Chair for Computer Science 10 (Media Computing and Human-Computer Interaction)

RWTH AACHEN UNIVERSITY

Design and Implementation of an Interactive Interface for EEG Data Analysis

Bachelor's Thesis at the
Media Computing Group
Prof. Dr. Jan Borchers
Computer Science Department
RWTH Aachen University

by
Vanessa Julia Bangert

Thesis advisor:
Prof. Dr. Jan Borchers

Second examiner:
Prof. Dr. Ulrik Schroeder

Registration date: 15.10.2025
Submission date: 16.02.2026

Zentrales Prüfungsamt/Central Examination Office

Eidesstattliche Versicherung
Declaration of Academic Integrity

Bangert, Vanessa Julia

Name, Vorname/Last Name, First Name

433824

Matrikelnummer (freiwillige Angabe)
Student ID Number (optional)

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Arbeit/Bachelorarbeit/ Masterarbeit* mit dem Titel
I hereby declare under penalty of perjury that I have completed the present paper/bachelor's thesis/master's thesis* entitled

Design and Implementation of Interactive Interface for EEG Data Analysis

selbstständig und ohne unzulässige fremde Hilfe (insbes. akademisches Ghostwriting) erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt; dies umfasst insbesondere auch Software und Dienste zur Sprach-, Text- und Medienproduktion. Ich erkläre, dass für den Fall, dass die Arbeit in unterschiedlichen Formen eingereicht wird (z.B. elektronisch, gedruckt, geplottet, auf einem Datenträger) alle eingereichten Versionen vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.
independently and without unauthorized assistance from third parties (in particular academic ghostwriting. I have not used any other sources or aids than those indicated; this includes in particular software and services for language, text, and media production. In the event that the work is submitted in different formats (e.g. electronically, printed, plotted, on a data carrier), I declare that all the submitted versions are fully identical. I have not previously submitted this work, either in the same or a similar form to an examination body.

Aachen, 16.02.2026

Ort, Datum/City, Date

V. Bangert

Unterschrift/Signature

*Nichtzutreffendes bitte streichen/Please delete as appropriate

Belehrung:
Official Notification:

§ 156 StGB: Falsche Versicherung an Eides Statt
Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.
§ 156 StGB (German Criminal Code): False Unsworn Declarations
Whosoever before a public authority competent to administer unsworn declarations (including Declarations of Academic Integrity) falsely submits such a declaration or falsely testifies while referring to such a declaration shall be liable to imprisonment for a term not exceeding three years or to a fine.
§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt
(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.
(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.
§ 161 StGB (German Criminal Code): False Unsworn Declarations Due to Negligence
(1) If an individual commits one of the offenses listed in §§ 154 to 156 due to negligence, they are liable to imprisonment for a term not exceeding one year or to a fine.
(2) The offender shall be exempt from liability if they correct their false testimony in time. The provisions of § 158 (2) and (3) shall apply accordingly.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:
I have read and understood the above official notification:

Aachen, 16.02.2026

Ort, Datum/City, Date

V. Bangert

Unterschrift/Signature

Contents

List of Figures and Tables

Abstract

HCI research often uses questionnaires and observations to study user experience. These methods give useful results, but they do not show users' mental states in a direct and continuous way. Electroencephalography (EEG) can provide such information, but it is difficult to use in practice. EEG interfaces are designed for experts and are hard to use. They require deep knowledge of signal processing and offer little support for typical HCI workflows. This makes it difficult for HCI researchers with limited EEG experience to use EEG in user studies.

This thesis presents the design and implementation of an interactive interface for EEG data analysis. The interface is designed for HCI researchers with little prior EEG knowledge. It focuses on easy setup, transparent preprocessing, and clear visualizations. The interface was developed using a user-centered design process. Interviews with HCI researchers were conducted to understand their needs and challenges. Based on these insights, design principles and interface requirements were defined. The final interface lowers the entry barrier for EEG-based user studies. It supports HCI researchers in recording, analyzing, and understanding EEG data in a clear and accessible way.

Überblick

In der HCI-Forschung werden häufig Fragebögen und Beobachtungen eingesetzt, um die Interaktionen zwischen Mensch und Computer zu untersuchen. Diese Methoden liefern zwar nützliche Ergebnisse, zeigen jedoch nicht direkt und kontinuierlich die mentalen Zustände der Benutzer. Die Elektroenzephalographie (EEG) kann solche Informationen liefern, ist jedoch in der Praxis schwer anzuwenden. EEG-Software ist für Experten konzipiert und schwer zu bedienen. Sie erfordern fundierte Kenntnisse der Signalverarbeitung und bieten wenig Unterstützung für typische HCI-Studienabläufe. Dies erschwert es HCI-Forschern mit begrenzter EEG-Erfahrung, EEG in Benutzerstudien zu integrieren.

Diese Arbeit präsentiert den Entwurf und die Implementierung eines interaktiven Interfaces für die EEG-Datenanalyse. Das Interface ist für HCI-Forscher mit geringen Vorkenntnissen im Bereich EEG konzipiert. Der Schwerpunkt liegt auf einer einfachen Nutzung, transparenten Verarbeitung und klaren Visualisierungen. Die Interface wurde unter Verwendung eines nutzerzentrierten Designprozesses entwickelt. Es wurden Interviews mit HCI-Forschern durchgeführt, um ihre Bedürfnisse und Herausforderungen zu verstehen. Basierend auf diesen Erkenntnissen wurden Designprinzipien und Interfaceanforderungen definiert. Das finale Interface senkt die Einstiegshürde für EEG-basierte Nutzerstudien. Sie unterstützt HCI-Forscher dabei, EEG-Daten auf klare und zugängliche Weise aufzuzeichnen, zu analysieren und zu verstehen.

Acknowledgments

First, I would like to thank Prof. Dr. Jan Borchers for examining my thesis and Prof. Dr. Ulrik Schroeder for being my second examiner.

Many thanks to Kristina Hörrmann for supervising my work. Thank you for your helpful feedback, for answering even my smallest questions, and for exploring the research field together with me. I am very grateful for the opportunity to work on something so visual and creative, which I enjoyed a lot. I hope that this research field will continue to grow and develop within the research group.

I would also like to thank everyone who took part in my user study and shared their thoughts and experiences with me.

Finally, I would like to thank my family and friends for their continuous support and motivation during my thesis.

Conventions

Throughout this thesis we use the following conventions:

- The thesis is written in American English.
- The first person is written in plural form.
- Unidentified third persons are described in female form.

Where appropriate, paragraphs are summarized by one or two sentences that are positioned at the margin of the page.

This is a summary of a paragraph.

Chapter 1

Introduction

Human–Computer Interaction (HCI) research is largely based on empirical user studies that aim to understand, evaluate, and improve interactive systems [Lazar et al., 2017]. These studies typically rely on self-report methods, such as questionnaires or think-aloud protocols, as well as behavioral measures observed during interaction, including reaction times or error rates, to assess usability, user experience, and satisfaction. Although these methods are well established, they often provide only indirect or delayed insights into users' cognitive and emotional states and are prone to bias [Frey et al., 2016].

HCI is often based on self-report methods with indirect or delayed insights.

As interactive systems become more adaptive and cognitively demanding, HCI researchers increasingly seek methods that allow for a more continuous and objective assessment of users' mental states during interaction [Cherng et al., 2016; Frey et al., 2013; Kumar and Kumar, 2016]. In this context, physiological sensing and in particular brain sensing techniques have gained relevance for evaluating user interfaces and experiences, where neural data can serve as an additional source of information [Aggarwal et al., 2014; Frey et al., 2016; Putze et al., 2022]. Within neuroergonomics, such approaches are, for example, applied to study how task difficulty, mental effort, and fatigue are reflected in brain activity [Müller-Putz et al., 2015].

Brain sensing techniques gain relevance.

Among brain sensing techniques, electroencephalography

EEG is used as an evaluation tool in HCI research.

(EEG) is the most widely used method in HCI research. A large-scale review of HCI publications reports EEG usage in 78% of studies involving neural data, which makes it the dominant technique in the field [Putze et al., 2022]. EEG is increasingly used not as a direct control mechanism as in classical brain computer interface (BCI) research, but as an evaluation tool to assess cognitive workload, attention, or responses during interaction [Cherng et al., 2016]. This reflects a shift from classical BCI research toward using EEG as a supportive evaluation tool in HCI [Putze et al., 2022].

EEG integration remains challenging.

Despite this growing interest, integrating EEG into HCI research remains challenging. It requires interdisciplinary expertise in neuroscience, signal processing, and experimental methodology [Putze et al., 2022].

Current EEG interfaces lack usability and workflow support for HCI researchers.

Many existing EEG recording and analysis interfaces are primarily designed for experts and do not prioritize usability, transparency, or integration into typical HCI workflows. Consumer-grade EEG devices only partially lower the entry barrier. Although these devices enable EEG studies to be conducted outside controlled laboratory settings and reduce cost and setup time, they do not address the challenges of data preprocessing, analysis, and interpretation. In particular, accessible and transparent interfaces that support HCI researchers within a user study with limited prior EEG knowledge remain largely absent. As a result, HCI researchers face difficulties when attempting to integrate EEG into user studies, despite the increasing availability of consumer-grade devices.

We designed a low-entry EEG recording and analysis interface

This thesis addresses this gap by designing and implementing a low-entry EEG recording and analysis interface tailored to the needs of HCI researchers with limited prior EEG experience. The proposed interface supports EEG-based user studies by emphasizing easy setup, transparent preprocessing, visualizations, and a clear workflow.

The focus lies on enabling HCI researchers to easily acquire, preprocess, and visualize EEG data from consumer-grade devices, compare conditions or participants within user studies, and understand the underlying processing steps in a transparent and reproducible way.

The research follows a user-centered design process informed by qualitative interviews with HCI researchers. The interviews explored current research practices, opportunities, and challenges of integrating EEG into user studies and expectations regarding an interface, data visualization, and analysis. The results informed the definition of design requirements and guided the development of the interface.

This thesis is structured as follows: Chapter 2 introduces the theoretical background of EEG relevant to this work. This includes preprocessing and feature extraction methods. Chapter 3 reviews related work, with a focus on the use of EEG in HCI research and existing EEG recording and analysis interfaces. Chapter 4 presents qualitative interviews with HCI researchers, which inform the design requirements for the implemented interface. Chapter 5 describes the design and implementation of the EEG interface and provides a detailed walkthrough of its core functionalities. Chapter 6 proposes an evaluation study planned for future work. Chapter 7 concludes the thesis and outlines directions for future work.

Chapter 2

Theoretical Background

This chapter provides the theoretical foundation for this work. Scetion 2.1 first introduces the physiological bases of EEH, the spacial characteristics, different EEG devices and the device used in this work. Section 2.2 describes different EEG analysis perspectives. Section 2.3 then presents the EEG data processing pipeline. Finally, the interpretation of EEG frequency bands and ratio-based metrics is discussed in Section 2.4.

2.1 EEG Data Collection

Brain sensing devices can be divided into invasive and non-invasive.

Following the classification by Tan and Nijholt [2010], methods for measuring brain activity can be divided into invasive and non-invasive methods. Invasive methods involve sensors that are implanted directly on or into the brain. They provide high temporal and spatial resolution, but they are impractical for HCI research because they require surgical implantation, involve medical risks, offer limited spatial coverage, and allow little flexibility in electrode placement. Non-invasive technologies use external sensors to measure brain activity.

FNIRS and EEG are main methods in HCI.

As argued by Tan and Nijholt [2010], electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) are the only methods that meet the requirements of cost, portability, and safety for HCI-focused research. FNIRS measures brain activity by detecting changes caused by variations in blood oxygenation and blood volume. It has a high spatial resolution but low temporal resolution. This work focuses on EEG, as it is the most widely used brain-sensing method in HCI research [Putze et al., 2022; Tan and Nijholt, 2010]

2.1.1 Physiological Basis of EEG

EEG measures summed cortical electrical activity at the scalp.

EEG uses electrodes to measure weak electrical potentials (5–100 µV) generated by synchronous cortical neuronal activity, at the scalp [Tan and Nijholt, 2010]. To understand the origin of these signals, it is necessary to review how neural activity generates electrical potentials.

Basic principles of neuronal signal transmission.

Neurons are the basic cells of the nervous system. They send and receive signals using electrical and chemical processes. A neuron consists of dendrites, a cell body called soma, and an axon. Dendrites receive signals from other neurons. These signals are combined in the cell body. If the signal is strong enough, the neuron produces an action potential, which is a short electrical signal. The action potential travels along the axon and causes the release of neurotransmitters at connections called synapses. When the neurotransmitter binds to receptors on the dendrites of the next neuron, they create postsynaptic potentials. Postsynaptic potentials are temporary changes in the electrical state of the receiving neuron [Rao, 2013].

EEG reflects synchronous postsynaptic activity of many neurons.

Action potentials are essential for neural communication, but they are too brief and localized to substantially contribute to scalp EEG. In contrast, postsynaptic currents last longer and occur synchronously across many neurons. Only when many neurons are temporally aligned and similarly oriented, their electrical fields sum and become detectable at the scalp [Jackson and Bolger, 2014]. Hence, EEG signals reflect the summation of postsynaptic potentials from many thousands of neurons [Rao, 2013].

2.1.2 Spatial Characteristics of EEG

EEG has a low spatial resolution because the signals recorded on the scalp come from a mixture of activity in many different brain areas. Because electrical signals spread through the brain and skull and become weaker with distance, EEG mainly detects activity from the cerebral cortex (the outer layer of the brain). Activity from deeper brain structures has only a small influence on the recorded signal [Holsheimer and Feenstra, 1977; Rao, 2013].

EEG has low spatial resolution.

As described by [Kumar and Bhuvaneswari, 2012], the brain is divided into the left and right hemispheres, each subdivided into four main lobes. The individual lobes are statistically linked to distinct cognitive functions.

EEG signals are interpreted using broad brain regions.

The frontal lobe, located behind the forehead, is primarily involved in higher cognitive functions such as planning, decision making, problem solving, attention, impulse control, and the regulation of behavior and emotions [Kumar and Bhuvaneswari, 2012]. The prefrontal cortex, in particular, plays an important role in research on cognitive load and engagement [Sarno et al., 2016; Sawangjai et al., 2019]. The temporal lobe, located on the sides of the brain, is involved in auditory processing, language comprehension, and memory functions [Kumar and Bhuvaneswari, 2012]. The parietal lobe, behind the frontal lobe, integrates sensory information from different parts of the body and is relevant for spatial awareness and body perception [Kumar and Bhuvaneswari, 2012]. The occipital lobe, located at the back of the head, is primarily responsible for visual processing and perception [Kumar and Bhuvaneswari, 2012].

This shows that the number and placement of electrodes define what can be measured. Which leads to a distinction between high-density and low-density EEG devices, which differ in spatial coverage, usability, and typical use cases.

Electrode placement determines spatial coverage

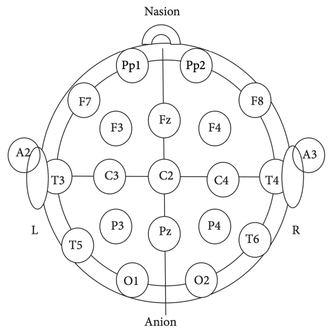


Figure 2.1: International 10–20 system for electrode placement which shows the standard electrode locations on the scalp [Al-Fahoum and Al-Fraihat, 2014].

2.1.3 High-Density, Low-Density, and Consumer EEG Devices.

Standard EEG electrode placement follows the 10–20 system.

The EEG electrodes are typically placed using the international 10–20 system, as can be seen in Fig. 2.1. It defines standard electrode positions based on 10% and 20% distances of the head size. Electrodes are labeled by letters indicating the brain region as in Section2.1.2 and by numbers or the letter z. Odd numbers refer to the left hemisphere, even numbers to the right, and z indicates midline positions [Müller-Putz et al.].

High-density EEG uses many electrodes for high spatial resolution.

High-density (HD) EEG devices use 64 to 256 electrodes. The high number of electrodes enables higher spatial resolution and enhances the localization of brain activity. However, the use of HD–EEG is limited by practical constraints like higher equipment costs, higher need for data processing, and longer setup and recording times. As a result, HD–EEG is primarily applied in specialized settings where high spatial precision is required [Stoyell et al., 2021].

Low-density EEG trades spatial resolution for usability and cost.

Low-density (LD) EEG devices use way fewer electrodes (typically 19 to 25) [Stoyell et al., 2021]. This limits spatial coverage but makes these devices more affordable and easier to use. Bach Justesen et al. [2019] suggests that increasing electrode density beyond standard low-density EEG provides only limited additional benefit, while recording duration may have a stronger influence on EEG diagnostic outcomes. That indicates that LD devices can be sufficient in many cases.

Consumer EEG devices balance usability and signal quality.

Recent technological advances have enabled the development of portable, affordable, and wearable consumer-grade EEG devices. They range from single-channel to multi-channel devices, which are mainly marketed for everyday and personal use. They represent a practical version of LD devices and typically focus on frontal and temporal regions. This reflects a practical trade-off between signal quality and usability [Putze et al., 2022; Sawangjai et al., 2019]. They are commonly used in applied research such as cognition research, brain–computer interfaces, and domains including education and gaming, in contrast to purely methodological studies [Sawangjai et al., 2019].

Consumer EEG devices have practical limits in data quality.

Consumer-grade EEG devices can record useful brain signals in some situations. However, they usually have fewer electrodes and are more sensitive to movement and environmental noise, which can reduce data quality and reliability. Therefore, it is important to consider these limitations and clearly report what the devices can and cannot do when interpreting results from consumer-grade EEG devices [Sawangjai et al., 2019].

2.1.4 Device Used in This Work: Muse

Muse has four dry electrodes and wireless data streaming.

This work uses the consumer grade EEG device Muse[1] from InteraXon Inc. . The device has a sampling rate of 256 Hz with a 12-bit resolution and supports wireless data transmission via Bluetooth. The Muse has four dry electrodes located at TP9, TP10, AF7, and AF8, with a reference electrode at FPz (Figure 2.2).

The frontal electrodes are positioned over scalp regions commonly associated with attentional processes. The temporal electrodes are placed over regions that are considered sensitive for capturing signals typically linked to memory and emotional processing [Sawangjai et al., 2019].

[1] `choosemuse.com/products/muse-2`, *last accessed 26.01.2026*

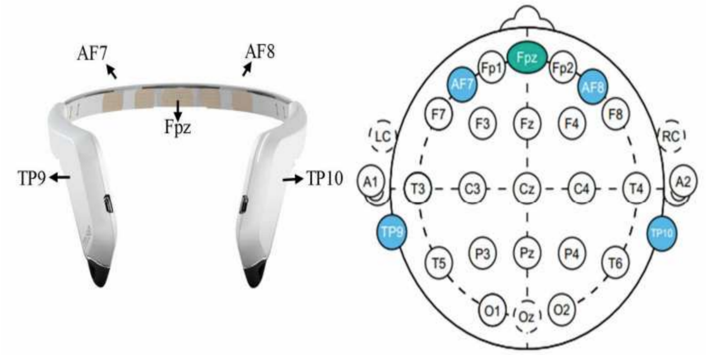


Figure 2.2: Muse device (left) and its electrode placement (right) [Sawan and Awad, 2023].

Muse is easy to use and suitable for HCI studies.

The Muse is a consumer EEG device originally designed for wellness applications, but it has been increasingly used in research due to its accessibility and open data streaming capabilities [Krigolson et al., 2017]. The use of dry electrodes allows for a fast setup without conductive gels. This makes the Muse suitable for exploratory usage outside controlled laboratory settings.

Muse has limitations in signal quality and stability.

Despite its practical advantages, the Muse device has several limitations. The use of dry electrodes and a small number of channels can lead to higher noise levels, reduced signal stability, and increased sensitivity to eye movements and blinks [Kyriaki et al., 2024; Sawangjai et al., 2019]. Also, individual factors such as head shape, hair characteristics, and electrode adjustment can further affect signal quality. In addition, wireless data transmission can introduce timing uncertainty, which should be considered when designing event-related experiments [Sawangjai et al., 2019].

2.2 EEG Analysis Perspectives

EEG can be divided into two main brain wave types, which are analysed in different ways. Continuous (also called spontaneous) EEG activity and event-related potentials (ERPs). These two approaches differ in how EEG signals are interpreted and which cognitive processes they capture [Müller-Putz et al., 2015].

Event-related potentials (ERPs) are EEG signals that appear after a stimulus. They show up as voltage changes at specific time points following the stimulus. ERPs are analyzed in terms of amplitude and latency, not frequency [Kosch et al., 2023]. The amplitudes of ERP components are usually much smaller than those of continuous EEG activity, so they are often not visible in the raw EEG signal. Therefore, ERPs are obtained by averaging EEG signals that are time-locked to repeated events or stimuli. This averaging reduces the continuous background EEG and makes the ERP components visible [Teplan et al., 2002]. Although ERPs are well suited for studying stimulus-driven cognitive processes, they require repeated, time-locked trials. This makes them less suitable for continuous or natural interaction settings [Nuamah et al., 2017].

ERPs measure brain responses to specific stimuli.

Continuous EEG activity refers to the ongoing electrical brain activity that is present at all times. It is not tied to a specific event or stimulus. Continuous EEG signals are commonly analyzed in the frequency domain by examining changes in spectral power across different frequency bands. Because it does not require repeated, time-locked stimuli, it is particularly suitable for studying sustained cognitive states and continuous interaction scenarios [Müller-Putz et al., 2015].

Continuous EEG is used to study ongoing cognitive states.

Besides ERPs, events can also cause changes in the frequency domain of the EEG signal. These changes are called event-related desynchronization (ERD) and event-related synchronization (ERS). Unlike ERPs, ERD and ERS are not obtained by trial averaging and do not appear as distinct voltage deflections in the time domain. Instead, they are analyzed as frequency-specific decreases or increases in spectral power relative to a baseline period. ERD and ERS are therefore conceptually positioned between continuous EEG and ERP analysis. They rely on event timing like ERPs, but are analyzed in the frequency domain like continuous EEG. They require a clearly defined baseline period right

ERD and ERS describe event-related changes in the frequency domain.

before the stimulus, which increases the complexity while recording [Pfurtscheller and Da Silva, 1999].

This work focuses on continuous EEG activity for capturing cognitive states, while ERD/ERS are very promising future work.

2.3 EEG Data Processing

EEG signals are preprocessed and then analysed in time, spectral, and spatial domain.

The following section outlines the signal processing pipeline used to extract features from continuous EEG recordings. EEG data analysis typically follows a multi-step pipeline (Figure 2.3). The raw data is preprocessed to reduce noise and artifacts. Then, features are extracted, including time, spectral, and spatial features [Rao, 2013]. As this work uses a consumer-grade EEG device with limited electrode coverage, only time- and frequency-domain features are included.

