

MENG FINAL YEAR PROJECT

FINAL REPORT

**Comparative Research on the Changes in Cognitive
Load following Auditory and Haptic Stimulation**

Student: Matilde Piccoli

CID: 01764158

Supervisor: Dr. Adam Spiers

Second Marker: Dr. Chen Qin

June 21, 2023

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING

Plagiarism Statement

I affirm that I have submitted, or will submit, an electronic copy of my final year project report to the provided EEE link.

I affirm that I have submitted, or will submit, an identical electronic copy of my final year project to the provided Blackboard module for Plagiarism checking.

I affirm that I have provided explicit references for all the material in my Final Report that is not authored by me, but is represented as my own work.

I have sporadically used ChatGTP v3.5 to improve the quality of my English and the Latex formatting; however, all technical content and references are my original work.

Acknowledgements

I would like to thank Dr. Ad Spiers for the opportunity to complete a project under his supervision and the guidance and encouragement he has offered me throughout.

I would also like to thank Dr. Katarina Poole for the valuable insight she has given me on neuroscientific topics involved in this project.

Abstract

The field of haptic technology and the exploration of its influence on the user state has only begun to emerge. This study aimed to lay the basis for this research by comparatively analysing the effects of auditory and haptic stimulus processing on the user's cognitive state and performance in a context requiring multi-tasking abilities. Our use case was to evaluate vibration and shape-changing-based haptics for navigation-assisting application, using a combination of physiological, task performance and subjective metrics. A passive Brain Computer Interface was used to record the participant's brain activity (using Electroencephalography) whilst performing audiovisual and tactile-visual tasks. The physiological results showed increased mental effort (lower β and γ power) for auditory compared to haptic stimulation. This was accompanied by lower engagement and stimulus receptiveness (greater δ power in the brain's frontal region, higher perceived effort and frustration and lower performance). The shape-changing-based haptic stimulation also outperformed the vibrotactile modality regarding task workload and stimulus awareness (lower perceived task demand and the higher α power). Nevertheless, we also recorded a more intense power in the θ band and lower time domain response in the frontal region, associated with increased mental load. Further research is needed to trace back the sources of these cognitive results with a more location-specific analysis of the brain activity. Overall, this study provided encouraging results on applying haptic-based stimulation to navigation-assisting interfaces and user-focused devices in general. Due to the novelty of this study, additional investigation of these results is expected.

Contents

1	Introduction	7
1.1	Project Specification	7
2	Background Research	9
2.1	EEG signals	9
2.2	Cognitive Load	11
2.3	Cognitive processing of sensory information	13
2.4	Haptic Stimulation Evaluation	13
2.5	Consumer-level Passive BCI validation	15
2.6	Standard Testing Methodologies	15
3	Experiment Design	17
3.1	Experiment Structure	17
3.2	Experiment Metrics	19
3.3	Ethics, Legal and Safety Measures	19
4	Hardware & Software Development	20
4.1	Brain Computer Interface	20
4.2	Auditory Stimulation	21
4.3	Vibro-Tactile Interface	21
4.4	Shape-Changing Interface	22
4.5	Software Development and Dataflow	23
5	Testing	24
5.1	Pipeline	24
5.2	Data Quality Checks	24
5.3	Mental State Classification	26
5.4	Visual Tests	27
5.5	Haptic Tests	27
6	Signal Processing & Analysis	28
6.1	Pre-processing and Denoising	28
6.2	Frequency domain Analysis	29
6.3	Time domain Analysis	30

6.4	Morlet Wavelets	33
6.5	Entropy Analysis	36
6.6	Independent Component Analysis	37
7	Results	40
7.1	Physiological Data Results	40
7.2	Task-Performance Results	45
7.3	Subjective Evaluation	47
8	Project Evaluation	48
9	Conclusions and Future Work	49
A	Appendix	55
A.1	Wavelets Analysis Results	55
A.2	ERP Analysis - RMS Results	58
A.3	ICA Results	58

1 Introduction

Recent studies have highlighted the improvement in task-specific performance when using haptic-based interfaces, both relative to visual-based ones[1] and when used in multi-sensory integration [2]. However, the current state of research lacks appropriate studies on the impact of different sensory stimulation on the cognitive load, especially when performing concurrent tasks, e.g. navigation and social interaction.

Furthermore, only a few studies have compared the impact of different haptic stimulation methodologies, e.g. pressure, vibration, shape-change-based feedback etc., on any task-related performance. To our knowledge, none of these so far have investigated the impact of these at the cognitive level.

Evaluating the effect of different sensory stimulation on our cognitive abilities is a fundamental step in understanding their impact on our daily lives and how they prevent us from focusing on concurrent tasks that require our attention. This represents a key element to consider when selecting a sensory stimulation methodology for devices that are meant to be used frequently and in situations requiring multitasking abilities. The applications range from choosing the type of sensory feedback for driving a smart car to enabling visually impaired people to better orient in a social environment.

1.1 Project Specification

Scope and methodology This study consisted of a comparative research on the effects of different sensory stimulation on multitasking performance and cognitive load, defined as the level of attention and mental fatigue. This project aimed to investigate three types of stimulation: an auditory, and two haptic ones, based on vibrating and shape-changing actuators.

Electroencephalography (EEG) was used to monitor the effects of the stimulation at the cognitive level using a passive Brain-Computer Interface (BCI). The measurements were performed using Muse 2, a low-cost, consumer-level BCI previously validated for research studies of similar nature (see section 2.5). The BCI detects the Event-Related Potentials (ERPs) manifesting in the frontal and temporal areas of the brain (AF7, AF8, TP9 and TP10 electrodes)¹ following changes in the cerebral activity, and reconstructs them in the form of electroencephalogram (EEG) signals.

These signals were streamed to a local device via Bluetooth and analysed to identify patterns in the time and frequency domain that could be indicative of changes in mental state, such as increased alpha power or variation in the P300 component, as discussed in section 2.2.2, 2.2.3, 2.4.

Research Hypotheses this project aimed to discover which, if any, of these stimulation methodologies constitutes a valid form of feedback (i.e. provokes an acceptable level of receptiveness in the subject) while having the smallest impact on the cognitive abilities when solving concurrent tasks.

The two main hypotheses investigated were:

- H_1 : In a multitasking environment, cognitive and task-related performance can improve by using haptic-based feedback instead of an auditory-based one.
- H_2 : Different types of haptic-based feedback influence the cognitive state and task performance in distinct ways.

Evaluation The target of this project was to successfully conduct the testing on 10 participants and obtain relevant comparative results on the effect of different sensory stimulation (possibly expanding to multiple haptic interfaces). The minimum goal was set to develop a fully functional pipeline, from comparing an auditory and a single type of haptic stimulation to the measurement of the cognitive load and the validation of the EEG data collected from the Muse device. The optimal project progression was to obtain novel results given the data from the physical (task-related), psychological (subjective), and cognitive performance evaluations.

These metrics were analysed individually and on a relative basis, i.e., examining the consistency/discrepancy in performance for each participant across all conditions and for each type of sensory stimulation or trial block number across all participants.

Given that this was a novel comparative study, it was not guaranteed that the experimental data would have confirmed the research hypotheses. Nonetheless, this project was expected to explore and give a better

¹Electrode nomenclature based on the expanded 10-20 system: [https://en.wikipedia.org/wiki/10%E2%80%9320_system_\(EEG\)](https://en.wikipedia.org/wiki/10%E2%80%9320_system_(EEG))

overview of the impact of haptic feedback modalities on the user’s cognitive abilities. Therefore, the project’s success will be evaluated based on the validity of the data collected and the level of development of the testing methodology, hardware-software integration, and data analysis.

Project structure The project flow was divided into four phases (Fig 1):

- Phase I consisted of a background literature review highlighting the current state of research in the field and a detailed project formulation.
- Phase II involved the hardware and software setup, familiarisation with the theoretical background, and the initial tests to evaluate the best tools to use. This phase also included formulating the experiment format and requesting ethics approval for the human trials.
- Phase III consisted of the development of the final experiment and data analysis tools, and data validation.
- Phase IV involved the human trials and parallel analysis of the data collected. This phase concluded with the research hypothesis and overall project evaluation.

Tasks	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June
Project Formulation	█								
Background Research		█							
BCI and Environment Setup			█						
Initial Testing			█						
Ethics Approval Process				█					
Experiment Development & Data Validation					█				
Experimental Trials							█		
Result Analysis & Evaluation								█	

Figure 1: Project Timeline

Applications This research study has a variety of applications, from Virtual Reality navigation for visually-impaired people [3] to concurrent performance of military and robotics tasks in simulated environments [4], or even grocery assisting systems [5]. However, the primary target of this project is to evaluate different haptic feedback options for a tactile interface that will aid visually-impaired people navigating in a social environment [6],[7].

This project represents an initial phase of the study that focuses solely on sighted people as a preliminary subject for this research; however, the natural progression will be to adapt and expand the experiment to visually-impaired people, given the target application.

2 Background Research

This section will cover the background theory and literature review required to understand the context and motivations of this project. Firstly, we will describe EEG signals and the current trends in EEG denoising and classification; then, we will dive into identifying the cognitive load, attention and mental fatigue from EEG data, after giving a formal definition of these concepts. Thirdly, there will be an overview of the current state of research on haptic-based feedback evaluation. And finally, we will be validating the BCI and Testing Methodology used in this project.

2.1 EEG signals

2.1.1 An overview

Electroencephalogramography is one of the most common and accessible ways of measuring the variation in cerebral activity. This non-invasive method requires placing several electrodes on the scalp to measure the potential caused by the electrical activity of the neurons [8]: every time a synapse occurs, a small current flows from the soma along the axon causing an action potential (AP); then, following the neurotransmitter release of ions at the axon's terminal boutons, this creates another current flow causing an excitatory (or inhibitory) postsynaptic potential (EPSP or IPSP, respectively).

When numerous neurons fire in a short amount of time, this causes an accumulation of potential in the pyramidal neurons, i.e. the ones closest to the cortical surface. If the potential is strong enough, this will be perceived by the electrodes on the scalp and amplified inside the BCI, resulting in spikes in the recreated EEG signal.

When the BCI is said to be detecting the Event-Related Potentials (ERPs), this indicates that the electrical activity detected from the brain is related to a cognitive activity following a specific event, e.g. sensory stimulation.

The electrical signals inside the brain have a frequency spectrum from 1 and up to 500Hz (including ripples), computed using Fast Fourier Transform (FFT); however, only the signals below 30Hz are conventionally used in clinical EEG applications [9]. For this reason, a bandpass filter is often used to limit the spectrum of the EEG between 0 and 40 Hz, and then separate further to obtain the four principal power bands: delta band (1-4Hz), theta band (4-8Hz), alpha band (8-13Hz), beta band (13-30Hz), and gamma band (30-40Hz).

As discussed by Fan et al. [10], the power spectrum can be analysed both in terms of the relative power of these frequency bands and in terms of power ratios (mainly, $\frac{\alpha+\theta}{\beta}$, $\frac{\alpha}{\beta}$, $\frac{\theta}{\beta}$, and $\frac{\alpha+\theta}{\beta+\alpha}$). These have been experimentally proven to be associated with different mental states and can be used as an indication of particular cognitive phenomena. For instance, alphas waves are usually associated with a normal awake state, whilst delta ones are recurring during deep sleep [9]. An example of EEG data and frequency-band extraction can be seen in Fig 2.

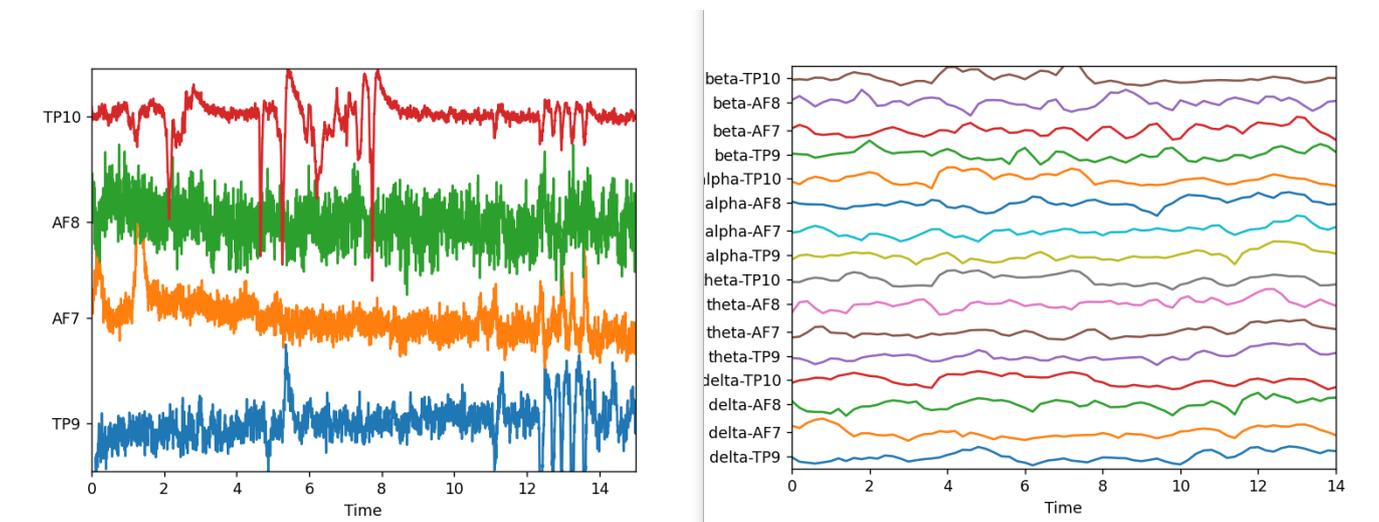


Figure 2: Example of raw EEG data and frequency band isolation for each of the four electrodes of the Muse 2 device

Regarding the ERP analysis, the standard nomenclature dictates that the component's name contains first the sign of its amplitude (P for positive peak and N for negative one) and then the approximate number of milliseconds after which the potential occurs with respect to the time of the event. For example, P300 is a positive deflection after 300ms from the event, or N170 correspond to a negative peak after 170ms. Similarly to the frequency bands, the amplitude and delay of these peaks, together with the brain region in which they manifest, can also be indicative of the mental state and overall cognitive behaviour. The analysis of these and other features is based on methodologies that will be discussed in section 2.1.3.

2.1.2 Noise in EEG signals

EEG signals are polluted by many types of noise and artefacts, mainly due to muscle movement, heart rate, power sources, environment and improper electrode positing. As explained in a literature review of EEG noise and artefact removal techniques by Lai et al. [11], the principal strategies involve the use of Independent Component Analysis (ICA), wavelet-based analysis, and other Machine Learning-based methods.

As a first standard denoising methodology, ICA is useful to decompose the signal into a new set of independent components (linear signals) that can ease the identification of artefacts. Some of these components can then be removed to reconstruct a cleaner signal. However, this often also causes a loss of valuable information; therefore, for some applications, it is recommended to use complementary techniques. ICA is also unable to remove the additive noise of each sensor, as opposed to the common component noise; for this reason, Factor Analysis (FA) can be helpful to remove these additive components in the first place, by considering the noise variance as well and ranking the principal components in order of signal decorrelation.

In addition, as proposed by Yin et al. [12], ICA and FA can be accompanied by Multivariate Empirical Mode Decomposition (MEMD) to reconstruct the original signal into an amplitude-frequency modulated set of components. This data-variant technique has been demonstrated to be particularly useful in identifying the high-oscillation components and isolating the ones most representative of the mental state features, as discussed in [13].

In parallel with MEMD, another common noise removal technique also used in statistical analysis is the FASTER method (Fully Automated Statistical Thresholding) for artefact removal proposed in [14]. This advanced method is characterised by high sensitivity to contaminated channels and epochs identification, as it interpolates between the different channels and compares with a grand average dataset created from all ERPs extrapolated; however, this technique works best with a high number of channels to extract from, as opposed to the four available in the Muse 2 device used in this project.

Finally, the current state of research aims to develop an automated way of EEG denoising based on Machine and Deep Learning techniques. Although this currently represents a significant technical challenge, an essential step in this field has been achieved by Zhang et al. [15] with EEGdenoiseNet, a dataset of clean EEG and artefacts recordings that enables to train and evaluate deep learning networks performance in EEG noise detection.

2.1.3 EEG feature extraction and classification

EEG signals contain various information that can be extracted from their temporal and time-frequency representation. For this purpose, it is common to use dimensionality reduction techniques, such as Principal Component Analysis (PCA) [16], or other decomposition methodologies, such as Wavelets Transforms, to ease the identification of the most informative components of the signal. In particular, Morlet Wavelets are a common technique used in neuroscience due to the ease of the physiological interpretability and the ability to highlight non-stationary features of the EEG data, as will be seen in section 6.4. Furthermore, if we predict that the data is the combination of several unknown non-Gaussian and independent underlying components to be separately analysed [17], Independent Component Analysis can also be used for the feature extraction, similarly to as described in the denoising section.

As described in [16], the typical procedure for EEG analysis and feature classification to identify the cognitive state consists of the following:

1. Data collection in a standardised setting.

2. Signal denoising and feature extraction (e.g. PCA, Wavelets, ICA, Statistical-based Analysis).
3. Classifier creation and training based on N time and frequency-based features selected (e.g. LDA or Support Vector Machines (SVM)).
4. Classifier testing to evaluate the performance on a global and single-participant level, usually against other objective (task performance) and subjective (questionnaires) metrics indicative of the mental state.

One of the most used dimensionality reduction and feature extraction techniques is Linear Discriminant Analysis (LDA), which has the benefit of having a relatively low computational cost and assumes Gaussian distributions, as in the case of the EEG features [18]. This classifier is characterised by a long history of successful BCI-related applications [19].

Support Vector Machines have also been widely used with EEG data processing, particularly for pattern recognition applications.

Statistical analysis of EEG signals is also useful for evaluating the amount of information in an EEG. A typical meter is, for example, the statistical entropy (Shannon's Entropy, SE) of the signal, indicative of its randomness and complexity both in the time and frequency domain [20]. This provides a measure of the uncertainty and allows the identification of specific patterns that could correspond to different brain states or conditions. For instance, higher entropy levels generally indicate a more relaxed or sleepy state, while lower entropy levels often reflect heightened focus or alertness. This type of analysis can assist in diagnosing and treating neurological disorders such as epilepsy and sleep disorders.

Additionally, a novel feature extraction technique for EEG data is the Artificial Neural Network based Feature Learning (ANNFL) method proposed by Gao et al. [21]. This model was evaluated against other PCA and auto-encoder-based architecture and resulted in being more effective in feature representation and single-channel temporal feature extraction while having the advantage of being an unsupervised learning method. However, this is limited to a narrow range of applications as its performance is strictly related to the number of electrodes and represents a computationally expensive option.

Moreover, an interesting approach was taken by Bird et al. [22], who developed a Multilayer Perceptron (MLP) classifier using an evolutionary algorithm for feature selection and hyper-tuning of the network. This was evaluated against an LSTM (Long Short-Term Memory) model using Adaptive Boosting (an ensemble learning method to boost the accuracy of classifiers). Both models were tested on various tasks, including attention level prediction. The evolutionary-optimised MLP resulted in a similar or greater performance compared to LSTM whilst requiring a lower training time.

2.2 Cognitive Load

2.2.1 Definition and identification

To measure the impact of sensory stimulation on the cognitive process and multi-tasking abilities, we have to find a formal definition of cognitive load and a way to quantify it.

Unfortunately, there are different definitions of cognitive load, and its effect manifests in various ways, from attention level to long-term fatigue.

One of the most consolidated systems in this field is the Cognitive Load Theory (CLT), which states that the human working memory, responsible for the manipulation of temporary information, is capable of processing and storing only a finite amount of data [23].

According to the CLT, there are three types of cognitive load: intrinsic (to the content of the learning process), extraneous (due to the quality of the format of the material learnt), and germane (dependent on the difficulty of the learning process itself) [24] [25]. This continuously-developing theory aims to find ways to focus these limited mental resources on the active learning process whilst minimising external sources of cognitive load.

According to this theoretical framework, when we are in a multi-tasking environment, the amount of cognitive load will be influenced by the task's difficulty and accumulated across all sources of interference, i.e. the different tasks requesting our attention. Therefore, by measuring the performance and the cognitive load, we have a dual quantitative metric to understand how much a specific event (in this case, the sensory stimulation) will impact our abilities to deal with other concurrent tasks that require our attention.

An empirical demonstration of the correlation between mental state indicators and task difficulty is presented in the study by Guame et al. [18]. This shows how measuring a combination of different features related to the cognitive load (prefrontal θ power, broad spatial range γ power, frontocentral β power, and frontocentral α power) and then using these to train a classifier, led to an accuracy in the prediction of task difficulty of up to 85% (over 30 epochs).

For this type of evaluation specifically, the cognitive load was associated with the variation in attention levels; however, the level of mental fatigue has also been shown to be a valid indicator of the working memory load [26]. Furthermore, both these factors have been shown to rely on features recurrent across participants [27] (posterior alpha band features for attention and frontal alpha and beta for mental fatigue), validating their use for cognitive state detection.

The following two sections will give an overview of how to identify these two elements.

2.2.2 Attention level

Attention can be defined as “the ability of individuals to select relevant/interesting stimuli while ignoring the other stimuli in the surrounding environment” [18] [28]. From this, Gaume et al. [18] distinguish between two types of attention: sustained, i.e. the ability to remain focused on a task for a period of time, and selective, i.e. ability to focus on relevant stimuli in a distracting environment.

In EEG analysis, Myrden et al. [27] explain how attention-related features are most commonly associated with the alpha frequency band detected in the posterior area of the brain. These can be used to classify the attention degree with an accuracy of up to 70% [29], [30]. However, it has to be noticed that, depending on the nature of the task requesting the participant’s attention, changes in cerebral activity can be measured in different areas of the brain. For instance, the frontocentral and central lobes have been shown to be the most critical area for accurately identifying attention level when performing motor tasks [16]; this is different if the task would involve, for instance, only visual stimulation [31].

2.2.3 Mental Fatigue

To have a complete picture of the cognitive load, it is not sufficient to focus solely on the attention level, as this is influenced not only by task difficulty but also by other external conditions. However, measuring mental fatigue can help interpret the attention level results by considering the effect of prolonged cognitive activity. Mental fatigue is, in fact, often associated with tiredness, lower attention level, reduced cognitive control [32] and decision-making [33].

A paper by Roy et al. [26] describes how increased mental fatigue affects both the cognitive workload and the Time on Task (TOT), with significant changes both in task performance and in EEG components (decrease in alpha power with an increased level of fatigue). This study showed how, based on features extracted from the EEG spectrum, it is possible to achieve an accuracy of 98.04% in the classification of the level of mental fatigue. Despite it depending on the individual abilities, this can also be used as an indicator of task difficulty.

Moreover, Myrden et al. [27] demonstrated how fatigue-indicative features are mainly detected by the frontal electrodes in all the primary frequency bands (delta, theta, alpha, and beta) [34], [35]. In particular, they showed how it is possible to classify mental fatigue with an accuracy above 70%, even on a single-trial basis.

Another study by Krigolson et al. [36] also explored using a combination of time and frequency features to accurately predict the level of mental fatigue against the subjective value reported by the participant. This highlighted how the ERP and EEG data representations contribute to independent features for the classification process. An example of P300 and N170 ERP components analysis from this paper can be seen in Fig 3. This paper also confirmed the proportional relation between increased alpha and theta power and a higher level of fatigue, as previously reported by Fan et al.[10]. The latter also investigated other effects of long visual search-based tasks on the cognitive load, concluding that the most representative indicators of increased fatigue were the following:

- a consistent increase in the $\frac{\alpha+\theta}{\beta}$ and $\frac{\alpha}{\beta}$ power ratios.
- an increase in the Shannon’s Entropy (SE), indicating the level of disorder in the brain’s neural activity.

- a slower increase in the θ and δ activity.
- a significant reduction in the β activity.

Finally, Käthner et al.[37] investigated the effect of mental fatigue on ERP components specifically, demonstrating how at increasing load levels, the amplitude of the P300 indicator decreases, together with the expected increase in alpha band activity and the subjective perception of fatigue (measured with NASA Task Load Index).

