Preprints.org

Article

Brain Wearables: Validation Toolkit for Ear-Level EEG Sensors

Guilherme Correia , Michael J Crosse , Alejandro Lopez Valdes [*]

Article

Brain Wearables: Validation Toolkit for Ear-Level EEG Sensors

Guilherme Correia [1] , **Michael J. Crosse** [2,6] **and Alejandro Lopez Valdes** [3,4,5,6,*]

1   Department of Physics, NOVA School of Science and Technology, Lisbon, Portugal
2   Segotia, Galway, Ireland
3   Department of Electronic and Electrical Engineering, Trinity College Dublin, Dublin, Ireland
4   Global Brain Health Institute, Trinity College Dublin, Ireland
5   Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland
6   Trinity Centre for Biomedical Engineering, Trinity College Dublin, Dublin, Ireland
*   Correspondence: alejandro.lopez@tcd.ie

Abstract: Electroencephalography (EEG) enabled earbuds represent a promising frontier in evaluating brain activity beyond traditional laboratory settings. However, these devices need more comprehensive characterization before widespread health related usage. A toolbox was developed to facilitate and expand the assessment of ear-EEG devices. The first component of this toolbox is a desktop application ("EaR-P Lab") that integrates EEG validation paradigms. This application is designed to be used with EEG amplifiers compatible with the Lab Streaming Layer (LSL) protocol. The second element of the toolbox introduces an adaptation of the phantom testing concept, to the domain of ear-EEG by using 3D ear scans of the tested individuals, allowing controlled assessment of ear-EEG devices. The EEG paradigms were validated for scalp acquisitions and applied on ear-EEG measurements, under wet ear electrode conditions. The ear-EEG phantom was successful in measuring useful metrics in EEG system characterization, highlighting performance differences within the electrodes on the tested earpieces. This knowledge was factored in to select an optimal electrode reference which resulted in increased response power on the auditory steady-state response (ASSR). Through this work, an ear-EEG evaluation toolkit is made available with the intention to enable future and systematic assessment of novel ear-EEG devices from the hardware to the neural signals.

Keywords: wearable electroencephalography; ear-EEG; validation, toolkit, eeg phantom, event-related potentials

1. Introduction

Conventional Electroencephalography (EEG) is an invaluable tool for assessing neurological disorders, with measurements performed in controlled clinical environments, utilizing full cap systems over the scalp with wet electrodes that provide low impedance for quality measurements with high temporal resolution. However, researchers and clinicians have been interested in measuring EEG outside the laboratory for better assessing neurological disorders [1].

The development of technology has led to the evolution of ambulatory EEG, resulting in research on smaller EEG measurement options. This has given rise to a new category of wearable EEG devices that are wireless, have aesthetic designs limited to only the area around the head, and use dry electrodes. The convenience of these devices enables individuals to incorporate EEG measurement into their everyday lives, making it easier to assess certain medical conditions such as epilepsy and sleep disorders and allowing for Brain-Computer Interface (BCI) applications outside of a laboratory environment. While these types of devices are commercially available and their capabilities acceptable for some current applications, the overall design of these wearables still constitutes a barrier to daily usage, being uncomfortable over long periods of time and making it obvious when a person is utilizing such an EEG device [2,3].

Quoting Looney et al., the next generation of wearable devices must be "discreet, unobtrusive, robust, user friendly and feasible". An in-ear-EEG approach checks all these boxes, trading the wide coverage across the scalp with an inconspicuous EEG recording ability, based on the ear [4]. The ability to acquire reliable brain recordings from inside the human ear is critical to accelerating the development of next-generation EEG-enabled earbuds. While ear-EEG is one of the best candidates for consumer BCI, being referred to as "beyond wearable" [2], real-world brain recordings are affected by multiple environmental factors that are not present in the laboratory. A systematic, well-established characterization of in-ear-EEG hardware and signal quality is crucial to understand the limitations and applications of newly developed devices.

When characterizing a novel EEG system, it is important to be able to evaluate all components of the system. This includes the biological neural signals, derived from validated EEG paradigms like the alpha block or the various Event-Related Potentials (ERPs); and the hardware and signal chain components (electrical and mechanical), which require testing in an controlled setting through a test-bench or phantom model [5].

In many areas of research and instrumentation, phantoms are utilized for testing, validating, and calibrating acquisition systems, namely in medical imaging like Positron Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) scans where standardized phantoms are available [6,7]. In EEG, a standardized phantom does not exist, although many prototypes from different materials and designs have been proposed.