2.3.1 Preprocessing

Preprocessing removes noise and non-neural artifacts.

Preprocessing reduces noise and non-neural activity in the EEG data and is therefore essential to improve signal quality and interpretability. Because EEG signals have very low amplitudes, recordings are highly susceptible to artifacts. Artifacts are signal components that do not originate from neural activity but contaminate EEG recordings. These include external artifacts like line noise or poor electrode contact, as well as internal artifacts caused by eye movements, muscle activity, clenching, chewing, swallowing, or body motion [Fatourechi et al., 2007; Kyriaki et al., 2024]. For more detailed information about different artifact types, see Fatourechi et al. [2007].

Resampling and re-referencing prepare EEG data for analysis.

Before removing artifacts, EEG data is often resampled to reduce data size. However, the sampling rate must remain high enough to prevent aliasing [Kyriaki et al., 2024; Landau, 2005]. EEG does not measure absolute voltage, but

EEG Data Processing Pipeline

Resampling/Re-referencing → Filtering → Artifact Removal (ICA) → Feature Extraction (Temporal, Spectral, Spatial) → Analysis

Figure 2.3: EEG Data Processing Pipeline

voltage differences between electrodes on the scalp. Therefore, choosing a reference is essential [Dhole et al., 2025]. Re-referencing means expressing each signal relative to a selected reference electrode. A common method is common average referencing (CAR), which subtracts the average signal across all channels and helps reduce overall noise [Kyriaki et al., 2024].

Filtering

Filtering is a fundamental step in EEG preprocessing. High-pass filters (commonly 0.1–1 Hz) are used to remove slow signal drifts caused by sweat, electrode movement, or baseline fluctuations. Low-pass filters (commonly 30–100 Hz) reduce high-frequency noise like muscle activity and environmental interference [Niedermeyer and da Silva, 2005]. Band-pass filters combine both approaches, and notch filters are commonly applied at 50 Hz or 60 Hz to suppress line noise [Newson and Thiagarajan, 2019].

Filtering removes slow drifts and high-frequency noise.

Although filtering can improve data quality, overly aggressive settings may distort neural signals or remove relevant information. Artifacts often overlap with high-frequency EEG components, which makes it difficult to separate noise from brain activity using frequency-based filtering alone [Müller-Putz et al.]. An optimal and automated filtering process across cognitive load studies remains an ongoing debate [Kyriaki et al., 2024].

Filtering must balance noise removal and signal preservation.

Independent Component Analysis

Independent Component Analysis (ICA) is widely used

ICA separates neural signals from artifacts.

as a standard approach for artifact correction [Gkintoni and Halkiopoulos, 2025]. It decomposes multichannel EEG recordings into a set of components that are as statistically independent as possible, where some components primarily reflect artifacts such as eye movements or muscle activity. By identifying these components (either through visual inspection or automated), they can be removed from the data [Rao, 2013].

ICA has been shown to be particularly effective for removing eye artifacts and muscle activity, which are common in computer-mediated and interactive tasks [Kyriaki et al., 2024].

ICA relies on assumptions and must be used carefully.

ICA is based on assumptions such as linear signal mixing and statistical independence between sources. These assumptions are only approximately valid for EEG data. As a result, careless or fully automated component rejection can lead to unintended removal of neural signals [Dhole et al., 2025].

2.3.2 Time Features

Time features describe signal amplitude over time.

After preprocessing, features can be extracted. The simplest approach is to describe the signal directly in the time domain.

Basic statistical descriptors such as mean, variance, standard deviation, RMS, and amplitude range summarize the distribution of EEG signal amplitudes and are commonly used as baseline features in EEG analysis [Stancin et al., 2021]. Amplitude-based measures, including minimum and maximum values, have been applied in clinical EEG research to capture pathological alterations, for example, by showing systematic differences between patient groups and healthy controls. Variance and standard deviation describe how strongly signal amplitudes fluctuate around the mean. Lower variability has been associated with reduced brain activity dynamics, while unusually high variance can indicate artifacts. [Rodrigues and Rodrigues, 2024].

Metric	Formula	Meaning
Min / Max	$x_{\min}, x_{\max}$	Range of amplitudes
Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$	Central tendency, baseline shift
Variance	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$	Dispersion of the signal around the mean
Standard Deviation	$\sigma = \sqrt{\sigma^2}$	Strength of fluctuations
RMS	$\sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$	Magnitude of the EEG signal
Peak-to-Peak	$A_{pp} = \max(x) - \min(x)$	Amplitude range

Table 2.4: Basic time-domain metrics [Rodrigues and Rodrigues, 2024].

In addition to descriptive time-domain metrics, the SNR is used as a simple indicator of overall recording quality.

SNR indicates overall EEG recording quality.

The Signal-to-Noise Ratio (SNR) is a measure used to quantify data quality by comparing the desired information (the signal) to unwanted interference (the noise). The SNR is essential for determining how useful a signal is and how much noise limits its performance. On a linear scale, the SNR is defined as the ratio of the root-mean-square (RMS) amplitude of the signal to the RMS amplitude of the noise [Semmlow, 2012]:

$$\mathrm{SNR}_{\mathrm{linear}} = \frac{\mathrm{RMS}_{\mathrm{signal}}}{\mathrm{RMS}_{\mathrm{noise}}}.$$

The SNR is commonly expressed in decibels (dB):

$$\mathrm{SNR}_{\mathrm{dB}} = 20\log_{10} \times \mathrm{SNR}_{\mathrm{linear}}.$$

2.3.3 Spectral Estimation

Frequency analysis links EEG rhythms to cognitive states.

EEG signals are complex time-series data. When viewed in the time domain, they show voltage fluctuations over time. However, many cognitive processes are not defined

Spectral Estimation Pipeline

Preprocessed EEG → Segmentation/Windowing → Fast Fourier Transformation → Power Spectral Density → Bandpower → Features

Figure 2.5: EEG Spectral Estimation Pipeline

by single voltage peaks, but by rhythmic activity in specific frequency ranges [Müller-Putz et al.]. Therefore, instead of only analyzing amplitude changes over time, it is often more informative to examine the signal in the frequency domain. The transformation from time-domain EEG to frequency-domain features consists of several steps as shown in Figure 2.5.

Step 1: Segmentation

Segmentation divides EEG into short, stable time windows.

EEG signals are continuous and change over time. However, most spectral estimation methods assume that the analyzed signal is approximately stationary, which means that its statistical properties do not change within the analyzed time window [Rao, 2013]. To meet this assumption, the continuous EEG recording is divided into short time segments. Each segment is treated as approximately stationary [Müller-Putz et al.]. The segment length determines a trade-off between temporal and frequency resolution [Al-Fahoum and Al-Fraihat, 2014]. With longer segments having a better frequency resolution and shorter segments having better temporal resolution. In practice, segment lengths between 1 and 2 seconds are common in cognitive EEG research. In addition, overlapping segments reduces variability in spectral estimates and allows smoother tracking of changes over time [Kyriaki et al., 2024; Müller-Putz et al.].

Step 2: Windowing

Windowing reduces spectral leakage at segment boundaries.

When dividing a signal into segments, abrupt edges are created at the boundaries of each segment. These sharp edges can introduce artificial frequency components, a phenomenon known as spectral leakage Al-Fahoum and Al-Fraihat [2014]. To reduce this effect, each segment is multiplied by a smooth weighting function before frequency transformation. This process is called windowing. Windowing reduces distortion at segment boundaries and improves the reliability of the subsequent spectral estimate [Müller-Putz et al.].

Step 3: Transformation to the Frequency Domain (FFT)

Once segmentation and windowing are applied, each segment can be transformed from the time domain into the frequency domain using the Fast Fourier Transform (FFT), which is the standard method [Al-Fahoum and Al-Fraihat, 2014; Rao, 2013].

FFT converts EEG signals from time to frequency domain.

The EEG signal is decomposed into its underlying frequency components, represented as a sum of sine and cosine waves, each characterized by a specific frequency and amplitude. For each frequency, it provides the Amplitude (how strong the oscillation is) and the Phase (its timing relative to other oscillations) [Rao, 2013].

However, the raw FFT output cannot be directly used for cognitive analysis. Instead of amplitude alone, signal power is more relevant because it represents the energy at each frequency. The FFT, therefore, serves as a computational step that enables power-based analysis [Al-Fahoum and Al-Fraihat, 2014; Rao, 2013].

Step 4: Power Spectral Density

Although the FFT provides information about amplitude and phase, cognitive EEG analysis is primarily concerned with signal power. Therefore, the next step is to compute the Power Spectral Density (PSD), which describes how signal energy is distributed across frequencies. It is typically expressed in units of power per Hertz (V^2/Hz) [Al-Fahoum and Al-Fraihat, 2014; Rao, 2013].

PSD shows how signal power is distributed across frequencies.

A simple PSD can be computed from a single FFT segment (periodogram). However, this estimate can be noisy. Therefore, a more robust method is often used: Welch's method [Welch, 2003]. Welch's method estimates the power spectral density by first splitting the signal into overlapping segments like in Section 2.3.3. Each segment is smoothed with a window function like in Section 2.3.3 and then transformed into the frequency domain using the FFT. From this, a power spectrum is calculated for each segment. These power spectra are then averaged, which reduces variability and leads to a more stable estimate of the PSD [Welch, 2003].

The PSD is a common intermediate result in spectral estimation, and many subsequent frequency-domain features can be derived from it [Rao, 2013].

Step 5: Bandpower

Bandpower summarizes power within specific frequency ranges.

The bandpower summarizes the power within specific frequency ranges and is computed by integrating the PSD over a chosen frequency band $f_1 - f_2$:

$$\text{Bandpower}_{[f_1,f_2]} = \int_{f_1}^{f_2} P(f)\, df\,.$$

where $P(f)$ denotes the PSD at frequency f [Al-Fahoum and Al-Fraihat, 2014; Rao, 2013].

Bandpower reduces the detailed frequency spectrum into a single value per band. This simplification makes it easier to compare conditions and interpret results in relation to cognitive processes. Absolute bandpower refers to the raw power within a band. Relative bandpower expresses bandpower as a proportion of total power across a broader frequency range and can be computed as $RBP = \text{Bandpower}_{[f_1,f_2]}/\text{Bandpower}_{[f_{min},f_{max}]}$.

The interpretation of bandpower depends on the frequency range over which it is computed. Bandpower is therefore calculated with respect to predefined EEG frequency bands that are associated with distinct neurophysiological and cognitive processes [Kosch et al., 2023; Kumar and Kumar, 2016].

2.4 EEG Data Interpretation

After extracting spectral features from the preprocessed EEG (Section 2.3.3), the next step is to interpret what these values mean in terms of cognitive states. The following sections summarize the main frequency bands and their cognitive associations and introduce the ratio indices used in this work.

2.4.1 Frequency Bands and Cognitive Associations

Frequency bands differ in their typical frequency range, amplitude, and functional associations. Changes in the bandpower of these bands have been associated with mental states such as attention, cognitive load, fatigue, and emotional engagement [Müller-Putz et al., 2015]. The frequency band boundaries are not fixed constants. Individual differences in brain anatomy, age, and cognitive strategies can lead to variability in peak frequencies, and exact band limits therefore vary across studies and are subject to ongoing debate in the literature [Norman Donald, 2013].

Frequency bands differ in range and meaning and vary across individuals.

Delta (δ) and gamma (γ) activity are not considered in further detail in this work. Delta activity is dominant during deep sleep and unconscious states and is largely irrelevant to task-related cognitive processing in awake users [Müller-Putz et al., 2015]. Gamma activity is sometimes linked to arousal and attention but has very low amplitude, and is highly susceptible to noise [Müller-Putz et al., 2015]. This makes it difficult to measure reliably with consumer-grade EEG devices [Cannard et al., 2021]. Therefore, this work focuses on theta, alpha, and beta power, which are dominantly used for workload assessment [Kosch et al., 2023].

Delta and gamma are excluded due to limited relevance and reliability.

Theta (θ) power has been linked to increased mental effort and focused attention on task-relevant stimuli, which

Theta power increases with mental effort and working memory load.

Band	Frequency Range	Primary Functional Associations
Delta (δ)	0.5–4 Hz	Deep sleep, unconsciousness, pathological states, motivational salience
Theta (θ)	4–8 Hz	Memory encoding, cognitive control, drowsiness, emotional processing
Alpha (α)	8–13 Hz	Relaxed wakefulness, attention, concentration, mental workload
Beta (β)	13–30 Hz	Active thinking, motor planning, focus, anxiety, cognitive processing
Gamma (γ)	>30 Hz	Perceptual binding, consciousness, attention, sensory processing

Table 2.6: EEG frequency bands and their primary functional associations. Adapted from Gkintoni and Halkiopoulos [2025] and expanded.

makes it a reliable indicator of working memory demands [Gevins and Smith, 2000]. An increase in theta power is therefore commonly interpreted as reflecting increased cognitive load [Gkintoni and Halkiopoulos, 2025]. In particular, frontal–midline theta power (measured at Fz) is associated with mental effort and focused attention during task performance and increases with the number of items maintained in working memory [Gkintoni and Halkiopoulos, 2025; Müller-Putz et al., 2015].

Alpha power decreases during cognitive demand.

Alpha (α) activity is most prominent during relaxed wakefulness, for example with eyes closed. Alpha power typically decreases during cognitively demanding tasks [Gevins and Smith, 2000; Gkintoni and Halkiopoulos, 2025]. Whereas higher alpha power has been associated with reduced concentration and higher error rates in detection tasks [Rao, 2013]. Reduced alpha power has also been linked to fatigue and declining alertness [Smith and Gevins, 2005]. Overall, alpha-band measures are commonly used to assess vigilance, attention, alertness, and mental workload [Fernandez Rojas et al., 2020; Puma et al., 2018].

Beta power reflects active concentration and cognitive processing.

Beta (β) activity is linked to active concentration, alertness, and task engagement. [Müller-Putz et al., 2015] Increased beta power has been associated with effortful attention [Müller-Putz et al.] and has been used to study visual attention and short-term memory [Fernandez Rojas et al., 2020]. It is also commonly associated with active cognitive processing, including visual attention, motor planning, and executive control [Baradari et al., 2025].

2.4.2 Ratio Indices

While individual EEG frequency bands are informative, many applied studies use ratios of bandpowers to summarize cognitive states. Ratio metrics can reduce inter-individual differences and are computationally efficient, which makes them attractive for applied and real-time settings [Gkintoni and Halkiopoulos, 2025; Raufi and Longo, 2022].

Bandpower ratios summarize cognitive states in a compact form.

Importantly, there is no single, universally accepted EEG ratio that uniquely represents engagement or cognitive load. Instead, the literature proposes a range of ratio definitions and naming conventions, and identical labels may refer to different formulations across studies. Commonly used ratios include $\beta/(\alpha + \theta)$, θ/α, and θ/β as indicators related to attention, workload, or cognitive demand [Fernandez Rojas et al., 2020].

There is no single standard EEG ratio.

In this work, two ratios are implemented an *Engagement Index* defined as $\beta/(\alpha + \theta)$ and a *Cognitive Load Index* defined as θ/α.

Cognitive Load Index

In HCI and applied workload research, the terms *cognitive workload*, *cognitive load*, and *mental workload* are often used interchangeably, even though they originate from different theoretical models and are not conceptually identical. This reflects the absence of a unified definition and the lack of a gold standard measurement for such constructs [Kosch et al., 2023].

Cognitive load terminology is inconsistent across disciplines.

The Cognitive Load Index is defined as theta divided by alpha power.

In this thesis, *cognitive load* is used as an umbrella term to describe the cognitive capacity allocated to accommodate the demands imposed by a task [Paas et al., 2016]. The implemented index is based on the ratio between theta and alpha band power, which has been widely used as a compact indicator of cognitive demand in EEG studies [Gkintoni and Halkiopoulos, 2025; Mazher et al., 2017; Müller-Putz et al., 2015]. The Cognitive Load Index is defined as:

$$\mathrm{CLI} = \frac{\theta}{\alpha}$$

where α and θ denote the band power in the respective frequency range. This formulation reflects the observation that cognitive load is associated with increased theta power and decreased alpha power as described in Section 2.4.1. Accordingly, a higher Cognitive Load Index indicates higher mental effort and increased task demands.

Engagement Index

The Engagement Index is defined as beta divided by alpha plus theta.

The Engagement Index is defined as:

$$\mathrm{ENG} = \frac{\beta}{\alpha + \theta}$$

where α, β, and θ denote the band power in the respective frequency range, and was introduced by Pope et al. [1995] in the context of adaptive systems to estimate task engagement from EEG. Higher values are commonly interpreted as increased attentional focus or engagement. Studies have validated the engagement index across a wide range of experimental examples [Apicella et al., 2022; Baradari et al., 2025; Eldenfria and Al-Samarraie, 2019; Hassib et al., 2017; Kosmyna and Maes, 2019]. This index combines increased beta power, associated with active concentration, with reduced alpha and theta power, associated with relaxed or drowsy states as described in Section 2.4.1.

2.4.3 Limitations of EEG Interpretation

EEG measures vary across individuals.

A fundamental limitation of EEG lies in its strong inter-individual variability. Differences in skull and scalp properties, brain anatomy, electrode placement, as well as demographic and physiological factors such as age, fatigue, or arousal level, affect how neural signals are measured at the scalp and limit direct comparability between participants [Müller-Putz et al.].

EEG reflects distributed brain activity, not precise locations.

EEG does not measure the activity of individual neurons but primarily reflects synchronized postsynaptic activity from large populations of cortical neurons as described in 2.1.1.

EEG results depend on preprocessing and interpretation choices.

EEG signals are highly susceptible to artifacts such as eye movements, muscle activity motion, and environmental noise [Kyriaki et al., 2024]. Consequently, EEG analysis strongly depends on preprocessing choices, including filtering, referencing, and artifact handling. As no universally accepted preprocessing standard exists, different methodological decisions can lead to different results [Trübutschek et al., 2024]. Moreover, EEG does not directly measure cognitive constructs such as attention or cognitive load. Instead, interpretations rely on statistically established associations between signal characteristics and mental states, which may vary across theoretical perspectives and research contexts [Putze et al., 2022].

Chapter 3

Related Work

This chapter reviews prior work relevant to designing accessible EEG systems for HCI research. First, Section 3.1 summarizes how EEG is used in HCI. Section 3.2 then discusses the role of consumer-grade devices in applied research. Section 3.3 reviews the current EEG software landscape from the perspective of HCI-oriented user studies. It includes commercial software, open-source frameworks, and graphical user interfaces. Finally, Section 3.4 examines existing systems that attempt to address parts of the identified gaps.

3.1 EEG as an Evaluation Tool in HCI

EEG in HCI is used mainly for evaluation, not control.

The use of brain signals in HCI is a research area that is related to, but clearly distinct from, classical brain–computer interface (BCI) research. Studies are often exploratory, involve small sample sizes, and are conducted under less controlled conditions [Putze et al., 2022]. The review by Putze et al. [2022] shows that most EEG-based HCI studies do not use EEG for control of interfaces. Instead, EEG is primarily used as an evaluation tool that provides continuous and objective insights into users' cognitive and experiential states during interaction. Within this evaluation-oriented research, EEG is commonly applied to examine constructs such as cognitive load, attention, and engagement.

Cognitive load measures in HCI are methodologically diverse.

It is important to note that those constructs are methodologically heterogeneous. A survey by Kosch et al. [2023] highlightes for example, that the terms used in cognitive load studies vary in their theoretical foundations and measurement approaches. EEG studies have a variety of evaluation approaches, including different frequency band ratios (see Section 2.4.2) or ERP measures. We will focus on frequency-domain approaches, particularly power changes in the theta, alpha, and beta bands.

Bandpower ratios are used as compact indicators in HCI.

In a wide range of studies, increased cognitive demands have consistently been associated with higher theta activity and reduced alpha activity, especially in the frontal and parietal regions [Anderson et al., 2011; Antonenko et al., 2010; Fernandez Rojas et al., 2020; Gevins and Smith, 2003; Holm et al., 2009; Tasmi et al., 2023]. Based on these findings the cognitive load index θ/α is used in applied studies [Ahmed et al., 2020; Baceviciute et al., 2020; Hassib et al., 2017]. Aswell as the engagement index $\beta/(\alpha + \theta)$ [Ahmed et al., 2025; Apicella et al., 2022; Baradari et al., 2025; Eldenfria and Al-Samarraie, 2019; Kosmyna and Maes, 2019; Nuamah et al., 2017].

3.2 Consumer-grade EEG Devices in Applied Research

Consumer-grade EEG devices are widely used in HCI research.

Consumer-grade EEG devices are widely used in HCI research due to their accessibility and low setup effort [Putze et al., 2022]. While these devices are limited by sparse electrode coverage and increased sensitivity to artifacts [Sawangjai et al., 2019], prior work indicates that they provide sufficient signal quality for frequency-domain analyses such as power spectral density and bandpower computation [Gkintoni and Halkiopoulos, 2025]. From an HCI perspective, the main challenge is not only data quality, but ensuring that the strengths and limitations of these devices are clearly understood and properly interpreted in the study context.

Validation studies indicate that the Muse can reliably capture low-frequency spectral features, including bandpower and power spectral density, while higher-frequency activity is less consistent [Cannard et al., 2021]. Beyond spectral analyses, large-scale and methodological studies have shown that Muse can also support event-related EEG measures [Krigolson et al., 2017, 2021]. Several studies have applied the Muse in classification-based settings to examine cognitive or mental states from spectral features [Bakshi, 2018; Bashivan et al., 2016; Rohit et al., 2017].

Muse supports frequency-domain analysis despite limitations.

While consumer devices make EEG technically accessible, the software for recording and analyzing the data is still a major limitation.

3.3 EEG Software Landscape

This section reviews existing EEG software from the perspective of HCI user studies. Instead of giving a complete overview of all tools, it focuses on common strengths and weaknesses that are important for researchers without expert knowledge. The section is grouped into commercial software, open-source frameworks, and GUI-based tools.

3.3.1 Commercial Software

Several commercial EEG software is available for data recording and analysis. Scientifically oriented software such as BESA Research[1] , BrainVision Analyzer[2] , Acq-

Commercial EEG software is powerful but closed and proprietary.

[1] `besa.de`, *last accessed 26.01.2026*

[2] `brainproducts.com/solutions/analyzer`, *last accessed 26.01.2026*

Knowledge[3] , LabChart[4] , NeuroPype[5] and Bitbrain[6] offer advanced signal processing, artifact handling, and visualization features. However, these softwares are not free, not open source, and closely tied to proprietary hardware, which limits their accessibility and flexibility. Commercial consumer-oriented applications provided with devices such as the Muse App[7] or NeuroSky[8] mainly target end users. They support only predefined use cases, such as meditation or basic feedback, and provide limited access to raw data or customizable analysis workflows.