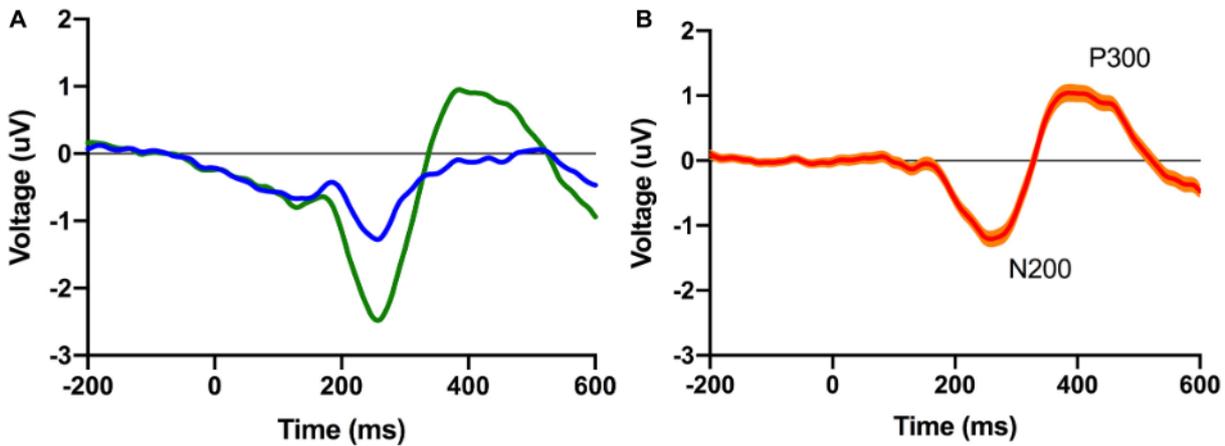


Figure 3: Grand average of ERP waveforms for the pooled posterior electrode (left); 95% confidence interval is plotted on the difference waveform (right). Picture and analysis by Krigolson et al. [36].

2.3 Cognitive processing of sensory information

It is important to distinguish what differences in cognitive response across various stimulation modalities are due to the levels of cognitive demand and what is simply due to the internal processing of the sensory information. Therefore, we will give a quick overview of the areas of the brain involved in processing haptic and auditory information.

Figure 4 shows the main regions of the sensory and motor cortex. We can localise the auditory processing in the auditory cortex, in the upper side of the temporal lobe [38], and of haptic vibrations in the primary somatosensory area, particularly in the anterior section (Brodmann area 3) [39]. Instead, the processing of haptically perceived shape information occurs at first in the primary somatosensory cortex and then propagates in the parietal lobe, including frontal cortex areas, for shape-based features and to the intraparietal sulcus, on the lateral surface of the parietal lobe, and frontal eye fields for feature location information [40]. Finally, spatial information related to navigation and orientation is also found in the vestibular system located in the inner ear, related to between perception of linear motion and passive linear displacement [41].

With the BCI available for this study, we can record the signal coming from the temporal and frontal areas of the brain; we will mainly use the temporal electrodes to detect the processing event, expecting a stronger response to auditory signals given the proximity of electrodes TP9 and TP10 to the auditory cortex, and use the frontal electrodes for the detection of mental state-related features.

2.4 Haptic Stimulation Evaluation

In recent years, there has been an increasing interest in developing haptic-based technologies, especially in human-computer interfaces and prostheses. For instance, Jafari et al. [43] review how recent developments have shown the positive impact of tactile integration in enhancing task performance and the learning process for children with disability. However, the same study also highlights how the literature lacks a quantitative assessment of the impact of haptic feedback on cognitive loading, including discomfort and fatigue.

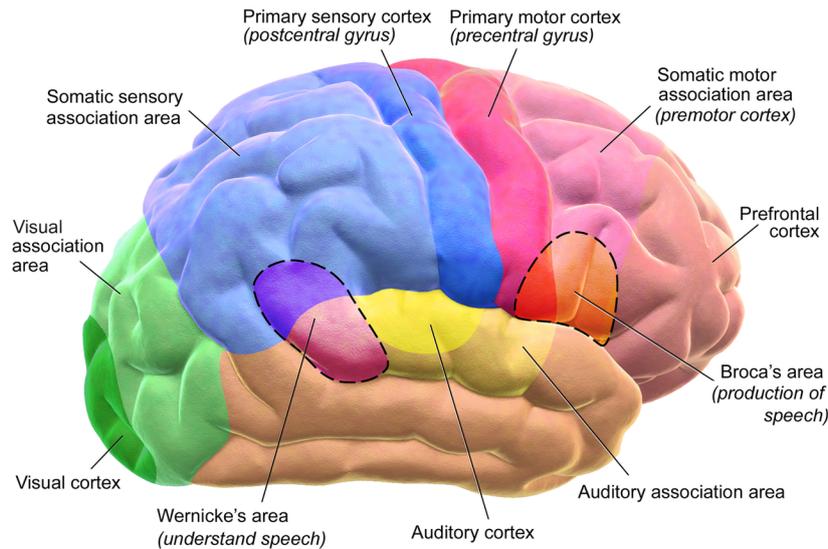


Figure 4: Motor and Sensory Regions of the Cerebral Cortex. Picture from Blausen.com [42].

An interesting research on this matter was pursued by Thomas et al. in 2021 [44], where haptic feedback-based prostheses were compared to standard visual feedback-based ones in stiffness discrimination tasks. As expected, they found that using vibrotactile feedback resulted in consistently better performance accuracy but also lower cognitive load (measured using functional near-infrared spectroscopy (fNIRS)) and was ultimately rated lower in terms of usage-induced frustration.

Recently, some initial studies have also measured the workload induced by haptic and combined multi-sensory integration using EEG specifically. An example is the research pursued by Marucci et al. [2], which evaluated the performance of haptic, auditory and visual integration in a virtual reality environment. This study concluded that integrating haptic information into auditory-visual feedback improved effective performance. This also resulted in higher amplitude of the P300 component, an ERP occurring around 300 ms after sensory stimulation and that previous studies have suggested being a potential indicator of higher attention level and lower fatigue [45] [46] [47].

A literature review by Fleury et al. [48] also discussed the utility of the P300 paradigm to evaluate the use of haptic feedback. This paper reviewed the different types of haptic interfaces, classifying them based on the sensation provoked (vibration, curvature, texture, contact, pressure, temperature, rigidity, and friction), and validated the use of P300 and SSSEP (steady-state somatosensory evoked potential) as haptic-related performance indicators.

Additionally, Burked et al. [49] performed a comparative literature review on the effects of visual-auditory versus visual-tactile-based feedback on user performance. This study highlights how visual-auditory feedback is most effective in normal workload conditions, such as single-task scenarios. However, using both auditory and visual channels can increase the cognitive workload for the user, making it less advantageous in situations where a high workload is already present. On the other hand, the visual-tactile modality was more effective when participants were performing multiple tasks with an increased workload. As explained in the paper, this is also supported by the Multiple Resource Theory [50], stating that the benefits of incorporating the tactile modality become more pronounced when the workload is high. These findings are particularly interesting as they highlight the necessity for further research in haptic technologies, especially when applied in a high-load, multitasking environment.

Finally, substantial progress in exploring different haptic methodologies has been achieved by Stanley et al. [51] by pursuing a detailed comparison of 10 different tactile feedback modalities used to perform un-sighted motor tasks. The team developed a series of 5 stimulating actuators (Tapper, Dagger, Squeezer, Twister, and Vibrator) and tested two algorithms used to provide tactile cues (Steady and Pulsing). These modalities were evaluated only in terms of task-related and subjective performance, leading to the conclusion that the choice of the tactile actuator and the algorithm controlling its activation play a crucial role in determining the user's success in following tactile guidance for motion. The following were the core results of this study:

1. The participants moved faster when the direction of the cue was indicated by the position of the tactile stimulus/skin stretch.
2. Feedback types that utilize a spatial tactile arrangement, such as Dragger Steady, Dragger Pulsing, and Vibration Pulsing, resulted in slower responses
3. The Vibration Steady feedback type resulted in significantly faster settling times compared to feedback types that utilize spatial patterns.
4. Pulsing cues, in general, performed more poorly, with the exception of the Tapper Pulsing, which had the best performance overall.

Although these results give us a better idea of the direction in which haptic technology should develop, this type of research would also benefit from an evaluation of the actuators' performance at cognitive level.

In conclusion, we are still at an early stage of understanding how tactile type of sensory stimulation influences the user, particularly in a multitasking environment; from this arises the motivation for the project presented in this report.

2.5 Consumer-level Passive BCI validation

Over the last few years, there has been an increased effort towards making neuroscientific experiments more accessible and developing consumer-level Brain-Computer Interfaces, such as devices for meditation and sleep tracking purposes or at-home medical devices used by patients with neurodegenerative diseases [52].

In 2022, Sabio et al. [53] conducted a thorough literature review of all the major papers that have used commercial-grade BCI for research purposes. As a result, they have concluded that BCI models from companies such as Emotiv, NeuroSky MindWave, InteraXon (Muse) and OpenBCI have played an increasingly important role in obtaining novel results; this study confirmed how these devices had met the research-level standards in various experimental and clinical contexts, validating their use for future applications.

Two recent studies have also been validating the use of Muse devices specifically for research on the mental state with similar conditions to the project's ones. First, Krigolson et al. [54] successfully demonstrated the validity of Muse for ERP research, using oddball visual tests to identify and quantify N200 and P300 visual components. Despite non-standard electrode positioning and the lack of event markers in the EEG signals, the study showed the experimental data was of good enough quality to quantify the ERPs. Even the limitations in the data synchronisation were considered irrelevant given the timing precision required, particularly since the random delay due to Bluetooth technology has a Gaussian distribution that approximately averages out over the trial duration.

In a following study, Krigolson et al. [36] used Muse again to research the link between mental fatigue and ERP components (visual N700 and P300). Not only did they demonstrate the validity of the BCI, but they also obtained novel results in the correlation between alpha power and mental fatigue.

2.6 Standard Testing Methodologies

A literature review of a range of neuropsychological tests has highlighted how the most well-validated methodologies for measuring the cognitive load involve the oddball paradigm².

This consists of a standardised format according to which the participant works on two concurrent tasks with different attention demands and frequencies: the Primary (Control) Task, requiring a medium-level of continuous attention, and the Complex Secondary Task (CST), requiring a diversion in the attention at occasional times during the experiment [16].

Some example applications of this methodology are: using different forms of visual stimulation to measure the mental fatigue of processing and moving between different tasks using the same sensory stimulation modality [36]; deploying an oddball auditory task concurrently with a motor task to measure the changes in the attention caused by the diverse and concurrent nature of the activities performed [16].

²Oddball Paradigm: https://en.wikipedia.org/wiki/Oddball_paradigm

Similarly to what is proposed in this project, Myrden et al. used this method to measure fatigue, attention and frustration using a passive BCI, achieving significant results even for single-trial detection across participants [27]. This study also suggested how this method could be used to measure “more complex aspects of user state, such as perceived loss of control over a system”, an important aspect to consider when testing navigation devices.

Another relevant example of an oddball paradigm application for measuring the mental workload is presented by Käthner et al.[37]. In this case, a dual task system was implemented: a primary spelling task and a secondary listening one, according to which two stories of different difficulty were told in the right and left ear. The performance was then measured on three levels: cognitive, based on the EEG indicators; task-related, asking questions on the stories listened to; subjective, using the NASA Task Load Index (TLX) questionnaire³.

The NASA TLX is a universal method to measure the perceived level of task difficulty and it is based on six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. This questionnaire is particularly effective when used to compare different tasks' difficulty and validate the cognitive results of a BCI-based evaluation. Many studies found a strong correlation between these two metrics, as discussed by Gaume et al. [18].

Finally, to ensure the results are statistically significant, the ANOVA (Analysis of Variance) Test is used as a standard methodology to reject a null research hypothesis or accept the alternative one [55].

³Nasa TLX at <https://humansystems.arc.nasa.gov/groups/tlx/>

3 Experiment Design

This chapter includes a description of the research methodology and evaluation metrics chosen for the experimental trials.

3.1 Experiment Structure

The experiment was designed around the oddball paradigm for reasons discussed in the background section 2.6. This consisted of the following:

- **Primary (Control) Task:** the participant stared at the computer screen, on which circles of different colours (blue/red) were displayed at regular intervals and in a randomised sequence. The viewer counted the number of blue circles and returns the value at the end of the experiment.
- **Secondary (Diversion) Task:** a short stimulation was performed, indicating to the participant to press the left or right arrow key on the computer keyboard (using their dominant hand); the manifestation of this request was occasional, and the timing was randomised.

An illustration of the experimental setup is shown in Fig 5.

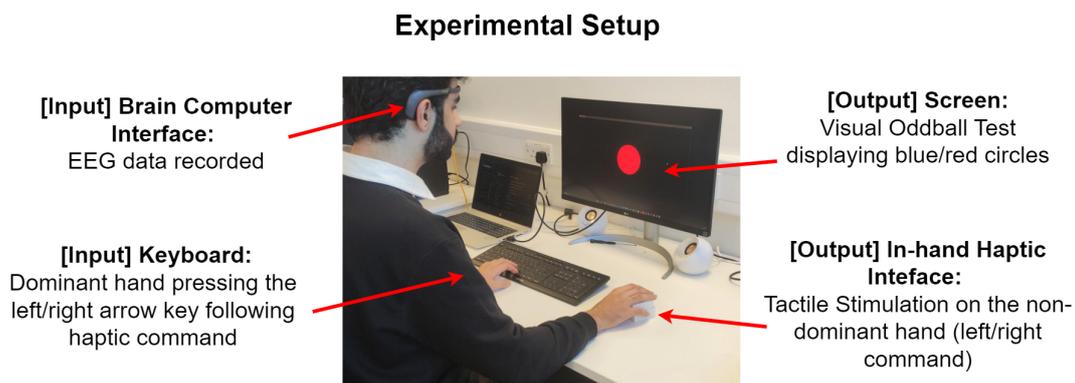


Figure 5: Experiment setup for the final test with haptic interface

The primary task has not been stopped whilst the disruptive one occurs, as this allowed to test the perceived level of intrusiveness of the sensory stimulation (a low level signifying the participant might not acknowledge the stimulus as they are focused on the primary task; a high level if the distraction harmed the performance of the primary task).

These tasks were chosen for their intuitiveness and simplicity whilst remaining relevant to the target application of the study (i.e. navigation-related multitasking activity). Other, more complex, motor tasks have also been considered; however, these would have led to motion artefacts in the EEG signal, more significant safety concerns, and the general impossibility of analysing the activity in the motor cortex due to the electrode disposition in the Muse2 device.

The experiment procedure for each participant comprised three separate tests of two 7mins trial blocks each (see Fig 6), differing only for the type of stimulation modality used:

- **Auditory stimulation:** a beeping sound in the left or right ear was associated with the left and right command; the sound generated was the same for both ears to represent a fair comparison with the similarity of left and right sensation used in the haptic modality.
- **Vibro-tactile stimulation:** a short vibration was output on the left/right side of the vibrating device with respect to the participant; the non-dominant hand of the participant was holding the device throughout the duration of the experiment.

- **Shape-changing stimulation:** a translation to the left/right of the participant was performed by the top of the in-hand, shape-changing device and then returned back to the initial position; the non-dominant hand of the participant was holding the device throughout the duration of the experiment.

To simplify, from this point of the report onward, these three experimental conditions will be referred to as “Audio”, “Vibro”, and “Shape”, respectively.

Experiment Progression for each participant

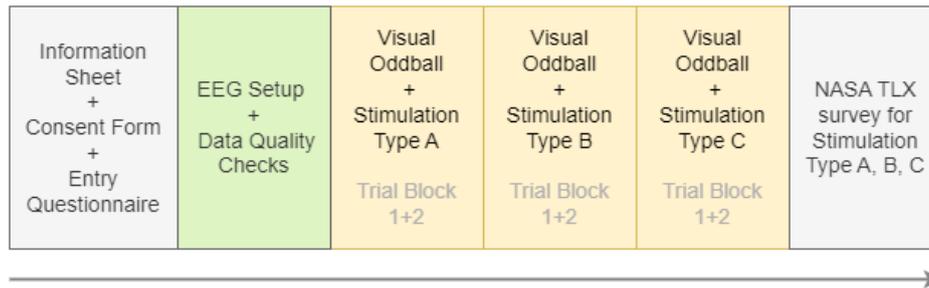


Figure 6: Experiment Format: the order of three stimulation modalities (auditory, vibration and shape-based ones) was randomised (A, B, C) across participants

Each experiment was conducted in a standardised manner and started by reading the participant information sheet, gathering written consent and filling up an entry questionnaire, and collecting basic information about the participant pool, e.g. gender, age, dominant hand and eligibility criteria. The latter consisted of the lack of hand injuries, visual or auditory impairments and the ability to stay focused and still for a prolonged amount of time (each of the 7 min trial blocks).

Then, the Brain Computer Interface was set up, and some data quality checks were run to ensure the electrodes were properly placed and there was sufficiently low external noise. Consequently, the experiment flow was explained to the participant, and they had the chance to familiarise themselves with the tasks and the stimulation modalities before the first trial block of the first stimulation modality started.

It is important to notice that the sequence of the three simulation types varied across participants to ensure the performance improvement was not only due to the participant growing accustomed to the dual task and to avoid biasing the experiment. At the end of the experiment, three separate NASA TLX questionnaires were presented to evaluate each stimulation modality from a subjective experience and difficulty level in relative terms to one another.

A single experiment lasted between 60 and 90 min, including the initial setup, all the 7 min trial blocks (two for each of the three stimulating modalities), and the exit questionnaires; each trial block contained between 40 and 50 trials (i.e. stimulation events), depending on the randomisation seed of the experimental sequence of first and second task stimuli. The specific timing was defined during the initial phase to ensure the test was long enough to gather sufficient data whilst not straining the participant. Choosing the appropriate timing increased the probability of detecting both the core features (e.g. mental fatigue arising from induced sustained attention over a long period of time) and side ones, such as the learning pattern and speed at which the subject grows accustomed to the specific feedback.

The test structure was designed to easily adapt to any sighted participant without requiring specific ability, previous knowledge or specific training. It has to be noticed that although the variation in performance between participants has been considered, the focus of this study was to compare the different types of stimulation, meaning that the deviation of performance between different trial blocks for the same participant had been at the core of the analysis. Therefore, the experimental data have been normalised across participants.

Due to the high rate of noise susceptibility of the Brain Computer Interface, the location of the experiment was chosen to limit the amount of electrical interference. At the beginning of each experiment, all unused electronic devices were turned off or removed from the room. The participants were asked to stay as still as possible during the recording. Moreover, during the initial familiarisation phase, the participants were shown their performance in terms of eye-blinking and other movement-related artefacts to raise awareness

of the expected standard for the participant's behaviour.

3.2 Experiment Metrics

The comparison between sensory modalities has been conducted following three criteria:

- **EEG-based Cognitive Performance:** this has been evaluated in terms of absolute and relative variation of the individual EEG power bands and the analysis of time-domain features, comparing against the participant values at resting state and the current literature around this topic. The specific criteria highly depended on the patterns detected in the recordings and will be explained in detail in section 6.
- **Task-related Performance:** this comprised the performance in both tasks:
 - Primary Task, evaluated in terms of accuracy of the blue circles' count, including the number of times the participant restarts the count from zero.
 - Secondary Task, evaluated in terms of coherence between order commanded by the stimulation (left or right) and outcome of the execution (left or right arrow key), and general responsiveness to the command.
- **Perceived Workload Evaluation:** this complementary method represented a way of evaluating the perceived level of task demand and the induced effort, stress and fatigue at task completion. The NASA Task Load Index (TLX) questionnaire has been chosen as the most widely used criteria to measure the perceived task difficulty and task-induced workload, as explained in section 2.6.

3.3 Ethics, Legal and Safety Measures

The collection of physiological data and the use of haptic stimulation on human participants raised ethical and safety concerns that had to be addressed. Firstly, informed consent was gathered from the participants, ensuring they understood the purpose of the project and the risks involved. The participant's personal data and EEG recordings have been kept confidential and secure, following the UK General Data Protection Regulation (GDPR). All participants were free to withdraw their experimental data from the study up to one week after the experiment took place, and the data was pseudo-anonymised. If the data obtained in the study will be published in an academic research paper, the participants will be referred to only by letters and numbers in an anonymised fashion.

Secondly, it was crucial to ensure that all equipment and devices used were safe, met appropriate standards, and had adequate procedures to address potential risks or emergencies. Participants were informed of any potential risks associated with using BCIs and haptic stimulation, and health and safety assessments were completed. By using a consumer-level BCI, compliant with the UK regulations for Product and Safety and Standards, and by not including any major motor tasks that could involve physical risks, the project format was designed to limit the possibility of hazardous situations. In addition, the haptic interfaces deployed were low-risk devices which have received ethics approval for similar tests in the past and implicated similar risks to the ones involved when using a computer mouse.

The trials of this study have been conducted internally in the department and as a pilot study. An application for ethical approval has been submitted to the Research Governance and Integrity Team of Imperial College London. The risks of this study are minimal; however, since the haptic devices have been created ad hoc for the study and therefore are not CE-marked, this study has to be reviewed by the Science, Engineering and Technology Research Ethics Committee (SETREC), which is responsible for non-health related research involving human participants and their data and high-risk educational research that is undertaken by College staff or students.

4 Hardware & Software Development

This chapter includes an overview of the devices and software tools developed.

4.1 Brain Computer Interface

The passive Brain-Computer Interface used was the consumer-level Muse 2 [56]. This was selected over a range of others due to its low cost and suitability for research purposes, as discussed in the background section 2.5. As shown in Fig 7, the model includes the following features [56]:

- EEG: 5 electrodes, including one reference electrode (FPz), two frontal electrodes (AP7, AP8), and two posterior electrodes (TP9, TP10), see Fig 8.
- Gyroscope and Accelerometer: movement sensors to help measure body movement and breadth.
- PPG and Pulse Oximetry: Photoplethysmography sensor to monitor the heart rate using optical measurement method.

For this project, the initial plan was to use only the features from the EEG sensors, with the possibility of expanding during future work.

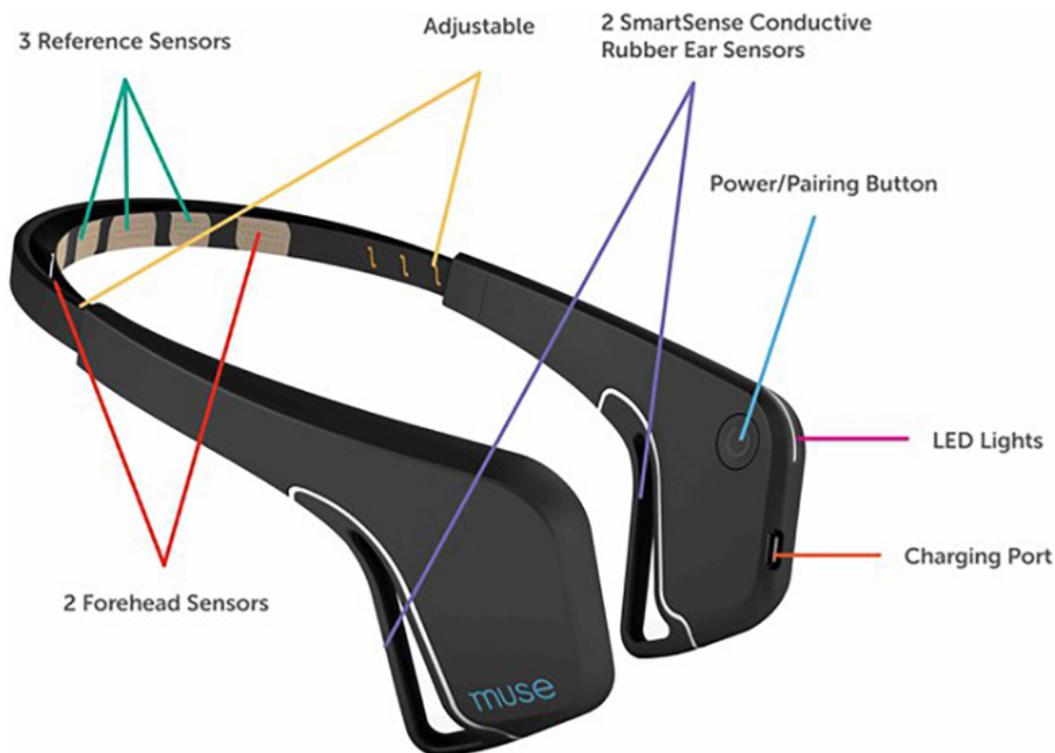


Figure 7: Muse EEG system made by InterAxon Inc., picture from [36]

Muse2 is easy to use and has good software support but does not have high precision in signal localisation due to the reduced number of electrodes. In fact, this BCI can detect brain activity in a specific part of the brain but cannot identify signals at the neuron level. Furthermore, in terms of data synchronisation, the device is characterised by limited levels of precision due to the jitter and lag inherited from Bluetooth communication [54], [36]. The data synchronisation procedure accounted for this by placing the event markers (associated with the time of the stimulation event) in the appropriate timestamp of the EEG stream, considering an estimate of the average lag as recorded in the initial tests.

Concerning the jittering, this has been taken into account by similar experiments [54] [36] by considering that its distribution is Gaussian and will average out throughout each trial. In fact, it has to be noticed that



Figure 8: Muse 2 electrode placement; picture by Tsiakas et al.[57]

the experiment does not require an extreme level of precision in terms of the timing of the ERP peaks, as the focus was on studying how the stimulation influences the overall cognitive state of the participant comparatively across modalities, as opposed to the investigating it from a strictly-neuroscientific perspective. Therefore, the precision in the data synchronisation provided by Muse 2 was sufficient to obtain significant results.

