sing P300 and virtual reality: A gaming approach for treating ADHD. 2014 3 *6th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. https://doi.org/10.1109/embc.2014.6944403

Rostami, M., Farashi, S., Khosrowabadi, R., & Pouretemad, H. (2020). Discrimination of ADHD subtypes using decision tree on behavioral, neuropsychological, and neural markers. *Basic and Clinical Neuroscience*, 11(3), 359–368. https://doi.org/10.32598/bcn.9.10.115

Sabbagh, A. J., Bajunaid, K. M., Alarifi, N., Winkler-Schwartz, A., Alsideiri, G., Al-Zhrani, G., Del Maestro, R. F. (2020). Roadmap for Developing Complex Virtual Reality Simulation Scenarios: Subpial Neurosurgical Tumor Resection Model. *World Neurosurgery*, *139*, e220–e229. https://doi.org/10.1016/j.wneu.2020.03.187

Seesjärvi, E., Puhakka, J., Aronen, E. T., Lipsanen, J., Mannerkoski, M., Hering, A., Zuber, S., Kliegel, M., Laine, M., & Salmitaival, J. (2021). Quantifying ADHD Symptoms in Open-Ended Everyday Life Contexts with a New Virtual Reality Task. https://doi.org/10.31234/osf.io/2vksh

Sonkusare, S., Breakspear, M., & Guo, C. (2019). Naturalistic stimuli in neuroscience: Critically acclaimed. *Trends in Cognitive Sciences*, 23(8), 699–714. https://doi.org/10.1016/j.tics.2019.05.004

St. Louis, E., Frey, L., Britton, J., Hopp, J., Korb, P., Koubeissi, M., Lievens, W., & Pestana-Knight, E. (2016). Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants. http://dx.doi.org/10.5698/978-0-9979756-0-4.

Sugden, N., Thomas, M., & Kiernan, M. (2021). A scoping review of the utility of self-report and informant-report prospective memory measures. *Neuropsychological Rehabilitation*, *32*(6), 1230–1260. https://doi.org/10.1080/09602011.2021.1875851

Sweeney, K. T., McLoone, S. F., & Ward, T. E. (2013). The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel Artifact removal technique. *IEEE Transactions on Biomedical Engineering*, 60(1), 97–105. https://doi.org/10.1109/tbme.2012.2225427

Szuromi, B., Czobor, P., Komlósi, S., & Bitter, I. (2010). P300 deficits in adults with attention deficit hyperactivity disorder: A meta-analysis. *Psychological Medicine*, *41*(7), 1529–1538. https://doi.org/10.1017/s0033291710001996

Thurley, K. (2022). Naturalistic neuroscience and virtual reality. *Frontiers in Systems Neuroscience,* 16. https://doi.org/10.3389/fnsys.2022.896251

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods,* 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wang, J.-W., Zhang, D.-W., & Johnstone, S. J. (2023). Portable EEG for Assessing Attention in Educational Settings: A Scoping Review. https://doi.org/10.31234/osf.io/ckbmw

Watson, Todd D. et al. "Event-Related Potentials as an Index of Similarity between Words and Pictures." *Psychophysiology* 42.4 (2005): 361–368.

Willcutt, E.G., Doyle, A.E., Nigg, J.T., Faraone, S.V., & Pennington, B.F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, 57, 1336–1346.

Wolf, D., Mittelberg, I., Rekittke, L.-M., Bhavsar, S., Zvyagintsev, M., Haeck, A., Cong, F., Klasen, M., & Mathiak, K. (2018). Interpretation of social interactions: Functional imaging of cognitive-semiotic categories during naturalistic viewing. *Frontiers in Human Neuroscience*, 12. https://doi.org/10.3389/fnhum.2018.00296

Zani, A., & Zani, A. (2020). *ERP and EEG Markers of Brain Visual Attentional Processing*. Basel, Switzerland: MDPI - Multidisciplinary Digital Publishing Institute.

Xu, J., Yazicioglu, R., Van Hoof, C., Makinwa, K. Low power active electrode ICS for wearable EEG Acquisition. (2019). *Springer international publishing.*

Appendices

Appendix 1 Spectral plot comparison of pipeline 0b and pipeline 1

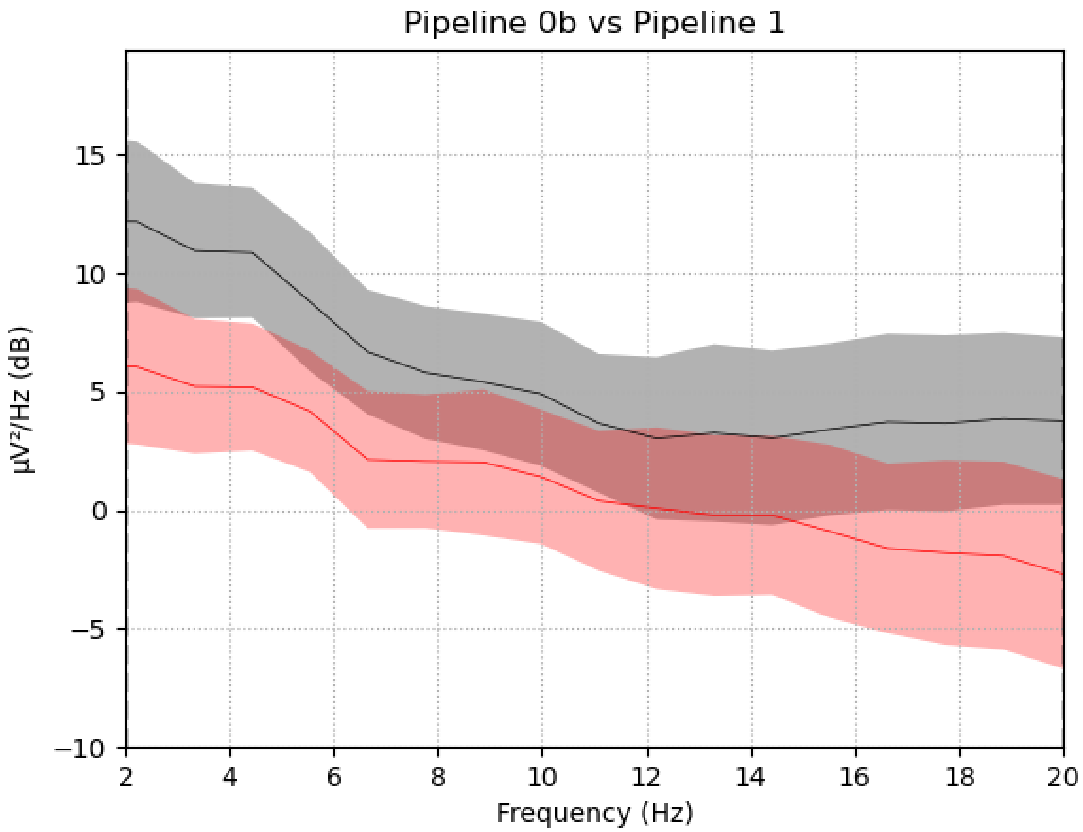


Figure 12. Pipeline 0b and pipeline 1 spectral density plots were plotted together. In grey, pipeline 0b amplitudes in each frequency. In pink, pipeline 1 amplitudes in each frequency, and in brown the overlap of the two. Figure shows the effect of preprocessing. Y axis shows spectral density amplitude in decibels, while X-axis shows frequency in hertz. Epochs used for this plot were baseline corrected.