A common approach used by research groups to test novel EEG systems or new types of electrodes consists of using salt-doped ballistic gelatin (BG) head models onto which scalp EEG sensors can be mounted for testing [8,9]. Nonetheless, phantoms can also be made from different organic (e.g., agar) or synthetic (e.g., carbon-doped thermoplastics) materials [10,11]. The basic principle for an EEG phantom prototype is to obtain a conductive physical model with the shape of the human head. This is usually achieved through image processing of anatomical scans (i.e., MRI or CT) and 3D printing of the resulting phantom's cast for molding [12].

An EEG phantom allows for the playback of a previously recorded known EEG signal. This recognizable signal is referred to as the "ground truth". Then, measurements can be done on the phantom for comparison with this known signal, which constitutes a way of characterizing the acquisition system without the need to account for the inherent variability and lack of repeatability of neural signals present when recording from human subjects, as well as identifying external sources of noise [13,14]. For a known previously recorded signal to be played out of the phantom, electrode antennas are driven inside it. These can be simple screws, exposed wire tips, or a coaxial cable to create dipoles [11,15,16]. The driven antennas can be either attached to the interior side of the scalp layer in hollow-shaped phantoms or are put in place during a phase of assembly where the constitution of the filling material allows for this procedure [11,17]. Additionally, EEG phantoms allow for a controlled analysis of electrode contact impedance measurements and signal noise floor characteristics [17,18].

Despite the existence and use of EEG phantoms for evaluating scalp based EEG systems, there is a lack of literature for ear-EEG validation via appropriate phantom models. To the best of our knowledge, appropriate EEG phantoms for this purpose do not exist and all previous EEG phantoms neglect the structures of the outer ear and ear canal systematically.

The feasibility of measuring brain signals through ear-EEG devices has been validated in recent publications [4,19,20]. For an in depth review of the field and overview of the technological state-of-the-art around ear-EEG we please refer to [21,22]. A commonality across the ear-EEG literature, is that the validation paradigms utilized tend to be known ERP paradigms, namely the alpha modulation (or alpha blocking) paradigm, the Auditory Steady-State Response (ASSR), the Steady-State Visual Evoked Response (SSVEP), Auditory Evoked Potentials (AEPs), Visual Evoked Potentials (VEPs) and oddball type paradigms to elicit responses linked to higher processes of the brain, like the P300 and Mismatch Negativity (MMN) responses. Similarly, the interest of utilizing

electro-oculography (EOG) measured by ear-EEG as a possible input for BCI applications as also been explored [20].

Given the predominance of these select group of paradigms utilized to derive ear-EEG responses, and the lack of a dedicated phantom model suitable for ear-based systems, here we propose a complete ear-EEG validation toolkit to contribute and expand to the characterization of this technology through:

1. A software framework ("EaR-P Lab") that allows the user to readily make a validation test battery for the characterization of ear-EEG devices at the neural signal acquisition level.
2. The design and prototyping of an ear-EEG suitable physical phantom for systematic characterization of in-ear sensors, allowing controlled comparison of fit form factors for ear-EEG acquisition.

This contribution to the field aims to allow for a more reliable assessment of out-of-the-box ear-EEG devices by providing proper benchmarking tools for comparing systems.

2. Materials and Methods: Ear-EEG Toolkit Design and Validation

2.1. EaR-P Lab - Design and Validation

A custom application, programmed with nine commonly utilized EEG recording paradigms, was created for the assessment of ear-EEG devices at the neural signal level. This application was written in the Python programming language and consists of a collection of scripts (requiring a local Python installation), pre-synthesized stimuli, and other auxiliary files and folders wrapped in an usable Graphical User Interface (GUI), based on the Tkinter Python package. The stimuli delivery and precise experiment timings are handled via PsychoPy, a Python based open source package for experiment control. [23].

We dubbed this software script as "EaR-P Lab" from the amalgamation of the words "ear" and "ERP", as per the motivation of streamlining the measuring of ERPs (as well as other paradigms) from the ear. The organizational structure and main functionalities of EaR-P Lab are depicted in Figure 1.