These experiment precision and resolution limitations could be overcome by using a more expensive, research-level BCI; nonetheless, Muse 2 has been validated for experiments of similar types and was the best choice given the project budget.

4.2 Auditory Stimulation

Wired earphones have been used for audio stimulation, playing a beeping sound on the right or left ear depending on the directional command. To isolate the participant from external noise and recreate similar conditions across all modalities, the earphones were used across all trials and white noise was played when testing the two haptic interfaces so that the sound of the moving devices was not perceived as a source of further stimulation. The sound type was chosen to be intrusive yet not highly disruptive, given the dual nature of the task to be performed.

4.3 Vibro-Tactile Interface

For the vibrotactile stimulation, a vibrating device has been developed, as shown in Figure 9. The device was designed to be held like a computer mouse. It was developed in Solidworks and 3D printed in PLA in two halves connected to two independent vibrating disk motors⁴. Separating material such as rubber and foam was used to isolate the vibration between the two parts of the device, and voltage levels of the input circuit were adjusted to recreate similar sensations with both stimuli. A biasing circuit was developed to step up the current and meet up the 70mA requirements of the motors from the max 20mA provided by the Arduino. A controller has been implemented in Arduino, tuning the voltage levels and the timing of the vibration to produce a perceivable yet not highly disruptive stimulus. The vibration command is triggered from the Python backend of the experiment framework to the Arduino, which varies the voltage levels across the motors accordingly.

⁴<https://uk.robotshop.com/products/vibrating-disk-motor>

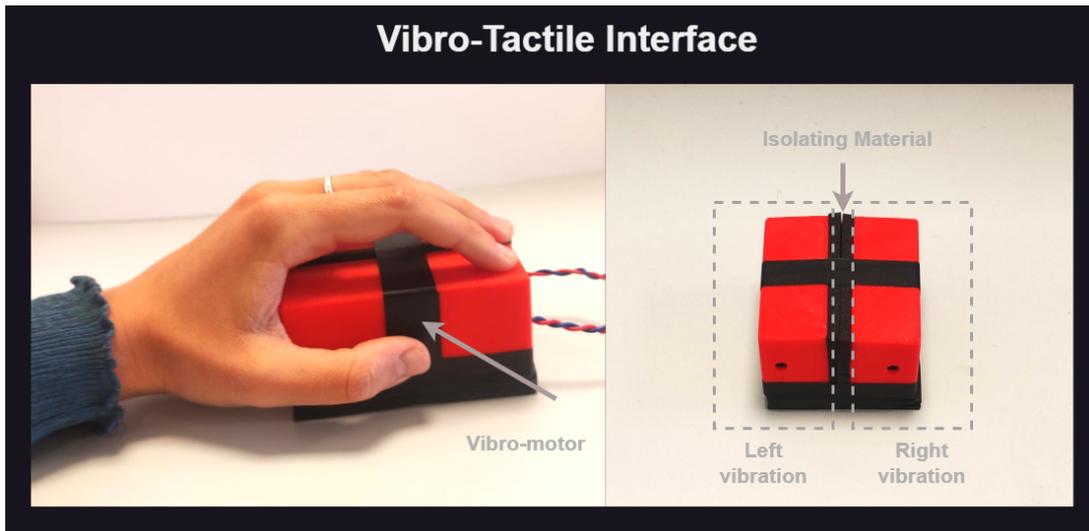


Figure 9: Vibro-tactile interface capable of representing left/right directional commands through the independent vibrating stimuli on the right or left side of the device.

4.4 Shape-Changing Interface

Inspired by the design developed by Spiers et al. [58] for pedestrian navigation assistance of visually-impaired people, the second haptic device comprises two parts: a base containing a linear rail, and a moving top, sliding along it. The parts were 3D printed, and an HS-82MG Metal Gear Micro Servo⁵ was used to actuate the left/right movement (see figure 11). The amount and speed of rotation were tuned and controlled through the Arduino to create a perceivable yet smooth motion. When the haptic command was sent from the Python script in the main experiment, this was received by the Arduino, which initialised the servo rotation, engaging with the rack inside the top half of the device and resulting in its translation. The haptic command of the left (or right) was represented by the movement to the left (or right) direction and the "return-to-centre" motion. This was also specified during the experiment to ensure that the second half of the movement was not interpreted as a second, opposite stimulus.

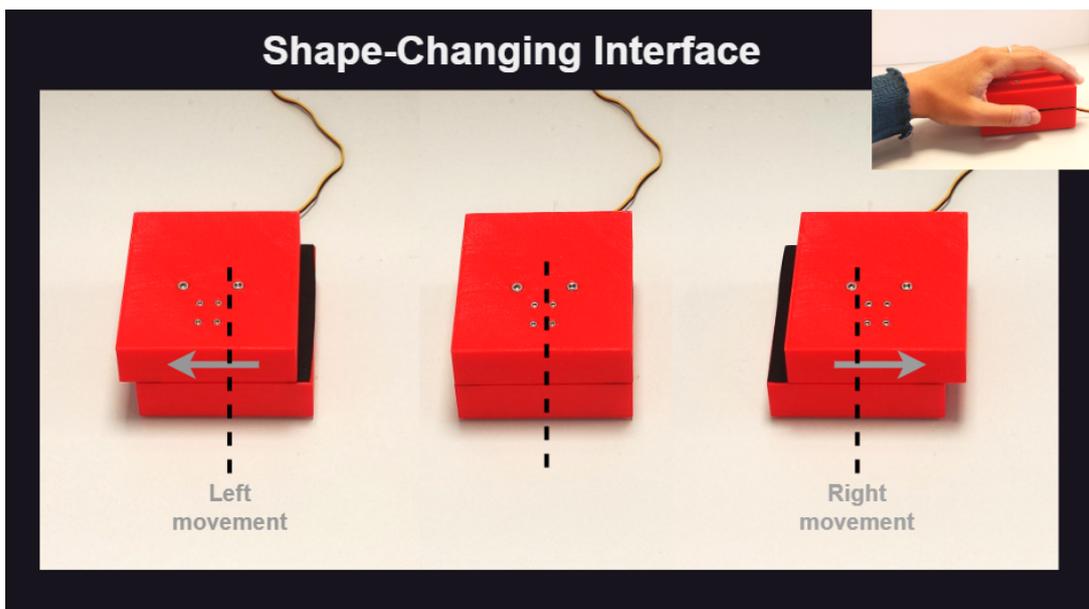


Figure 10: Shape-changing haptic interface capable of representing left/right directional commands.

⁵<https://uk.robotshop.com/products/hs-82mg-metal-gear-micro-servo>

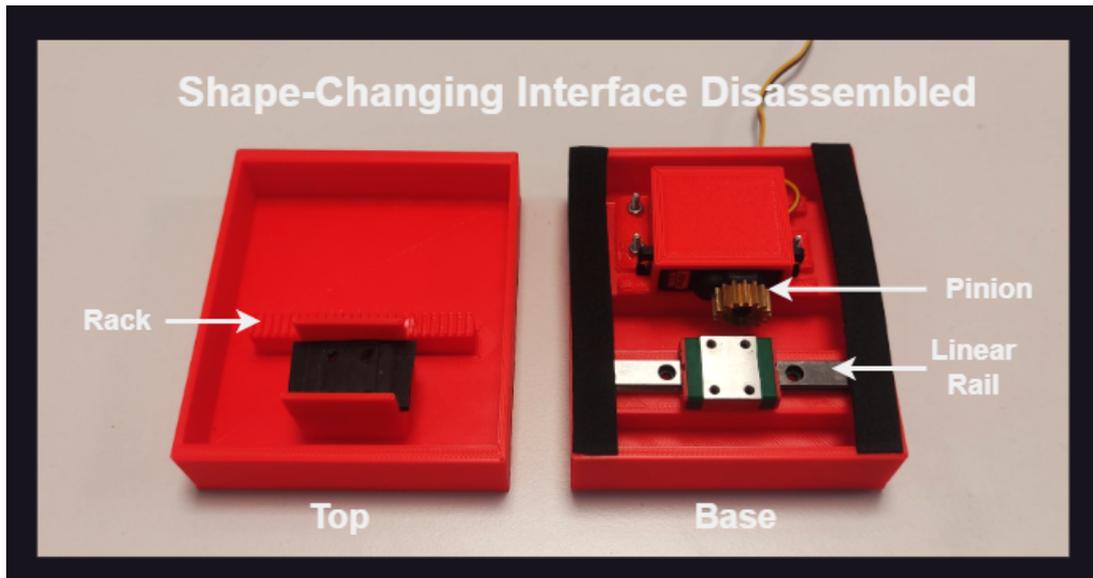


Figure 11: Disassembled Shape-changing haptic interface

4.5 Software Development and Dataflow

Python has been used for experiment environment and test development as the most compatible with all the pre-existing packages and tools supporting communication with Muse devices and EEG signals processing. The followings are the main libraries and software tools used for the experiment development and data analysis:

- BlueMuse [59]: software used for starting the stream between Muse 2 and the local device, working via Bluetooth.
- muse_ls1 [60]: a Muse extension of the python Lab Streaming Layer, a library used to interface with sensor/device and record/manipulate multiple data streams.
- eeg-notebook [61]: notebook of packages and scripts, including visual and audio experiments compatible with Muse, used for the initial verification of the device functionality.
- Psychopy [62]: python library for the creation of the visual part of the experiment and general framework for psychological and neuroscientific experiments.
- MNE [63]: python library for EEG signal processing and data analysis.

The experiment was developed in the format of a Jupiter notebook and a library of functions for the visual, haptic and auditory stimulation, data collection and recordings' synchronisation. The timing of the stimulation was synchronised with the experiment flow, and the timestamp of the auditory/haptic stimulation and the one of any arrow-key pressed are inserted in the annotation field of the EEG data for the analysis.

5 Testing

This section summarises the initial tests and validation of Muse, haptic devices and experimental setup.

5.1 Pipeline

The implementation and testing of each component run in parallel, following the steps below:

- **Data Quality Checks:** automated tests on the Muse data stream aimed to validate the data quality and synchronisation between EEG and event markers (i.e. timestamp at which a stimulus occurs).
- **Mental State Classification:** initial test providing a proof of concept for the development of a classifier able to detect simple changes in mental state (relaxed vs active state).
- **Visual Test:** visual test in the format of the oddball methodology later used for the Primary Task of the experiment, as described in the experiment section 3.1.
- **Auditory Test:** validation of the Complex Secondary Task (CST), including testing the implementation of the auditory stimulus and input processing from the user's keyboard, synchronising the stimulation and reaction events.
- **Haptic Test 1,2:** checks targeting the development of the two tactile interfaces, supporting circuits, controllers and data flow, similar to Auditory Test 1.
- **Overall Experiment Testing:** tests run after gathering Data Quality checks and Visual, Auditory and Haptic components into a cohesive structure. All steps were automated to obtain a standardised experiment flow and to test various experiment formats, including different trial timings and sequences between stimulation modalities.
- **Experimental Trials:** experimentation of the three stimulation modalities on 11 participants, following the format described in section 3.1.

5.2 Data Quality Checks

As explained in the background section 2.1.2, many types of noise and artefacts characterise EEG signals; some of them can be removed by correctly positioning the electrodes and limiting the user's movements during the recording. Before each experiment, as an initial test on the signal quality, the Muse App was used to check whether the BCI had been properly fitted and all electrodes were appropriately positioned, making contact with the scalp.

Subsequently, the Muse device was connected to the laptop, and the EEG stream was displayed in real-time using the muse-lsl library [60]. An example of a clear EEG stream is shown in Fig 12. Based on the live stream, the BCI was then readjusted until the expected characteristics were obtained. However, some artefacts were unavoidable (e.g. eye blinking, see Fig 13), and this has then been considered during the data processing stage, as discussed in 6.

Moreover, to check the time-frequency behaviour of the data, the EEG stream was also filtered and the $\alpha, \beta, \theta, \delta$ power bands were isolated and plotted together with the raw data, in real-time (see Fig 14). A script automating all these tests has then been used at the beginning of each experimental trial to monitor the BCI positioning and level of artefacts and external noise.

The power spectrum of the initial EEG recording was plotted after waiting for signal stabilisation, and the quality of the recording was evaluated by inspection of the relevant peaks in the spectrum and the overall magnitude of the plot (see examples of frequency spectra shown in figure 15, 16).

Recordings with magnitude below $20dB$ across the whole spectrum, in combination with real-time voltage levels below $15\mu V$ for time domain data from all channels, were considered usable. These thresholds were defined following the guidelines in the literature, considering the limitation of the Muse BCI and that a clean EEG signal consists of potentials of a few micro Volts, as perceived at the scalp level.

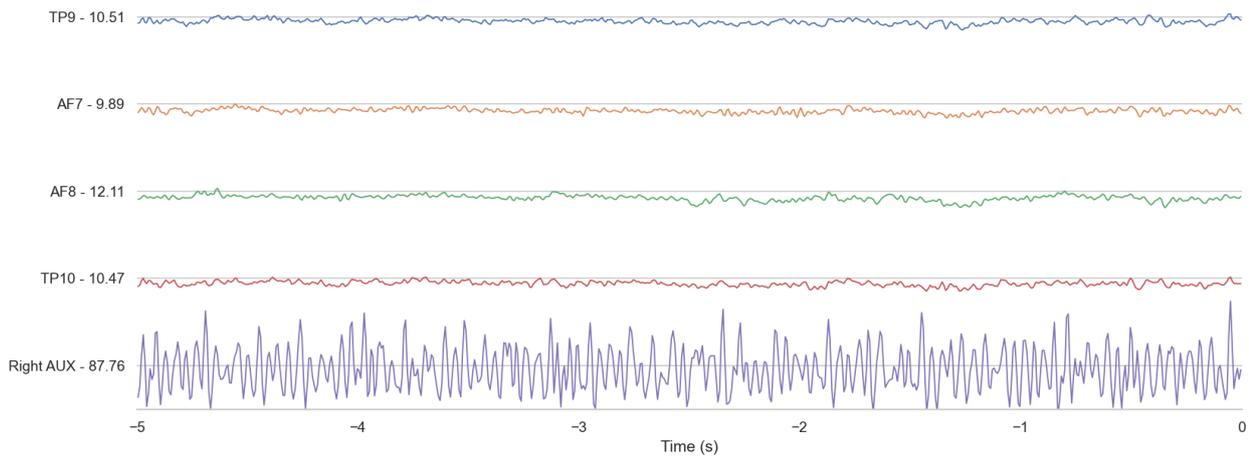


Figure 12: Standard EEG stream with clean data, with voltage levels for each electrode expressed in micro Volts and unused extra electrode from the Right AUX channel; visualisation created using muse-1s1[60]

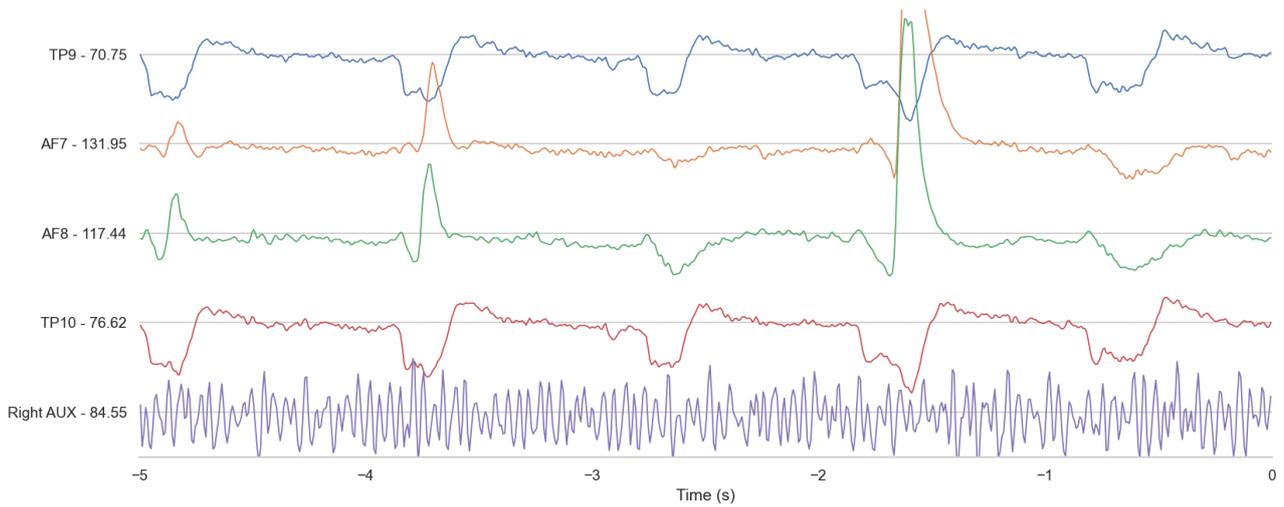


Figure 13: EEG with a high level of noise and artefacts due to eye blinking, with voltage levels for each electrode expressed in micro Volts and unused extra electrode from the Right AUX channel; visualisation created using muse-1s1[60]

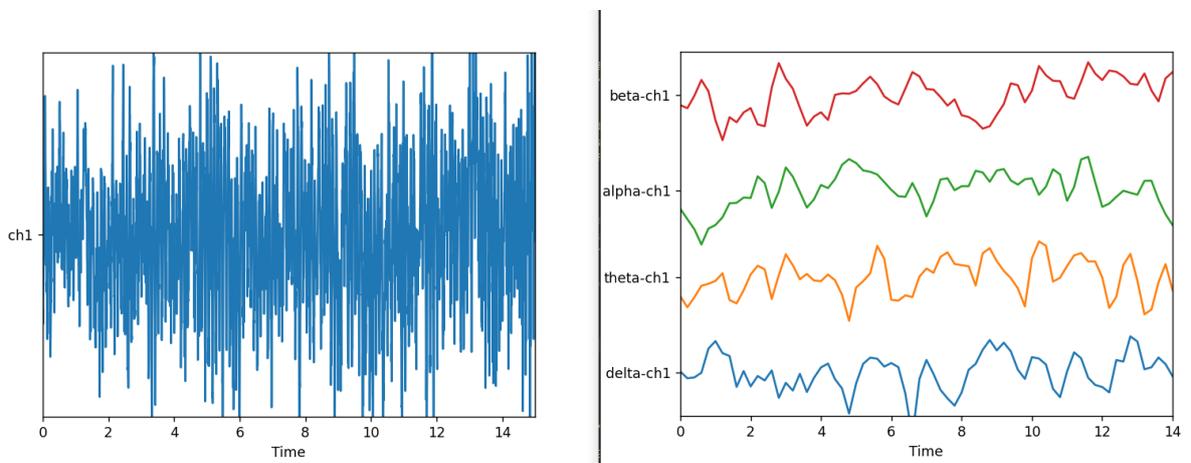


Figure 14: $\alpha, \beta, \theta, \delta$ frequency bands (right plot) computed in real-time from the raw data (left plot) of channel 1 (electrode TP9)

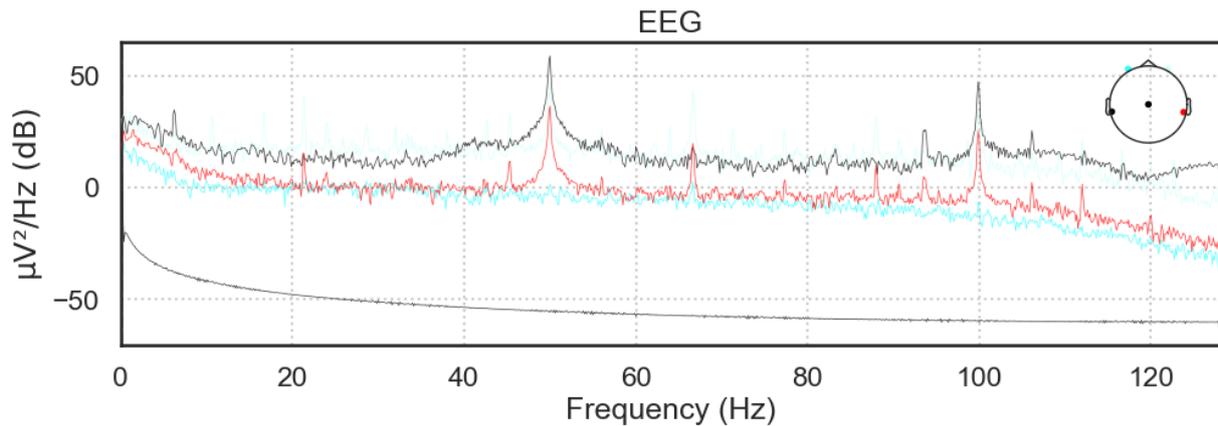


Figure 15: Complete EEG power spectrum (unfiltered); results for the P300 visual experiment from the eeg-notebook[61]

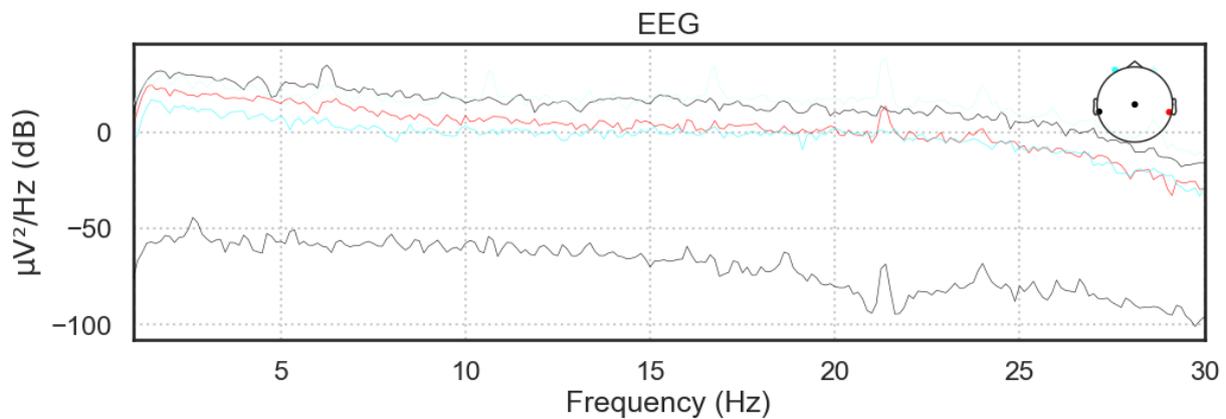


Figure 16: Power spectrum filtered between 0-30Hz; results for the P300 visual experiment from the eeg-notebook[61]

5.3 Mental State Classification

This test was based on the framework proposed in the BCI workshop by the NEuroTechX community⁶. The format of the experiment consisted of three phases: during the first, the BCI-user was asked to close their eyes and relax for 20 seconds; then, they were asked to open their eyes and perform mental arithmetic for another 20 seconds; a simple classifier was trained on the labelled recordings from the first 40 seconds; finally, the participant could decide what to do between the two activities, while the classifier was tested in its prediction of the user state in real-time.

The reported accuracy of the classifier was up to 80%. This test was used to understand how to implement a working pipeline from data streaming to data classification, although it led to unstable and inconsistent results depending on the electrode positioning and amount of electrical interference in the laboratory. This test was useful to highlight the performance fluctuation based on where the tests were taking place. Following this observation, the experimental trials were then conducted in locations with lower levels of electronic interference.

⁶BCI Workshop by NeuroTechX: <https://github.com/NeuroTechX/bci-workshop/blob/master>

5.4 Visual Tests

Multiple versions of the visual task were created, with various levels of task difficulty and colour contrast to ensure ease of stimulus distinction. Static visual tasks were also considered (e.g. counting the circles in a distribution on a static screen); however, this would have required only intermittent focus (with the possibility to stop the task during the occurrence of the secondary stimulus), as opposed to the sustained attention this study was targeting.

The task that led to the best trade-off between sustained attention and low training required was the blue circles counting task across a sequence of blue and red circles occurrences at random intervals. An illustration of the task is displayed in figure 17.

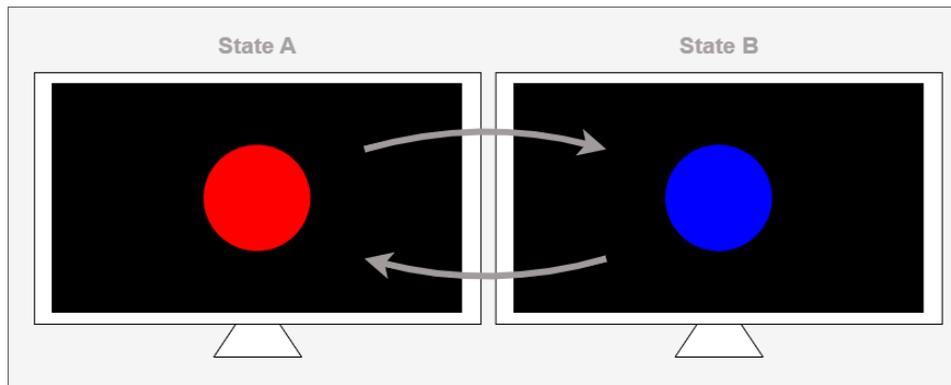


Figure 17: Visual Stimulus of the Experiment as alternating sequence of state A and B at random intervals.