Proprietary software limits transparency and reproducibility.

So a central limitation of commercial EEG software lies in its restricted transparency and reproducibility. Several studies note that proprietary software often provide insufficient documentation of their internal signal processing, filtering, artifact correction, or classification methods, making it difficult for researchers to fully understand or validate the results [Chevallier et al., 2024; Gilmore et al., 2017; Paul et al., 2021]. Closed data formats and limited export options further complicate independent verification and cross-study comparison. As a result, reproducing analyses or extending existing workflows is often not feasible for other researchers [Niso et al., 2022]. These characteristics stand in contrast to current open science principles, which emphasize transparency, accessibility, and well-documented analysis pipelines as prerequisites for reproducible and trustworthy research [Brito et al., 2020; Gorgolewski and Poldrack, 2016].

3.3.2 Open Source Frameworks

Open source EEG frameworks offer flexible, code-based pipelines.

In addition to commercial software, a large number of open-source EEG and BCI frameworks are available. Prominent examples include FieldTrip [Oostenveld

[3] `biopac.com/product/acqknowledge-software`, *last accessed 26.01.2026*

[4] `adinstruments.com/products/labchart`, *last accessed 26.01.2026*

[5] `neuropype.io`, *last accessed 26.01.2026*

[6] `bitbrain.com/neurotechnology-products/software/sennslab`, *last accessed 26.01.2026*

[7] `choosemuse.com/pages/app`, *last accessed 26.01.2026*

[8] `store.neurosky.com/pages/mindwave`, *last accessed 26.01.2026*

et al., 2011], BCILAB [Kothe and Makeig, 2013], BioSig [Vidaurre et al., 2011], BrainFlow, MNE-Python [Gramfort et al., 2013], BCI2000 [Schalk et al., 2004] and NeuroKit2[9]. These frameworks offer functions for EEG signal processing, feature extraction, classification, and statistical analysis. The frameworks are designed as programming libraries rather than end-user applications with graphical interfaces. Users are expected to build custom analysis pipelines using code. While basic visualization is often available (see Figure 3.1), none of these frameworks provides a fully integrated, visual workflow that covers guided setup, structured recording, and comparative analysis within a single system.

Pipeline choices influence EEG results.

As a result, many studies use custom preprocessing and analysis pipelines tailored to individual projects [Cowley et al., 2017; Schummer et al., 2025]. Large-scale replication efforts such as EEGManyPipelines by Trübutschek et al. [2024] show that differences in preprocessing and analysis choices can directly influence reported results.

Open frameworks lack guided workflows for non-expert users.

Those findings highlight the importance of transparency and clear documentation of all steps in EEG data processing. Overall, open source EEG frameworks are powerful and offer alot of options and functions. However they mainly target technically skilled users and rely on script-based workflows. While this supports flexible and customizable analysis, it provides limited support for non-expert users, which are common in EEG oriented HCI research.

3.3.3 Graphical User Interfaces

GUI simplify analysis but separate recording from analysis.

Graphical User Interface (GUI) based EEG interfaces aim to make EEG recording and analysis more accessible by providing visual and interactive workflows. In contrast to script-based frameworks, these interfaces reduce the need for programming and support direct interaction with EEG data through graphical interfaces.

[9] `pypi.org/project/neurokit2/`, *last accessed 01.02.2026*

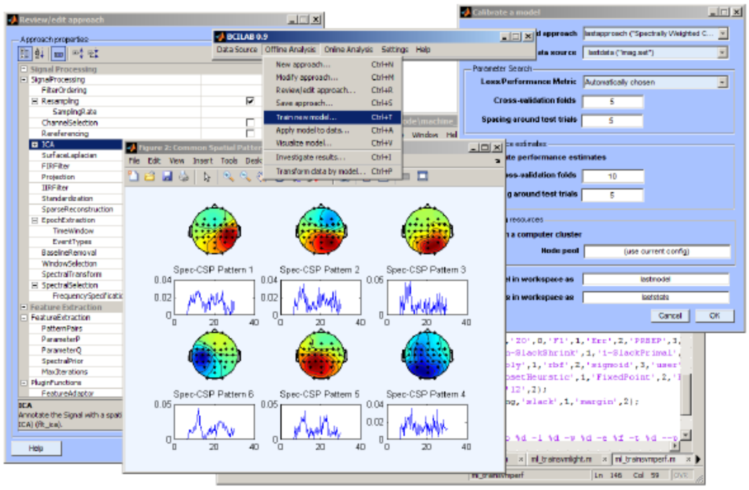


Figure 3.1: The BCILAB graphical user interface, including the main menu, a model visualization window, a parameter configuration dialog for a BCI approach, a method configuration window, and the MATLAB editor workspace. [Brunner et al., 2012].

Some GUI based platforms focus primarily on the analysis of previously recorded EEG data. Most popular examples include EEGLAB [Delorme and Makeig, 2004] and Brainstorm [Tadel et al., 2011]. These interfaces provide visual analysis and support a wide range of preprocessing and feature extraction methods, but they require that recordings are conducted externally and imported afterwards. As a result, they do not support integrated study workflows that combine recording, condition management, and analysis within a single interface.

Recording GUIs lack integrated analysis and comparison.

Other GUI based interfaces focus on EEG recording and real-time monitoring, but provide little or no built in analysis functionality. LabRecorder[10] for example supports synchronized recording of multiple data streams but does not include signal processing. OpenSignals[11] and the Open-

[10] `github.com/labstreaminglayer/App-LabRecorder`, *last accessed 26.01.2026*

[11] `pluxbiosignals.com/pages/opensignals`, *last accessed 26.01.2026*

BCI GUI[12] (see Figure 3.2) similarly support mainly data recording and not including analysis.

Figure 3.2: Overview of the OpenBCI GUI.

Although the OpenBCI GUI has similar visuall components to the interface presented in this work, both interfaces differ in scope. The OpenBCI GUI is designed primarily for device control, real-time monitoring, and raw data acquisition for subsequent processing in external environments such as MATLAB or Python. It provides live visualizations (e.g., time series and FFT plots) but does not include structured preprocessing, condition-based comparison, or integrated cognitive metric computation within the same interface. Furthermore the OpenBCI GUI is officialy only compatible with OpenBCI hardware[13].

Fragmented GUI provide little guidance for HCI researchers.

Simply combining recording and analysis GUIs does not address core methodological challenges. Existing interfaces lack a shared representation of conditions, explicit links between recording and analysis, and support for systematic comparison across sessions. For inexperienced HCI

[12] docs.openbci.com/Software/OpenBCISoftware/GUIDocs, *last accessed 26.01.2026*

[13] docs.openbci.com/Software/OpenBCISoftware/GUIDocs, *last accessed 26.01.2026*

researchers, this fragmentation offers little guidance and leaves the complexity of EEG analysis largely unchanged.

3.3.4 Identified Gaps

High Entry Barrier for HCI Researchers.

Current EEG softwares are primarily designed for expert users rather than HCI researchers. Open-source frameworks require programming and many methodological decisions. Commercial software reduces coding effort, but it is expensive and often tied to proprietary hardware. GUI-based tools lower technical effort, yet they still offer little step-by-step guidance. As a result, non-expert HCI researchers face a high entry barrier when setting up EEG studies and collecting reliable data.

Limited Support for Transparent, GUI-Based Processing.

A second gap is the lack of interfaces that combine transparency with ease of use. Commercial tools often function as black boxes, while code-based pipelines are transparent but difficult for non-experts and vary across studies. Existing GUIs typically focus either on recording or on post-hoc analysis, without integrating preprocessing and metric computation in one clear workflow. As a result, few tools provide a visual and understandable processing pipeline that users can inspect and reproduce.

Missing Study Logic for HCI User Studies.

Finally, existing tools are not structured around the needs of HCI user studies. There is little support for representing study conditions consistently or for comparing results across participants and sessions. Instead, many tools focus on device control, BCI development, or specific applications. This leaves a gap for a general-purpose interface that supports structured HCI study workflows.

3.4 Existing Systems Addressing Identified Gaps

Several prototypes and systems improve accessibility, transparency, or workflow integration. However, none of them solves all identified gaps. They still offer useful design ideas and show alternative ways to approach these challenges.

3.4.1 Interfaces to Lower the Entry Barrier

A first group of interfaces focuses on making EEG studies more accessible through web-based or remotely usable workflows. Sans Tracas by Desai et al. [2022] is a cross-platform interface for conducting online EEG experiments and aims to enable participants to run studies outside the lab with guided instructions. This supports low-threshold data collection. MYND by Hohmann et al. [2019], for example, also focuses on easier data collection, enabling smartphone recordings in field settings. EEG Collect by Stingl and Knierim [2024] is an open source, web-based platform that reduces the need for expert supervision during setup and recording including time feedback and impedance checks that help participants adjust electrodes until signal contact is acceptable.

Web-based interfaces lower setup effort for EEG studies.

Those interfaces do not provide preprocessing or feature of the data analysis but align with the goal of lowering setup friction for non-experts and, therefore, lowering the entry barrier.

3.4.2 Interfaces with Open and Transparent Data Processing

While open-source frameworks offer transparency through code-based pipelines, a smaller number of systems attempt to provide transparent EEG data processing within graphical interfaces.

OpenViBE is an open source platform that enables the construction of EEG processing pipelines through a fully graphical interface. Its core concept is the Scenario Designer, a visual programming environment in which users connect modular processing blocks for data acquisition,

OpenViBE offers transparent pipelines but targets BCI use.

filtering, feature extraction, classification, and feedback. This approach supports very visual, transparent, and modular signal processing without requiring direct programming and makes individual processing steps explicitly visible[Brunner et al., 2012; Renard et al., 2010]. However, OpenViBE is primarily designed for BCI development rather than for structured HCI user studies. Study elements such as participant handling, condition-based recording, and standardized comparison must be set up manually, which is difficult for non-expert users [Brunner et al., 2012].

3.4.3 Application Specific Systems

Applied systems demonstrate EEG use in concrete application contexts.

Application-specific systems demonstrate practical value in concrete HCI contexts. Aggarwal et al. [2014] presents a prototype that combines EEG recording with a visual analysis of screen recording and mouse cursor postitions to evaluate user experience during interface interaction, and links EEG-derived measures to the interface context. The EngageMeter by Hassib et al. [2017] is used on measure group engagement during presentations. Research by Ahmed et al. [2020, 2025] investigates cognitive load and engagement in learning environments and video games. NeuroChat by [Baradari et al., 2025] integrates EEG-based metrics directly into an adaptive tutoring system.

Existing systems are tailored to specific use cases and goals.

While these systems demonstrate the potential of EEG-driven evaluation and adaptation, they are built around predefined use cases and fixed analysis goals. The processing logic and computed metrics are tightly coupled to the specific application context.

Interface gap motivates a qualitative study of HCI researchers' needs.

Taken together, the reviewed literature shows that EEG is widely used in HCI as an evaluation tool and that accessible hardware is available. However, existing software solutions either require expert knowledge, lack transparent processing workflows, or are tailored to specific application contexts. To better understand the practical needs of HCI researchers, the next chapter presents qualitative interviews with HCI researchers. These interviews aim to complement the literature-based analysis with first-hand perspectives in order to derive concrete design requirements for an HCI oriented EEG interface.

Chapter 4

Empirical Study and Design Implications

This chapter presents the empirical foundation of the interface design. To understand how HCI researchers without prior EEG experience perceive, interpret, and imagine integrating EEG into their research practice, five semi-structured interviews were conducted.

The chapter first describes the participants in Section 4.1, interview design in Section 4.2, and the analytic approach in Section 4.3. It then presents the key themes derived from the interviews in Section 4.4. Based on these findings, higher-level design principles are defined in Section 4.5, which translated into concrete interface requirements in Section 4.6.

4.1 Participants

Participants were five HCI researchers from different HCI research fields with no prior experience in EEG, BCI, or neuroscience. Their research fields include shape-changing interfaces, deceptive patterns, spatial computing and virtual reality, generative AI in creative processes, and design for textile interfaces. All participants were PhD candidates.

Participants represent non-expert HCI researchers as target users.

This participant group represents the interface's target audience: HCI researchers who conduct user studies but have limited familiarity with EEG methodology.

4.2 Interview Design

Semi-structured interviews explore perceptions and expectations of EEG.

Semi-structured interviews were used to combine open exploration with some level of consistency across participants. Because participants had no prior EEG experience, open-ended questions were chosen to reveal implicit needs and possible use cases that a fixed questionnaire might miss. The interview guide included four main topics: (1) participants' research background and data practices, (2) their familiarity with and perception of EEG, (3) possible areas of application and (4) their views on trust and transparency regarding EEG. The whole guide can be found in Appendix A.

All interviews were conducted individually and in person. They lasted 25–40 minutes, and were audio recorded with prior consent. The interviews were conducted in german to avoid language barriers. The recordings were transcribed, anonymized, and assigned participant codes (P1–P5). They were then translated into English while preserving the original meaning.

4.3 Analytic Approach

Reflexive Thematic Analysis was used for qualitativ analysis.

The interviews were analyzed using Reflexive Thematic Analysis (RTA) following Braun and Clarke [2006]. The analysis focused on how participants described their expectations and views, rather than on how often certain statements were made.RTA was chosen because it makes it possible to find common patterns across interviews while acknowledging that the researcher plays an active role in interpreting the themes. This approach was suitable for exploring how HCI researchers without EEG expertise understand EEG and express their needs regarding recording and analysis tools.

All interviews were designed, conducted, and transcribed by the author. That ensured close familiarity with the data. No second coder was involved. Instead, quality was ensured through repeated checks for consistency and critical reflection. Transcription supported the identification of recurring patterns. Themes were iteratively reviewed, compared with the data, and refined to reduce overlap. The codebook can be found in appendix B.

4.4 Interview Findings

The following section presents the resulting themes based on the RTA analysis of the interviews. Each of the following themes describes a key aspect of how participants view the use of EEG in HCI research and what they expect from a recording and analysis interface.

Findings summarize key themes.

1. Competence boundaries and resulting uncertainty
2. Motivation driven by perceived usefulness
3. EEG as an supportiv source of insight
4. Low-friction integration into user study routines
5. Trust through transparency and methodological grounding
6. Comprehensibility and accessibility of the interface
7. Interest profiles regarding cognitive states

Each theme is explained in detail below with representative quotations and interpreted in relation to the overall research context.

4.4.1 Competence Boundaries and Resulting Uncertainty

EEG is mainly associated with medical contexts, not with HCI research.

Most participants associated EEG with medical contexts. "I would definitely have thought of a hospital" [P1], "of those caps [...] and somehow of those tubes people are put into" [P2]. No one knew what EEG actually measures or how the data can be interpreted: "I don't really have a clear idea" [P1], "because I simply have no clue about this kind of data" [P5], "What the data itself actually says [...] I can't tell" [P3].

Participants do not understand what EEG measures or how to interpret the data.

Because of this uncertainty, participants found it difficult to see the value of the EEG for their own research. As P1 explained: "I hadn't considered it for my studies before, and I still find it hard now. I don't really know yet what we could do with it in the future." [P1]. This uncertainty reflects less a lack of interest than a lack of understanding of how EEG could be integrated into their research practice.

EEG is not seen as a common method in HCI literature.

One possible reason is that participants were not familiar with HCI literature, which includes EEG: "I've seen very few papers in the literature [...] especially about brain measurements" [P1]. P4 replied when asked whether he had seen any work involving brain activity, "None at all, I think." [P4]. Similarly, P5 said: "From what I know of the literature so far, it hasn't really been done much, or at all, in that way." [P5].

Overall, this theme shows a clear knowledge gap between HCI researchers and EEG methods. Participants did not reject EEG, but they did not know what it measures, how the signals are processed, or what conclusions can be drawn from it. EEG was seen as medical, technically complex, and not part of their usual research practice. This made it hard for them to imagine how it could be used in their own work. Researchers need clear explanations of how the EEG works, how the data is processed, and what its limits are.

4.4.2 Motivation Driven by Perceived Usefulness

Participants are open to EEG, but only if it clearly benefits their research.

Even though there were uncertainties, participants were generally open to working with EEG data. However, their openness seemed tightly connected to the perceived usefulness. P3 described this relationship clearly: "I'm always open. I think it depends on whether I feel it's worth it. [...] If I knew it could really give me new perspectives that I didn't have before and couldn't get otherwise, then definitely." [P3] P5 emphasized: "Absolutely. Because if it's data that makes me feel like it gives us a new perspective on a topic, then yes." [P5] Similarly, P2 said: "... to see whether it would lead me in directions I couldn't measure or didn't consider before." [P2]

Interest in EEG depends strongly on whether it offers clear new insights.

At the same time interest could quickly decline if the entry barrier was too high or the purpose unclear. Some participants also expressed doubts about whether they would actually integrate EEG into their own work. P1 said: "I'd like to say yes right away, but you realize there are alternatives beyond questionnaires that I've known about for years and still haven't used." [P1] This shows that openness alone does not automatically lead to adoption. Existing routines and uncertainty act as barriers, even when the general interest is present. P3 said: "I probably wouldn't read up on it for hours, but if I see potential for my studies, then definitely, I'd invest more time." [P3]

Participants want clear examples and realistic expectations.

To assess early on whether EEG fits their work, participants wanted clear and realistic guidance. P3 described this need: "A list of possibilities or just an example so I know what's even possible with it." [P3] Across interviews, participants emphasized the importance of realistic expectations and avoiding overselling. They wanted to understand not only what an EEG interface can do, but also what it cannot.

The interface must quickly show value and clearly explain its limits.

This theme shows a practical, usefulness-focused attitude toward EEG integration. Participants are motivated by the expectation of clear benefits. Learning was seen as an investment that needs to pay off, rather than something driven by curiosity alone. This shows the need for an interface that offers guidance, quick initial benefits, and clear communication about what it can and cannot do.

4.4.3 EEG as an Supportiv Source of Insight

Participants describe EEG as a complementary and objective data source.

The interviews were designed to support open reflection on possible use cases, without adding technical or conceptual limits. Across the interviews, the usage of EEG was mainly described as an additional and objective source of data for user studies.

Only one participant imagined applications beyond user studies.

Only P5 mentioned applications outside of user studies: "If you had a measurement that detects when someone is highly focused, you could, for example, have a small light at the door turn on "focus zone." for office enviornment and they also imagined using EEG in their own data analysis workflow, by defining commands to trigger quick visualizations on screen. All other participants mentioned possible applications only in the context of user studies. They were not encouraged to think beyond the ideas they brought up themselves. This was intentional, as the goal was to understand their current and realistic view of EEG. Their answers reflect their existing research practices rather than future or speculative ideas. Since all participants regularly conduct user studies, it is not surprising that they mainly placed EEG within this context.

EEG may capture states that questionnaires cannot measure.

Within the context of user studies, EEG was primarily seen as an observational, objectiv and complementary measurement method: "Where such data could be added and somehow complement what we already have" [P4]. P2 described EEG as an opportunity "to really gain study results or insights [...] things that you can't easily ask with a questionnaire."[P2]. So, EEG could help research "rely on more than just the subjective perception of a person." [P1]. And capture "Something the participant might not even be aware of [...] that you gain new perspectives from that [...] I find that very exciting." [P3]. P1 also highlighted EEG's potential to overcome recall bias and retrospective uncertainty: "People forget a lot or have problems in places they aren't even aware of themselves. [...] Seeing a bit more continuously what's going on in people, I find that interesting." [P1].

A recurring motiv was the interest in time points: "Timing points are probably the most relevant." [P4] Similarly, P2 described the desire "to see over time when the person was in flow and how confident the interface is about those specific moments." [P2].

Time-based information was of interest.

Several participants suggested combining EEG data with other data sources. P2 proposed displaying EEG data alongside video recordings, describing "a kind of brain intensity track underneath, appearing in the same timeline". P5 suggested linking EEG with eye tracking data "to relate specific observed states to individual interface elements." These ideas further support the view of EEG as an additinal data source.

Participants want EEG to be combined with video or eye tracking data.

Only P2 imagined EEG as a potential feedback mechanism for active system control in user studys: "It would be very interesting to let the system react to that." [P2] In contrast, "Whether we really need to control shape-changing interfaces with the brain or with thoughts. I don't think that's so interesting." [P3] Among all other participants, the topic of active control or feedback loops did not emerge.

Active system control is rarely considered and not a main interest.

This theme presents EEG as a supporting source of evidence. It is seen as an extension of existing research methods rather than replacing them. Researchers described EEG as having a validating role, especially for mental states that are hard to express or remember later.

EEG should extend existing methods, not replace them.

4.4.4 Low Friction Integration into User Study Routines

For all participants, it was important that an EEG interface could be easily integrated into existing study workflows. Most participants described their studies as a series of tasks carried out under different conditions. An interface therefore should be quick to set up, unobtrusive, and require little configuration. Time and effort were repeatedly men-

EEG must integrate smoothly into existing study workflows.

tioned as key factors for acceptance. P3 noted that using an EEG interface would become problematic "if it significantly increases the overall study effort." [P3]

EEG should not distract participants or change their behavior.

Several participants emphasized that the device should not cause distraction or irritation, as this could influence participant behavior and thereby compromise data quality. In this context, P5 expressed concern about a possible reinforcement of the Hawthorne effect, that is, behavioral changes resulting from participants' awareness of being observed. EEG should therefore not become the focus of the study but rather be included into an existing setup.

During the study, only basic feedback is needed.

There was also a clear interest in immediate feedback during data collection. Researchers wanted to ensure that the interface was functioning properly without having to perform detailed analyses during the session. P2 said: "that it's not a complete blind flight, but that while the study is running, I can already see some parameters showing that I'm measuring meaningful values and a short visualization, like what the person's current emotional or concentration state is." [P2] Similarly, P4 emphasized interest in a partial real-time view, as "it could be exciting to have some kind of semi-live stream of it," enabling "more specific follow-up questions if needed." [P4] There was agreement that detailed analysis should take place only after data collection. During the study, a "quick and dirty estimate" [P2] was considered sufficient to ensure that the measurements were valid.

The interface should minimize interaction during recording.

This theme shows that acceptance of an EEG interface depends on how smoothly it fits into existing research practice. The EEG interface should be reliable, unobtrusive, and easy to integrate without changing existing study designs. It should required little setup time and minimal attention during experiments.

4.4.5 Trust through Transparency and Methodological Grounding

Trust depends on transparency and methodological grounding.

Insight into data processing and data quality was seen as essential. P1 said they need all the processing infos and "want to be able to calculate it myself." Similarly, P2 explained: "If it gives me a result, it has to explain how it got there, because I have to explain that in the end too." These statements show a strong rejection of black box behavior and a clear demand for verifiable analysis steps.