**EaR-P Lab - validation stimuli and markers**
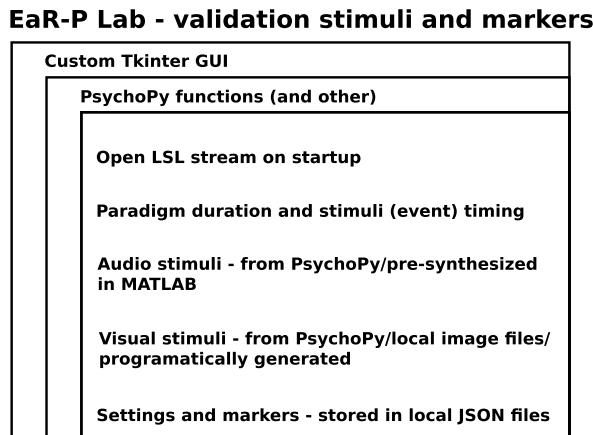
Figure 1. EaR-P Lab - structure and main attributes.

When initialized, EaR-P Lab provides a user-friendly GUI to present the necessary stimuli and insert the relevant markers that delimit different paradigms or indicate the start of a given stimulus in the validation test battery. The marker's timestamps are synchronized with the EEG data through the Lab Streaming Layer (LSL) protocol that handles the synchronization between the EEG data and maker streams through a local network (an EEG amplifier compatible with LSL must be used) [24,25]. The marker and EEG data streams can then be recorded with another software (i.e., LabRecorder) into a single *.xdf* file on the hard disk of the computer for offline analysis, as illustrated in Figure 2.
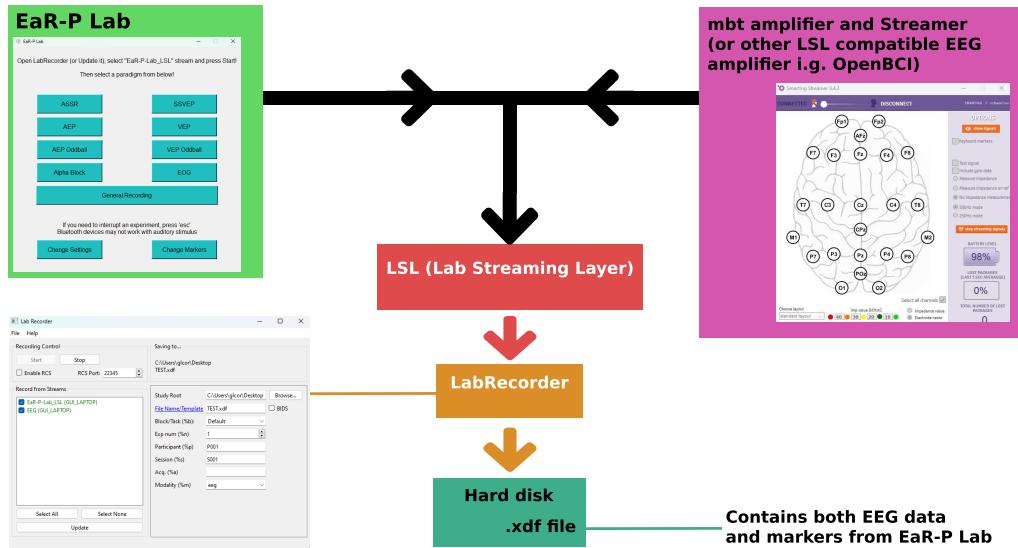
Figure 2. EaR-P Lab - functioning framework.

2.1.1. GUI - Main Menu

When EaR-P Lab is launched, the Main Menu in Figure 3 appears. This menu gives the user access to the nine pre-loaded EEG acquisition paradigms, with the script automatically opening the LSL marker stream using the *pylsl StreamOutlet* function at startup. Eleven buttons are presented: nine that direct to the different paradigms and two (Change Settings and Change Markers) that allow users to modify specific experiment parametres. Some disclaimer information on how EaR-P Lab operates is also shown in this window. Refer to Appendix A for a full overview of the features and functions of each button and settings menus.
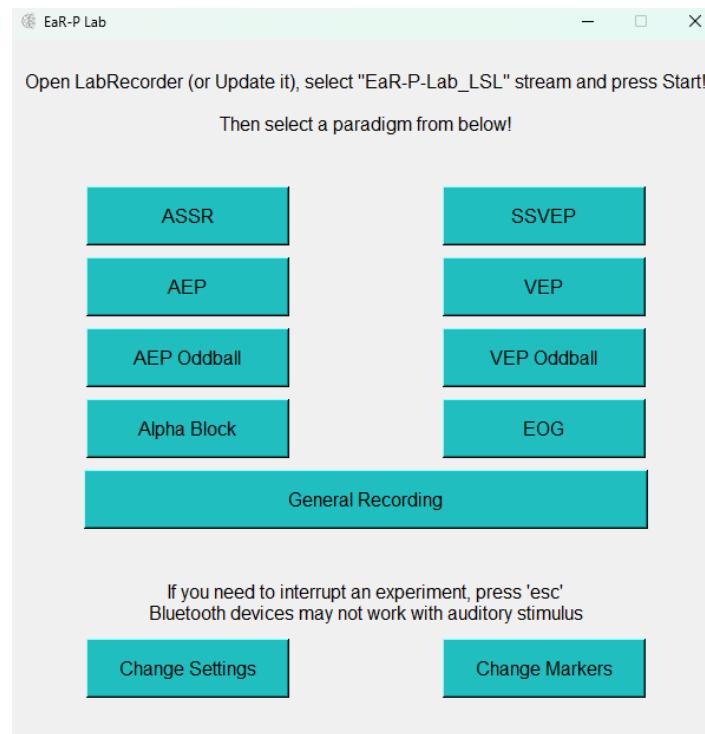