5.5 Haptic Tests

Stress tests were performed on the haptic interfaces to assess their reliability and consistency across the duration of one experiment. The voltages feeding the vibrating motor and the timing for both interfaces were tuned to recreate similar sensations for right and left commands and the same stimulation time across modalities. Multiple iterations of the shape-changing haptic device were designed to adjust the rack to the shape and movement of the servo motor and ensure smooth translations. Multiple configurations of the vibrating device were produced to improve the isolation between the vibration of the left and right side, with the final design composed of separated halves only in contact with the isolation rubber and foam sheets.

The final implementation for the shape-changing achieved satisfactory functionality, producing consistent and clear directional simulations across a prolonged period of time.

Given that the application of this study was closely related to the development of an in-hand navigation device, we opted for a single in-hand interface instead of two separate vibrating devices stimulating different body parts. However, due to this choice, the final design resulted in varying performance depending on the gripping modality, with increased vibration propagation between the right and left side of the hand with a tighter grip around the whole device. To limit this variation and ensure clear directional commands, the gripping configuration was standardised across all participants and the functionality was explained at the beginning of each experiment.

6 Signal Processing & Analysis

This section aims to give a description of the signal processing and data analysis techniques used on the EEG recordings, together with the expectation of the results from the experimental data; instead, in the next section, we will dive deeper into the results themselves.

As discussed in the background section 2.1.3, there are various features to consider when analysing EEG signals. The data have been represented in the frequency domain, analysing the magnitude of the different power bands; in the time domain, studying the Event-Related Potential (ERP) response to stimulation; finally, Morelt wavelets, Entropy and Independent Component analyses have also been used to identify further trends across each of the three stimulation modalities.

6.1 Pre-processing and Denoising

The data obtained from the EEG recording has been synchronised to include the annotations from the events recorded during the experiment, corresponding to the manifestation of blue and red circles, of left and right directional commands, and the input of the right or left arrow key. All EEG data was transformed into the frequency domain using a 1024-point fast Fourier Transform (FFT) and a Butterworth filter applied to cut off the electrical interference and other types of high-frequency noise ($f > 40\text{Hz}$).

Then, for the time domain analysis, the data has been epoched, i.e. the time window around the presentation of an event has been extracted for each event manifestation, grouped by event type. Based on the literature review on the manifestation of Event-Related Potentials following visual, haptic and auditory stimulation (see 2.1.1), different ERP responses usually appear between 0.3 and 0.6 sec from the event; therefore, the time window for the epoch was chosen between -0.3 and 0.7 sec around the time of the stimulus presentation, so to have a baseline voltage prior the event and not to cut off any possible action potential in response to it.

Then, an outlier detection function was implemented to identify the epochs with high levels of noise; following a similar pre-processing procedure as used by Rezaee et al. [64], the noisy epochs were defined as any time window with more than 30% outliers, i.e. voltages above $70\mu\text{V}$, indicating a predominance of noise.

Each trial block was analysed, obtaining a total of 19 out of 56 passing the threshold. However, it has been noticed that the noise performance varied hugely depending on the specific electrodes, with a higher presence of noise in the temporal electrodes (37.5% clean trial blocks for TP9 and TP10) compared to the frontal ones (89.3% clean trial blocks for AF7 and AF8). This could be due to the less standardised positioning and higher unreliability of contact with the scalp for the electrodes positioned behind the ears.

Following this observation, the data analysis was conducted separately for each channel to maintain a larger number of usable trial blocks for each of them.

6.2 Frequency domain Analysis

The average frequency spectrum across all trial blocks has been computed for each of the three stimulation modalities. The results of the pre-processing stage can be seen in figure 18, where the power mean and standard deviation across trials have been plotted against the frequency. As expected, the denoised data resulted in a much cleaner spectrum compared to the raw (bandpass filtered) one, with lower standard deviation and increase visibility of the main peaks. The quality of the denoised data is in accordance with the literature-based threshold, with a magnitude approximately between 10 and -10dB for all spectra.

We observe a greater noise level in the haptic modalities, with increased standard deviation between trials of the same type and higher overall power across the whole spectrum of the Shape modality. These factors will be considered in the comparative analysis by normalising the values and accounting for the heterogeneous standard deviation in the statistical evaluation of the results.

Moreover, from these initial plots, different distributions of the power across the spectrum can already be observed for each modality. For instance, the spectrum of Audio and Shape modalities are characterised by higher magnitude at low frequency compared to Vibro one. We can also identify three main peaks across stimulation modalities at approximately 2, 12, and 22Hz.

To investigate the meaning of these initial results, it is recommended to analyse each channel separately, consider their topological characteristics, and isolate the main frequency bands, allowing us to compare the result obtained with the literature around them in a standardised way. Therefore, the main EEG power bands, defined as delta (0.1-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz), and gamma (30-40Hz), were extracted for each electrode and analysed both in terms of absolute power and relative power variation (ratio of powers between the signal band and the whole spectrum).

Following the background review and the result presented by Fan et al. [10], our analysis also focused on the $\frac{\alpha+\theta}{\beta}$ ratio, with the expectation to find an increase in the α , θ and δ bands and reduction in β in correspondence of a change in cognitive load; the relative power of each frequency band has been computed as a percentage of the absolute value of the total power in the frequency spectrum (0-40Hz); statistically relevant differences in each frequency band were investigated across different stimulation modalities using Welch's ANOVA test, to account for multiple unmatched set with heterogeneous variance. All results have been reported in section 7.1.1.

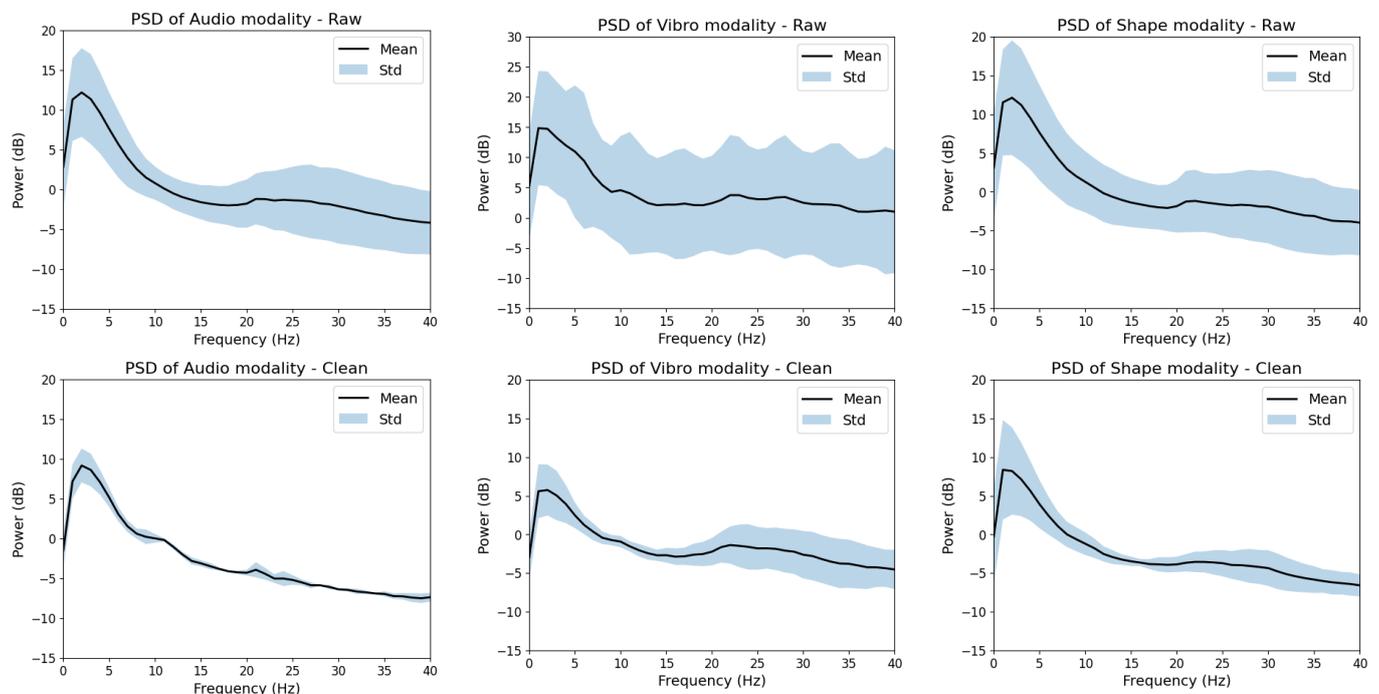


Figure 18: Power Spectrum of the raw (top) and pre-processed (bottom) data for each of the three stimulation modalities, Audio, Vibro and Shape respectively (left to right)

6.3 Time domain Analysis

The feature extraction of time-domain features has focused on the research of Event-Related Potentials (ERPs) following the presentation of a stimulation event. To obtain more detailed results, the responses to ‘left’ and ‘right’ commands for each stimulation modality have been separately analysed, although their effect at the cognitive level is expected to be closely related.

Furthermore, the position of the electrodes is an important element to consider when analysing time domain features, as discussed in 2.3. Whilst the temporal electrodes are close to the auditory and somatosensory cortex and are expected to detect the reception of sensory information (with distinct emphasis on auditory ones), the frontal electrodes are expected to detect small potentials linked to the effect of the response on the working memory load and fatigue levels [27]. Therefore, the response from these will represent different underlying processes that must be considered individually.

Unfortunately, much of the process of haptic and visual information happens in the central and posterior regions of the brain, respectively. These areas are not monitored by the electrode disposition of the BCI used; however, this study is focused on the cognitive load and mental state rather than the general processing of sensory information. Therefore, on the one hand, we can confirm the effectiveness and quality of the recording by using the temporal electrodes to see different levels of response in the temporal region according to the various stimulation modalities, expecting a higher potential following the Audio commands. On the other hand, we can use the frontal electrodes to analyse the effect of the stimulation on the mental state and memory load in a comparative way between modalities, as per our goal.

Firstly, all trial blocks have been pre-processed and epoched; all evoked responses to each event type have been extracted and combined; then, the mean and standard deviation across all events for all trials of the same stimulation modality have been computed. Figure 19 shows example plots of Event-Related Potential to the auditory event ‘left’ for a single participant’s average response, together with the power distribution across the scalp at key response moments and the absolute power across channels, or Global Field Potential (GFP), respectively. As expected, we observe a clear potential at around 0.2sec from the time of the event in the temporal region due to the processing of the stimulus in the auditory cortex.

Consequently, we compute a grand average response across all participants and for each stimulation to

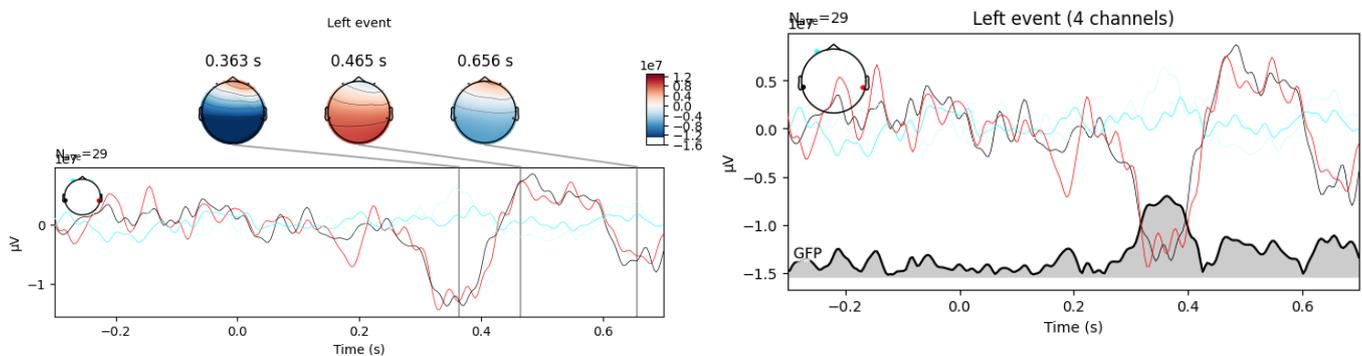


Figure 19: Event-Related Potential to the auditory event ‘left’ for a single participant’s average response; power distribution across the scalp at key moments of the response (left) and Global Field Potential (right); the presentation of the event occurs at time 0.

identify whether this response is generalisable and how it compares across different modalities. Figure 20 and 21 respectively show the raw and pre-processed data obtained from all evoked responses to the event ‘left’ command for all Audio stimulation trials and participants; the event presentation occurs at time 0 of the plot, with time window extricated between -0.3 and 0.7s from the event time.

Although some noise is still present in the filtered data, it can be observed that the overall magnitude and standard deviation are substantially decreased thanks to the pre-processing step. The latter also helped achieve the threshold of $|x| > 10\mu V$, increasing the probability of identifying micro-Volts responses in all channels. Some generalised trends across all participants can already be identified in the plots and confirm similar patterns to the single-participant data. An example is the negative peak at around 0.2ms after the occurrence of left audio stimulation, detected by the temporal electrode TP10. All results for the three stimulation modalities will be analysed in section 7.1.2.

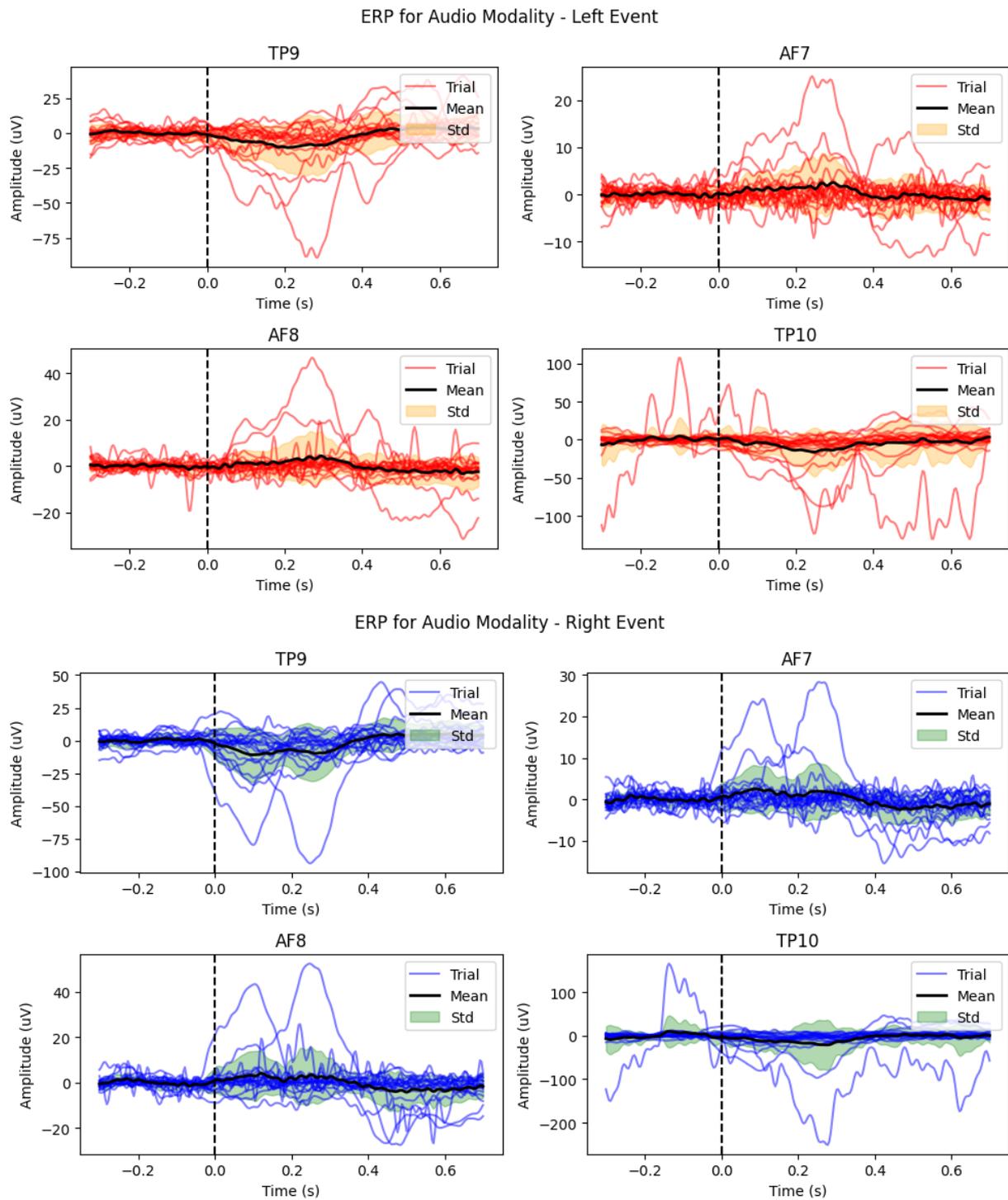


Figure 20: ERP extraction for the Audio stimulation trials from raw data for 'left' and 'right' events types (top and bottom respectively); the presentation of the event occurs at time 0.

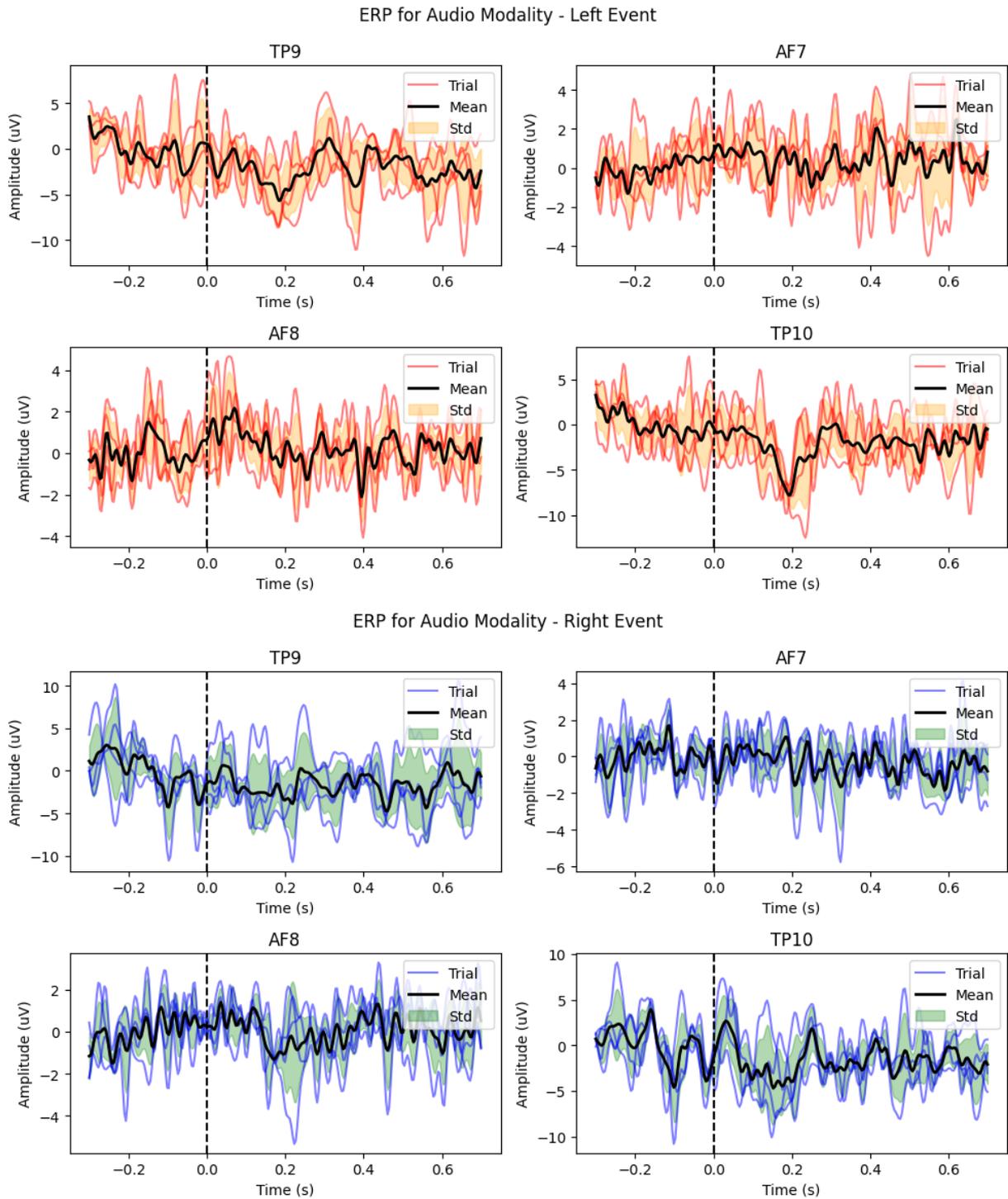


Figure 21: ERP extraction for the Audio stimulation trials from pre-processed (right) data for ‘left’ and ‘right’ events types (top and bottom respectively); the presentation of the event occurs at time 0.

6.4 Morlet Wavelets

To visualise time and frequency features in a joint way and identify ulterior non-stationary features of the EEG, we computed the Morlet Wavelet representation of the data.

Following the definition and application described in [65], this transform is computed by convolution of the signal with the kernel:

$$\psi(t) = c_\omega * e^{-\frac{t^2}{2}} * e^{j\omega t} \quad (1)$$

Where the normalisation factor c_ω was set to $c_\omega = \pi^{-\frac{1}{4}}$.

In our case, the evoked responses for each event type and each channel were filtered, combined and transformed into the wavelet domain. Then, their power was averaged and plotted firstly in terms of absolute values for each stimulation modality. An example plot of the single-modality absolute power is provided in figure 22 for the evoked response to the event 'left' Audio stimulation. Due to the presence of substantial noise within the spectrum, it is hard to identify the response, although we can recognise a low-frequency peak at around 0.2sec from the event in the temporal region and a slower response at around 0.5sec in the frontal one. All plots of the absolute power for the other event and stimulation types can be found in the appendix A.1.

To more clearly identify substantial patterns in the response, we then computed the difference between the spectra of two modalities for all combinations of Audio, Vibro and Shape stimulation. Figure 23, 24, and 25 shows the comparative results for the responses to the 'left' event, plotted in the linear domain to highlight the differences in the low-frequency components.

Regarding the data from the temporal region, we observe a consistent power difference between Audio and Shape conditions in the 0.5 sec following the event, with peaks at around 0.2 and 0.4 sec; this is likely due to the ERP detected in the auditory cortex being stronger than the somatosensory one, given the electrode positioning. Unsurprisingly, similar results are found in the comparison between Audio and Vibro conditions, although the peak seems stronger in magnitude in this case, suggesting that the vibrating stimulus elicits a less similar response to the auditory one that the shape-changing did. These results found correspondence with the frequency band analysis, as will be described in section 7.1.1. Finally, the two haptic conditions seem to have a more similar power distribution (with less intense negative peaks given); in this case, the two responses are slower, with peaks at approximately 0.25 and 0.45sec from the event presentation. If this is due to a slower reaction time for the processing of the vibrating stimulus, this result could be related to the lower task-performance results this modality is characterised by, as later discussed 7.2.

Regarding the frontal electrodes, the response difference is less evident, also due to the lower overall power of the signals from this brain region. However, we can still identify a more intense low-frequency power in the AF8 electrode at 0.4 sec from the Audio response compared to the haptic ones. Similarly, a stronger low-frequency response in the frontal area is also registered for the Shape compared to the other two modalities, as can be deduced from the negative peak at around 0.5 sec in both AF7 and AF8 electrodes.

Similar results are found in response to the 'right' event, with plots in the appendix A.1.

Overall, the wavelet analysis led to some interesting observations in the variation of low-frequency components across time; however, the high noise level makes it challenging to clearly attribute these differences in the time-frequency spectrum to a clear pattern in the mental state variation across conditions. Therefore, some complementary analysis is needed.

Since focusing on non-stationary features of the EEG data could lead to a more noise-robust analysis of the overall results, an investigation of the level of disorder and complexity of the data was pursued, as described in the next section.