P2 also proposed "something like a certainty score" [P2] to better judge the reliability of results and reduce uncertainty. Trust was further linked to established scientific practice. P1 noted that trust increases "if people from outside HCI maybe from a more medical or professional background use this methods. That would build a lot of trust for me." and added that he would "mainly rely on external resources to build basic trust." P2 similarly stated feeling "most confident when the analysis is based on published work by experts in the field, ideally in reputable journals, that show how such data are analyzed." P5 added that trust arises "when using a device that is already known from other research fields to produce reliable results."

Participants reject black-box systems and demand traceable analysis steps.

Participants also expressed a need for predefined, recommended configurations as starting points: "I usually like having some defaults that are recommended, what the professionals do, so I don't have that fear, but I can still change and inspect things if I want more control." [P3]

Participants want recommended defaults.

Overall, this theme shows that trust in EEG analysis interfaces is built through transparency, traceability, and clear documentation.

4.4.6 Comprehensibility and Accessibility of the Interface

Regarding the visual interface itself P1 preferred minimal displays: "Basically just a line that shows, for example, mental demand or stress." [P1]. Several participants stressed the importance of controlled information density. P2 said: "If the visualization doesn't overload you with data, but first shows something simple that can become more complex once you understand it." [P2]. "Not too

Participants prefer simple visuals with optional depth.

much displayed at once, maybe separate tabs for different variables." [P3]. A reduced and well-structured interface was therefore seen as important, especially for first-time users: "I'm definitely in favor of even professional software being a bit simpler at the beginning. [. . .] There always has to be the possibility to then look a bit deeper.", "If you suddenly have to configure everything, I would feel a bit less comfortable with it, because I simply have no idea about that stuff." [P5].

A low entry barrier is important.

P2 described the ideal interface as "easy to learn, hard to master," suggesting a low entry barrier with optional depth. P5 added that an interface should have a short "time to blink," meaning it should provide immediate visible feedback after interaction. Both statements emphasize the importance of fast responsiveness and progressive complexity.

Participants prefer short tooltips.

To support onboarding, participants preferred short, context sensitive help. P2 suggested "that when you hover over certain areas, it has a small help window in the lower left corner, and there's always an explanation of what you're currently hovering over."

Comprehensibility directly influences usability and trust.

Overall, this theme shows that comprehensibility is central to both usability and trust. Participants favored interfaces that are simple at first, avoid overload, and still allow deeper analysis when needed.

4.4.7 Interest Profiles Regarding Cognitive States

Participants mention both general cognitive states and context-specific emotions.

Certain cognitive and emotional states were especially relevant to the interviewed researchers. The most frequently mentioned state was cognitive load, also referred to as mental effort or mental workload. Other states were more context specific, including stress, flow, creativity, affect, creepiness, fatigue, or motion sickness, as well as emotions such as nervousness or surprise.

P5 suggested that EEG could be used to "obtain a direct measurement to determine baseline stress levels". Similarly, P1 expressed interest in establishing a baseline of general arousal to better track changes across sessions.

4.5 Design Principles based on Interview Findings

The seven themes reveal recurring patterns that extend beyond individual statements. Based on the themes derived from the interviews, design principles are defined (Table 4.1) by clustering the seven interview themes into higher-level, goals for the interface. Themes related to limited EEG knowledge and the need to understand results (competence boundaries, trust through transparency) were consolidated into *Transparency and Methodological Trust* (DP1). Themes describing the need for a low entry barrier and gradual complexity (motivation driven by usefulness, comprehensibility, and accessibility) informed *(Progressive) Accessibility* (DP2). Concerns about reliable recordings and uncertainty during data collection (low friction integration, data-related uncertainty) led to *Reducing Uncertainty in Data Collection* (DP3). Finally, the repeated focus on fitting EEG into existing study routines (EEG as supportive insight, workflow integration) resulted in *Seamless Integration into User Study Workflows* (DP4).

Themes are summarized into higher-level design principles.

These principles define the foundation of the interface and guide the definition of concrete functional requirements.

Principle	Design Implication
(DP1) Transparency and Methodological Trust	Analysis steps, parameters, and assumptions are visible, documented, and reproducible. The interface provides scientifically grounded defaults that users can inspect and change
(DP2) (Progressive) Accessibility	The interface supports quick first results and optional depth through progressive disclosure
(DP3) Reducing Uncertainty in Data Collection	The interface supports data quality with electrode prechecks
(DP4) Seamless Integration into User Study Workflows	The workflow minimizes interaction during recording and avoids distraction

Table 4.1: Interface design principles derived from the interviews.

4.6 Interface Requirements based on Design Principles

Design principles are translated into concrete requirements.

Based on the design principles, concrete interface requirements were derived. These requirements translate the principles into implementable features and system properties. A focused subset was selected to guide the first functional implementation of the interface (Table 4.2). Additional requirements identified during analysis but not implemented in this work are included in Appendix C.

Core requirements implement the design principles.

The requirements directly support the defined design principles in Table 4.1. Electrode contact check (R2) and artifact handling (R3) support DP3. Visible processing steps (R6), export of processed data (R9) and recommended defaults (R5, R7) support DP1. A step-by-step workflow mode (R8) supports DP2, while easy device connection (R1) and event marking (R4) support DP4.

Together, these requirements form the bridge between the empirical insights and the interface design described in the next chapter. The interviews informed design principles. These principles guided the selection of core requirements.

Requirement	Rationale
R1: Fast device connection	Reduces setup time and lowers the barrier for non-expert users during study preparation
R2: Electrode contact quality check	Ensures collected data is usable and reduces uncertainty about electrode placement
R3: Basic artifact detection and cleaning	Detection of common artifacts such as eye blinks and motion to remove obvious noise and improve the usability of recorded data
R4: Event marking during live recording	Allows mapping EEG activity to specific events, supporting time-based analysis
R5: Recommended defaults for processing parameters	Offers validated presets that reduce configuration effort and user uncertainty
R6: Display of data (pre)processing steps	Increases methodological transparency by showing how results are generated
R7: References to established methods	Builds confidence through methodological grounding and alignment with validated approaches
R8: Guided step-by-step workflow	Guides non-expert users through EEG usage
R9: Export of (processed) data	Enables further analysis and integration with questionnaires or log data
R10: Clear communication of interface capabilities and limitations	Prevention of overinterpretation and support for realistic expectations.

Table 4.2: Interface core requirements derived from the design principles.

The final interface implements these requirements. This chapter provided the empirical grounding for the interface design. The derived design principles and requirements serve as the conceptual and functional foundation for the implementation presented in the following chapter.

Chapter 5

Design and Implementation of the Final Interface

This chapter presents the final designed and implemented Interface based on the findings in Chapter 3 and Chapter 4. Section 5.1 first defines the scope of the final interface. Section 5.2 then describes the prototyping process and explains how the workflow structure and navigation concept were developed and refined. Section 5.3 outlines the final system architecture and data management. Section 5.4 presents the final interface workflow from the user's perspective and describes the five main screens of the interface. Section 5.5 concludes with implementation details of the data processing and Section 5.6 discusses the methodological and technical limitations of the final interface.

5.1 Scope of the Final Interface

The final interface is designed as a guided EEG workflow for conducting user studies. It primarily targets researchers in Human–Computer Interaction (HCI), but can also be used by researchers from other fields who apply EEG for study evaluation.

Interface supports an end-to-end workflow for non-experts in HCI study contexts.

The system is intended for non-expert users with limited prior knowledge of EEG. It supports the full study workflow: defining an experiment, connecting a device, checking electrode quality, recording data, and analyzing results. The application context is the evaluation of user studies, not clinical diagnosis or pure BCI control.

5.1.1 Intendet Experiment Condition

Interface assumes low-noise, stationary recordings.

The current implementation is designed for low-noise recordings. Participants should be seated and should move as little as possible. The study should avoid strong motion, typing, and speaking during recording. In high-noise contexts, muscle activity and motion can dominate the signal. In such cases, simple frequency-domain metrics become less reliable [Kyriaki et al., 2024].

5.1.2 Preprocessing and Data Analysis

Interface implements a basic preprocessing pipeline with filtering and ICA.

For low noise experiments, a common preprocessing approach includes manual rejection of noisy segments, band-pass filtering, and ICA as summarized by Kyriaki et al. [2024]. So the interface supports two preprocessing steps: frequency-based filtering and ICA.

Analysis focuses on continuous frequency-domain features.

The analysis of the data focuses on the continuous frequency-domain. Frequency-based features are comparatively robust to noise and temporal variability, which is especially important when working with consumer-grade EEG devices or in less controlled study environments [Kyriaki et al., 2024].

Interface computes engagement and cognitive load indices.

Building on the frequency-domain analysis, the interface computes a cognitive load index and an engagement index as defined in Section 2.4.2. These measures were selected because they can be calculated with a small number of electrodes [Conrad and Bliemel, 2016; Kumar and Kumar, 2016] and are widely used in applied EEG research, as shown in Chapter 3.

Overall, the interface is intended for exploratory and comparative analysis across tasks, conditions, and participants. Its scope prioritizes accessibility, transparency, and methodological clarity over analytical breadth. The limitations are stated explicitly in Section 5.6 to reduce the risk of overinterpretation.

Scope prioritizes accessibility and transparency.

5.2 Prototyping of the Interface

The interface was developed using a human-centered design approach following Norman Donald [2013]. The prototyping phase translated the defined design principles and core requirements from Chapter 4 into a clear and guided workflow suitable for HCI user studies. The prototypes are intentionally simple and, due to time constraints, focused on validating the workflow structure rather than the visual design.

Prototyping translated interview-based requirements into a guided workflow structure.

The first prototype explored the overall workflow of the interface using low-fidelity paper sketches. Based on the interview findings, it became clear that the interface should support all major phases of an EEG study in a structured way. First, users need a setup phase where they can connect the device and define basic information about the recording. Before starting the actual recording, a electrode contact check is required to reduce uncertainty and ensure usable data (DP3, R2). After that, the recording itself should be simple and require minimal interaction so that it does not disturb the user study (DP4). Since the interface aims to support an end-to-end workflow, the interface should also include an analysis phase where recorded data can be inspected and preprocessed. This structure implements requirement R8 (Guided step-by-step workflow) Based on these considerations, the prototype resulted in a linear workflow with five main screens: start screen, experiment setup and device connection, electrode contact check, recording, and analysis. The fixed order was chosen to ensure that critical steps occur in the correct sequence, such as verifying signal quality before recording. Figure 5.1 shows the start, recording and analysis screens of this initial pa-

Early prototype includes a linear, step-by-step workflow.

per prototype. The whole prototype can be found in the Appendix D.

Exploratory walkthrough suggested the workflow is understandable for non-expert users.

The prototype was used in an exploratory walkthrough with two HCI researchers. Although the small sample size does not allow general conclusions, the walkthrough did not reveal any major problems in the workflow. It showed that a clear step-by-step process helps users work correctly and reduces uncertainty, especially for non-experts.

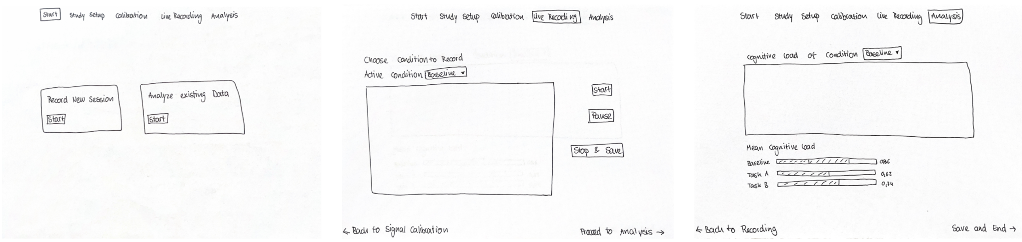


Figure 5.1: Start screen (left), recording screen (middle) and analysis screen (right) of the initial paper prototype.

Workflow adapted to support consecutive recording sessions.

The walkthrough also revealed the need to support consecutive recording sessions. In real studies, multiple participants are often recorded in sequence. So they do not need to analyse the data right away and want to avoid unnecessary repetition. At the same time, from a technical point of view, a new device connection should be required for each new recording to avoid issues caused by device inactivity or unplanned disconnection. Therefore the workflow was adapted so that users do not have to analyse the dataset after recording, but can navigate back to the start screen to start a new recording.

Users requested optional detail and time marker options.

The analysis view was perceived as clear and easy to understand. However, participants requested optional access to more detailed information and explicit event markers to support time-based analysis.

Navigation changed from top bar to sidebar to improve orientation and prevent errors.

During the walkthrough, one participant clicked on the "Start" item in the top bar in Figure 5.1 instead of the main start button that was meant to move to the next step. This showed that the navigation structure was confusing and could lead to errors. Different navigation concepts were explored using low-fidelity digital mockups. The top navigation was replaced with a left-aligned sidebar, as shown in Figure 5.2. The sidebar is no longer meant for active navigation. Instead, it indicates the current screen of the workflow. This helps users understand where they are in the process and reduces the risk of skipping important steps. In addition, the top bar can now be used to display constant information, such as the experiment name and recording number.

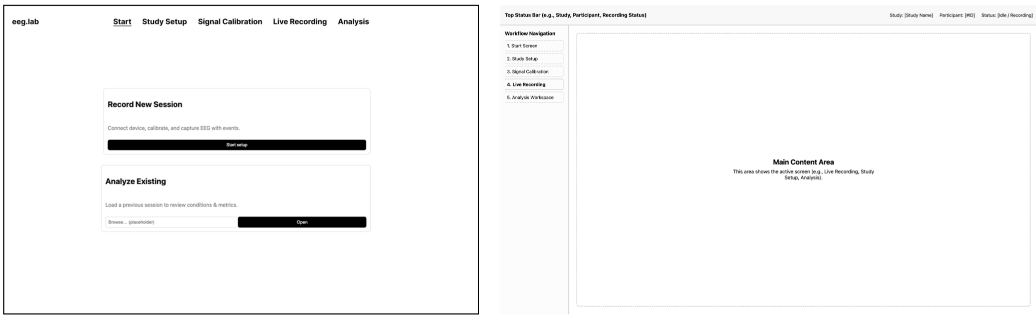


Figure 5.2: Start screen mockups with top navigationbar (left) and side navigationbar (right).

Overall, the prototyping phase validated the linear workflow structure and informed key interaction and navigation decisions. Building on this validated workflow and navigation concept, the following section presents the implementation of the final interface.

Prototyping informed key navigation decisions and validated the linear workflow.

5.3 Final Interface System Architecture and Data Management

This section describes the technical foundation of the interface, including its software architecture, data acquisition, and storage model. The interface is implemented as a web-based application building on related work from Section 3.4.1 to lower setup effort and improve accessibility. The frontend is built with Vue 3 and TypeScript. The backend is implemented using FastAPI, which was chosen for the asynchronous processing. The code is publicly available to support reproducibility.

LSL acquisition runs in a separate process to ensure stable synchronized streaming.

The EEG data acquisition is handled through the Lab Streaming Layer (LSL)[1] and runs in a separate operating system process. This splitting isolates the data streaming from the frontend and backend load and ensures stable, time-synchronized acquisition. LSL also allows the interface to be extended to additional devices in the future without adding big changes to the interface logic.

Offline processing uses MNE library.

The data is offline processed and analysed using the MNE Python library[2] which is widely used in research [Gramfort et al., 2013]. It provides all the main processing steps, from preprocessing to statistical analysis. It also supports advanced analysis methods like supervised learning. This allows the analysis pipeline in the interface to be extended without changing the software framework. Because MNE is fully based on Python and does not rely on MATLAB or other proprietary software, it improves accessibility and reproducibility. In addition, MNE data objects are compatible with other Python-based libraries such as NeuroKit2[3], which allows additional analyses to be added in the future.

Local CSV-based storage keeps data readable.

All recorded data is stored locally in a fixed directory structure:

```
/data/
 experimentName/
     RecordingN/
         metadata.csv
         raw_samples.csv
         events.csv
         cleaned_samples.csv
         metrics_output.csv
```

Separate files for raw, events, cleaned data, and metrics improve traceability and comparison.

Each experiment contains one or more recordings, which could be logically mapped to a participant. Each recording includes raw data, metadata, event markers, cleaned data, and a metrics file. The interface runs locally, so the data is

1 `labstreaminglayer.org/#/`, *last accessed 05.02.2026*
2 `mne.tools/stable/index.html`, *last accessed 05.02.2026*
3 `pypi.org/project/neurokit2/`, *last accessed 01.02.2026*

not shared externally, which supports data privacy and security. The metadata file stores experiment and recording parameters such as device name, sampling rate, channel names, and defined conditions. The raw EEG samples are stored separately from event markers. Event files contain recording starts and stops, and manually set markers. After preprocessing, the cleaned data is saved separately. Derived metrics and the processing parameters used to compute them are stored in a metrics output file. This separation supports transparency and enables clear comparison between raw, cleaned, and derived data. CSV files were selected as the main format for storing data since they can be read by humans and are simple to review and export. This supports accessibility for HCI researchers who may not be familiar with specialized EEG formats.

Export supports reproducibility.

The whole folder containing the above files can be exported at the end of the analysis. This makes the data and analysis process transparent and reproducible, which supports design principle DP1 (Transparency and Methodological Trust) and implements requirement R9 (Export of (processed) data).

5.4 Final Interface Workflow

Interface is split into five screens.

After outlining the technical foundation, this section describes the interface from the user's perspective. The final interface is organized into five main screens that follow the workflow evaluated through the prototypes in Section 5.2:

Start Screen: Create a new recording or open and analyze an existing dataset.

Setup Screen: Define the experiment and conditions and connect the Muse device.

Electrode Contact Check Screen: Record a baseline and verify electrode contact quality with live feedback.

Recording Screen: Record EEG for each condition.

Analysis Screen: Preprocess data and inspect bandpower, PSD, and index results.

Each screen displays only the functions required at that stage, which implements requirement R8 (Guided step-by-step workflow). Once all actions on a screen are completed, users can proceed to the next screen. This provides a clear sense of progress and completion, which supports Shneiderman [2003] rule to design dialogs to yield closure.

Tooltips support guidance.

While the clear workflow supports first-time users, it is intentionally designed to avoid becoming obstructive for recurring use. So instead of always visible explanations, the guidance is primarily delivered through contextual tooltips that appear when hovering over headings and key interface elements. These tooltips explain the purpose of functions and provide a brief background, which should help users to understand what actions are required.

Continuous feedback and safeguards reduce errors and uncertainty.

The interface also provides continuous status feedback. For example, displaying the remaining recording time. Buttons are also automatically disabled while required inputs are missing or while processing is running, which prevents invalid actions and reduces errors. The interface is responsive and adapts to different screen sizes, for example, by collapsing the sidebar into a burger menu on smaller displays. Error handling is implemented through clear and informative messages that notify users when problems occur and explain how they can be resolved. This supports robust interaction and reduces uncertainty during use.

Documentation explains interface, processing choices and limitations to manage expectations.

The documentation screen (Figure 5.3) provides an overview of the interface workflow and detailed descriptions of the implemented data processing functions. It also explicitly states what the interface can and cannot do, which manages users expectations and reduces the risk of overinterpretation, implementing R10 (Clear communication of interface capabilities and limitations). All default settings of the underlying MNE functions and processing parameters are documented, supporting DP1 (Transparency and Methodological Trust) and implementing requirement R6 (Display of data (pre)processing steps). It also includes a section with references to papers and books, which could be useful, and on which the processing pipeline is oriented, implementing R7 (References to established methods).

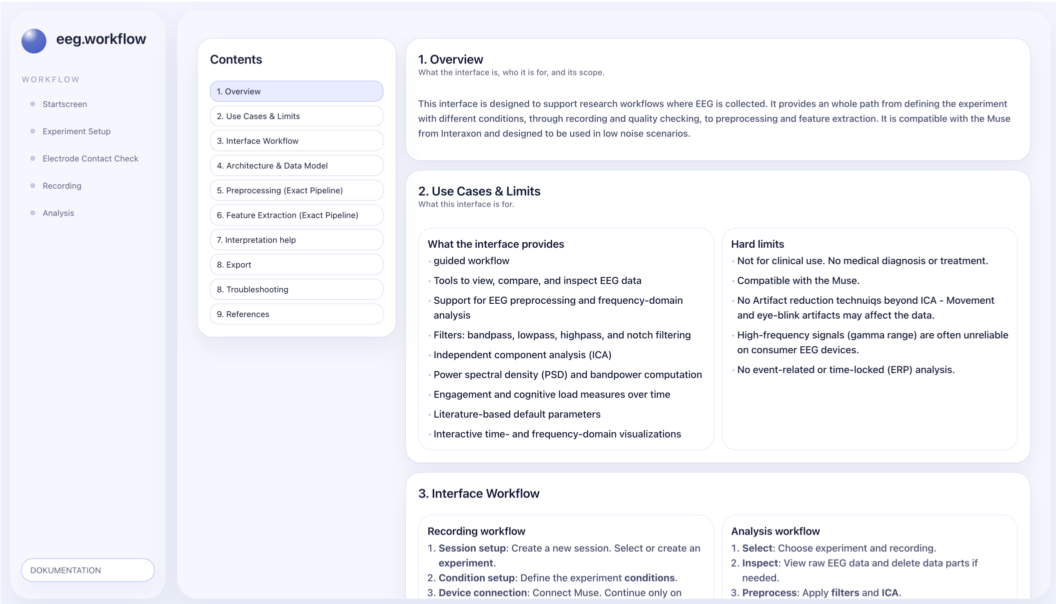


Figure 5.3: The interface dokumentation screen includes.....

5.4.1 Screen 1: Start

The Start screen (Figure 5.4) serves as the entry point of the interface. Users can either start a new recording or analyze an existing dataset. The screen provides an overview of stored recordings on the local server and allows users to upload data as a ZIP file. A demo data set for a first data analysis exploration is also provided.

Start screen supports creating new recordings and loading or uploading datasets.

5.4.2 Screen 2: Experiment Setup

During the setup, users must create a new experiment with a name and conditions (with optional duration) or select an experiment and its conditions from the server. This supports consistent experiment and condition naming across recordings and prevents label drift.

Setup screen defines experiments and conditions and manages device connection.