Figure 3. EaR-P Lab - Main Menu.

2.1.2. Stimuli and Trigger Latency

To accurately evaluate Event-Related Potentials (ERPs), good synchronization between data and markers sent at stimulus onset (triggering) is crucial [26]. The PsychoPy library allows sounds to be pre-scheduled with the high precision "PTB" settings and the *callOnFlip* method to start visual stimuli promptly, both used in EaR-P Lab [27,28]. However, unavoidable latency and jitter issues related to screen refresh rates and monitor syncing may still arise [29].

To measure the timing characteristics of the EaR-P Lab we used an amplifier specific accessory (i.e., mBrainTrain Delay/Jitter (DJ) box). This accessory includes an audio input, output, and photo-diode sensor that allowed us to test the marker synchronization with the audio and visual stimuli. The device was placed at the latest stage possible in the audio delivery setup, with the photo-diode centered and 30 cm away pointing to the center of the screen.

During the early testing phase, we observed an increasing cascading effect on the latency delay when recording different blocks of transient responses within the same *.xdf* file, illustrated in Figure 4.
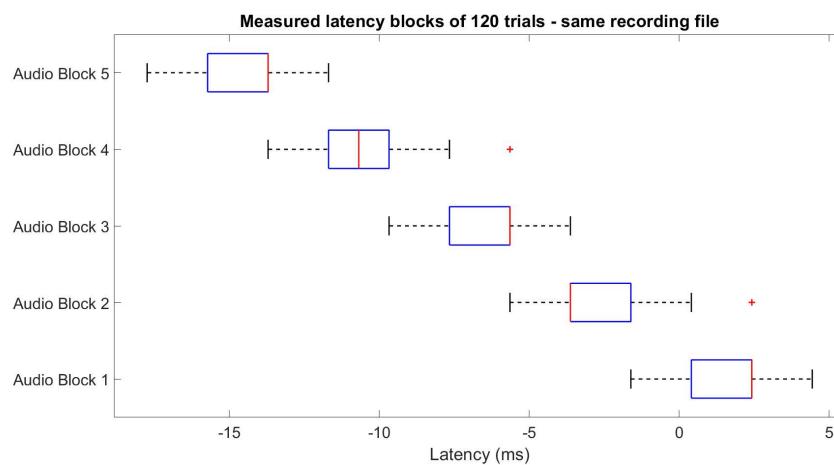

Figure 4. Escalating latency when recording multiple Event-Related Potential (ERP) blocks on the same file, exemplified for auditory stimuli - a similar effect happens for visual stimuli.

Our troubleshooting showed that to correct this effect, the data streaming from the amplifier software (mBrainTrain Streamer) had to be restarted after each EEG recording paradigm. Hence, each paradigm must be recorded in separate *.xdf* files to ensure reliable trigger latencies, as shown in Figure 5.
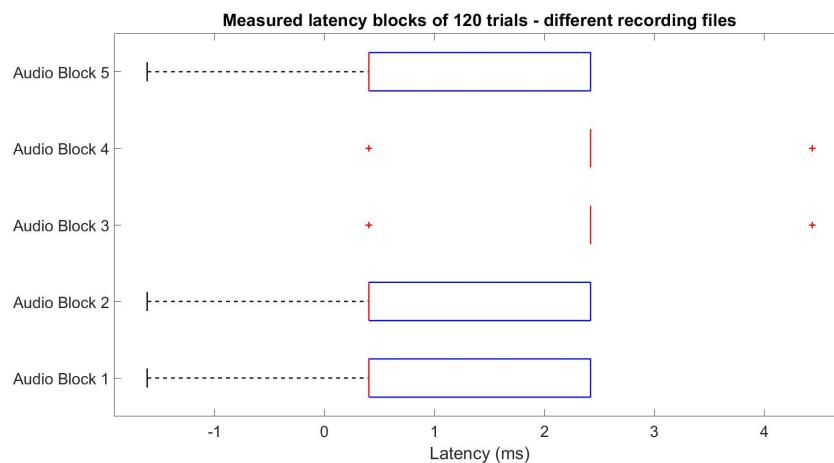