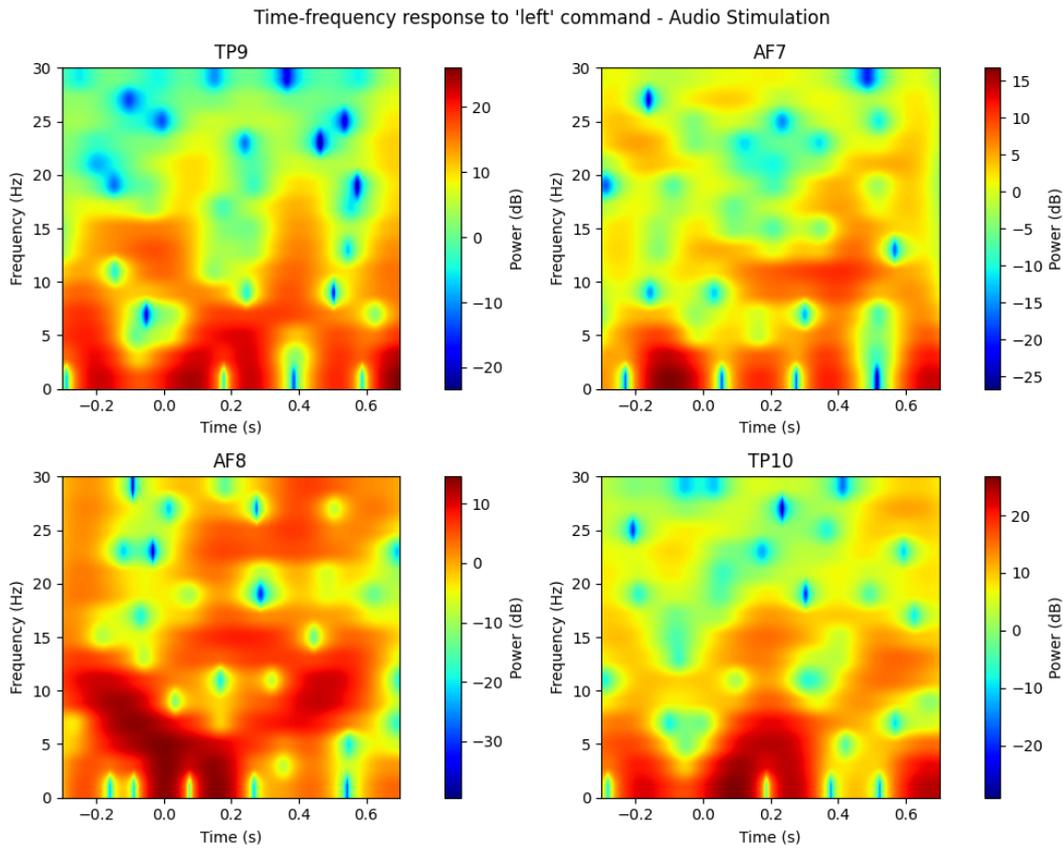


Figure 22: Morlet Representation of Time-Frequency response to 'left' command for Audio stimulation type; event occurrence at time 0; average across all trials and participants.

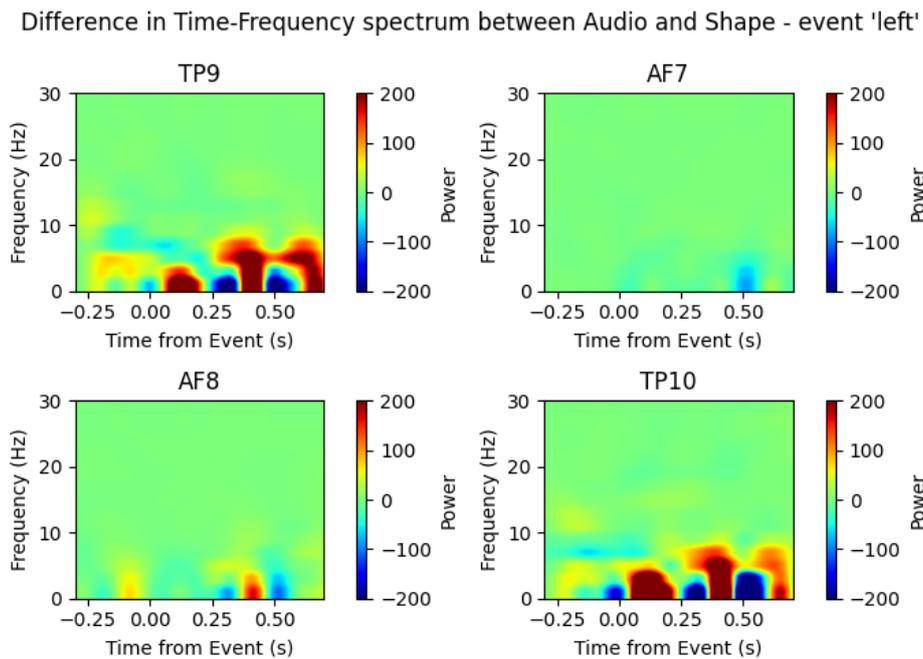


Figure 23: Morlet Representation of the difference in Time-Frequency response between Audio and Shape stimulation type; response to 'left' command, with event occurrence at time 0; average across all trials and participants

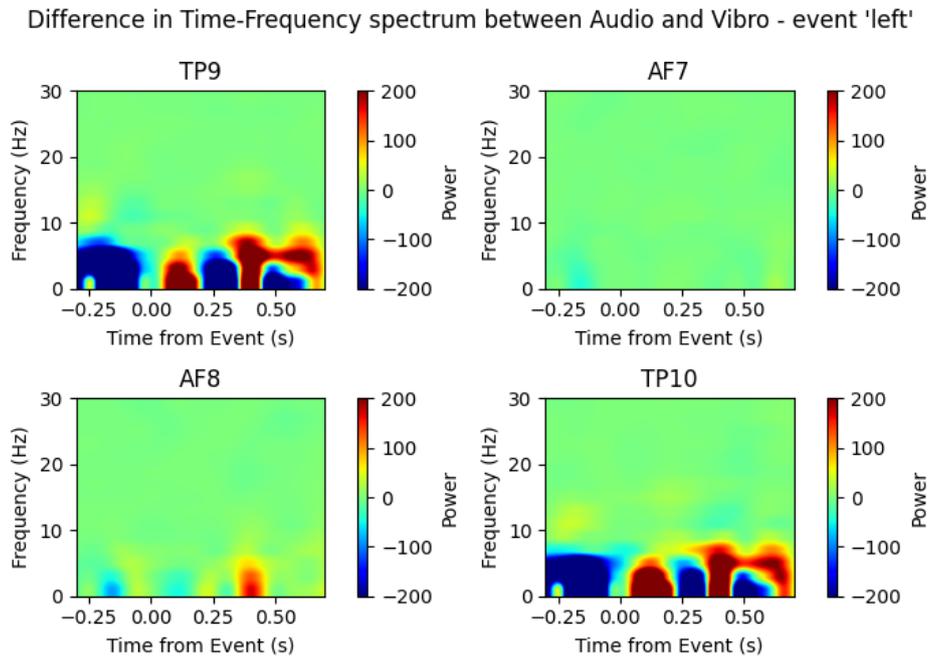


Figure 24: Morlet Representation of the difference in Time-Frequency response between Audio and Vibro stimulation type; response to 'left' command, with event occurrence at time 0; average across all trials and participants

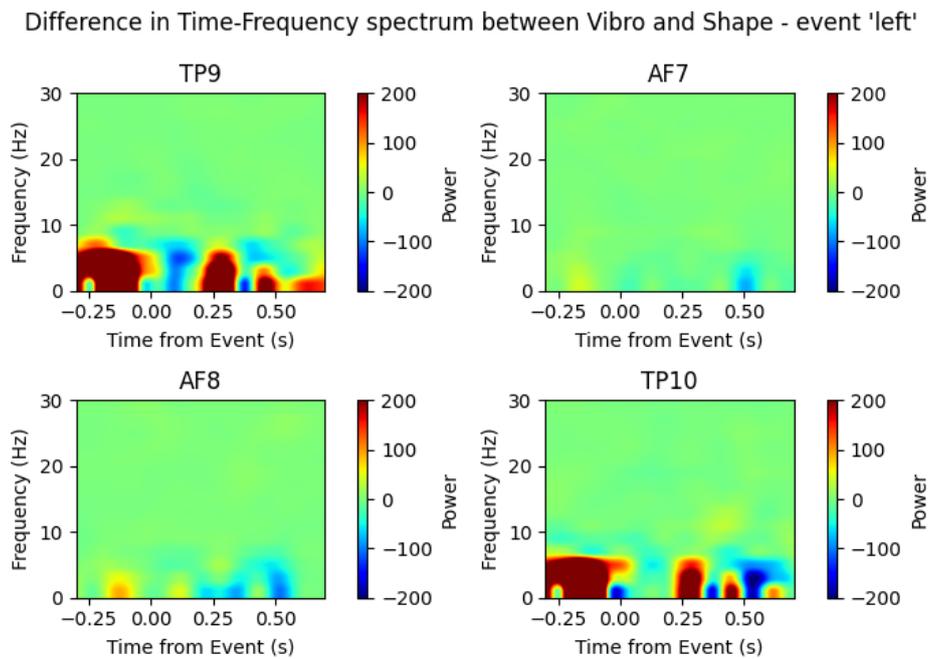


Figure 25: Morlet Representation of the difference in Time-Frequency response between Vibro and Shape stimulation type; response to 'left' command, with event occurrence at time 0; average across all trials and participants

6.5 Entropy Analysis

Following the results presented in the literature, the level of entropy in EEG has been hypothesised to indicate mental state variation when analysed in a comparative way between epochs or conditions (see 2.1.3). In particular, the study of two types of entropy has led to promising results in previous studies: Shannon Entropy (SE), denoting the amount of randomness and complexity in a signal, was used as an indicator of mental fatigue as identified by Fan et al. [10]; Permutation Entropy (PE), measuring of the unpredictability of the ordinal patterns in the time series, was used to highlight underlying patterns in ERP responses due to its robustness to noise and low computational power [66], [67].

Based on the work of [68], [69], SE was used as a disorder quantifier and defined as:

$$SE = - \sum_j p_j \log(p_j) \quad (2)$$

SE levels were computed for each trial block, and the results were compared against the set of trials for each stimulation modality.

Instead, following the work of [70] and [71], PE was used to quantify the complexity of the EEG time series by considering the order relations between its values (or sequence of their ordinal value given their relative magnitude), and was computed as follows:

Given a time series $X = x(1), x(2), \dots, x(N)$, we define an embedding dimension n and construct the vector:

$$X_{mi} = [x(i), x(i+1), \dots, x(i+n-1)] \quad (3)$$

With i between 1 and $N - n + 1$.

Then, if π is a permutation of order n of the ordered arrangement of the values in X_{mi} , we obtain $n!$ permutations and compute their relative frequency in the time series as $p(\pi)$.

Finally, the permutation entropy will be the Shannon Entropy of these frequency values:

$$PE(X) = SE(p(\pi)) = - \sum_j p(\pi)_j \log(p(\pi)_j). \quad (4)$$

Similarly to SE, PE was used to measure the unpredictability, although in this case, this was computed for ordinal patterns in the EEG time series instead of on the raw values themselves.

Although some SE and PE variations were recorded, with higher mean entropy across Shape-changing haptic stimulation trials, the statistical analysis of the results (using standard Analysis of Variance) highlighted no relevant difference between modalities. This result is likely due to the high noise level in the recording, which contributes to the randomness and complexity of the signal, consequently resulting in a high entropy across all trials.

6.6 Independent Component Analysis

Independent Component Analysis was investigated with a dual goal: firstly, for denoising purposes, to determine the components containing motion artefacts; secondly, for feature extraction purposes, to identify additional differences across conditions that might not be evident from the simple time and frequency analysis.

ICA was performed on the time domain representation of the epoched data. Four components were extracted across all channels and plotted in the form of the scalp topography, i.e. the distribution of its power across the whole area of the head based on each electrode contribution to that component.

Following the guide on source identification of EEG IC components provided by Pion-Tonachini et al. [72], we were able to distinguish between artefact and brain activity-related components. For example, in the plot in figure 26, we identify eye-movement artefacts in the fourth component (ICA003), characterised by two frontal areas with higher and opposite power. Instead, the first component (ICA000) is indicating a consistent difference in power between the frontal and temporal areas of the brain. However, it has to be noticed that having only four electrodes limits the amount of information and precision in source localisation; therefore, all ICA results must be considered with care.

A further analysis arises from the plot of the ICA component contribution in the time domain. In figure 27, we select an epoch-long time series, individually removing one independent component in each plot and comparing the cleaned and raw data. In particular, this process leads to satisfactory artefact-removal results when excluding the first component.

However, removing the first component for such a low resolution and high noise signal could also mean losing part of the useful information contained in the epoch not affected by the artefact. Therefore, in a second moment, we decided not to use this technique for artefact removal and limit our denoising process to excluding the epochs where distinct motion artefacts such as the one in the plot are present and by setting a threshold for the maximum signal amplitude.

On the other hand, for feature extraction purposes, each component was analysed by plotting its frequency spectrum and time domain average, highlighting the distribution and similarity for each epoch extracted by a single recording. Figure 28 shows an example result of this analysis for the first component extracted by recording in the Audio modality; all other components for all three stimulation modalities are provided in the appendix A.3.

Although the time features did not appear to be informative, given the non-periodic occurrence of stimulation events across the stimulation, it was possible to identify new frequency-domain features, consistently varying depending on the stimulation modality. In particular, a peak around 10 Hz, and a corresponding increased relative power in the alpha band, were observed in the Audio and Shape modality compared to the Vibro one. This observation, although not significantly informative if taken individually, helps support the results previously found in the frequency analysis and further discussed in section 7.1.1.

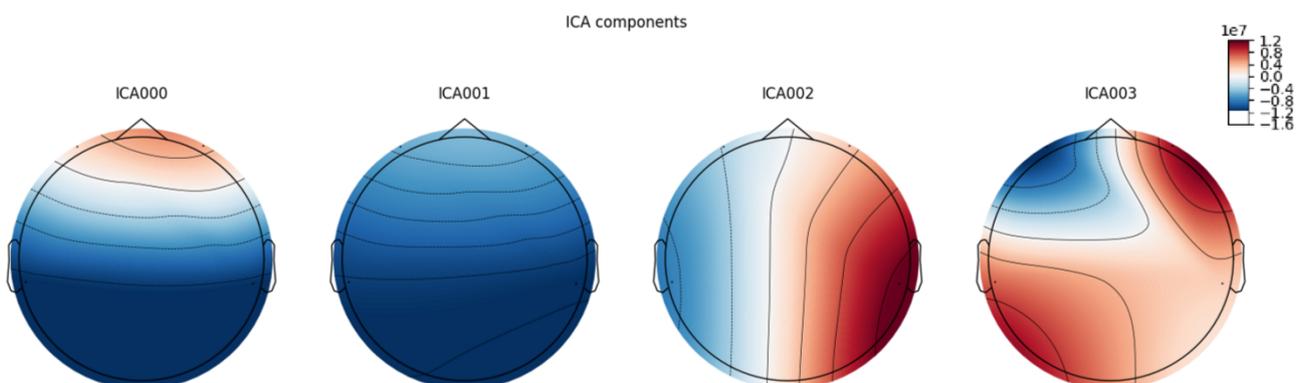


Figure 26: Scalp Topography of four Independent Components extracted from the 4 electrodes (top-down view of the head, with frontal electrodes at the top and temporal electrodes at the sides)

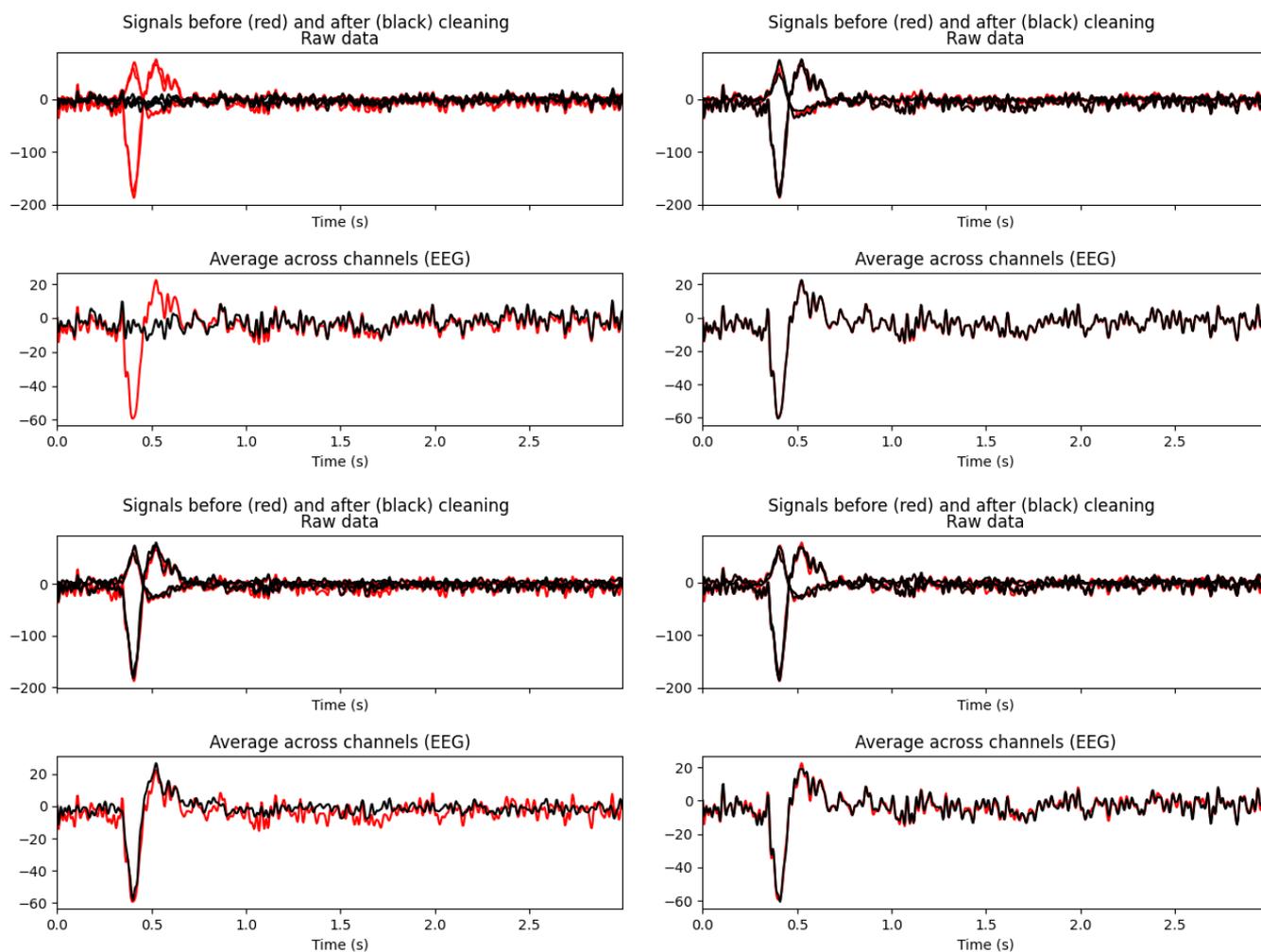


Figure 27: Example of four independent components contribution (top to bottom and left to right, removing the components 1 to 4 respectively) extracted from a single epoch (amplitude in μV). The artefact is removed by cleaning the data from the contribution of the first component (top-left plot).

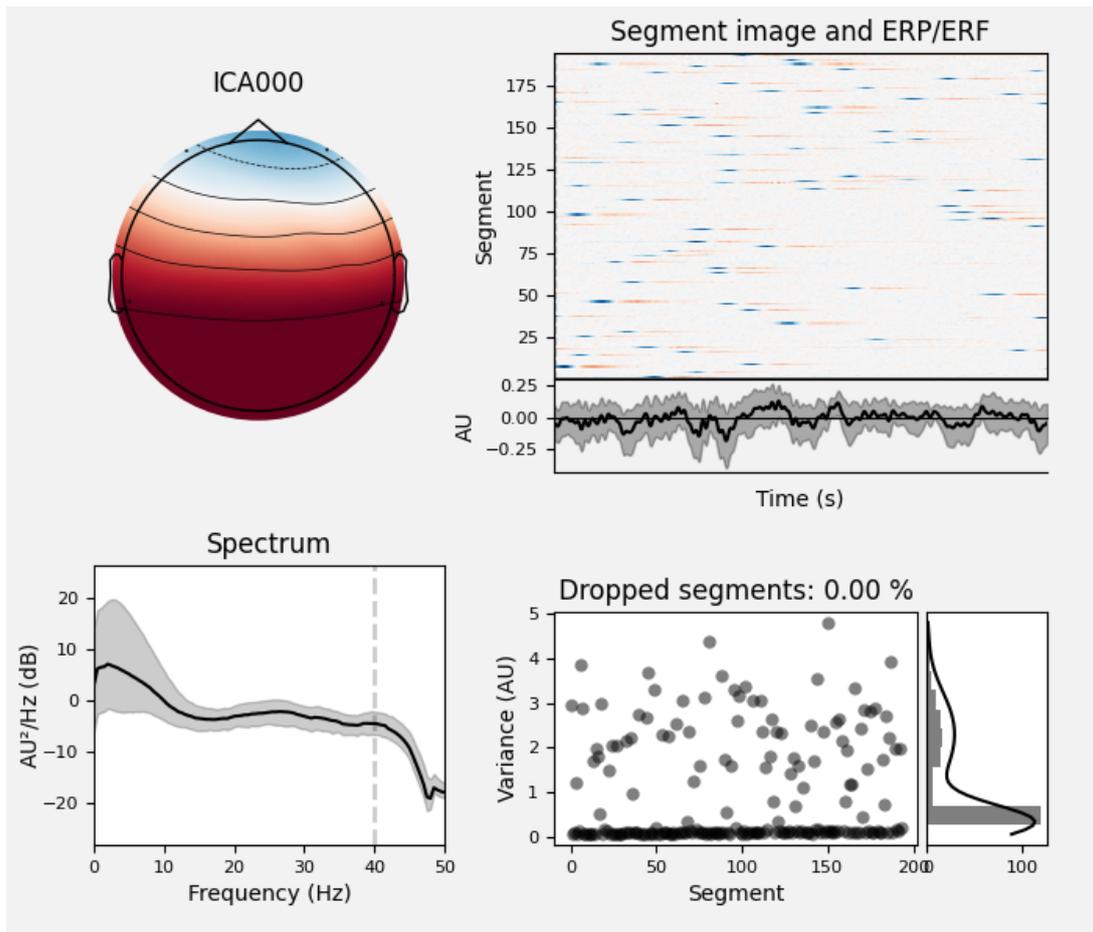


Figure 28: Example of ICA diagnostics for the first independent component extracted from a recording in the Audio modality. Top-left: scalp topography; Bottom-left: mean and standard deviation of the frequency spectrum across all epochs; top-right: ERP distribution for each epoch (segment) and mean amplitude of the epochs in the time domain; bottom-right: distribution of the variance for each of the epoch, indicating a proportional noise contribution.

7 Results

This section covers the main results obtained from the experimental trials, presenting the analysis of EEG, task performance and questionnaire-based data.

7.1 Physiological Data Results

7.1.1 Frequency-domain Features:

Figure 29 shows the mean and standard deviation of the frequency spectrum for all trials of each stimulation modality, separately extracting the data from each channel.

As previously observed, we identify three peaks at approximately 2, 12, and 22Hz; these correspond to increased power in the delta, alpha, and beta frequency bands, and are more evident in the temporal region, indicating their link with the processing of sensory information; this is supported by the fact that their magnitude is higher for auditory stimulation, followed by the shape stimulus and lastly the vibration one. In fact, as explained in background section 2.3, different haptic sensations are processed in separate regions of the brain; the higher magnitude of responses to translation motions of the Shape-changing device, compared to vibrations, is in accordance with the fact that the first ones are linked to orientation and navigation features, which are processed in the vestibular system in the inner ear, whilst the latter are mainly processed in the Brodmann areas in the primary sensory cortex, further from the TP electrodes.

Additionally, this magnitude difference in the mid bands of the spectrum meets the experimental results obtained by [73], [39], indicating how vibrotactile stimulation induces suppression of both alpha and beta band activity in the sensorimotor areas.

Regarding the analysis of each frequency band, figure 30 shows the mean and standard deviation of the relative power distribution for each stimulation modality. The relative power is expressed as a percentage of the total spectral power for each channel, normalised to account for the different total power of each modality. The data were analysed to identify the statistically relevant differences between conditions, obtaining the following results:

- Higher Beta power in all regions for both haptic modalities compared to auditory one, previously linked to sensory-motor tasks that require active cognitive processing.
- Lower Gamma power for Audio in all regions compared to both haptic modalities, related to an increase in mental effort [74].
- Lower Delta power for Shape and Vibro in the frontal electrodes, associated with the sustained attention and inversely proportional to fatigue or disengagement levels [75].
- Lower Theta for Vibro in frontal electrodes, observed in response to lower workload and a decrease in sustained attention to new information [76], and a marker of multisensory divided attention [77].
- Higher Alpha for Shape in AF8, indicating lower levels of “alpha resynchronisation”, i.e. a decrease of the alpha power proportional to the cognitive demand of a task compared to a resting state and relating to the inhibitory role of alpha-band oscillations discussed by [76].
- Greater overall statistical difference of each haptic modality with the auditory one, compared to the more similar power distribution between the two tactile modalities themselves.

These results highlight a consistent difference in the cognitive processes, with similar stimulus processing responses across haptic modalities due to their association with sensorimotor information, and analogous mental load between Shape and Audio conditions. As supported by the task performance and subjective results later discussed in 7.2 and 7.3, the Shape condition appears to involve lower cognitive demand, and the Audio one correlates with features linked to mental effort and low attention levels. Furthermore, a higher alpha resynchronisation for the vibration modality matches the study of Kim et al. [39], who observed significant alpha suppression in sensorimotor regions and other cortical areas in correspondence with the vibrotactile stimulation of the non-dominant hand.