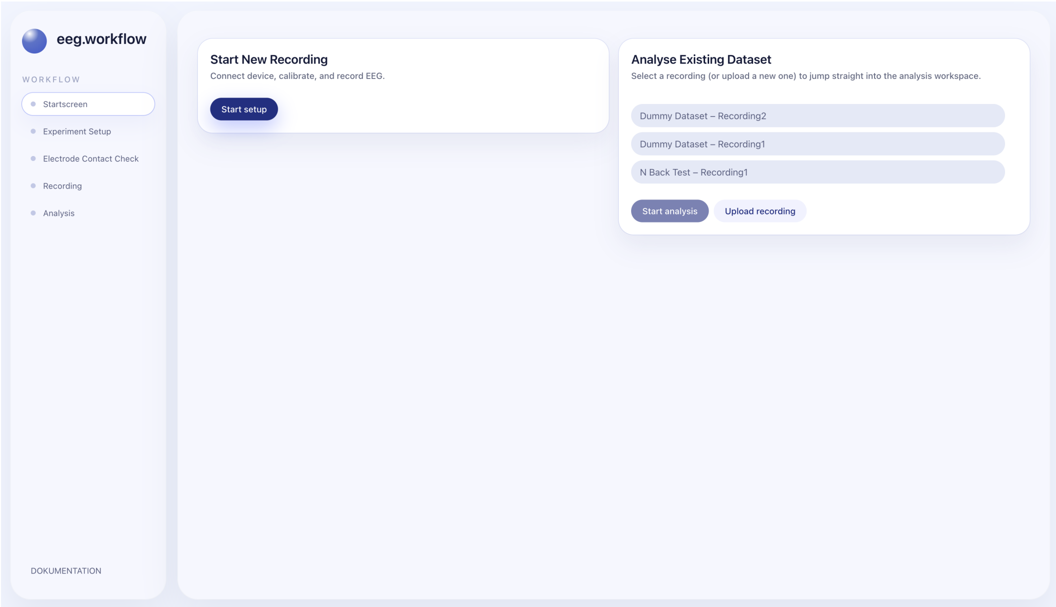


Figure 5.4: On the Start Screen, users can start a new recording session, analyze an existing dataset, or upload previously recorded data.

A baseline condition is added by default for every recording and is used as a reference for metric comparison and for computing the electrode contact check on the next screen. That reduces setup effort and prevents missing comparisons. The term baseline follows Nielsen [1994] heuristic of matching the interface to the real world by reflecting established terminology that is familiar to HCI and EEG researchers. Although the primary use cases are user studies, the term experiment was chosen intentionally. This more general wording makes the interface applicable to a wider range of contexts and follows Shneiderman [2003] rule to cater to universal usability.

To connect a device, the users have to first scan for available devices and select the target device from all devices listed. Once the device is successfully connected, key device parameters such as sampling rate and channel configuration are stored as part of the recording in the metadata csv file. The clear scan–select–connect flow minimizes setup overhead and supports DP4 (Seamless Integration into User Study Workflows) and implements R1 (Fast device connection).

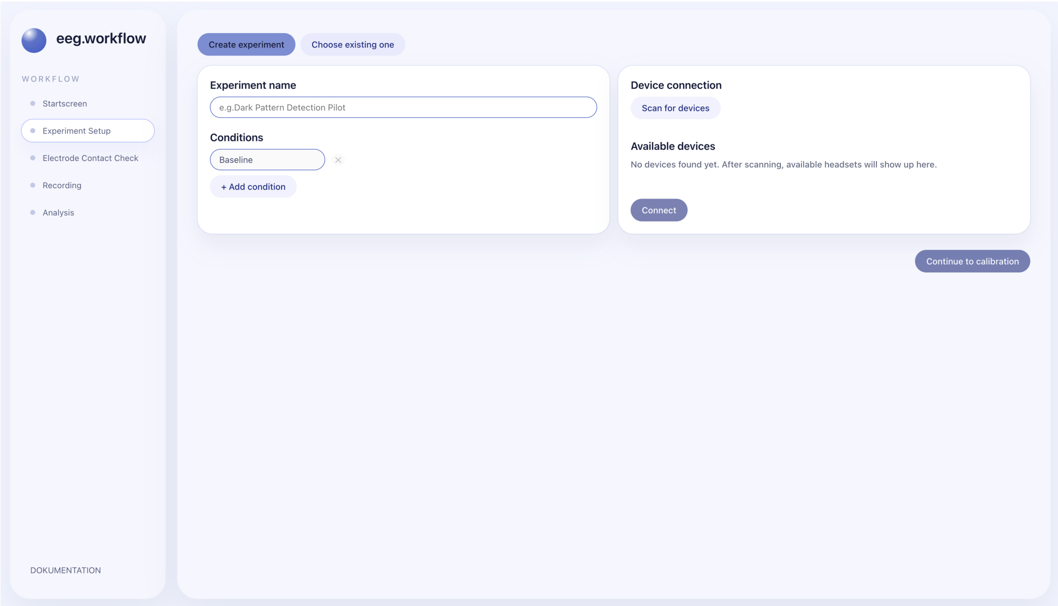


Figure 5.5: On the experiment setup screen the experiment can be defined and the EEG device can be connected.

5.4.3 Screen 3: Electrode Contact Check

Electrode contact check screen records a baseline and provides feedback on electrode contact quality.

The electrode contact is the next mandatory step before data can be recorded. The screen (Figure 5.6) provides a live EEG plot that allows users to verify that a data stream is present and stable. It also includes a short video tutorial[4] from InteraXon's YouTube channel that explains how to properly fit the Muse headset. This supports users with little EEG experience and reduces setup errors.

On this screen, users are required to record a baseline. This baseline serves as the data source for computing the electrode contact quality, which implements R2 (electrode contact quality check). The results are shown via a graph and color-coded indicators. That provides an immediate and easy-to-understand feedback, and users can quickly identify problematic channels, adjust the headset if necessary, and rerecord the baseline. By separating the electrode contact check from recording, the interface reduces the risk of

[4] www.youtube.com/watch?v=v8xUYqqJAIg, *last accessed 01.02.2026*

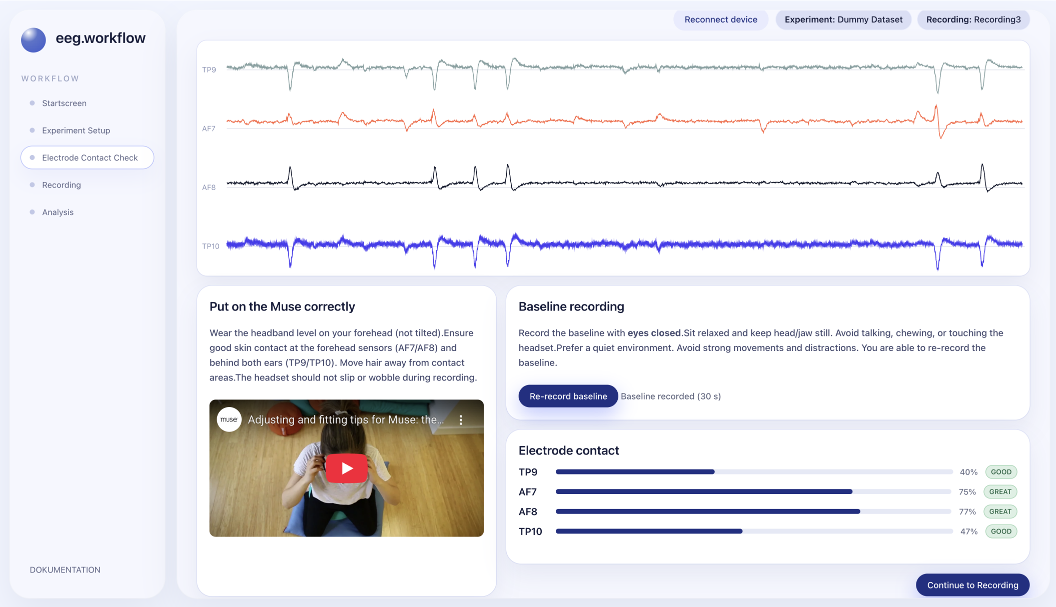


Figure 5.6: The electrode check screen shows a live stream of the EEG data and includes the baseline recording.

collecting unusable data and supports DP3 (Reducing Uncertainty in Data Collection).

5.4.4 Screen 4: Live Data Recording

Recording screen minimizes interaction and supports condition-based recording with markers.

On the recording screen, users can record EEG data for the different predefined conditions (Figure 5.7). The screen is intentionally simple to minimize interaction and distraction during the recording. It displays a live view of the incoming EEG signal, the currently selected condition, start and stop controls, a marker button, and the current recording time. By plotting live data and showing the recording time, it follows Nielsen [1994] principle of visibility of interface status. A lost device connection is immediately visible because the live signal stops updating. The device can then be reconnected. Also, users can navigate back to the Electrode Contact check if a recheck is needed.

Disabling non-essential elements during recording reduces operator load and prevents errors.

Once a recording starts, non-essential interface elements are visually disabled. Only functions required for monitor-

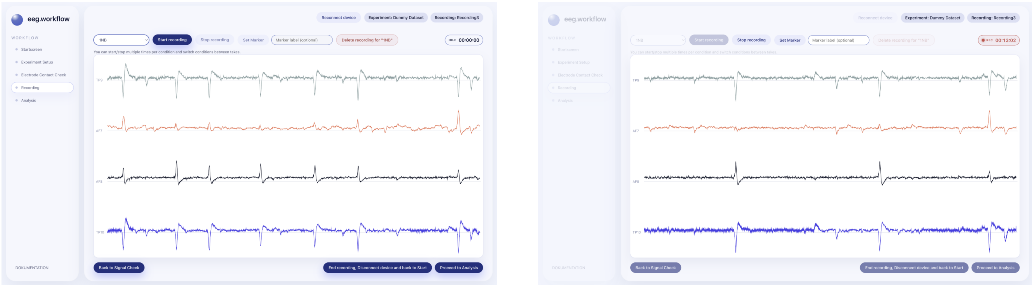


Figure 5.7: The recording screen (left) and during active EEG recording (right). The screen provides a live view of the EEG signal, displays the current condition and recording status, and supports manual event marking. During active recording, non-essential screen elements are visually disabled.

ing and event marking are then available (as can be seen in Figure 5.7). This reduces the risk of accidental interaction and lowers operator load during the study, which supports DP4 (Seamless Integration into User Study Workflows). If a recording duration is defined, the recording stops automatically. This supports standardized recording and reduces the risk of inconsistent recording lengths. Automatic stopping also removes the need for manual interaction at the end of a recording, also supporting DP4.

The interface only includes Start and Stop controls and does not provide a pause function. This reduces complexity and avoids unclear recording states during live studies. Instead of pausing, users can start and stop recordings multiple times within the same condition. This allows the same flexibility as a pause function, while keeping each recording segment clearly defined and easier to manage during analysis.

Simplified control logic avoids ambiguous recording states.

During recording, users can place event markers either by clicking a button or by using a keyboard shortcut. Optional event labels can be defined in advance. These markers allow later alignment of EEG data with events and implement requirement R4 (Event Marking During Live Recording).

Event markers enable alignment of EEG and study events.

EEG data is only stored in the CSV while a recording is active. Data received from the EEG device outside an explicit

Data storage is strictly bound to active recording phases.

recording phase is not stored. This prevents the accumulation of irrelevant data and ensures that all saved samples are clearly associated with a specific condition and recording session. If a recording was started too early, assigned to the wrong condition, or something happened in the recording, the recorded data for a condition can be deleted directly within this screen and then be rerecorded.

Automatic device disconnection ensures clean session boundaries.

After the recording is finished and the user leaves the recording screen (either to return to the start screen or to proceed to data analysis), the EEG device is automatically disconnected. This prevents unintended background streaming and ensures that each recording session starts with a fresh device connection. This behavior is intentional, as the Muse headset tends to disconnect after longer periods. Forcing a clean disconnect helps ensure reliable reconnection for the next recording.

5.4.5 Screen 5: Analysis of the Data

Analysis screen includes preprocessing and feature computation.

The analysis screen supports the EEG data processing. The theoretical processing pipeline described in Chapter 2 is translated into the following steps. The analysis screen is designed to help users understand how the data is processed and how the metrics are computed. Preprocessing and feature computation are intentionally placed on the same screen. This design allows users to visually connect preprocessing steps with their effects on the data and derived features. At the same time, it enables users to revisit and adjust preprocessing parameters after inspecting the computed features.

Step-wise structure enforces deliberate preprocessing and supports transparency.

The preprocessing part is organized into clear steps that lead to an overview of the frequency-domain features and computed metrics. This explicit step structure supports DP1 (Transparency and Methodological Trust) because users can clearly see how each processing step affects the data. At the same time, it supports progressive accessibility (DP2) and implements R3 (Basic Artifact Detection and Cleaning). By making the processing steps explicit and requiring users' confirmation, the interface makes it clear that preprocessing changes the data and should be handled consciously.

Users can rely on the provided default settings to get quick results. To further support exploration without risk, all preprocessing and feature computation settings provide a *reset to default* option.

Parameter defaults are available.

Preprocessing Step 1: Raw Data Inspection

In the first step (Figure 5.8), users can inspect the raw EEG data for each recorded condition. Manual data cleaning is a common but subjective and time-consuming EEG preprocessing method based on visual inspection, often used to remove segments affected by artifacts such as eye movements or muscle activity [Kyriaki et al., 2024]. The interface supports manual deletion of user-defined time ranges, for example, when disturbances occur during a study.

Manual inspection enables controlled artifact rejection before processing.

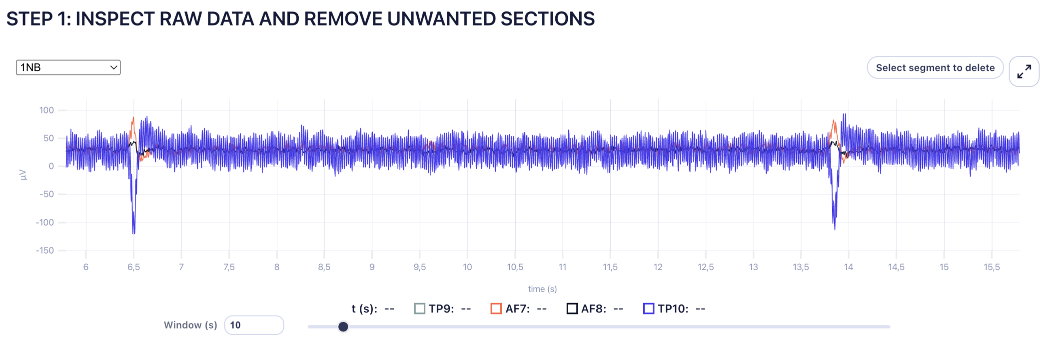


Figure 5.8: Interface Preprocessing Step 1 - Inspection. In this step, users can inspect the recorded raw EEG data for each condition and manually remove time ranges, as well as open the graph in full screen.

Preprocessing Step 2: Filter Application

In the second step (Figure 5.9) users can apply frequency filters to the raw EEG data. Recommended filter settings are enabled by default, implementing requirement R5 (Rec-

Default filters offer guidance.

ommended Defaults for Processing Parameters). Users can adjust these settings if needed.

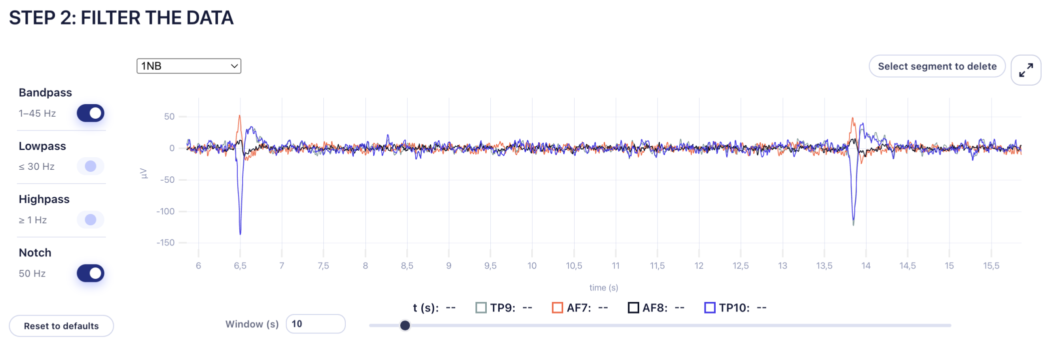


Figure 5.9: Interface Preprocessing Step 2 - Filter. In this step, users can apply filters and see the changes immediately.

Preprocessing Step 3: ICA Application

In the third step, the filtered EEG data is automatically decomposed into independent components using ICA. Users then have to inspect the resulting components in the image, as can be seen in Figure 5.10 and manually decide which components to exclude. Tooltips provide guidance to support this decision.

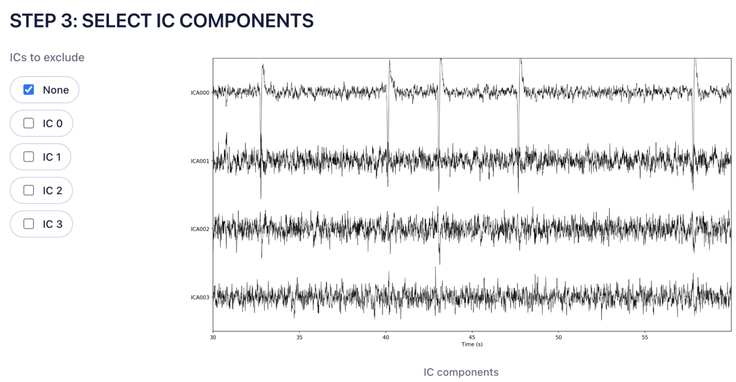


Figure 5.10: Interface Preprocessing Step 3 - ICA. Here the IC components are displayed to visually examine them and choose the ones to be excluded.

Manual component exclusion was chosen because the effectiveness of ICA depends on the number of electrodes and is therefore limited for consumer-grade devices such as the Muse. In such settings, fully automatic component rejection is often unreliable [Kyriaki et al., 2024]. Despite this limitation, ICA was included because Blind Source Separation methods, and especially ICA, are the dominant artifact removal technique in EEG-based cognitive load research [Kyriaki et al., 2024].

Preprocessing Step 4: Preprocessing Settings Application

After Steps 1 to 3 are completed, the user must apply the processing settings by pressing the "Apply preprocessing settings" button (see Figure 5.11). The step also shows which settings are about to apply to the data. This design ensures that preprocessing is performed deliberately. When Filters or ICA changes, the users have to reapply the processing settings and receive a clear notification for that (Figure 5.12) supporting Shneiderman [2003] rule of informative feedback.

An explicit application step ensures that data changes are applied consciously while keeping the interface responsive.

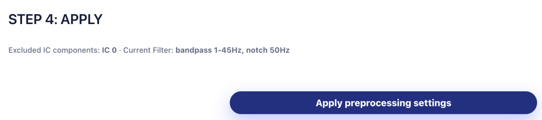


Figure 5.11: Interface Preprocessing Step 4 — Apply settings. In this step users have to aplly the settings

The preprocessing pipeline displayed in the sidebar (Figure 5.12) makes the sequence of processing steps explicit and implements requirement R6 (Display of Data (Pre)processing Steps).

Sidebar pipeline visualizes the preprocessing steps.

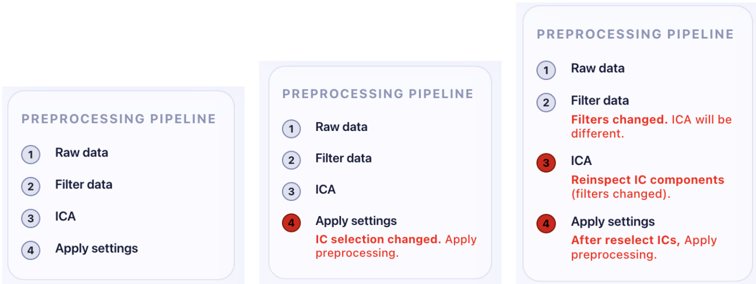


Figure 5.12: Overview of the processing pipeline across three stages.

Separation of raw and cleaned data ensures transparency.

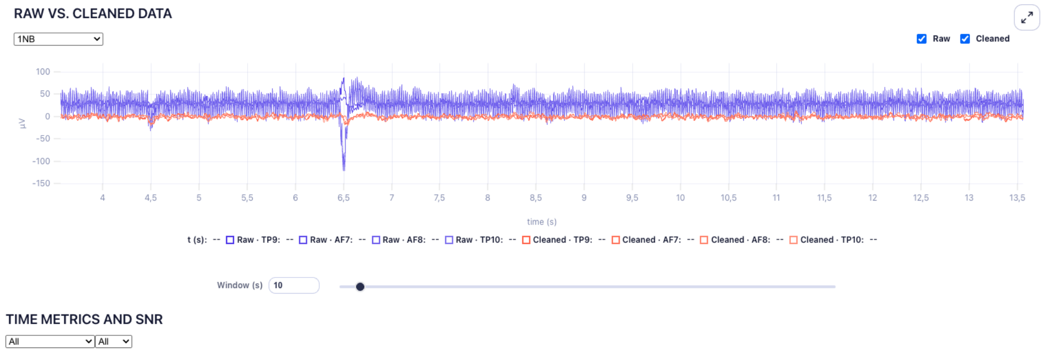
The decision to require explicit application of preprocessing settings is motivated by both performance and data integrity considerations. To keep the raw data, the interface generates a separate cleaned data file. So, preprocessing EEG data from the CSV can be computationally expensive for larger datasets. Automatically reapplying preprocessing would introduce unnecessary latency and reduce interface responsiveness.

Once preprocessing is applied, the interface computes frequency-domain features and ratio metrics as described in Chapter 2 and visualizes them at different levels of detail.

Raw vs. Cleaned Data Visualization

Direct raw–cleaned comparison supports transparency.

After preprocessing is complete, the user can inspect the cleaned EEG data here. For an easy visual comparison, the raw data and cleaned data are plotted in one graph (Figure 5.13). This allows the users to clearly see how filtering and artifact removal affect the signal.



	Samples	Min	Max	Peak-to-Peak	Mean	Variance	Std	RMS	SNR (dB)	SNR (lin)
RAW	67481	-136.719	138.672	275.391	28.265	196.841	14.030	31.599	0.776	1.196
CLEANED	67481	-48.697	55.959	104.656	-0.013	16.414	4.051	4.051	21.857	153.363

Figure 5.13: Graph which enables a comparison of the raw and cleaned EEG data. Visually and through the displayed time metrics.

In addition to the visual comparison, time-domain metrics are calculated for both the raw and the cleaned data. These include the number of samples, minimum and maximum values, peak-to-peak amplitude, mean, variance, standard deviation, and root mean square (RMS). The SNR is also computed and shown here for direct comparison between raw and cleaned data in both dB and linear units.

Time-domain metrics and SNR provide quantitative quality indicators.

Engagement and Cognitive Load over Time

The engagement index and the cognitive load index can be viewed over time (Figure 5.14). Both indices are calculated from the cleaned EEG data and shown as continuous curves for the selected condition. This view helps users see how engagement and cognitive load change during the task. It makes it easier to identify trends, peaks, or changes across different phases.

Time-resolved indices enable interpretation of cognitive changes.