Figure 5. The cascading effect is nullified when recording multiple ERP blocks in different files after restarting data streaming, exemplified for auditory stimuli - a similar effect happens for visual stimuli.

However, to the best of our knowledge, this may only apply to the mBrainTrain Smarting Mobi recording platform. We recommend testing the system's latency before recording and having separate *.xdf* files for the AEP, VEP, EOG, and oddball type paradigms to avoid latency synchronization issues.

Once our latency issues were resolved, the latency delay was measured for 1000 trials of both auditory and visual events, showing a mean delay of 0 ms (2.2 ms std./jitter) and 21 ms (2.8 ms std./jitter), respectively. The processing of the EEG paradigms accounted for these measured latencies. These values are specific to the used stimuli presenting laptop and are likely to vary between systems.

2.1.3. Acquisition Setup and Test Battery

In order to validate EaR-P Lab as tool to assess ear-EEG (and EEG in general) standard scalp EEG data were acquired for each of the paradigms of this tool. EEG data were collected from five individuals with normal or corrected vision/hearing (age avg. 27.0 ± 4.3 std.). A custom EEG cap with 12 channels was utilized for the EEG control group acquisitions. The cap uses conventional Ag/AgCl electrodes at F3, Fz, F4, T7, C3, Cz, C4, T8, P3, Pz, P4, and Oz positions. The reference electrode (CMS) and ground (DRL/GND) were situated at FCz and AFz, respectively. Data were collected using the mBrainTrain Smarting Mobi amplifier, capable of high-quality recordings at a sampling rate of 500 Hz with wireless data transmission, providing 24 bits of resolution in the analog-digital converter (ADC) [30].

A laptop (ASUS VivoBook 15, Windows 11 OS, Intel Core i7-10750H Processor, NVIDIA GeForce GTX 1650 Ti Graphics Card, 12GB RAM) ran the essential recording software (mBrainTrain Streamer and LabRecorder) and stimulus presentation (EaR-P Lab). Data was transmitted via Bluetooth from the amplifier at the back of the subject's head to a USB dongle connected to the laptop. Research-grade ER-2 earphones (Etymotic Research,INC.), shielded with copper tape to reduce interference, delivered sound stimulus to the subjects. To ensure audible stimuli without distortion, an external sound card (TASCAM US-100) and digital-analog converter (DAC, FiiO Alpen2) were used (see Figure 6). Laptop and the DAC volumes were set to the maximum, with adjustments made on the sound card only. Earphones were fitted with foam ear tips. Laptop brightness settings were maximized for visual stimuli.
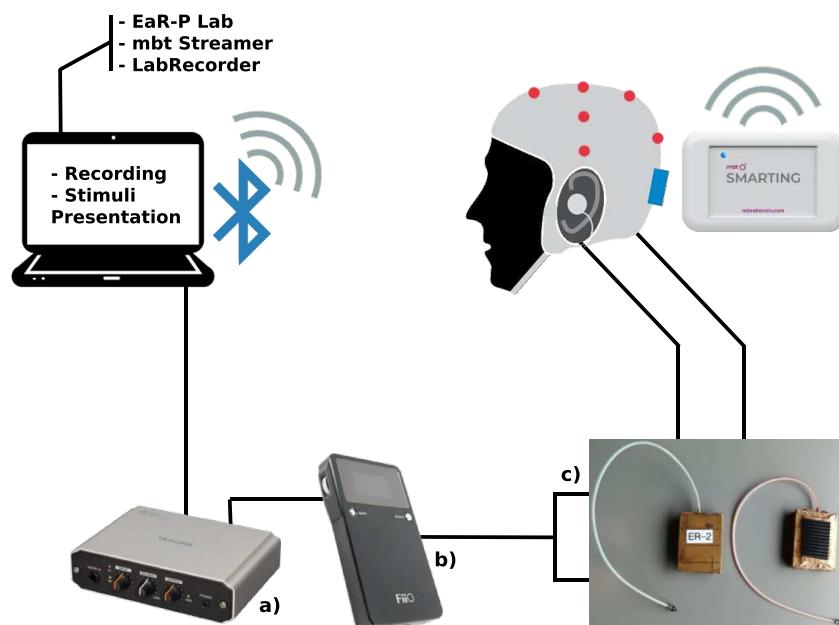