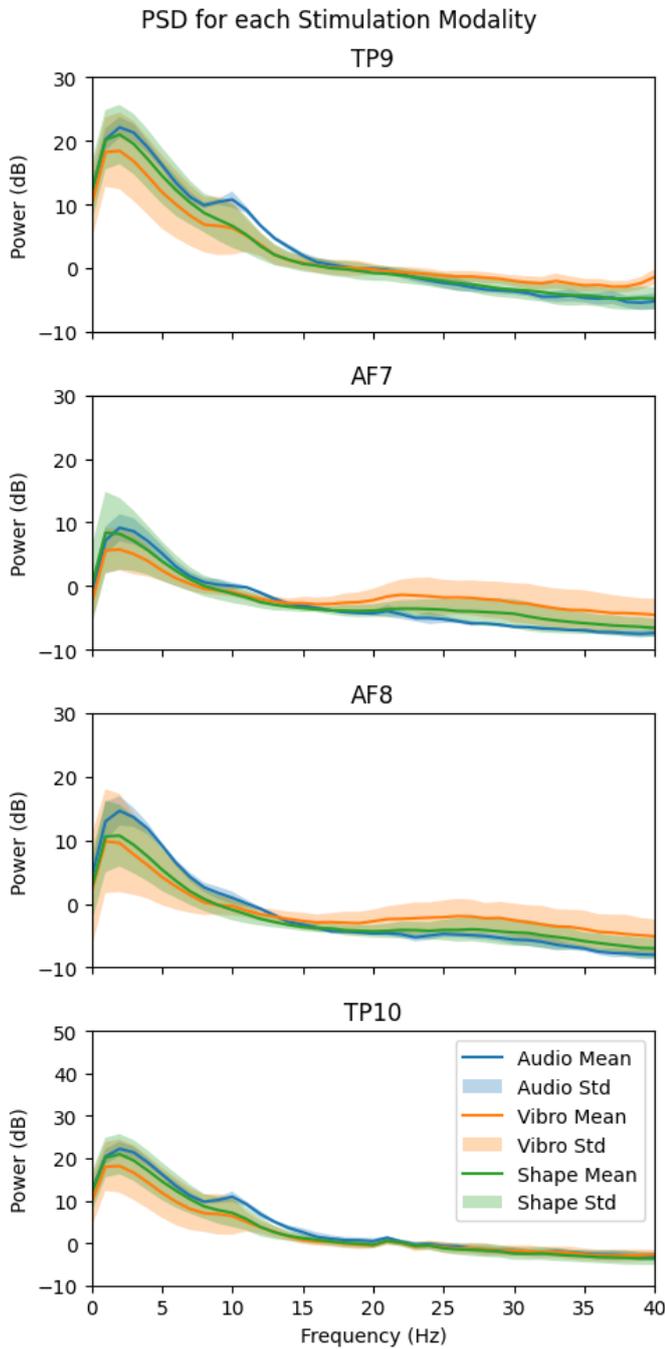


Figure 29: Mean and standard deviation of single channel frequency spectrum across all trials of each stimulation modality.

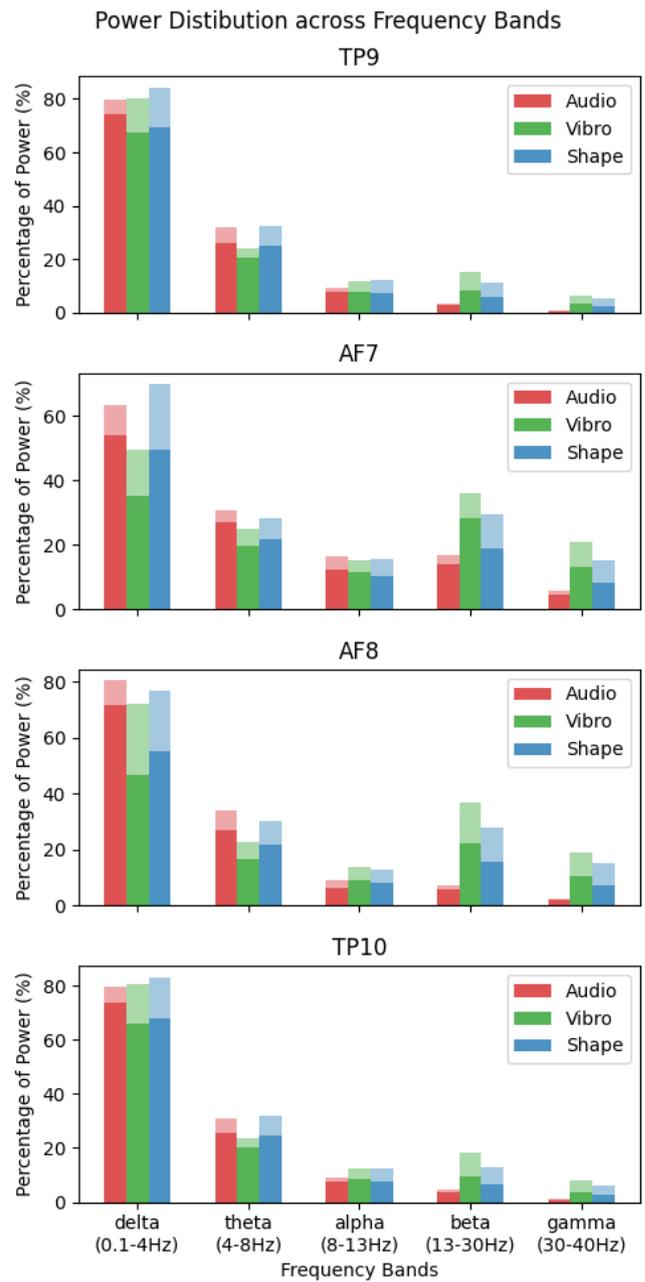


Figure 30: Mean and standard deviation of the relative power distribution across frequency bands for each stimulation modality. Relative power is expressed as a percentage of the total spectral power for each channel.

7.1.2 Time-domain Features:

The results of the ERP analysis on a cross-participant basis are shown in figure 31. Although the noise influence is substantial, we can identify some patterns in the mean ERP across trials depending on the brain region and stimulation modality. Firstly, a high correlation between electrodes of the same region is observed across all stimulations and events, with more elevated power in the temporal region compared to the frontal one. Similarly to what was found during the wavelet analysis as discussed during the wavelets analysis 6.4, a strong negative deflection at around 0.3 sec for the event is recognisable in the temporal region for both Shape and Audio modality, whilst the frontal electrodes show signs of a lower amplitude response ($5\mu V$ and $2\mu V$ for each modality respectively) with similar timing.

However, the overall data is characterised by a high level of noise, especially in the Vibro modality, hampering our ability to expand the comparative analysis of these results.

To overcome this limitation, we used a more noise-robust strategy: the Root Mean Squared (RMS) between the mean amplitude of a baseline time window (0.2-0s before the event) and of a response-timed one (0.2-0.4s after the event) was computed for all three conditions. Some relevant results were observed when comparing the RMS across different stimulation for trial blocks and channels individually (see appendix A.2). The analysis showed a weaker response amplitude in the temporal region of the brain for the Shape condition compared to the Auditory one and a greater potential of the response in the frontal region following Vibro stimulation. Similarly to what was uncovered by the other metrics in 7.1.2 and 7.3, this could suggest a more intense processing activity for auditory stimuli compared to haptic ones and high load levels in the frontal cortex for Vibro modality.

Instead, as common practice in EEG time domain analysis, the data were inspected separately for each participant, identifying the cleanest trials and drawing some individual-specific conclusions. It has to be noticed that, although these results cannot be generalised, they give an initial direction of investigation on a type of feature that is inherently participant-dependent. Furthermore, each participant's response results from an average across approximately 50 trials for each event type, meaning that some generalisation is still possible, at least in the comparative analysis of different experimental conditions.

An example of the average response to 'left' and 'right' commands for each stimulation modality is plotted from the data of Participant 04 and displayed in figure 32; this was analysed separately for each of the two trial blocks in order to identify possible variation due to long-term fatigue. We observe how, despite a non-negligible deviation between trials of the same block, clear differences can be noticed due to the variation in stimulation modality and the time progression (trial block 1 vs 2).

As observed before, Audio trials are characterised by a slow ERP response ($t \approx 0.3s$ from the event) with similar amplitude between block trials, suggesting a low effect of mental fatigue throughout the experimental progression. Instead, the Vibro stimulation resulted in a faster response at around 0.1s from the event, with a shorter peak duration and amplitude in the second trial block. This could indicate how the cognitive response to the vibration-based stimulus is subjected to a greater variation throughout the experimental progression, hinting at a time-based change in the mental state (lower ERP amplitude implying a higher cognitive load, as previously found by [37]).

Similar effects were noticed for different trial blocks of the Shape condition, indicating that both haptic modalities could have stronger time-varying features than the auditory one. It is hard to establish if this could be due to lower familiarity with these types of commands, indicating a relationship with the amount of learning involved or due to the processing load inherited to the stimulation modality. Further experiments are needed to document the time variation over longer training periods and with lower noise levels.

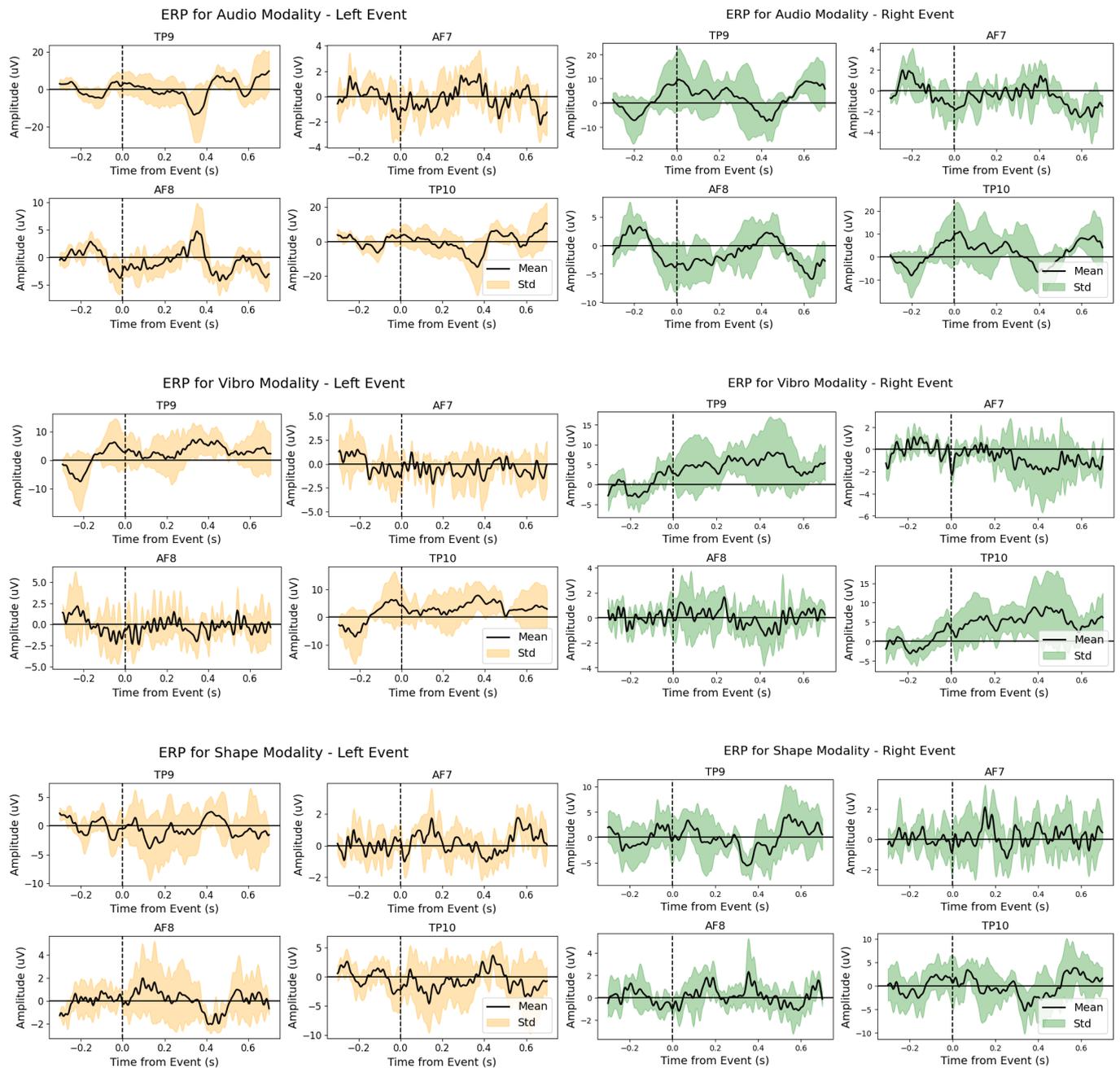


Figure 31: Average ERP response to ‘left’ and ‘right’ commands of each stimulation modality; Mean and Standard Deviation computed across all trails and all participants, with $t = 0$ corresponding to the stimulus occurrence

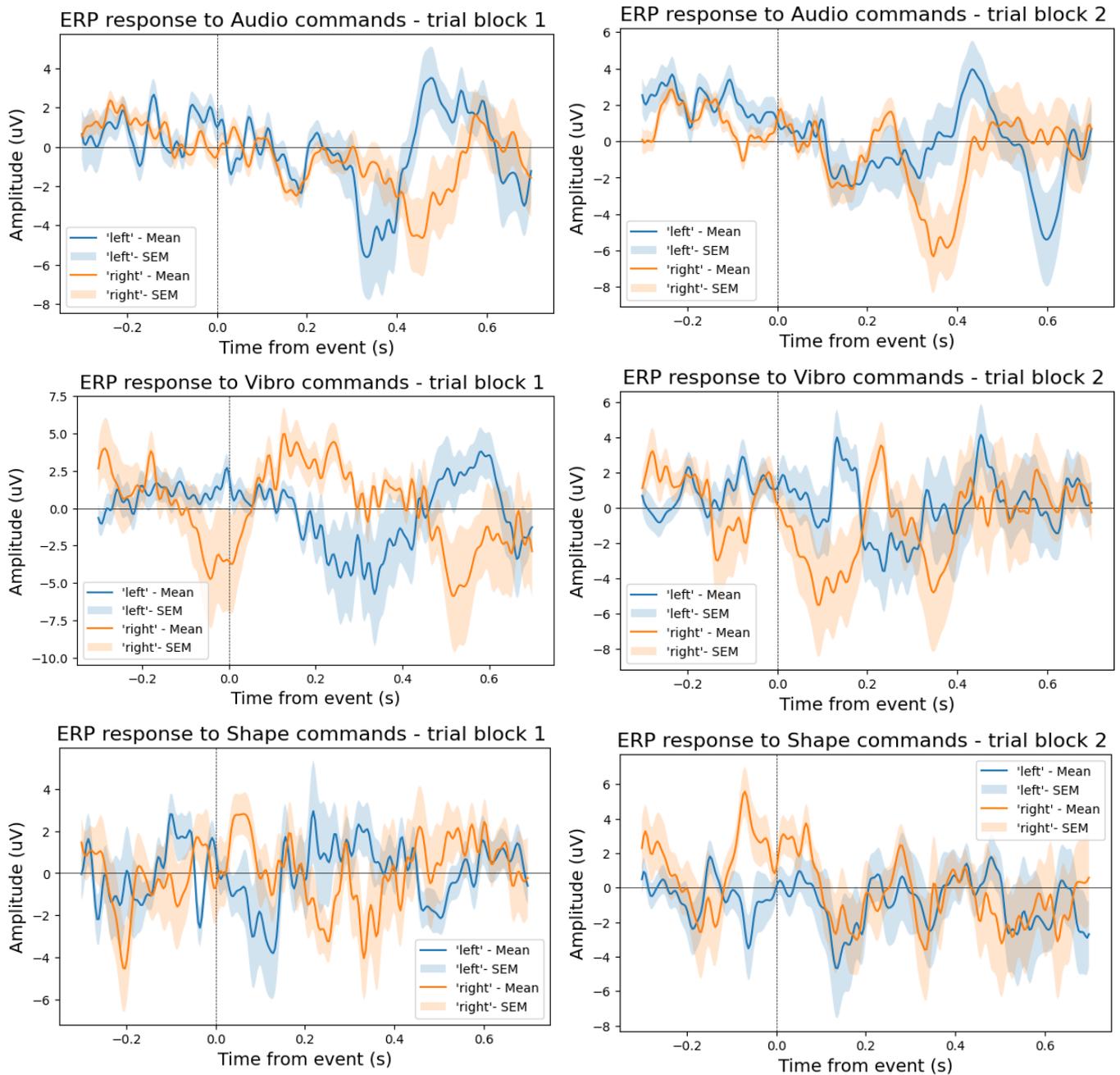


Figure 32: Average ERP response from all channels of Participant 04 to 'left' and 'right' commands of each stimulation modality; Mean and Standard Error Mean (SEM) across all trials of each trial block plotted in time, with $t = 0$ corresponding to the stimulus occurrence.

7.2 Task-Performance Results

Metrics: Three parameters were considered when evaluating the dual task objective performance:

- **Counting Error (%)**: the error in counting the blue circles on the screen as a percentage of the total number of blue circles appearances across the trial block.
- **Missed Direction (%)**: the percentage of missed responses (no arrow key pressed) following directional stimulation out of the overall number of stimulation events. The stimulation event was considered missed if no arrow key was pressed when a new stimulus was presented, considering a minimum lag between stimulation events of 2-3 seconds.
- **Wrong Direction (%)**: the percentage of wrong responses following directional stimulation (i.e. right arrow key pressed following left command, and vice-versa) out of the overall number stimulation events for each trial block.

Statistical Analysis Results: To reject the null hypothesis, statistical analysis was performed to establish the relevance of the differences in performance across different stimulation types and trial blocks.

First, the Brown-Forsythe test was used to evaluate whether the variance of each condition was homogeneous. Since the p-values from Bartlett's test resulted in being less than the 0.05 threshold, we rejected the null hypothesis that each group had the same variance. We, therefore, performed Welch's Analysis of Variance (ANOVA) instead of the standard ANOVA method to account for the heterogeneous variance. The results from Welch's ANOVA are reported in table 1, where it can be seen a statistical difference ($p < 0.05$) between the different stimulation modalities for the mean errors of the missed and wrong responses, while the difference in mean circle miscount can be disregarded as not statistically relevant.

Similar tests were used to assess the difference in results between trials block 1 and 2; in this case, due to the homogeneity of the variance (assessed using Levene's test based on the absolute deviations of the data from the group means), a simple ANOVA test was performed. This led to high p-values (also in table 1), implying the statistical irrelevance of the difference between trial blocks. This could be due to the difficulty of the dual task being low, meaning that little to no learning affected the results between the first and the second trial block. The little variation in performance could also be due to a randomly distributed variation in the attention level of the participants. This is an expected result as the target of the test was to assess the participant's ability and stimulation responses, limiting the task difficulty and the amount of learning required in exchange for a higher dependence upon the mental state (mental fatigue and attention drop) of each participant.

Since no statistical difference in the performance was found even between the first and second trials of each modality individually, we hypothesised that the difference in EEG signals across trials was a consequence of the variation in the attention level and mental fatigue due to the accumulating load of using a specific stimulation modality, rather than due to a variation in the difficulty caused by a learning curve between trial blocks.

<i>Varying Condition</i>	Missed Directions	Wrong Directions	Counting Error
<i>Stimulation Modality</i>	0.000487	0.000035	0.087875
<i>Trial Number</i>	0.506276	0.813120	0.568053

Table 1: p-values following ANOVA and Welch's ANOVA tests on the statistical difference in performance results across different stimulation modalities (Auditory, Vibro-Tactile, Shape-Changing), and between trial numbers (First VS Second Trial). P-values < 0.05 are highlighted, signifying a statistical difference between conditions.

Auditory vs Haptic Task Performance: Figure 33 shows the mean and standard deviation of each stimulation modality (Auditory, Vibro-Tactile, Shape-Changing) grouped by the performance parameters. Some interesting results can be observed: firstly, auditory stimuli led to the highest percentage of missed responses.

This low receptiveness could be due to the higher familiarity with stimulation type (auditory stimulus being the standard modality for navigational guidance) and the consequent difficulty of maintaining a higher level of focus in a repetitive and familiar task, as opposed to testing novel interfacing modalities. However, even with prolonged usage, the receptiveness to haptic modalities did not converge to the same levels of errors as the audio one, indicating that the auditory task was inherently a more difficult modality to be focused on. This result was confirmed by the questionnaire-based subjective evaluation and the verbal feedback 7.3.

As per the experiment design, the dual task required increased levels of sustained attention, highly varying across participants; in this regard, we notice a consistently higher standard deviation of the performance in the Audio modality, suggesting a stronger dependence between the stimulus receptiveness and participant’s mental state in the auditory compared to the haptic conditions.

When considering the implications of this on the deployment of navigational devices, we hypothesise that this would lead to higher variation in performance and increase mental load depending on the ongoing mental state of the participants. Conversely, the user might find it harder to ignore a haptic stimulus, even if in a state of low alertness. We also hypothesise that this could be linked to the perceived proximity of the stimulus, with higher levels of “danger” presented by a haptic sensation compared to an auditory one.

On the other hand, we also observe a consistently lower error due to incorrect responses following audio commands compared to haptic ones. This could suggest that the spatial mapping between the origin of the sound and the direction commanded is more immediate and robust than the relationship between direction and in-hand haptic sensation. However, these results should be investigated to test the effect of longer training periods with the two haptic devices.

Finally, promising results have been obtained for linear motion-based stimulation (i.e. Shape modality) compared to vibration-based ones. The coherence between command and response direction and the receptiveness to the stimulus were far greater for the former, with consistent results across trial blocks. This suggests that it is more intuitive to associate a directional command with a linear movement, such as in the case of shape-changing haptic sensations, compared to linking it to a spatial mapping of the vibrating ones.

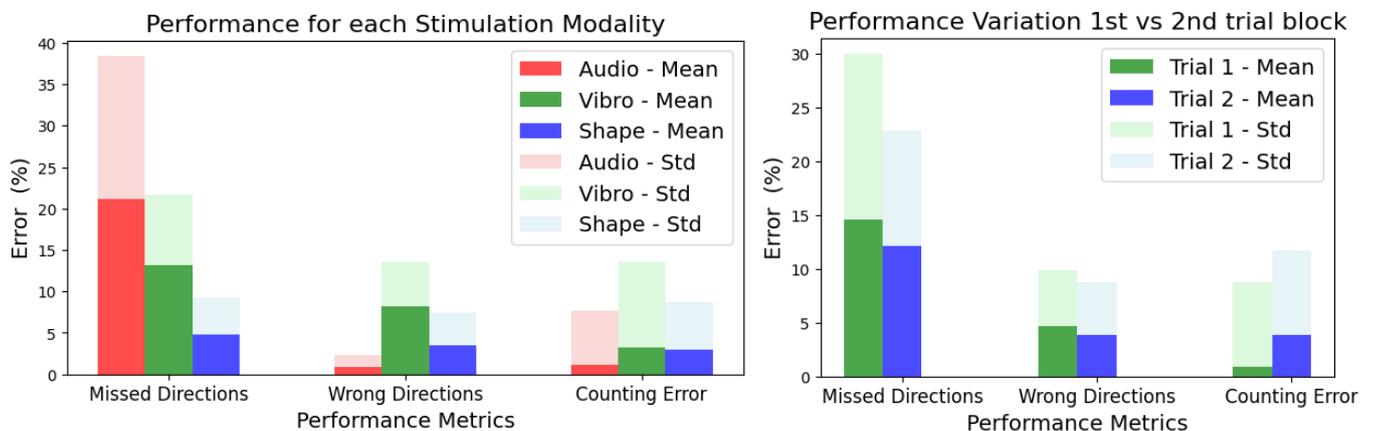


Figure 33: Dual Task Performance for different stimulation modalities (left) and for first and second trial (right) across all participants. Mean and standard deviation of the error expressed as a percentage of missed and wrong responses to directional command and circle miscounting

7.3 Subjective Evaluation

In accordance with what was found in the objective performance, the Shape modality performed consistently lower in terms of work and mental load compared to the vibration-based device. Surprisingly, the mental, physical, and temporal demand from responding to this modality was ranked invariably lower even compared to the more familiar auditory stimulation, with the exception of the perceived level of effort, which was similar across the two.

The results from the statistical analysis of the difference between the three modalities can be summarised as follows:

- The perceived effort of responding to the vibrating command was vastly higher than the other two stimulating modalities, supporting the observations drawn from both EEG and task-performance analyses.
- Higher physical demand was recorded in the vibration modality compared to the other two, confirming the hypothesis on the difficulty of spatial mapping explained in the task-performance evaluation.
- Higher mental demand and frustration levels were reported for auditory modality, supporting the hypothesis that lower responsiveness corresponds to an increased level of fatigue and frustration due to a higher difficulty of the multitasking task; the verbal feedback of the participants matched this, as the majority of them reported difficulty in distinguishing between auditory and visual stimulation when having to respond to the concurrent stimuli.