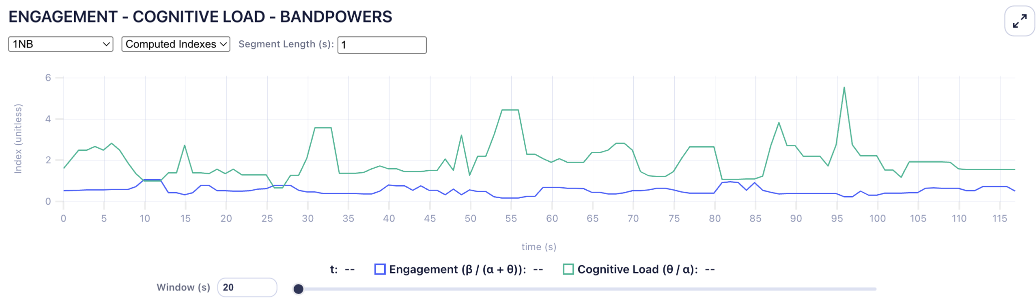


Figure 5.14: This graph shows the time-resolved engagement and cognitive load indices.

Window length parameter controls temporal resolution and noise trade-off.

The time resolution depends on the Segment Length parameter which defines the length of the analysis window. Shorter windows react faster to changes but are more noisy. Longer windows create smoother curves but may hide short-term fluctuations.

Indices Summary per Condition

Aggregated summaries support condition-level comparison.

Below the time graph, an aggregated summary of the engagement and cognitive load indices is shown for each condition (Figure 5.15). In addition to the mean index value per condition, the interface also displays the difference relative to the baseline. This allows users to compare cognitive states across conditions.

INDEXES SUMMARY

	Engagement		Cognitive Load	
Condition	Mean	Δ to baseline	Mean	Δ to baseline
Baseline	0.4994	–	0.7620	–
1NB	0.6740	+0.1745	1.7047	+0.9428
2NB	0.7422	+0.2427	1.6145	+0.8525

Figure 5.15: Overview of engagement and cognitive load indices per condition.

Furthermore, the indices can be compared across multiple recordings within the same experiment (Figure 5.16).

Cross-recording comparison supports experiment-level analysis.

EXPERIMENT WIDE COMPARISON

Condition	Recording1		Recording2	
	ENG	CL	ENG	CL
Baseline	0.3139	0.5653	0.4994	0.7620
1NB	0.5386	1.9863	0.6740	1.7047
2NB	0.6500	1.8422	0.7422	1.6145

Figure 5.16: Comparison of engagement and cognitive load indices across recordings

Bandpower over Time

By switching from *Computed Indices* to *Frequency Bands* in the time graph, the bandpower values for individual frequency bands over time (Figure 5.17) are displayed. This makes it possible to inspect the individual frequency bands in detail.

Bandpower level time series allow inspection beyond computed indices.

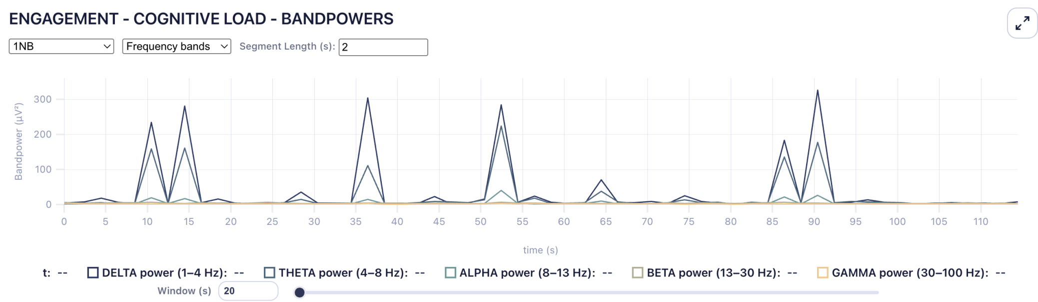


Figure 5.17: This graph shows the bandpower over time for all frequency bands.

PSD and Bandpower Exploration

The collapsible exploration section includes a PSD graph and a bandpower summary and is intentionally hidden by

Collapsible exploration section implements progressive disclosure for advanced analysis.

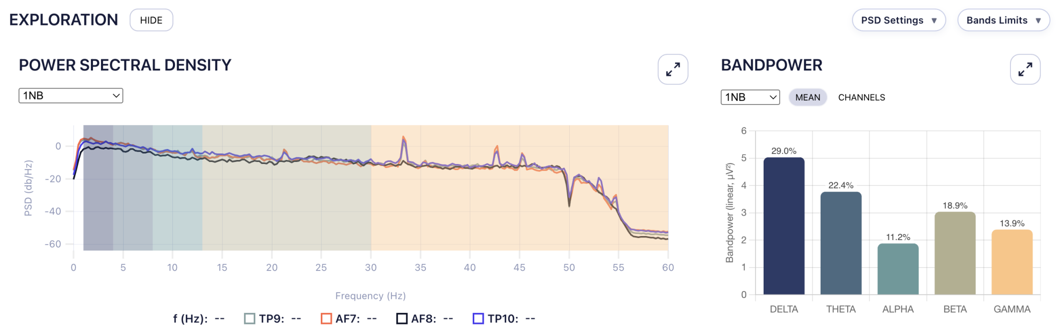


Figure 5.18: The left graph shows the PSD for the selected condition. The right graph shows the overall bandpower.

default to support DP2 (Progressive Accessibility). The PSD and bandpower shown here summarize longer segments of data (the full condition), whereas the index-over-time and bandpower-over-time plots operate on short non-overlapping windows.

PSD visualization supports spectral analysis and channel comparison.

The PSD graph (Figure 5.18) shows the power spectral density across frequencies for the selected condition and for each EEG channel.

Users can adjust the PSD parameters which are hidden by default and can be expanded when needed, supporting DP2 (Progressive Accessibility). When users change these PSD settings, both the PSD plot and the calculated bandpower values update automatically. This direct connection helps users see how changes in spectral estimation influence the resulting bandpower measures.

BANDPOWER

All

Condition	delta	theta	alpha	beta	gamma
Baseline	7.01 37.3%	3.34 17.7%	4.46 21.9%	3.89 21.0%	0.67 3.5%
1NB	4.69 32.3%	3.91 27.3%	2.04 14.2%	3.10 22.6%	0.56 3.8%
2NB	5.67 36.0%	3.39 22.9%	1.92 13.1%	3.45 24.5%	0.57 3.6%

Figure 5.19: This table displays the bandpower for every frequency band for every condition.

The bandpower is shown either as bar charts (for one selected condition) or as a table for comparison between conditions as shown in Figure 5.19. Absolute bandpower shows the total power within each frequency band. Relative bandpower divides the bandpower by the overall power, which makes it easier to compare results.

Bandpower summarie provide absolute and relative spectral measures.

By default, bandpower values are averaged across channels to provide a robust and easy-to-read overview for non-expert users. For more detailed inspection, channel-wise visualizations are available as shown in Figure 5.20. This option is included because, as discussed in Chapter 2, certain oscillatory patterns are more commonly reported over specific scalp regions. With Muse's limited electrodes the interface therefore offers channel-wise views to help users assess whether observed effects are localized to specific electrodes.

Channel-wise visualization supports spatial interpretation with limited montages.

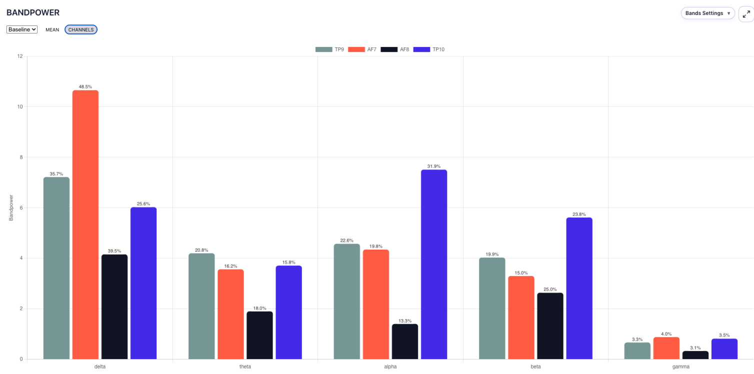


Figure 5.20: This graph displays the bandpower for every individual channel for the selected condition.

A consistent color scheme is used throughout the interface. The same colors are applied for each channel, each frequency band, and for the engagement and cognitive load indices across all visualizations. This consistency supports visual continuity, reduces cognitive load when comparing plots, and helps users quickly recognize corresponding signals and features in different views.

Consistent color coding supports cross-view comparability.

Overall, the analysis screen combines guided preprocessing with transparent feature computation, enabling both

Interpretation help provides guidance and warns against overinterpretation.

quick inspection and deeper exploration. To help users interpret the shown bandpower and ratio-based indices, the analysis screen includes a small "Interpretation Help" button at the bottom of the page. On hover, it shows short qualitative hints and reminds users that EEG metrics are correlational and depend on the context.

5.5 Final Interface Functionality Implementation details

This section explains how the main functions of the interface are implemented on a technical level. It describes the concrete methods and parameter choices that are used for the electrode contact check, preprocessing, and feature computation.

5.5.1 Electrode Contact Check

Impedance-based quality checks are replaced by signal-based heuristics due to hardware constraints.

Direct impedance-based electrode checks are the standard in EEG systems but require specialized hardware and interrupt recording [Casadei and Ferrero, 2022; Tautan et al., 2014; Zhao et al., 2018]. They are not available on the device used in this work. When impedance measures are unavailable, EEG systems typically rely on indirect signal-quality indicators such as signal variability or noise characteristics.

Variance and standard deviation serve as pragmatic stability indicators.

Each signal metric captures different aspects. Signal variance and standard deviation describe amplitude fluctuations around the mean and are commonly used as simple indicators of signal stability and data quality in applied EEG systems [Kyriaki et al., 2024]. Several studies discard data segments when channel values deviate strongly from their typical distribution, often defined relative to the mean and standard deviation [Nakamura et al., 2021; Ortiz et al., 2020].

Baseline variability is used as a qualitative contact indicator.

Following this practice, the interface uses baseline signal variability as an indicator of electrode contact quality. Assuming the participant remains still, good electrode contact produces stable signals, while poor contact increases variability. Since variability can have multiple causes, this check is treated as a qualitative indicator rather than a strict acceptance criterion.

The computation method and threshold ranges are adopted from Desai et al. [2022]. Let $x_c[t]$ be the raw EEG signal of channel c at sample t, with N samples in the baseline recording. Each channel is first normalized by removing its mean:

Contact quality is computed via mean removal and channel-wise standard deviation.

$$\tilde{x}_c[t] = x_c[t] - \frac{1}{N}\sum_{t=1}^{N} x_c[t].$$

This removes constant offsets. The standard deviation of the zero-mean signal is then computed:

$$\sigma_c = \sqrt{\frac{1}{N}\sum_{t=1}^{N} \tilde{x}_c[t]^2}.$$

Each channel is then mapped to a qualitative contact category using fixed thresholds: *Great* ($\sigma_c \leq 15\mu$V) , *Good* ($15 < \sigma_c \leq 30\mu$V), *Okay* ($30 < \sigma_c \leq 45\mu$V), and *Poor* ($\sigma_c > 45\mu$V). Channels with $\sigma_c \leq 45\mu$V are considered acceptable for recording.

The fixed thresholds are hardware dependent and influenced by sampling rate, referencing, electrode type, and amplification. They should therefore be understood as setup-specific heuristics. Future work should investigate adaptive or device-independent quality metrics.

5.5.2 Processing Pipeline

The following section provides details on the technical implementation of the processing pipeline. It is important to note that there are no standardized optimal processing pipelines across tasks, devices, and study contexts for

Processing pipeline reflects literature-based defaults but no universal standard exists.

preprocessing [Kyriaki et al., 2024] and feature extraction [Rabbi et al., 2009].

Parameter defaults provide orientation but do not define a gold standard.

Instead, studies report a wide range of pipelines adapted to specific constraints [Putze et al., 2022]. So default parameters can give a first direction, but should still be used carefully, as well as the implemented feature extraction pipeline is just one of many, and should be extended in the future.

Resampling and re-referencing are excluded due to fixed device configuration.

Resampling and re-referencing as explained in 2.3.1 are not implemented, as the Muse headset uses a fixed sampling rate of 256 Hz and a predefined reference configuration that is sufficient for the frequency ranges analyzed in this work.

Filter

Filtering relies on MNE's zero-phase FIR implementation.

A band pass filter is applied by default with cutoff frequencies of 1–45 Hz which is based on related work [Ahmed et al., 2025; Stingl and Knierim, 2024]. If the band pass filter is disabled, optional high pass and low pass filters can be applied. Also, a notch filter at 50 Hz is applied by default. All filters rely on MNE's filter functions mne.filter.filter_data[5] and its default settings (zero phase FIR filtering, automatic filter length, and automatic transition bandwidth selection). For notch filtering, MNE defaults uses zero-phase FIR filtering with automatically determined notch width and transition bands centered on the specified line noise frequency. Related work identified FIR as the dominant filter in cognitive load studies and for artifact reduction [Dhole et al., 2025; Kyriaki et al., 2024; Stingl and Knierim, 2024].

ICA

ICA relies on MNE's implementation.

ICA is implemented using mne.preprocessing.ICA[6]. The parameters used for the ICA computation are:: `ICA(n_-`

5 `mne.tools/stable/generated/mne.filter.filter_data.html`, *last accessed 01.02.2026*

6 `mne.tools/stable/generated/mne.preprocessing.ICA.html`, *last accessed 01.02.2026*

`components=4, max_iter="auto", random_state=97)`. The parameter `n_components=4` defines the number of independent components to be estimated. For the Muse device, which provides four EEG channels, this uses the maximum possible number. The setting `max_iter="auto"` lets MNE automatically control convergence by stopping the algorithm once convergence is reached or an internal iteration limit is exceeded, with a max of 1000. The fixed `random_state=97` ensures reproducible results by keeping the random initialization of the ICA solver constant across runs. All other MNE defaults were retained. After that, the function mne.preprocessing.ICA.apply[7] is used to exclude the components selected by the user.

Time Metrics

The time-domain metrics are computed directly from the EEG time series. For each channel, the following metrics are calculated based on the full data of one condition: minimum value, maximum value, mean, variance, standard deviation, root mean square (RMS), and peak-to-peak amplitude. The variance is computed as the population variance, and the standard deviation is derived as its square root. The RMS value is calculated as the square root of the mean squared signal amplitude. All metrics are first computed individually for each channel and each condition. Aggregate metrics are derived by averaging the channel-wise values except of min, max, and peak-to-peak which are computed jointly across all channels or all conditions.

Time-domain metrics summarize amplitude and variability characteristics.

Signal to Noise Ratio

The SNR is computed separately for each EEG channel c. The signal component S_c is defined as the band-pass filtered EEG in the 1–30 Hz range, which covers the main frequency bands relevant for cognitive load and engagement as explained in Chapter 2. The noise component N_c is ob-

SNR is calculated via RMS ratio.

[7] `mne.tools/stable/generated/mne.preprocessing.ICA.html#mne.preprocessing.ICA.apply`, *last accessed 01.02.2026*

tained by subtracting the filtered signal from the original data. The linear SNR is calculated as the ratio of the root-mean-square (RMS) amplitudes:

$$\mathrm{SNR}_c = \frac{\mathrm{RMS}(S_c)}{\mathrm{RMS}(N_c)}.$$

and converted to decibles:

$$\mathrm{SNR}_{c,\mathrm{dB}} = 20 \log_{10} \times \mathrm{SNR}_c.$$

PSD Computation

PSD estimation uses MNE Welch's method.

The PSD is computed using Welch's method with segmentation and windowing as described in Chapter 2 and is implemented with the MNE function: mne.time_frequency.psd_array_welch[8]

User-adjustable parameters control spectral resolution and range.

In the PSD graph, users can adjust `fmin` (default: 0 Hz), `fmax` (default: 45 Hz), `Segment length` (default: 5 seconds), and `Overlap` (default: 50%). The parameter `Segment length` defines the Welch segment length, and `Overlap` controls how much Welch segments overlap. The choice of fmax=45 Hz reflects the limited reliability of consumer-grade EEG at higher frequencies and the focus on theta, alpha, and beta bands. Beyond these parameters, the MNE defaults are used. These include a Hann window, removal of the DC offset, constant detrending of each segment, and averaging PSD estimates across segments using the mean.

Bandpower Computation

Bandpower integrates PSD values within configurable frequency limits.

Bandpower is computed from the PSD by integrating power within a defined frequency band. Default frequency band limits are provided based on Teplan et al. [2002], but users can adjust them in the Bandpower graph. This is important because frequency band boundaries can vary between individuals (see Chapter 2) and differ in related work like in Hassib et al. [2017].

[8] `mne.tools/stable/generated/mne.time_frequency.psd_array_welch.html`, *last accessed 01.02.2026*

All frequency bins within the selected band are extracted, and the PSD values are integrated over this range using: scipy.integrate.simpson[9] with a spacing of integration points defined as `dx=freq_res`. For the bandpower bar chart, bandpower is computed from the PSD (using the PSD settings described above) over the full data of the selected condition. This yields one bandpower summary per band and condition (and optionally per channel), which is then displayed in the bar chart or in the condition table.

Engagement, Cognitive Load and Frequency Bands Over Time

For the time-series visualization (indices over time and bandpower over time), the cleaned EEG data is divided into non-overlapping time windows with the length of the Segment Length parameter set by the user in the graph. The default value is 1 second, following Hassib et al. [2017]. For each time window, the PSD is computed, following related work such as Ahmed et al. [2020, 2025].

Time-resolved features are computed using fixed-length non-overlapping windows.

The PSD is computed using the same function as in Section 5.5.2, but with fixed settings chosen specifically for time-resolved estimation: `fmin=1`, `fmax=45`, and FFT windows the same size as the segment length parameter. This means that the PSD is computed on each window without an additional inner-window Welch overlap (as in Hassib et al. [2017]).

From the PSD of each window, bandpower is calculated using the current frequency band limits (the same limits that are configurable for the bandpower bar chart). Bandpower values are then averaged across all channels for each window and displayed as bandpower-over-time when the user selects *Frequency Bands*.

When the user selects *Computed Indices*, the interface uses the same window-wise bandpower values to compute two

Engagement and cognitive load indices are computed as bandpower ratios.

[9] `docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.simpson.html`, *last accessed 01.02.2026*

ratio metrics for each time window:

$$\text{Engagement} = \frac{\beta}{\alpha + \theta}$$

$$\text{Cognitive Load} = \frac{\theta}{\alpha}$$

Median filtering stabilizes index against short artifacts.

To reduce the effect of short artifacts, a moving median filter with a window length of 5 seconds is applied to both indices, following Szafir and Mutlu [2013] and Hassib et al. [2017]. The filtered index values are the values shown in the index-over-time plot.

Condition-level summaries aggregate window-wise index values.

For the values shown in the INDEX SUMMARY card, the mean of all time-resolved index values over the full condition duration is computed, based on similar aggregation in Hassib et al. [2017]. This yields one engagement value and one cognitive load value per condition.

5.6 Final Interface Limitation

Design principles are implemented through deliberate scope limitations and trade-offs.

The final interface translates the previously defined design principles and core requirements from Chapter 4 into a coherent and functional system. However, these principles, such as accessibility, transparency, and methodological clarity, also imply deliberate trade-offs. The focus on non-expert HCI researchers, consumer-grade EEG devices, and frequency-domain metrics results in a limited analytical scope. The following section therefore outlines the conceptual, technical, and methodological limitations of the current implementation.

System is not intended for clinical use or advanced spatial analysis.

The interface is not intended for clinical use or diagnosis. Due to the focus on consumer-grade EEG devices (currently only supporting the Muse) and does not support source localization or spatial analyses. It also does not support event-related potential (ERP) analysis. ERP analysis requires precise stimulus timing, strict control of noise, and careful trial-based processing. This is outside the target setting of this work.

System is limited to low-noise recording environments.

The system is designed for low-noise recordings and therefore provides only limited support for high-noise scenarios (e.g., speaking, typing, walking, or mobile studies). In such contexts, motion and muscle artifacts can dominate the signal, and simple bandpower and ratio-based metrics risk misinterpretation even after basic filtering and ICA.

Real-time indicator excluded due to methodological uncertainty and transparency concerns.

The current design does not include a real-time (live) indicator even though mentioned in the interviews in Chapter 4. Real-time indicators have been proposed in previous work (e.g., Hassib et al. [2017]) and are also included in tools such as the OpenBCI GUI. However, OpenBCI documentation does not explain how this indicator is calculated [10] and there is only a community forum post explaining it [11], which also raises uncertainty about the implementation. Several decisions have to be made for a live indicator. For example, whether averaging amplitudes or bandpower is a good approach. How sampling rate or FFT length affects frequency resolution and stability. Since there is no clearly validated and transparent implementation that works across devices and study contexts, adding a fixed live indicator could lead to overinterpretation and would conflict with the transparency goals of this interface. Therefore, such features should be tested extensively and are left for future work.

No machine learning or classification models are implemented.

The interface does not provide classification or machine learning approaches. These often require training data and should also be tested extensively. Leaving it for future work.

Fully automated preprocessing was considert but excluded in favor of guided interaction.

Fully automated end-to-end analysis pipelines were considered during development. However in low-noise settings with consumer-grade devices, artifact handling often depends on visual inspection and manual rejection, because automatic correction is difficult when only a few electrodes are available [Cannard et al., 2021]. For this reason, the interface does not apply a fixed, fully automated preprocessing pipeline. Instead, it guides users step by step through the preprocessing process.

[10] `docs.openbci.com/Software/OpenBCISoftware/GUIDocs/`, *last accessed 26.01.2026*

[11] `openbci.com/community/focus-visualization-widget`, *last accessed 26.01.2026*

Performance is not optimized for large datasets or high-density EEG.

The interface is not optimized for runtime performance or large-scale data processing. Since it works with CSV files, some actions can take a relatively long time to complete. For longer recordings or EEG devices with more channels and higher data density, additional performance optimizations would be necessary.

Default settings reflect common practice but require critical reflection by researchers.

Most importantly, there is no clear gold standard for measuring engagement or cognitive load from EEG data. Studies employ varying preprocessing steps, frequency band limits, ratio formulas, window lengths, and normalization methods. Because of this, a single default setting cannot be correct for all tasks, devices, or research questions. The default settings in this interface are based on commonly used methods in the literature, but represent only one possible way to calculate these measures. This is a general challenge in the research field. Therefore, researchers should carefully consider the chosen parameters and adjust them if necessary to fit their specific study.

Despite these limitations, the interface provides a transparent and accessible starting point for EEG-based user studies in HCI contexts.

Chapter 6

Planned Evaluation Study

This chapter presents the planned evaluation study for the interface. The goal of this evaluation is to examine whether the interface is usable, understandable, and transparent for HCI researchers with limited EEG experience.

First, the evaluation goals and research questions are defined in Section 6.1. Next, the target participants are described in Section 6.2. The evaluation method is then explained in Scetion 6.3. After that, the two main tasks representing the core workflows of the interface are presented in Section 6.4. Section 6.5 then describes the study setup and procedure. Finally in Section 6.6, the planned data analysis is outlined.

6.1 Evaluation Goal

The evaluation assesses the usability, clarity, and workflow transparency of the interface. The evaluation focuses on the following research questions:

Evaluation focuses on usability, clarity, and workflow transparency.

RQ1: Can HCI researchers complete the full workflow (experiment setup to data analysis) without help?

RQ3: Do users understand the preprocessing and analysis pipeline (e.g., segment deletion, filtering, ICA, PSD, bandpower) and see it as transparent?

These questions directly reflect the design principles and interface requirements from Chapter 4.

6.2 Participants

Study targets HCI researchers with limited EEG experience.

Participants are HCI researchers who have experience conducting user studies but little or no prior experience with EEG. This group reflects the intended users of the interface. A sample size of five participants was chosen following Virzi [1992]. Although the "rule of five" is debated [Lazar et al., 2017], usability sample sizes depend on study goals and constraints [Lazar et al., 2017]. Because this evaluation is exploratory and qualitative, it does not aim for statistical generalization. Instead, it focuses on identifying recurring usability issues and misunderstandings in the workflow.

6.3 Evaluation Method

Mixed-method evaluation combines tasks, SUS, and interviews.

Participants will complete two tasks from Section 6.4, then complete the System Usability Scale (SUS), and be interviewed in semi-structured interviews. The whole interview guide can be found in appendix E. The SUS was chosen because it is a short, widely validated measure of overall usability and fits the goal of evaluating whether the workflow is understandable, learnable, and usable. Broader user experience questionnaires and workload measures (NASA-TLX) were not considered because they focus on hedonic aspects or task workload, which are not the focus of the study [Lazar et al., 2017].

Think Aloud captures understanding of the interface.

During task execution, participants are encouraged to think aloud and verbalize their thoughts as they interact with the interface. The Think-Aloud approach is particularly suitable for early-stage design evaluations [Lazar et al., 2017]. Think Aloud is used as it is a well-established HCI method for identifying usability problems, misunderstandings, and users' mental models that are difficult to capture with quantitative measures alone [Lazar et al., 2017]. It was chosen in addition to the SUS because the research questions focus on understanding, transparency, and perceived control rather than efficiency or performance. Measures such as task completion time would not capture whether users understand, for example, the EEG preprocessing steps.

6.4 Tasks

Two tasks represent the core recording and analysis workflows.

The two workflows represent the two main phases of the interface: the recording phase and the data analysis phase. To avoid order effects, the order of the two tasks will be counterbalanced [Lazar et al., 2017].

Task A: Experiment Setup and Recording

1. Start the setup process from the Start screen.
2. Create a new experiment named Cognitive Load Pilot Study.
3. Define two experimental conditions:
   - **Task1**: fixed duration of 60 seconds.
   - **Task2**: no fixed duration.
4. Connect the Muse EEG headset.
5. Perform the electrode contact check until all electrodes reach at least the status Good.
6. Start a recording for the **Task1** condition and wait until it stops automatically.
7. Start a recording for the **Task2** condition, set one event marker named "Test Marker", and manually stop the recording.

8. End the recording session without proceeding to data analysis.

Each step is considered successfully completed if the participant reaches the intended interface state without help from the evaluator. This workflow directly maps to Phases 1–3 of the interface (Sections 5.4.2,5.4.3,5.4.4).

Task B: Data Analysis

1. From the start screen, open the dataset 'Dummy Dataset – Recording 01'.
2. Change the bandpass filter to 1–30 Hz.
3. Inspect the IC components and exclude on. Solution: IC0
4. Apply the preprocessing settings.
5. Inspect the resulting analysis output.
6. Answer the following questions:
    - Under which condition was the cognitive load the highest? Solution: 1NB
    - Which frequency band has the highest bandpower at 10 seconds of the baseline recording? Solution: beta
    - What are the interface frequency band limits for the alpha band? Solution: 8-13

A step is considered successfully completed if the participant can do the action and explain its purpose in their own words and get the right solution. This workflow evaluates whether participants understand and can meaningfully interact with the analysis pipeline from Section 5.4.5. In addition to task completion, the answers to the analysis questions will be evaluated for correctness and for how participants justify their responses during the think-aloud protocol. This allows distinguishing between correct answers based on genuine understanding and answers based on guessing or superficial interpretation.

6.5 Setup and Procedur

The evaluation will be conducted in a quiet usability laboratory comparable to standard HCI study environments. The setup includes a desk and chair, a laptop running the interface, a Muse EEG headset, a screen recording of the screen running the interface, an audio recording for Think Aloud verbalizations and the semi-structured interview, and screen mirroring that allows the evaluator to observe the participant's interaction in real time.

Standardized lab setup ensures consistency.

The evaluation follows a structured procedure to ensure consistency across participants and reproducibility.

Structured procedure supports reproducibility.

1. **Welcome and consent:** The participant is welcomed and informed about the study purpose. The participant reads and signs the informed consent form, including consent for screen and audio recording.

2. **Demographic questionnaire:** The participant completes a short questionnaire covering academic status, research field, experience with user studies, and prior experience with EEG hardware or data analysis.

3. **Think Aloud instruction:** The evaluator explains the Think Aloud protocol and asks the participant to verbalize thoughts, decisions, and uncertainties during interaction.

4. **Workflow execution:** The participant performs Workflow A and Workflow B while thinking aloud. The evaluator observes passively, takes notes on task execution, errors, and comments, and records task completion. No assistance is given unless the participant cannot continue.

5. **Usability questionnaire:** After completing both workflows, the participant fills out the System Usability Scale (SUS).

6. **Semi-structured interview:** After that a semi-structured interview is conducted to reflect on the

participant's experience, focusing on difficulties, unclear interface elements, trust in the results, and perceived transparency of the EEG processing pipeline.

6.6 Data Analysis

Qualitative (RTA) and descriptive analysis evaluates performance and understanding.

The Think-Aloud data and interview transcripts will be analyzed using reflexive thematic analysis. The aim is to identify recurring usability problems, misunderstandings, and patterns in how participants understand the workflow and the EEG analysis steps.

For Task B, each analysis question will be evaluated on two aspects: **Correctness** (correct / incorrect) and **Understanding**, based on how participants explain their answer during the Think-Aloud session. This makes it possible to see whether participants truly understand the analysis pipeline or simply arrive at the correct answer without deeper understanding. For each step of the task, it will be noted whether it was completed without help, completed with help, or not completed. Descriptive summaries will show how many participants completed each step and answered each question correctly. Recurring misunderstandings will be compared with the themes from the Think-Aloud and interview data.

This combined analysis allows assessing both what participants were able to do and how well they understood the underlying analysis process, directly addressing the evaluation goals.

Evaluation results will inform future work.

The planned evaluation will provide insights into usability problems, evidence of whether the overall workflow structure is understandable, and feedback on whether the EEG processing and analysis steps are perceived as transparent. In addition, the evaluation will reveal which functionalities are missing or insufficiently supported. The resulting findings will guide future work on the interface.

Chapter 7

Summary and Future Work

This chapter concludes the thesis with a summary and gives an outlook into possible future work.

7.1 Summary and Contributions

This work examined the needs of HCI researchers regarding the integration of EEG into their research and resulted in the design and implementation of an interactive EEG recording and analysis interface for applied HCI studies using consumer-grade devices. The focus was on continuous frequency-domain analysis methods.

Work investigates HCI researchers' needs and presents an EEG interface.

The thesis includes a review of existing EEG interfaces in Chapter 3 as well as qualitative interviews conducted with HCI researchers in Chapter 4. The interviews provided insights into current practices, challenges, and expectations regarding EEG use in user studies and informed the definition of design principles and interface requirements.

Literature review and interviews informed design principles and requirements.

Based on these requirements, a interface was designed and implemented as described in Chapter 5. The interface supports the complete workflow of an EEG-based user

Designed and implemented interface supports the full EEG study workflow.

study, including experiment setup, recording, preprocessing, and analysis. In summary, this work contributes an EEG recording interface design that balances accessibility, transparency, and usability for HCI research and demonstrates how EEG analysis can be integrated into user studies.

7.2 Future Work

A first future step should be to evaluate the interface in a full empirical user study. The evaluation concept is outlined in Chapter 6.

Future work can extend the interface in several directions, but should still focus on transparency, accessibility, and workflow-oriented EEG analysis. Additional requirements based on the interviews can be found in Appendix C.

Future versions may extend preprocessing and analysis pipeline.

One important direction is to extend the preprocessing and analysis pipeline. Future versions could include features such as bad channel detection, resampling, and re-referencing. The preprocessing steps could also adapt to different noise conditions by adjusting filters or artifact handling depending on the recording situation. Methods such as an exponentially weighted moving average, as used in previous work [Hassib et al., 2017; Szafir and Mutlu, 2012, 2013], and a normalization approach [Baradari et al., 2025; Hassib et al., 2017] could also be added. In addition, machine learning–based classifiers may be explored [Nuamah et al., 2017].

Signal quality metrics should become more adaptive and device-independent.

This also includes the development of adaptive or device-independent signal quality and electrode placement metrics that are less dependent on fixed hardware-specific thresholds, like in the current interface.

Interface could be extended to ERD/ERS and ERP-based analyses.

The interface could also be extended beyond purely continuous frequency-domain analysis. An ERD approach could enable more time-resolved interpretations of cognitive demand while remaining compatible with the current frequency domain focus [Antonenko et al., 2010]. For more controlled experimental settings, support for ERP analysis could also be explored, as used in Cherng et al. [2016]

Another possible extension is the addition of a simple live indicator for cognitive load or engagement during recording, similar to approaches proposed in prior work or implemented in tools such as the OpenBCI GUI. Such an indicator could provide a rough overview during data collection without distracting participants. However, as discussed in Section 5.6, there are still open methodological questions regarding how to compute a reliable live indicator. Therefore, future research should compare different live indicator approaches within the current interface.

Future research should evaluate reliable live indicator approaches.

Beyond cognitive load and engagement, future versions could support additional cognitive metrics. One example is cognitive absorption, which describes states of deep engagement and flow. Such states were mentioned in the interviews and have been linked to physiological signals, including EEG [Conrad and Bliemel, 2016].

Additional cognitive metrics could be supported.

Finally, the interface could support synchronization with other data sources. Future versions could combine EEG with video or screen recordings, as well as other physiological signals such as eye tracking, electromyography, or electrocardiography. Libraries like NeuroKit2[1] could be used to analyze these signals together with EEG, for example, to compute heart rate or skin conductance measures. This would support richer interpretations and reflect the interview finding that EEG should complement, rather than replace, other methods.

Future versions may integrate EEG with additional physiological signals.

Overall, future work should not simply increase analytical complexity. Any extensions should remain optional, transparent, and clearly structured. By adding advanced features on top of a simple and guided core workflow, the interface could support a wider range of study contexts without reducing usability or trust.

Extensions should remain optional, transparent, and workflow-oriented.

[1] `neuropsychology.github.io/NeuroKit/`, *last accessed 07.02.2026*

Appendix A

Interview Guide

This appendix presents the semi-structured interview guide used for the empirical study described in Chapter 4. The guide was designed to explore the requirements, challenges, and expectations of HCI researchers regarding the integration of EEG data into their research workflows.

A.1 Background of the Interviewee

- What is your current research area, or what projects are you working on?
- To what extent do you work with data?
- Do you currently use any tools or software to analyze data? If so, which ones?
- Is there anything about complex data visualizations that frustrates you?
- How important is visual exploration to you compared to numerical analysis?

A.2 Perceptions of EEG

- What comes to mind when you hear “EEG data”?

- Can you imagine an application area in your current HCI research where you would use an EEG device?
- What might be exciting or difficult about working with brain data?

A.3 EEG as a Data Source

- Assuming we measure cognitive states, which ones would interest you?
- Which comparisons would be key for you (e.g., points in time, conditions, groups)?
- What would be important to you to have confidence in the data or analyses?
- How far do you want to be able to understand the processing and analysis?

A.4 Topic of User Studies

- What would be the best way to integrate EEG into a user study?
- How exactly do you envision an EEG-based user study (e.g., setup, participants, measurement conditions)?
- When will the data analysis take place (how much during the study and how much afterwards)?
- Let's assume there is an interface that can analyze EEG data from user studies. How would that work perfectly for your needs?

A.5 General Closing Questions

- Do you prefer guided steps or free experimentation with complex interfaces?

- What would motivate you to use such a tool in your own research or teaching?
- Is there anything that would prevent you from using an EEG analysis tool?
- Do you have any further thoughts or ideas that we have not discussed yet?

Appendix B

Codebook

The following codebook gives an overview of the themes and subcodes that were developed during the Reflexive Thematic Analysis of the interviews in Chapter 4. It shows the main themes that were used to derive the design implications.

For transparency, the table also includes frequency counts, which show how often each code was applied in the dataset. These numbers are only meant to document how the data was distributed across codes. Following Reflexive Thematic Analysis, themes were identified based on shared patterns of meaning and their relevance to the research question, not only on how often they occurred.

Code	Frequency
Competence Boundaries and Resulting Uncertainty	22
No knowledge what EEG measures and what data means	8
Too limited knowledge for self validation	4
Medical Framing of EEG	5
Unfamiliarity with EEG Applications in HCI	5
Motivation Driven by Perceived Usefulness	19
Need for Realistic Framing, No overselling	7
Selftest as first step to evaluation of usage	3
Learning Motivation Depends on Relevance and Usefulness	9
EEG as a Supportive Source of Insight	32
Limited Interest in Active Control	1
Interest in Feedback Loops or active control	3
EEG as objective datasource for Data Triangulation	23
Temporal Resolution and Event Marking	5
Low Friction Integration into Study Routines	17
Low Setup Effort and Non-Disruptiveness	6
Concerns About Influence on Participants	3
Post-Hoc Analytical Depth	5
First small feedback in studie	3
Trust through Transparency and Methodological Grounding	22
Transparency of Processing important	8
Scientific Referencing and External Legitimization	7
Defaults based on related work	5
Certainty score	1
Data Protection Concerns	1
Comprehensibility and Accessibility of the Interface	20
Progressive Disclosure	9
Visual reprensentation insted of just numerical	4
Contextual Learning Support	7
Interest Profiles Regarding Cognitive States	17
Baseline Measurement Interest	2
Cognitive Load / Mental Effort	3
Stress / Arousal	2
Flow	1
Creativity (divergent/convergent)	1
Affect	1
Fatigue	1
Motion Sickness	1
Enyoment/ Wellbeeing	2
Concentration	1
Physical Demand	1
Anger	1

Table B.1: Codebook

Appendix C

Additional Requirements for Future Work

This appendix summarizes additional requirements that emerged from the interview study in Chapter 4 but were not implemented in the current interface.

Requirement	Interview grounding
Epoching around stimuli/tasks	Supports time-based analysis
Import and align non-EEG data (video, audio)	Integration enables data triangulation
Batch processing after data collection	Supports preference for post hoc analysis
Analysis audit log	Supports traceability and reproducibility of analyses
Certainty or confidence score for computed outputs	Helps assess the reliability of derived metrics
Option for additional participant data (demographics)	Supports easyer integration into study workflow

Table C.1: Future requirements based on the Interviews

Appendix D

Paper Prototype

The following paper prototype was used in the walk-through mentioned in Section 5.2.

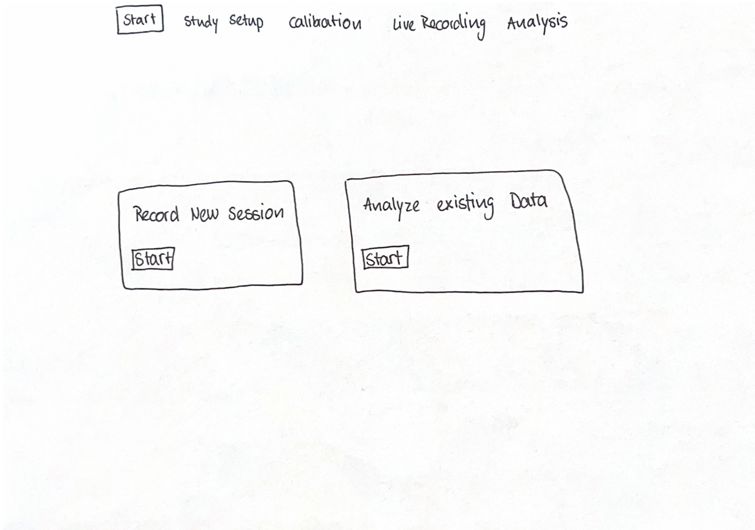


Figure D.1: Paperprototyp - Start Screen

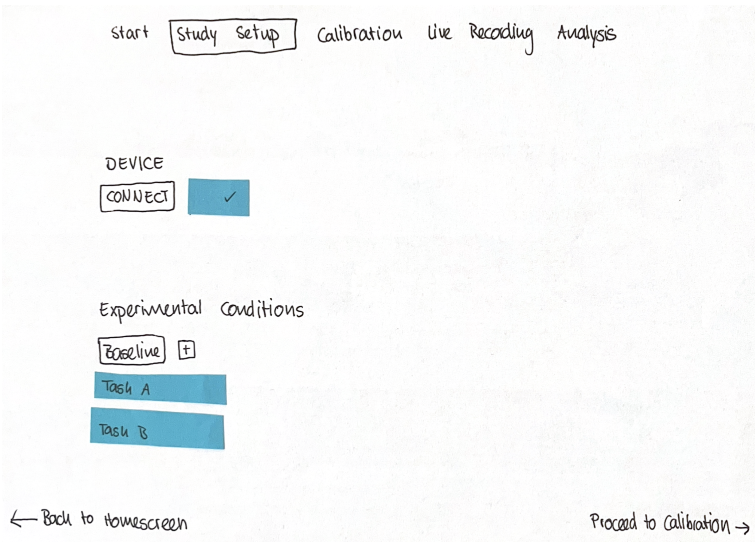


Figure D.2: Paperprototyp - Setup Screen

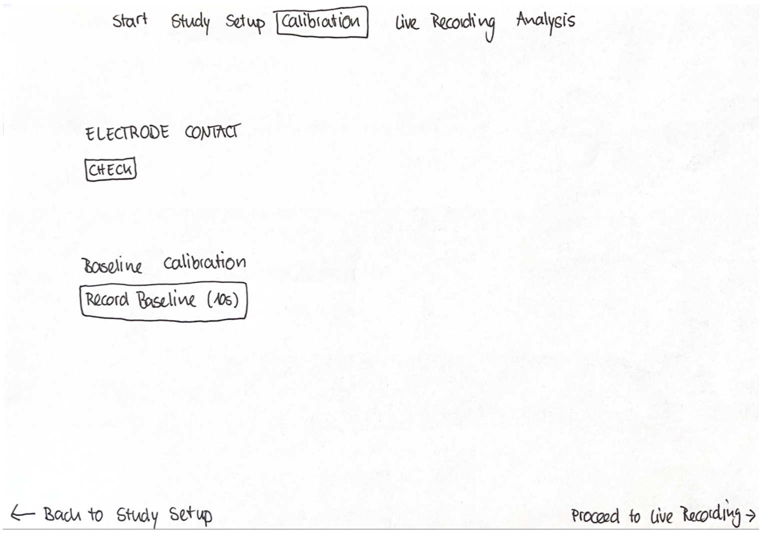


Figure D.3: Paperprototyp - Calibration Screen

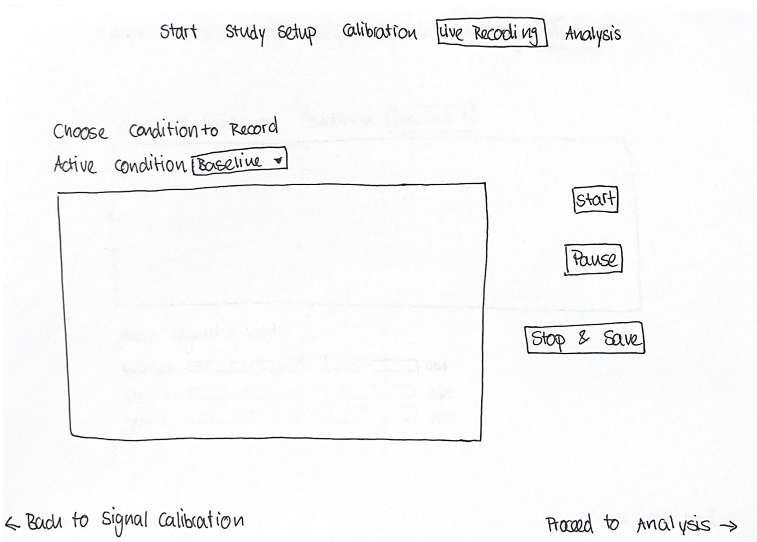


Figure D.4: Paperprototyp - Recording Screen

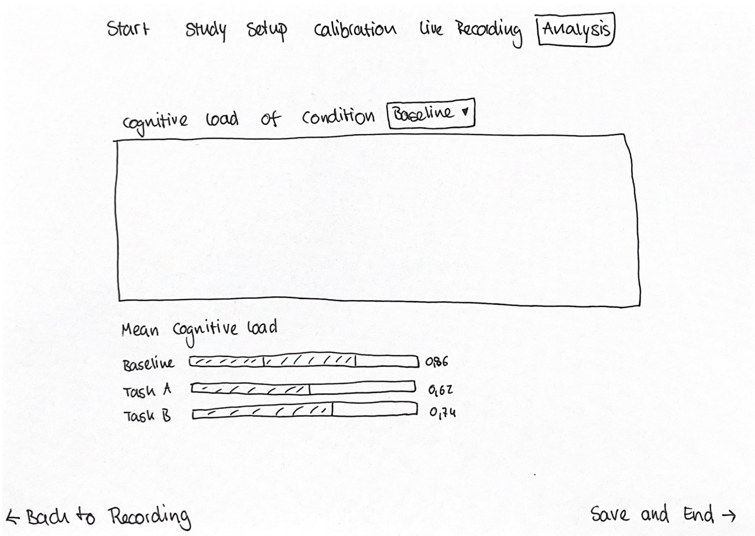


Figure D.5: Paperprototyp - Analysis Screen

Appendix E

Interview Guide for Evaluation

This appendic presents the semi-structured Interview Guid which should be used for the Evaluation. The interviewer should adapt the follow-up questions depending on the observed behavior, errors, or uncertainties during the study.

E.1 Overall Impression

- How would you describe your overall experience with the interface?
- Did the interface match your expectations for a tool supporting EEG-based user studies? Why or why not?