These results match the observations by previous studies, such as the one pursued by Spiers et al. [58], which also reported a better performance evaluation of the shape-changing interface against the vibrating one, particularly regarding the intuitiveness, user comfort and level of distraction.

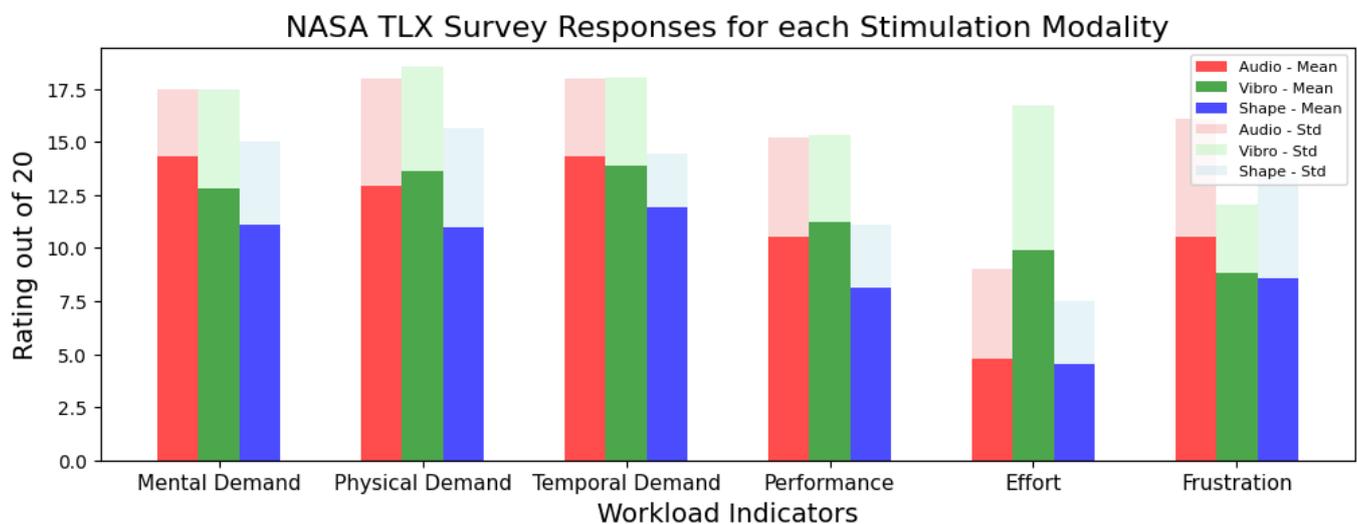


Figure 34: Results from the NASA TLX Questionnaire evaluating the perceived level of workload for each of the stimulation modalities; rating out of 20 point scale.

8 Project Evaluation

Given the successful experimental trials with 10 participants and the novel results found on two types of haptic stimulation, the project has met its target objective, as initially laid out in section 1.1.

The experiment design targeted the creation of a multitasking environment with appropriate levels of task difficulty and attention levels required. In particular, given that there was no statistically relevant difference in the primary task performance (counting the circles) across different stimulation types, this implies that the task design reached its goal, requiring sustained attention (with immediate effects on the performance when out of focus) whilst also ensuring a stimulation modality-independent distracting activity. In fact, since the project application consists of navigation-aiding devices, the core of this study was to analyse the results from the secondary task performance. Therefore, the experiment's success lies in the discovery of statistical differences in directional task performance between stimulation modalities, as discussed in the result section 7.

Regarding the EEG data, the physiological results obtained can be said to validate the research methods pursued and the use of the Muse device for similar studies. However, evident limitations arose from the use of such low-cost BCI, these being the reduced number of electrodes available, the high level of noise in the EEG, and the non-standardised setup, which ultimately led to lower data quality. As forecasted, this could have been avoided using a more precise research-level BCI if the project budget had allowed it.

Furthermore, the quality of experimental results would have also benefitted from a longer experiments duration, which was limited by the time availability of the participants; by running more extended trial blocks, we could have better targeted the effect of long-term mental fatigue and performed real-time feature extraction to check the quality of the recording in a more thorough and immediate way. Lastly, recording the participant's EEG at resting state would have also helped create a baseline for comparing the results on a trial-by-trial basis. These considerations will be taken into account in the next steps of the study.

Concerning the functionality of the haptic devices, both implementations met the requirements of the target experimental application; however, some improvement could be made to the vibration device to stimulate a narrower area of the hand and further separate the sensation provoked by the left and right commands, as suggested in the feedback from the participants of the experiment.

9 Conclusions and Future Work

Reprising the hypotheses laid down at the beginning of the study 1.1, we conclude that both of them have found some confirmation in the experimental data, although some consideration must be taken into account. The study provided supporting results to H_1 , which stated that in a multitasking environment, cognitive and task performance can improve by using haptic-based feedback instead of auditory-based one. However, we also showed how this depends on the type of haptics used, validating the H_2 hypothesis predicting that vibrating and shape-changing tactile stimuli would have influenced the user state and behaviour in distinct ways.

In summary, haptic stimuli appeared harder to ignore and corresponded to higher levels of user alertness and ability to discern between them and a visual stimulus. This might seem to imply that haptic-based devices are more disruptive; however, when looking at the results for the concurrent primary tasks, no decrease in performance can be found compared to Audio trials. In particular, we documented a lower power in higher frequency bands (β and γ) and higher power in the δ band from the frontal region under Audio conditions (7.1.1), indicating an increased effort and fatigue combined with a lower engagement.

If we combine these observations with the participant's subjective evaluation of the auditory stimulus, which was characterised by higher mental demand and frustration levels than any other stimulation, we can conclude that this modality is less effective in a multitasking context.

Furthermore, the lower number of wrong responses in the Audio trials, accompanied by higher overall cognitive load, also suggests how the better performance in this specific objective metric could be due to the higher familiarity with the stimulus, rather than because the audio directional command is inherently easier to process than haptic ones. In this regard, further exploration of the effect of prolonged training with haptic devices is required before drawing any further conclusions.

Concerning the comparative analysis of the two haptic modalities, the linear motion-based stimulus (Shape condition) represented a more intuitive type of feedback than the vibration-based one, possibly thanks to the inherent spatial mapping between the direction of the motion and the command. This was also characterised by lower stimulation-induced cognitive loading, both in terms of subjective and objective (task-performance) evaluation. However, we obtained contradicting results in the physiological analysis, with higher α power in the Shape modality compared to the Vibrating one, indicating lower active processing and cognitive demand, but higher θ power and time domain response in the frontal region (6.4), suggesting a higher workload and multisensory division of the attention. Furthermore, the vibrotactile stimulation was characterised by higher time-variation of its responses (7.1.2), indicating the possible influence of learning or long-term fatigue.

Ultimately, some encouraging results were found, supporting the development of new shape-changing haptic technologies over vibrotactile ones, but additional research on their cognitive effects will be required to expand the interpretation of these results. Using a higher resolution Brain-Computer Interface and pursuing a more location-precise analysis of the cognitive responses, we could confirm the statistical differences found and distinguish between variations in mental workload and sensory processing, obtaining a more detailed view of the underlying cognitive patterns in the different brain regions.

Additionally, a greater and more diverse range of participants needs to be tested to assess the effectiveness of these devices, with the ultimate goal of verifying the benefit of using haptic technologies also for visually impaired individuals.

Lastly, a possible expansion of this study could be to explore other types of haptic stimulation, e.g. pressure, heat, and skin-stretch-based ones.

In conclusion, the field of haptic technologies and the study of their impact on user performance and mental state are still in their nascent stages. This study serves as a fundamental stepping stone in this emerging area, demonstrating the potential benefits of employing such technologies in user-assistive devices. However, given the novel nature of this research, the findings presented necessitate additional corroborative studies and further in-depth exploration.

References

- [1] C. Jeunet, C. Vi, D. Spelmezan, B. N’Kaoua, F. Lotte, and S. Subramanian, “Continuous tactile feedback for motor-imagery based brain-computer interaction in a multitasking context,” **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, vol. 9296, pp. 488–505, 2015. pages 7
- [2] M. Marucci, G. D. Flumeri, G. Borghini, N. Sciaraffa, M. Scandola, E. F. Pavone, F. Babiloni, V. Betti, and P. Aricò, “The impact of multisensory integration and perceptual load in virtual reality settings on performance, workload and presence,” **Scientific Reports**, vol. 11, 12 2021. pages 7, 14
- [3] Y. Zhao, C. L. Bennett, H. Benko, E. Cutrell, C. Holz, M. R. Morris, and M. Sinclair, “Enabling people with visual impairments to navigate virtual reality with a haptic and auditory cane simulation,” **Conference on Human Factors in Computing Systems - Proceedings**, vol. 2018-April, 4 2018. pages 8
- [4] J. Y. Chen and P. I. Terrence, “Effects of tactile cueing on concurrent performance of military and robotics tasks in a simulated multitasking environment,” **Ergonomics**, vol. 51, pp. 1137–1152, 8 2008. pages 8
- [5] S. Advani, P. Zientara, N. Shukla, I. Okafor, K. Irick, J. Sampson, S. Datta, and V. Narayanan, “A multitask grocery assist system for the visually impaired: Smart glasses, gloves, and shopping carts provide auditory and tactile feedback,” **IEEE Consumer Electronics Magazine**, vol. 6, pp. 73–81, 1 2017. pages 8
- [6] A. Spiers, E. Young, and K. J. Kuchenbecker, “The s-ban: Insights into the perception of shape-changing haptic interfaces via virtual pedestrian navigation,” **ACM Trans. Comput.-Hum. Interact.**, vol. 30, no. 1, mar 2023. [Online]. Available: <https://doi.org/10.1145/3555046> pages 8
- [7] M. P. Menikdiwela, K. M. Dharmasena, and A. M. S. Abeykoon, “Haptic based walking stick for visually impaired people,” 2013. pages 8
- [8] A. Biasiucci, B. Franceschiello, and M. M. Murray, “Electroencephalography,” **Curr Biol. 2019 Feb 4;29(3):R80-R85. doi: 10.1016/j.cub.2018.11.052. PMID: 30721678.**, 2018. pages 9
- [9] C. S. Nayak and A. C. . Anilkumar, “Eeg normal waveforms - statpearls - ncbi bookshelf,” Jul 2022. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK539805/> pages 9
- [10] X. Fan, Q. Zhou, Z. Liu, and F. Xie, “Electroencephalogram assessment of mental fatigue in visual search,” **Bio-Medical Materials and Engineering**, vol. 26, pp. S1455–S1463, 2015. pages 9, 12, 29, 36
- [11] C. Q. Lai, H. Ibrahim, M. Z. Abdullah, J. M. Abdullah, S. A. Suandi, and A. Azman, “Artifacts and noise removal for electroencephalogram (eeg): A literature review,” in **2018 IEEE Symposium on Computer Applications and Industrial Electronics (ISCAIE)**, 2018, pp. 326–332. pages 10
- [12] Y. Yin, J. Cao, and T. Tanaka, “Eeg energy analysis based on memd with ica pre-processing,” in **Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference**, 2012, pp. 1–5. pages 10
- [13] M. Islam and T. Lee, “Multivariate empirical mode decomposition of eeg for mental state detection at localized brain lobes,” 2022. [Online]. Available: <https://arxiv.org/abs/2206.00926> pages 10
- [14] H. Nolan, R. A. Whelan, and R. B. Reilly, “Faster: Fully automated statistical thresholding for eeg artifact rejection,” **Journal of Neuroscience Methods**, vol. 192, pp. 152–162, 2010. pages 10
- [15] H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, and Q. Liu, “Eegdenoisenet: A benchmark dataset for deep learning solutions of eeg denoising,” **Journal of Neural Engineering**, vol. 18, 10 2021. pages 10
- [16] S. Aliakbaryhosseinabadi, E. N. Kamavuako, N. Jiang, D. Farina, and N. Mrachacz-Kersting, “Classification of eeg signals to identify variations in attention during motor task execution,” **Journal of Neuroscience Methods**, vol. 284, pp. 27–34, 6 2017. pages 10, 12, 15

- [17] F. Yao, J. Coquery, and K. Lê Cao, "Independent principal component analysis for biologically meaningful dimension reduction of large biological data sets," **BMC Bioinformatics**, vol. 13, 2012. pages 10
- [18] A. Gaume, G. Dreyfus, and F. B. Vialatte, "A cognitive brain-computer interface monitoring sustained attentional variations during a continuous task," **Cognitive Neurodynamics**, vol. 13, pp. 257–269, 6 2019. pages 11, 12, 16
- [19] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," **J Neural Eng**, vol. 4, no. 2, pp. R1–R13, Jan. 2007. pages 11
- [20] N. P. Subramaniam, "Lab talk," Feb 2020. [Online]. Available: <https://sapienlabs.org/lab-talk/measuring-entropy-in-the-eeeg/> pages 11
- [21] W. Gao, J. A. Guan, J. Gao, and D. Zhou, "Multi-ganglion ann based feature learning with application to p300-bci signal classification," **Biomedical Signal Processing and Control**, vol. 18, pp. 127–137, 2015. pages 11
- [22] J. J. Bird, D. R. Faria, L. J. Manso, A. Ekárt, and C. D. Buckingham, "A deep evolutionary approach to bioinspired classifier optimisation for brain-machine interaction," **Complexity**, vol. 2019, 2019. pages 11
- [23] J. J. V. Merriënboer and J. Sweller, "Cognitive load theory and complex learning: Recent developments and future directions," **Educational Psychology Review**, vol. 17, pp. 147–177, 2005. pages 11
- [24] P. A. Kirschner, "Cognitive load theory: implications of cognitive load theory on the design of learning," **Learning and Instruction**, vol. 12, no. 1, pp. 1–10, 2002. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0959475201000147> pages 11
- [25] F. Paas, A. Renkl, and J. Sweller, "Cognitive load theory and instructional design: Recent developments," **Educational Psychologist**, vol. 38, no. 1, pp. 1–4, 2003. [Online]. Available: https://doi.org/10.1207/S15326985EP3801_1 pages 11
- [26] R. N. Roy, S. Bonnet, S. Charbonnier, and A. Campagne, "Mental fatigue and working memory load estimation: Interaction and implications for eeg-based passive bci," 2013, pp. 6607–6610. pages 12
- [27] A. Myrden and T. Chau, "A passive eeg-bci for single-trial detection of changes in mental state," **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, vol. 25, pp. 345–356, 4 2017. pages 12, 16, 30
- [28] P. Diez, A. Garces, L. Orosco, E. Laciari, and V. Mut, "Attention-level transitory response: A novel hybrid bci approach," **Journal of neural engineering**, vol. 12, p. 056007, 08 2015. pages 12
- [29] N.-H. Liu, C.-Y. Chiang, and H.-C. Chu, "Recognizing the degree of human attention using eeg signals from mobile sensors," **Sensors (Basel, Switzerland)**, vol. 13, pp. 10 273–86, 08 2013. pages 12
- [30] B. Hamadicharefet and al., "Learning eeg-based spectral-spatial patterns for attention level measurement," in **Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)**, May 2009, pp. 1465–1468. pages 12
- [31] A. List, M. D. Rosenberg, A. Sherman, and M. Esterman, "Pattern classification of eeg signals reveals perceptual and attentional states," **PLoS ONE**, vol. 12, 4 2017. pages 12
- [32] K. Mizuno, M. Tanaka, K. Yamaguti, O. Kajimoto, H. Kuratsune, and Y. Watanabe, "Mental fatigue caused by prolonged cognitive load associated with sympathetic hyperactivity," **Behavioral and Brain Functions**, vol. 7, no. 1, p. 17, May 2011. pages 12
- [33] D. Gaba and S. Howard, "Gaba dm, howard sk: Fatigue among clinicians and the safety of patients. n engl j med 347: 1249-55," **The New England journal of medicine**, vol. 347, pp. 1249–55, 11 2002. pages 12
- [34] B. Jap, S. Lal, P. Fischer, and E. Bekiaris, "Using eeg spectral components to assess algorithms for detecting fatigue [part 1]," **Expert Systems with Applications**, vol. 36, pp. 2352–2359, 03 2009. pages 12

- [35] S. K. L. Lal and A. Craig, "Reproducibility of the spectral components of the electroencephalogram during driver fatigue," *Int J Psychophysiol*, vol. 55, no. 2, pp. 137–143, Feb. 2005. pages 12
- [36] O. E. Krigolson, M. R. Hammerstrom, W. Abimbola, R. Trska, B. W. Wright, K. G. Hecker, and G. Binsted, "Using muse: Rapid mobile assessment of brain performance," *Frontiers in Neuroscience*, vol. 15, 1 2021. pages 12, 13, 15, 20
- [37] I. Käthner, S. C. Wriessnegger, G. R. Müller-Putz, A. Kübler, and S. Halder, "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*, vol. 102, pp. 118–129, 2014. pages 13, 16, 42
- [38] G. Celesia and G. Hickok, "The auditory cortex: A synthesis of human and animal research," 2015. pages 13
- [39] H. Kim, J.-H. Jo, J.-H. Lee, J.-j. Jung, K.-H. Kim, J.-S. An, Y.-J. Kim, M.-H. Choi, and S.-C. Chung, "Comparison of neural activation area in primary somatosensory cortex and brodmann area 3 according to finger and phalange high-frequency vibration stimulation," *Applied Sciences*, vol. 10, no. 17, p. 6142, 2020. pages 13, 40
- [40] K. Sathian, "Analysis of haptic information in the cerebral cortex," *Journal of Neurophysiology*, vol. 116, pp. 1795–1806, 10 2016. pages 13
- [41] Z. Britton and Q. Arshad, "Vestibular and multi-sensory influences upon self-motion perception and the consequences for human behavior," *Frontiers in Neurology*, vol. 10, p. 63, 2019. pages 13
- [42] B. staff, "Medical gallery of blausen medical 2014," *WikiJournal of Medicine*, vol. 1, no. 2, 2014, own work. pages 14
- [43] N. Jafari, K. D. Adams, and M. Tavakoli, "Haptics to improve task performance in people with disabilities: A review of previous studies and a guide to future research with children with disabilities," *Journal of Rehabilitation and Assistive Technologies Engineering*, vol. 3, p. 205566831666814, 1 2016. pages 13
- [44] N. Thomas, G. Ung, H. Ayaz, and J. D. Brown, "Neurophysiological evaluation of haptic feedback for myoelectric prostheses," *IEEE Transactions on Human-Machine Systems*, vol. 51, pp. 253–264, 6 2021. pages 14
- [45] C. C. Duncan-Johnson, "Young psychophysicologist award address, 1980: P300 latency: a new metric of information processing," *Psychophysiology*, vol. 18, pp. 207–215, 1981. pages 14
- [46] A. Magliero, T. R. Bashore, M. G. Coles, and E. Donchin, "On the dependence of p300 latency on stimulus evaluation processes," *Psychophysiology*, vol. 21, pp. 171–186, 1984. pages 14
- [47] J. Polich, "P300 development from auditory stimuli," *Psychophysiology*, vol. 23, pp. 590–597, 1986. pages 14
- [48] M. Fleury, G. Lioi, C. Barillot, and A. Lécuyer, "A survey on the use of haptic feedback for brain-computer interfaces and neurofeedback," *Frontiers in Neuroscience*, vol. 14, 6 2020. pages 14
- [49] J. Burke, M. Prewett, A. Gray, L. Yang, F. Stilson, M. Covert, L. Elliott, and E. Redden, "Comparing the effects of visual-auditory and visual-tactile feedback on user performance: A meta-analysis," *ICMI'06: 8th International Conference on Multimodal Interfaces, Conference Proceeding*, pp. 108–117, 11 2006. pages 14
- [50] C. Wickens, "Multiple resources and performance prediction," *Theoretical Issues in Ergonomics Science*, vol. 3, no. 2, pp. 159–177, 2002. pages 14
- [51] A. A. Stanley and K. J. Kuchenbecker, "Evaluation of tactile feedback methods for wrist rotation guidance," *IEEE Transactions on Haptics*, vol. 5, pp. 240–251, 2012. pages 14
- [52] E. C. McWilliams, F. M. Barbey, J. F. Dyer, M. N. Islam, B. McGuinness, B. Murphy, H. Nolan, P. Passmore, L. M. Rueda-Delgado, and A. R. Buick, "Feasibility of repeated assessment of cognitive function in older adults using a wireless, mobile, dry-ecg headset and tablet-based games," *Frontiers in Psychiatry*, vol. 12, 6 2021. pages 15

- [53] J. Sabio, N. S. Williams, G. M. McArthur, and N. A. Badcock, "A scoping review on the use of consumer-grade eeg devices for research," [bioRxiv](https://www.biorxiv.org/content/early/2022/12/05/2022.12.04.519056), 2022. [Online]. Available: <https://www.biorxiv.org/content/early/2022/12/05/2022.12.04.519056> pages 15
- [54] O. E. Krigolson, C. C. Williams, A. Norton, C. D. Hassall, and F. L. Colino, "Choosing muse: Validation of a low-cost, portable eeg system for erp research," [Frontiers in Neuroscience](https://doi.org/10.3389/fnins.2017.00015), vol. 11, 3 2017. pages 15, 20
- [55] S. Glen, "Anova test: Definition, types, examples, spss," Apr 2022. [Online]. Available: <https://www.statisticshowto.com/probability-and-statistics/hypothesis-testing/anova/> pages 16
- [56] InteraXon, "Muse 2: The Brain Sensing Headband," <https://choosemuse.com/product/muse-2/>, [Online; accessed Jan 28, 2023]. pages 20
- [57] K. Tsiakas, "Interactive learning and adaptation for personalized robot-assisted training designing a socially assistive robot for adaptive and personalized training view project reinforcement learning for interactive agents view project," 2018. [Online]. Available: <https://www.researchgate.net/publication/330761412> pages 21
- [58] A. J. Spiers and A. M. Dollar, "Outdoor pedestrian navigation assistance with a shape-changing haptic interface and comparison with a vibrotactile device," vol. 2016-April. IEEE Computer Society, 4 2016, pp. 34–40. pages 22, 47
- [59] J. Kowaleski and S. Wicklund, "BlueMuse: The Brain-Computer Interface App," , [Online; accessed Jan 28, 2023, version 2.0.0.0]. pages 23
- [60] A. Barachant, "Muse LSL Library," <https://github.com/alexandrebarachant/muse-lsl>, [Online; accessed Jan 28, 2023, version 2.0.2]. pages 23, 24, 25
- [61] J. Griffiths, J. Tredup, NeuroTechX, and Contributors, "EEG-notebooks," <https://github.com/EEG-notebooks/>, [Online; accessed Jan 28, 2023, version 0.2.0]. pages 23, 26
- [62] Peirce, J.W., "PsychoPy: An open-source software to create psychology stimuli in Python," <https://doi.org/10.5281/zenodo.3345931>, 2019, accessed: June 2023. pages 23
- [63] Gramfort, A. and Luessi, M. and Larson, E. and Engemann, D. and Strohmeier, D. and Brodbeck, C. and Parkkonen, L. and Hämäläinen, M., "MNE Software Suite," <https://mne.tools>, 2021, accessed: June 2023. pages 23
- [64] Q. Rezaee, M. Delrobaei, A. Giveki, N. Dayarian, and S. J. Haghighi, "Driver drowsiness detection with commercial eeg headsets," in [2022 10th RSI International Conference on Robotics and Mechatronics \(ICRoM\)](https://doi.org/10.1109/ICRoM49421.2022), 2022, pp. 546–550. pages 28
- [65] S. Kiebel, J. Kilner, and K. Friston, "Chapter 16 - hierarchical models for eeg and meg," in [Statistical Parametric Mapping](https://doi.org/10.1017/S0269972707003333), K. FRISTON, J. ASHBURNER, S. KIEBEL, T. NICHOLS, and W. PENNY, Eds. London: Academic Press, 2007, pp. 211–220. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780123725608500164> pages 33
- [66] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, "Permutation entropy and its main biomedical and econophysics applications: A review," [Entropy](https://doi.org/10.3390/entropy14081553), vol. 14, no. 8, pp. 1553–1577, 2012. [Online]. Available: <https://www.mdpi.com/1099-4300/14/8/1553> pages 36
- [67] Y. Cao, W.-w. Tung, J. B. Gao, V. A. Protopopescu, and L. M. Hively, "Detecting dynamical changes in time series using the permutation entropy," [Phys. Rev. E](https://doi.org/10.1103/PhysRevE.70.046217), vol. 70, p. 046217, Oct 2004. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevE.70.046217> pages 36
- [68] A. Chaudhuri, A. Routray, and S. Kar, "Effect of sleep deprivation on estimated distributed sources for scalp eeg signals: A case study on human drivers," in [2012 4th International Conference on Intelligent Human Computer Interaction \(IHCI\)](https://doi.org/10.1109/IHCI49421.2012), 2012, pp. 1–6. pages 36
- [69] C. E. Shannon, "A mathematical theory of communication," [The Bell System Technical Journal](https://doi.org/10.1002/j.1539-3099.1948.tb04376.x), vol. 27, no. 3, pp. 379–423, 623–666, July 1948. pages 36