E.2 Understanding and Control of the Analysis Pipeline

- Did you feel that you were in control of the preprocessing steps (e.g., filtering, ICA), or did you feel that the interface implicitly made decisions for you?

- Do you trust the default parameters provided by the interface?
- At which points would you have preferred more explanations or references to better understand or trust the processing steps?

E.3 Specific Decisions

- Can you describe why or why not you excluded this specific ICA component?
- The interface shows cognitive load relative to a baseline recording. How do you interpret a positive deviation value?

E.4 Interpretation of Results

- How confident did you feel when answering questions about cognitive load and engagement based on the displayed results?
- Did the visualizations and labels help you to understand what the computed metrics represent?

E.5 Recording and Electrode Contact Check

- How easy or difficult was the electrode contact check before recording?
- Was the time needed to reach Good electrode contact quality acceptable for a real user study?
- During recording, did the interface distract you?

E.6 Reflection

- Which parts of the interface felt most clear and understandable?
- Which parts felt most confusing or uncertain?
- What would you change first to improve the interface for your own research practice?
- Would you consider using such an interface in a real HCI user study? Why or why not?
- Is there anything else you would like to add that we did not discuss?

E.7 Follow Up Questions

- At this point, what did you think the system was doing?
- What outcome did you expect after performing this action?
- Looking back, why do you think this step did not work as expected?
- Was there anything in the interface that suggested to you that this action was correct?
- What information or feedback would have helped you avoid this mistake?

Bibliography

[1] Akshay Aggarwal, Gerrit Niezen, and Harold Thimbleby. User experience evaluation through the brain's electrical activity. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, pages 491–500, 2014.

[2] Sabbir Ahmed, Md Abul Ala Walid, and Monira Islam. EEG-based Cognitive Load assessment in Matlab GUI and impact on Learning System. In *2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)*, pages 484–487. IEEE, 2020.

[3] Yusuf Ahmed, Martin Ferguson-Pell, Kim Adams, and Adriana Ríos Rincón. EEG-Based Engagement Monitoring in Cognitive Games. *Sensors*, 25(7):2072, 2025.

[4] Amjed S Al-Fahoum and Ausilah A Al-Fraihat. Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. *International Scholarly Research Notices*, 2014(1):730218, 2014.

[5] Erik W Anderson, Kristin C Potter, Laura E Matzen, Jason F Shepherd, Gilbert A Preston, and Cláudio T Silva. A user study of visualization effectiveness using EEG and cognitive load. In *Computer graphics forum*, volume 30, pages 791–800. Wiley Online Library, 2011.

[6] Pavlo Antonenko, Fred Paas, Roland Grabner, and Tamara Van Gog. Using electroencephalography to measure cognitive load. *Educational psychology review*, 22(4):425–438, 2010.

[7] Andrea Apicella, Pasquale Arpaia, Mirco Frosolone, Giovanni Improta, Nicola Moccaldi, and Andrea Pollastro. EEG-based measurement system for monitoring student engagement in learning 4.0. *Scientific Reports*, 12(1):5857, 2022.

[8] Sarune Baceviciute, Aske Mottelson, Thomas Terkildsen, and Guido Makransky. Investigating Representation of Text and Audio in Educational VR using Learning Outcomes and EEG. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–13, New York,

NY, USA, 2020. Association for Computing Machinery. doi.org/10.1145/3313831.3376872.

[9] Anders Bach Justesen, Mette Thrane Foged, Martin Fabricius, Christian Skaarup, Nizar Hamrouni, Terje Martens, Olaf B. Paulson, Lars H. Pinborg, and Sándor Beniczky. Diagnostic Yield of High-Density versus Low-Density EEG: The Effect of Spatial Sampling, Timing and Duration of Recording. *Clinical Neurophysiology*, 130(11):2060–2064, November 2019. doi.org/10.1016/j.clinph.2019.08.007.

[10] Vipin Bakshi. *Towards Practical Driver Cognitive Workload Monitoring via Electroencephalography*. University of Toronto (Canada), 2018.

[11] Dünya Baradari, Nataliya Kosmyna, Oscar Petrov, Rebecah Kaplun, and Pattie Maes. NeuroChat: A Neuroadaptive AI Chatbot for Customizing Learning Experiences. *arXiv preprint arXiv:2503.07599*, 2025.

[12] Pouya Bashivan, Irina Rish, and Steve Heisig. Mental state recognition via wearable EEG. *arXiv preprint arXiv:1602.00985*, 2016.

[13] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.

[14] Jaqueline J Brito, Jun Li, Jason H Moore, Casey S Greene, Nicole A Nogoy, Lana X Garmire, and Serghei Mangul. Recommendations to enhance rigor and reproducibility in biomedical research. *GigaScience*, 9(6):giaa056, 2020.

[15] Clemens Brunner, Giuseppe Andreoni, Lugi Bianchi, Benjamin Blankertz, Christian Breitwieser, Shin'ichiro Kanoh, Christian A Kothe, Anatole Lécuyer, Scott Makeig, Jürgen Mellinger, and others. BCI software platforms. In *Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications*, pages 303–331. Springer, 2012.

[16] Cédric Cannard, Helané Wahbeh, and Arnaud Delorme. Validating the Wearable MUSE Headset for EEG Spectral Analysis and Frontal Alpha Asymmetry. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 3603–3610, December 2021. doi.org/10.1109/BIBM52615.2021.9669778.

[17] Cédric Cannard, Helané Wahbeh, and Arnaud Delorme. Validating the wearable MUSE headset for EEG spectral analysis and Frontal Alpha Asymmetry. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 3603–3610. IEEE, 2021.

[18] Valentina Casadei and Roberto Ferrero. Online Skin-Electrode Contact Quality Monitoring in Wearable Devices: An EEG Application. In *2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pages 1–6. IEEE, 2022.

[19] Fu-Yin Cherng, Wen-Chieh Lin, Jung-Tai King, and Yi-Chen Lee. An EEG-based Approach for Evaluating Graphic Icons from the Perspective of Semantic Distance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 4378–4389, New York, NY, USA, 2016. Association for Computing Machinery. `doi.org/10.1145/2858036.2858133`.

[20] Sylvain Chevallier, Igor Carrara, Bruno Aristimunha, Pierre Guetschel, Sara Sedlar, Bruna Lopes, Sébastien Velut, Salim Khazem, and Thomas Moreau. The largest EEG-based BCI reproducibility study for open science: the MOABB benchmark. *arXiv preprint arXiv:2404.15319*, 2024.

[21] Colin David Conrad and Michael Bliemel. Psychophysiological measures of cognitive absorption and cognitive load in e-learning applications. 2016.

[22] Benjamin Cowley, Jussi Korpela, and Jari Torniainen. Computational testing for automated preprocessing: a Matlab toolbox to enable large scale electroencephalography data processing. *PeerJ Computer Science*, 3, 03 2017. `doi.org/10.7717/peerj-cs.108`.

[23] Arnaud Delorme and Scott Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[24] Ronit Desai, Rita Orji, Eugenie Roudaia, and Allison B Sekuler. Sans Tracas: A Cross-platform Tool for Online EEG Experiments. In *Companion of the 2022 ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, pages 37–43, 2022.

[25] Pravin V Dhole, Deepali G Chaudhary, Vijay D Dhangar, Sulochana D Shejul, Snehal S Datwase, and Bharti W Gawali. A Review of EEG Artifact Removal Techniques for Brain-Computer Interface. *SN Computer Science*, 6(8): 1026, 2025.

[26] Atef Eldenfria and Hosam Al-Samarraie. Towards an online continuous adaptation mechanism (OCAM) for enhanced engagement: An EEG study. *International Journal of Human–Computer Interaction*, 35(20):1960–1974, 2019.

[27] Mehrdad Fatourechi, Ali Bashashati, Rabab K Ward, and Gary E Birch. EMG and EOG artifacts in brain computer interface systems: A survey. *Clinical neurophysiology*, 118(3):480–494, 2007.

[28] Raul Fernandez Rojas, Essam Debie, Justin Fidock, Michael Barlow, Kathryn Kasmarik, Sreenatha Anavatti, Matthew Garratt, and Hussein Abbass. Electroencephalographic workload indicators during teleoperation of an unmanned aerial vehicle shepherding a swarm of unmanned ground vehicles in contested environments. *Frontiers in neuroscience*, 14:40, 2020.

[29] Jérémy Frey, Christian Mühl, Fabien Lotte, and Martin Hachet. Review of the use of electroencephalography as an evaluation method for human-computer interaction. *arXiv preprint arXiv:1311.2222*, 2013.

[30] Jérémy Frey, Maxime Daniel, Julien Castet, Martin Hachet, and Fabien Lotte. Framework for electroencephalography-based evaluation of user experience. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2283–2294, 2016.

[31] Alan Gevins and Michael E Smith. Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral cortex*, 10(9):829–839, 2000.

[32] Alan Gevins and Michael E Smith. Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical issues in ergonomics science*, 4(1-2):113–131, 2003.

[33] Rick O Gilmore, Michele T Diaz, Brad A Wyble, and Tal Yarkoni. Progress toward openness, transparency, and reproducibility in cognitive neuroscience. *Annals of the New York Academy of Sciences*, 1396(1):5–18, 2017.

[34] Evgenia Gkintoni and Constantinos Halkiopoulos. Mapping EEG metrics to human affective and cognitive models: An interdisciplinary scoping review from a cognitive neuroscience perspective. *Biomimetics*, 10(11):730, 2025.

[35] Krzysztof J Gorgolewski and Russell A Poldrack. A practical guide for improving transparency and reproducibility in neuroimaging research. *PLoS biology*, 14(7):e1002506, 2016.

[36] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, and others. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroinformatics*, 7:267, 2013.

[37] Mariam Hassib, Stefan Schneegass, Philipp Eiglsperger, Niels Henze, Albrecht Schmidt, and Florian Alt. EngageMeter: A system for implicit audience engagement sensing using electroencephalography. In *Proceedings of the 2017 Chi conference on human factors in computing systems*, pages 5114–5119, 2017.

[38] Matthias R Hohmann, Michelle Hackl, Brian Wirth, Talha Zaman, Raffi Enficiaud, Moritz Grosse-Wentrup, and Bernhard Schölkopf. MynD: A platform for large-scale neuroscientific studies. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–6, 2019.

[39] Anu Holm, Kristian Lukander, Jussi Korpela, Mikael Sallinen, and Kiti MI Müller. Estimating brain load from the EEG. *The Scientific World Journal*, 9(1): 639–651, 2009.

[40] J Holsheimer and BWA Feenstra. Volume conduction and EEG measurements within the brain: a quantitative approach to the influence of electrical spread on the linear relationship of activity measured at different locations. *Electroencephalography and clinical neurophysiology*, 43(1):52–58, 1977.

[41] Alice F Jackson and Donald J Bolger. The neurophysiological bases of EEG and EEG measurement: A review for the rest of us. *Psychophysiology*, 51(11): 1061–1071, 2014.

[42] Thomas Kosch, Jakob Karolus, Johannes Zagermann, Harald Reiterer, Albrecht Schmidt, and Paweł W Woźniak. A survey on measuring cognitive workload in human-computer interaction. *ACM Computing Surveys*, 55(13s): 1–39, 2023.

[43] Nataliya Kosmyna and Pattie Maes. AttentivU: an EEG-based closed-loop biofeedback system for real-time monitoring and improvement of engagement for personalized learning. *Sensors*, 19(23):5200, 2019.

[44] Christian Andreas Kothe and Scott Makeig. BCILAB: a platform for brain–computer interface development. *Journal of neural engineering*, 10(5):056014, 2013.

[45] Olave E Krigolson, Chad C Williams, Angela Norton, Cameron D Hassall, and Francisco L Colino. Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research. *Frontiers in neuroscience*, 11:109, 2017.

[46] Olave E Krigolson, Mathew R Hammerstrom, Wande Abimbola, Robert Trska, Bruce W Wright, Kent G Hecker, and Gordon Binsted. Using Muse: Rapid mobile assessment of brain performance. *Frontiers in Neuroscience*, 15:634147, 2021.

[47] J Satheesh Kumar and P Bhuvaneswari. Analysis of electroencephalography (EEG) signals and its categorization–a study. *Procedia engineering*, 38:2525–2536, 2012.

[48] Naveen Kumar and Jyoti Kumar. Measurement of cognitive load in HCI systems using EEG power spectrum: an experimental study. *Procedia Computer Science*, 84:70–78, 2016.

[49] Konstantina Kyriaki, Dimitrios Koukopoulos, and Christos A Fidas. A comprehensive survey of eeg preprocessing methods for cognitive load assessment. *IEEE Access*, 12:23466–23489, 2024.

[50] HJ Landau. Sampling, data transmission, and the Nyquist rate. *Proceedings of the IEEE*, 55(10):1701–1706, 2005.

[51] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017.

[52] Moona Mazher, Azrina Abd Aziz, Aamir Saeed Malik, and Hafeez Ullah Amin. An EEG-based cognitive load assessment in multimedia learning using feature extraction and partial directed coherence. *IEEE Access*, 5:14819–14829, 2017.

[53] Gernot R Müller-Putz, René Riedl, Selina C Wriessnegger, and others. Electroencephalography (EEG) as a research tool in the information systems discipline: Foundations, measurement, and applications. *Communications of the Association for Information Systems*, 37(1):46, 2015.

[54] Gernot R. Müller-Putz, René Riedl, and Selina C. Wriessnegger. Electroencephalography (EEG) as a Research Tool in the Information Systems Discipline: Foundations, Measurement, and Applications. 37. `doi.org/10.17705/1CAIS.03746`.

[55] Akihiro Nakamura, Yasuyuki Suzuki, Shunpei Yano, and Taishin Nomura. EEG activity related to decrease in persistency of gait cycle variability during distracted walking. In *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*, pages 183–184. IEEE, 2021.

[56] Jennifer J Newson and Tara C Thiagarajan. EEG frequency bands in psychiatric disorders: a review of resting state studies. *Frontiers in human neuroscience*, 12:521, 2019.

[57] Ernst Niedermeyer and FH Lopes da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[58] Jakob Nielsen. Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 152–158, 1994.

[59] Jakob Nielsen. *Usability engineering*. Morgan Kaufmann, 1994.

[60] Guiomar Niso, Rotem Botvinik-Nezer, Stefan Appelhoff, Alejandro De La Vega, Oscar Esteban, Joset A Etzel, Karolina Finc, Melanie Ganz, Rémi Gau, Yaroslav O Halchenko, and others. Open and reproducible neuroimaging: From study inception to publication. *NeuroImage*, 263:119623, 2022.

[61] A Norman Donald. *The design of everyday things*. MIT Press, 2013.

[62] Joseph K Nuamah, Younho Seong, and Sun Yi. Electroencephalography (EEG) classification of cognitive tasks based on task engagement index. In *2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, pages 1–6. IEEE, 2017.

[63] Robert Oostenveld, Pascal Fries, Eric Maris, and Jan-Mathijs Schoffelen. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011(1): 156869, 2011.

[64] Oscar Ortiz, D Blustein, and Usha Kuruganti. Test-retest reliability of time-domain EEG features to assess cognitive load using a wireless dry-electrode system. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 2885–2888. IEEE, 2020.

[65] Fred Paas, Juhani E Tuovinen, Huib Tabbers, and Pascal WM Van Gerven. Cognitive load measurement as a means to advance cognitive load theory. In *Cognitive load theory*, pages 63–71. Routledge, 2016.

[66] Mariella Paul, Gisela H Govaart, and Antonio Schettino. Making ERP research more transparent: guidelines for preregistration. *International Journal of Psychophysiology*, 164:52–63, 2021.

[67] Gert Pfurtscheller and FH Lopes Da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110 (11):1842–1857, 1999.

[68] Alan T Pope, Edward H Bogart, and Debbie S Bartolome. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology*, 40(1-2):187–195, 1995.

[69] Sébastien Puma, Nadine Matton, Pierre-V Paubel, Éric Raufaste, and Radouane El-Yagoubi. Using theta and alpha band power to assess cognitive workload in multitasking environments. *International Journal of Psychophysiology*, 123:111–120, 2018.

[70] Felix Putze, Susanne Putze, Merle Sagehorn, Christopher Micek, and Erin T. Solovey. Understanding HCI Practices and Challenges of Experiment Reporting with Brain Signals: Towards Reproducibility and Reuse. *ACM Trans. Comput.-Hum. Interact.*, 29(4), March 2022. `doi.org/10.1145/3490554`.

[71] Ahmed F Rabbi, Kevin Ivanca, Ashley V Putnam, Ahmed Musa, Courtney B Thaden, and Reza Fazel-Rezai. Human performance evaluation based on EEG signal analysis: a prospective review. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1879–1882. IEEE, 2009.

[72] Rajesh PN Rao. *Brain-computer interfacing: an introduction*. Cambridge University Press, 2013.

[73] Bujar Raufi and Luca Longo. An evaluation of the EEG alpha-to-theta and theta-to-alpha band ratios as indexes of mental workload. *Frontiers in Neuroinformatics*, 16:861967, 2022.

[74] Yann Renard, Fabien Lotte, Guillaume Gibert, Marco Congedo, Emmanuel Maby, Vincent Delannoy, Olivier Bertrand, and Anatole Lécuyer. Openvibe: An open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments. *Presence*, 19(1):35–53, 2010.

[75] Sérgio Daniel Rodrigues and Pedro Miguel Rodrigues. Electroencephalogram-based time-frequency analysis for Alzheimer's disease detection using machine learning. *Journal of Biological Methods*, 12(1):e99010042, 2024.

[76] Fnu Rohit, Vinod Kulathumani, Rahul Kavi, Ibrahim Elwarfalli, Vlad Kecojevic, and Ashish Nimbarte. Real-time drowsiness detection using wearable, lightweight brain sensing headbands. *IET Intelligent Transport Systems*, 11(5): 255–263, 2017.

[77] Riyanarto Sarno, Muhammad Nadzeri Munawar, Brilian T Nugraha, R Sarno, M Munawar, and B Nugraha. Real-time electroencephalography-based emotion recognition system. *Int. Rev. Comput. Softw. IRECOS*, 11(5):456–465, 2016.

[78] Aktham Sawan and Mohammad Awad. Assessing the Stress-Reducing Effects of Quranic Listening: An EEG-Based Classification Study. In *2023 24th International Arab Conference on Information Technology (ACIT)*, pages 1–9. IEEE, 2023.

[79] Phattarapong Sawangjai, Supanida Hompoonsup, Pitshaporn Leelaarporn, Supavit Kongwudhikunakorn, and Theerawit Wilaiprasitporn. Consumer grade EEG measuring sensors as research tools: A review. *IEEE Sensors Journal*, 20(8):3996–4024, 2019.

[80] Gerwin Schalk, Dennis J McFarland, Thilo Hinterberger, Niels Birbaumer, and Jonathan R Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on biomedical engineering*, 51(6):1034–1043, 2004.

[81] Moira C Schummer, Robin Nehler, and Sarah Jessen. Methodological Challenges in EEG Research with 11-Month-Old Infants-A Pilot Study. In *Student Conference Proceedings*, volume 1, pages 2008–2008, 2025.

[82] John Semmlow. *Signals and systems for bioengineers: a MATLAB-based introduction*. Academic press, 2012.

[83] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization*, pages 364–371. Elsevier, 2003.

[84] Michael E Smith and Alan Gevins. Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator. In *Biomonitoring for Physiological and Cognitive Performance during Military Operations*, volume 5797, pages 116–126. SPIE, 2005.

[85] Igor Stancin, Mario Cifrek, and Alan Jovic. A review of EEG signal features and their application in driver drowsiness detection systems. *Sensors*, 21(11): 3786, 2021.

[86] Lukas Stingl and Michael T Knierim. OpenBCI Meets the Web: A Scalable, Customizable Platform for Open-Source EEG Data Collection. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 925–929, 2024.

[87] Sally M Stoyell, Janina Wilmskoetter, Mary-Ann Dobrota, Dhinakaran M Chinappen, Leonardo Bonilha, Mark Mintz, Benjamin H Brinkmann, Susan T Herman, Jurriaan M Peters, Serge Vulliemoz, and others. High-density EEG in current clinical practice and opportunities for the future. *Journal of clinical neurophysiology*, 38(2):112–123, 2021.

[88] Daniel Szafir and Bilge Mutlu. Pay attention! Designing adaptive agents that monitor and improve user engagement. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 11–20, 2012.

[89] Daniel Szafir and Bilge Mutlu. ARTFul: adaptive review technology for flipped learning. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1001–1010, 2013.

[90] François Tadel, Sylvain Baillet, John C Mosher, Dimitrios Pantazis, and Richard M Leahy. Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience*, 2011(1):879716, 2011.

[91] Desney Tan and Anton Nijholt. Brain-computer interfaces and human-computer interaction. In *Brain-computer interfaces: Applying our minds to human-computer interaction*, pages 3–19. Springer, 2010.

[92] Tamanna Tasmi, Mohammad Parvez, and Jia Uddin. Cognitive load measurement based on EEG signals. In *Applied AI and Multimedia Technologies for Smart Manufacturing and CPS Applications*, pages 246–255. IGI Global, 2023.

[93] Alexandra-Maria Tautan, Vojkan Mihajlovic, Yun-Hsuan Chen, Bernard Grundlehner, Julien Penders, and Wouter Serdijn. Signal quality in dry electrode EEG and the relation to skin-electrode contact impedance magnitude. In *International Conference on Biomedical Electronics and Devices*, volume 2, pages 12–22. SciTePress, 2014.

[94] Michal Teplan and others. Fundamentals of EEG measurement. *Measurement science review*, 2(2):1–11, 2002.

[95] Darinka Trübutschek, Yu-Fang Yang, Claudia Gianelli, Elena Cesnaite, Nastassja L Fischer, Mikkel C Vinding, Tom R Marshall, Johannes Algermissen, Annalisa Pascarella, Tuomas Puoliväli, and others. EEGManyPipelines: a

large-scale, grassroots multi-analyst study of electroencephalography analysis practices in the wild. *Journal of Cognitive Neuroscience*, 36(2):217–224, 2024.

[96] Carmen Vidaurre, Tilmann H Sander, and Alois Schlögl. BioSig: the free and open source software library for biomedical signal processing. *Computational intelligence and neuroscience*, 2011(1):935364, 2011.

[97] Robert A Virzi. Refining the test phase of usability evaluation: how many subjects is enough? *Human factors*, 34(4):457–468, 1992.

[98] Peter Welch. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73, 2003.

[99] Zhichun Zhao, Kamen Ivanov, Ludwig Lubich, Olatunji Mumini Omisore, Zhanyong Mei, Nan Fu, Jinying Chen, and Lei Wang. Signal quality and electrode-skin impedance evaluation in the context of wearable electroencephalographic systems. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4965–4968. IEEE, 2018.

Index

Typeset February 16, 2026