- [70] C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series," **Phys. Rev. Lett.**, vol. 88, p. 174102, Apr 2002. [Online]. Available: <https://link.aps.org/doi/10.1103/PhysRevLett.88.174102> pages 36
- [71] S. Berger, G. Schneider, E. F. Kochs, and D. Jordan, "Permutation entropy: Too complex a measure for eeg time series?" **Entropy**, vol. 19, no. 12, 2017. [Online]. Available: <https://www.mdpi.com/1099-4300/19/12/692> pages 36
- [72] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "Iclabel: An automated electroencephalographic independent component classifier, dataset, and website," **NeuroImage**, vol. 198, pp. 181–197, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1053811919304185> pages 37
- [73] M.-Y. Kim, H. Kwon, T.-H. Yang, and K. Kim, "Vibration alert to the brain: Evoked and induced meg responses to high-frequency vibrotactile stimuli on the index finger of dominant and non-dominant hand," **Frontiers in Human Neuroscience**, vol. 14, p. 576082, 2020. pages 40
- [74] T. S. Grummett, S. P. Fitzgibbon, T. W. Lewis, D. DeLosAngeles, E. M. Whitham, K. J. Pope, and J. O. Willoughby, "Constitutive spectral eeg peaks in the gamma range: suppressed by sleep, reduced by mental activity and resistant to sensory stimulation," **Frontiers in Human Neuroscience**, vol. 8, p. 927, 2014. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4240063/> pages 40
- [75] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni, "Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness," **Neuroscience & Biobehavioral Reviews**, vol. 44, pp. 58–75, 2014. pages 40
- [76] W. Klimesch, "-band oscillations, attention, and controlled access to stored information," **Trends in Cognitive Sciences**, vol. 16, no. 12, pp. 606–617, 2012. pages 40
- [77] A. S. Keller, L. Payne, and R. Sekuler, "Characterizing the roles of alpha and theta oscillations in multisensory attention," **Neuropsychologia**, vol. 99, pp. 48–63, 2017. pages 40

A Appendix

A.1 Wavelets Analysis Results

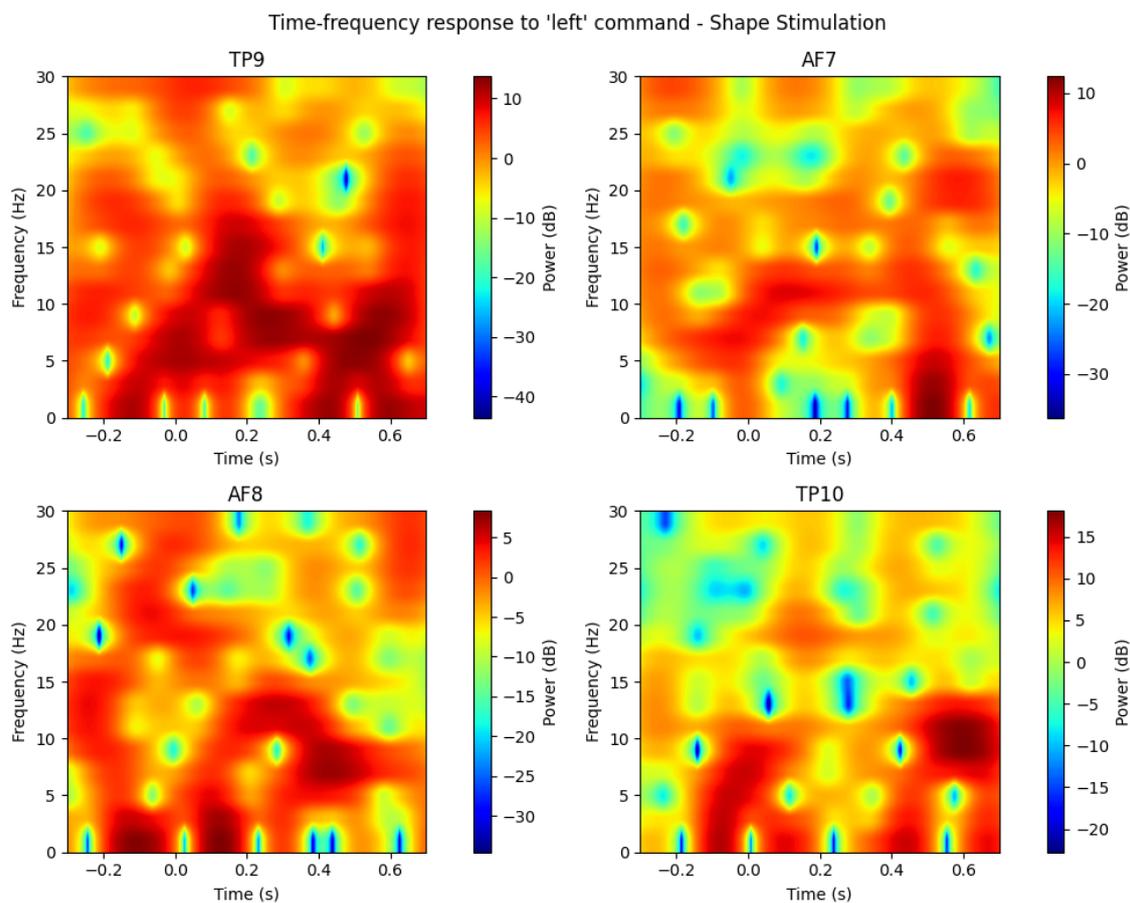


Figure 35: Morlet Representation of Time-Frequency response to 'left' command for Shape stimulation type; event occurrence at time 0; average across all trials and participants.

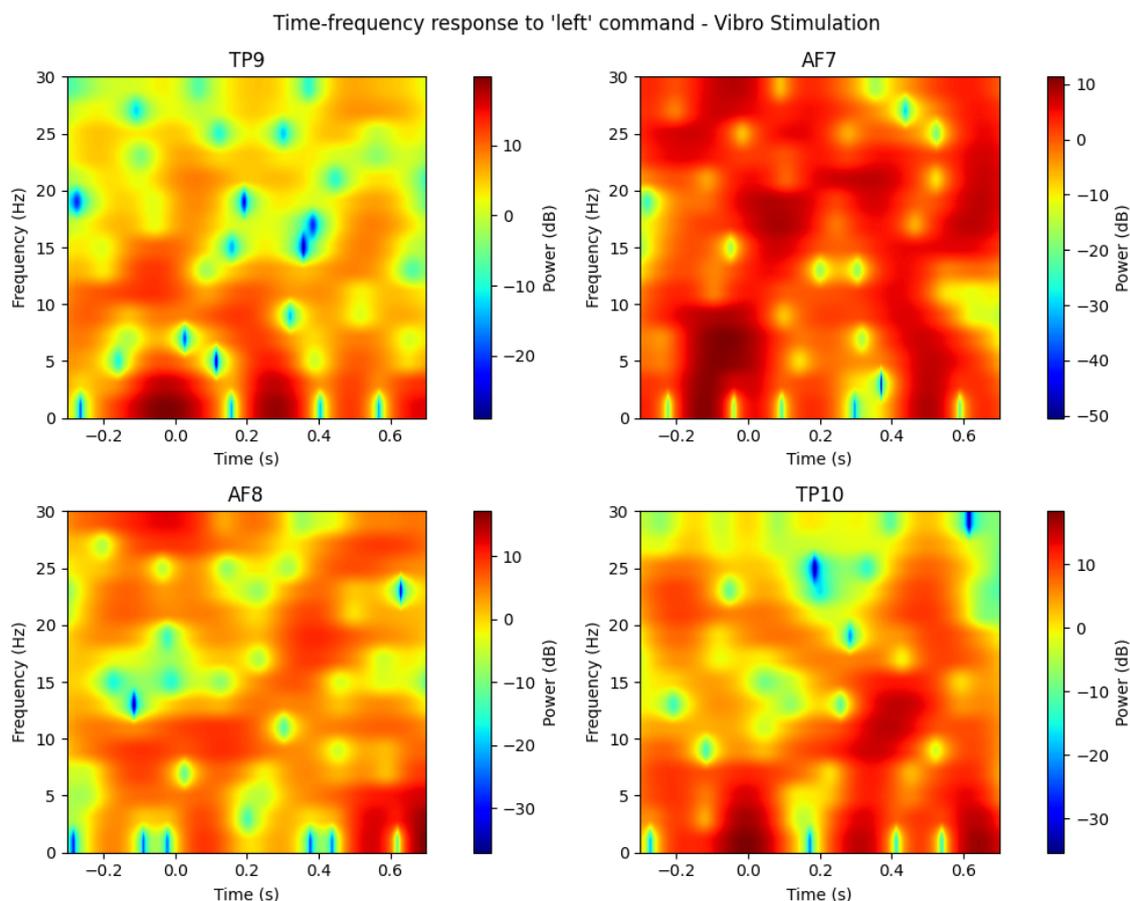


Figure 36: Morlet Representation of Time-Frequency response to 'left' command for Vibro stimulation type; event occurrence at time 0; average across all trials and participants.

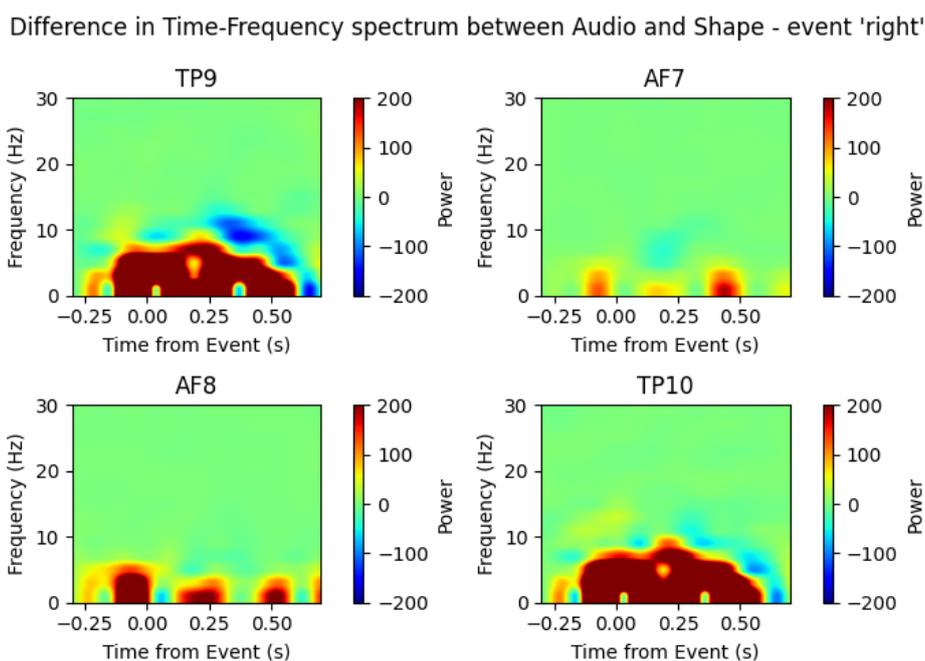


Figure 37: Morlet Representation of the difference in Time-Frequency response between Audio and Shape stimulation type; response to 'right' command, with event occurrence at time 0; average across all trials and participants

Difference in Time-Frequency spectrum between Audio and Vibro - event 'right'

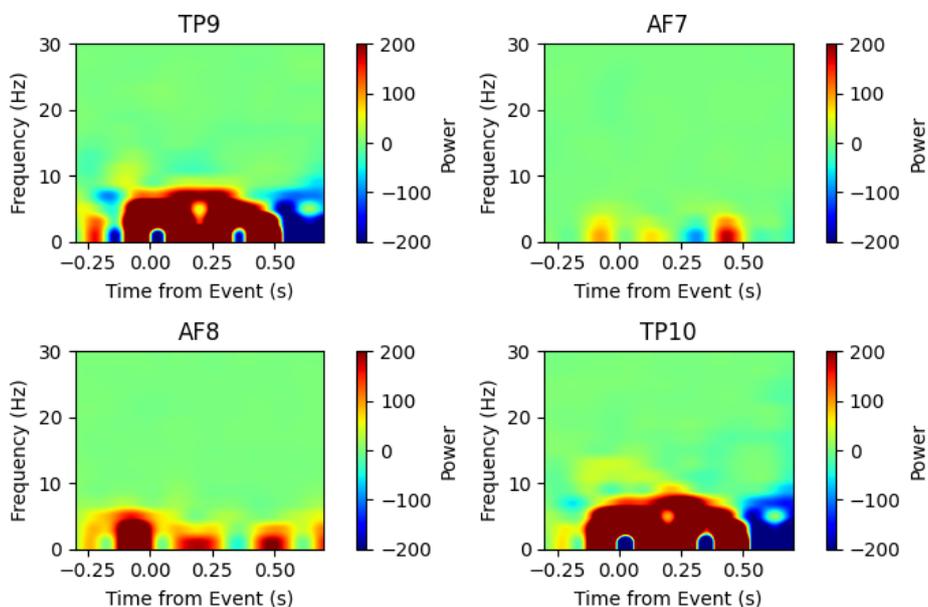


Figure 38: Morlet Representation of the difference in Time-Frequency response between Audio and Vibro stimulation type; response to 'right' command, with event occurrence at time 0; average across all trials and participants

Difference in Time-Frequency spectrum between Vibro and Shape - event 'right'

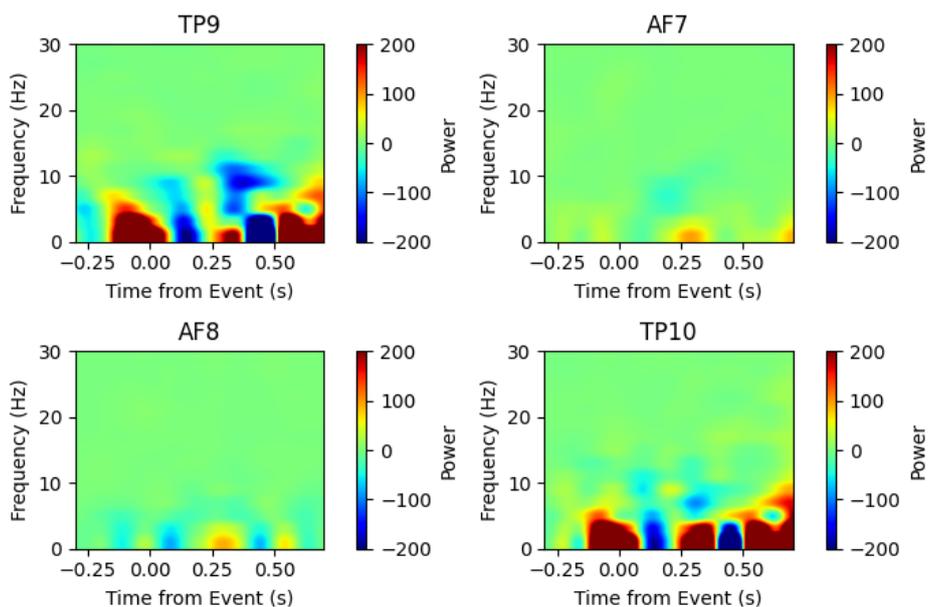


Figure 39: Morlet Representation of the difference in Time-Frequency response between Vibro and Shape stimulation type; response to 'right' command, with event occurrence at time 0; average across all trials and participants

A.2 ERP Analysis - RMS Results

Event	Channel	Audio	Vibro	Shape
Left	TP9	2.7599	3.0712	2.0638
	AF7	1.1562	1.1052	0.8844
	AF8	1.1673	0.9914	1.0674
	TP10	2.7399	2.7835	1.5581
Right	TP9	2.6627	2.1004	1.8765
	AF7	1.0129	1.2710	0.9425
	AF8	1.0636	1.1130	1.1873
	TP10	2.6993	2.1620	1.5717

Table 2: ERP Root Mean Square of the difference between the mean amplitude of a baseline time window (0.2-0s before the event) and of a response-timed one (0.2-0.4s after the event) for three conditions and event type.

A.3 ICA Results

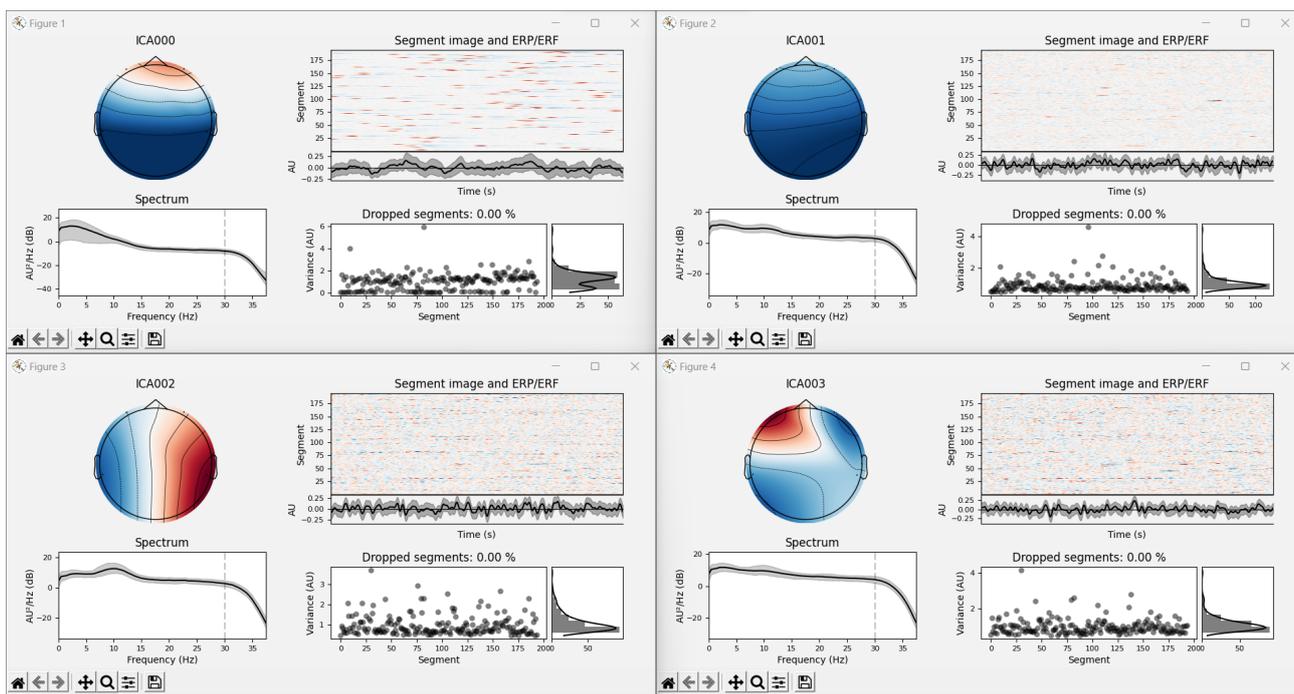


Figure 40: ICA diagnostics for the 4 principal components of EEG recording from Audio Stimulation

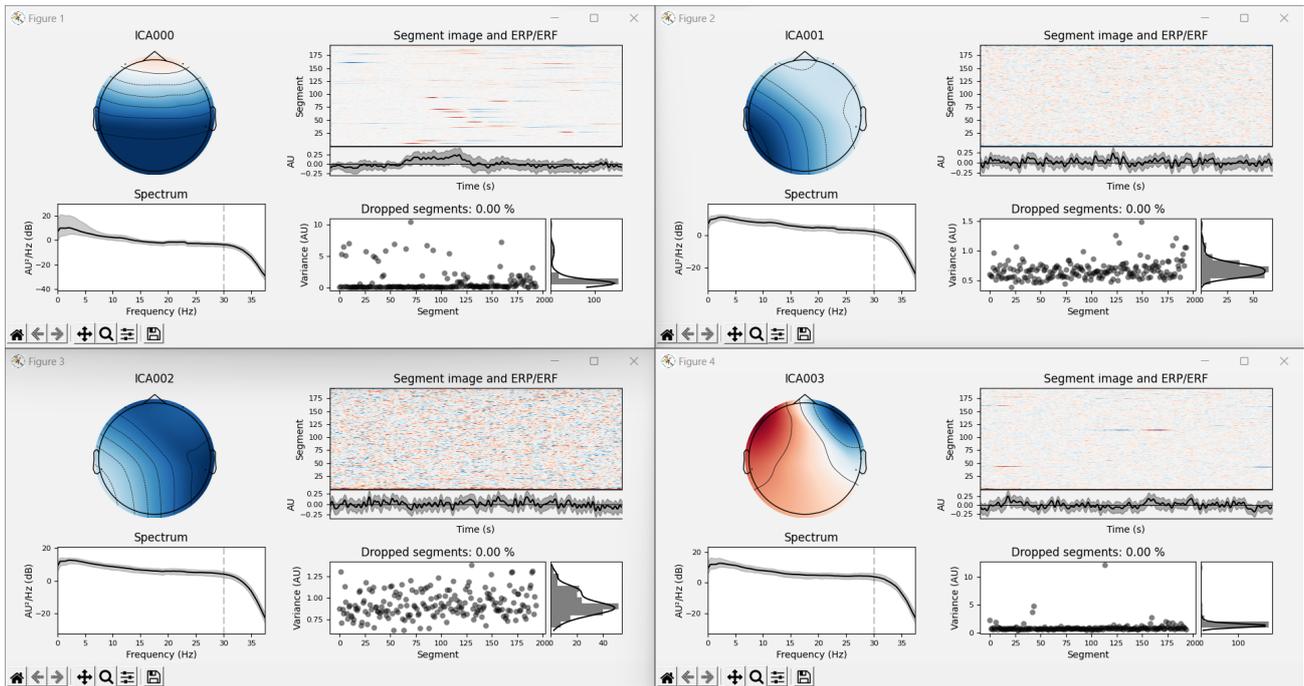


Figure 41: ICA diagnostics for the 4 principal components of EEG recording from Vibro Stimulation

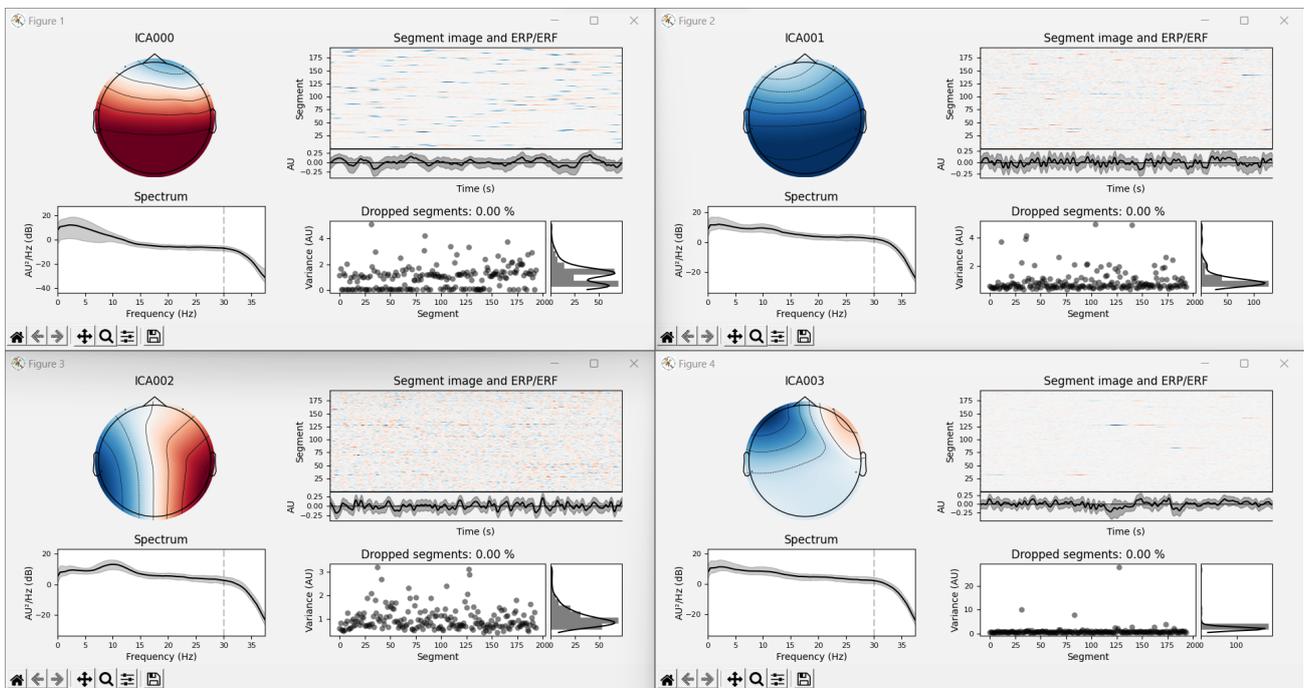


Figure 42: ICA diagnostics for the 4 principal components of EEG recording from Shape Stimulation