Figure 6. EEG acquisition setup schematic and equipment - **a)** USB Audio Interface TASCAM US-100 **b)** - Digital-Analog Converter (DAC) Amplifier FiiO Alpen 2 **c)** - ER2 Etymotic tubal insert research grade earphones

EEG recordings were done in a quiet environment with only the participant and the technician present. The participant was asked to sit in a comfortable chair and relax during the recording. The participant was given a brief explanation of the recording process at the start of the session, with thorough instructions provided before each task.

Conductive gel (Signagel, Parker) was applied on each scalp electrode, generally achieving impedances below 5 kOhm, measured by the mBrainTrain Streamer.

The EEG data standard and ordered acquisition protocol and settings for each EaR-P Lab paradigm employed were as follows:

1. **General Recording -** 4 minutes of resting state pre-testing
2. **Alpha Block -** 1 minute per section, 4 minutes total
3. **ASSR -** 4 minutes of continuous 1 kHz carrier 40 Hz modulated sound wave
4. **SSVEP -** 4 minutes of continuous 10 Hz visual stimulation, subject at a distance of 60 cm from the middle of the screen, measured from the point in between the eyes, during this paradigm, the room's light was turned off, to maximize the flickering
5. **AEP -** 200 trials of discreet 1 kHz sound stimulus, with 200 ms of duration and a 10 ms ramp up/fall off, Interstimulus Interval (ISI) between 1200 ms and 1800 ms, for a total duration of around 7 to 8 minutes
6. **VEP -** 200 trials of onset-offset pattern reversal with the dartboard target being shown for 500 ms of duration and a 500 ms ISI, taking about 5 minutes to complete
7. **AEP Oddball -** Standard: 440 Hz sound wave | Target: 880 Hz sound wave - a total of 200 oddball/target events were presented to the subject, with every sound having a duration of 100 ms with a 10ms ramp up/fall off and ISI between 1200 ms and 1800 ms, for a total length of around 15 minutes
8. **VEP Oddball -** Standard: Blue Square | Target: Red Circle - a total of 200 oddball/target events were presented to the subject, stimuli on screen for 500 ms, ISI between 600 ms and 700 ms, subjects were told beforehand to press "SPACE" once the target stimuli appear, for a total length of around 18 minutes
9. **EOG -** 80 trials per saccade, dot moved for 500 ms, ISI between 1000 ms and 1600 ms, subject at a distance of 30 cm from the middle of the screen, measured from the point in between the eyes, which translates to 16.2° angle for each saccade, for the follow-the-dot paradigm, the subject's head was stabilized using an adjustable chin rest and the monitor was leveled and centered with the subject's eyes by being placed on a box. This section had a total duration of about 10 minutes

Subjects were allowed to calibrate the volume themselves to the highest volume within their own comfort audible zone. Each acquisition session took about two hours to complete from start to finish, including setting up the EEG cap. The subjects were entitled to short breaks (2 to 5 minutes) at the end of each paradigm but were also allowed longer breaks when requested. For some subjects, the protocol's order wasn't strictly followed.

2.1.4. Processing and Statistical Analysis

All data processing and plotting of results was done via custom scripts in MATLAB (MATLAB. (R2020a). Natick, Massachusetts: The MathWorks Inc.). EEG data were first parsed and filtered channel by channel with the *filtfilt* function: first, a highpass filter (*butter* function) was applied, followed by a 50 Hz notch filter (*iirnotch* function), and finally, a lowpass filter (*butter* again). Frequency-based paradigms (ASSR, SSVEP, and Alpha Block) were filtered between 1 Hz and 100 Hz, and transient responses (AEP, VEP, MMN, P300, and EOG saccades) were filtered between 1 Hz and 20 Hz. To isolate the EOG blinks in the respective three second window, this data was filtered between 0.2 Hz and 3 Hz. The data were then referenced to the Cz electrode.

ASSR and SSVEP frequency responses were computed with the *pwelch* function on an 8 s window with 50% overlap, and the alpha block spectrograms were obtained with the *spectrogram* function on a 2 s window, 50% overlap.

All transient responses were baseline corrected to the 100 ms pre-stimulus interval and then averaged across trials. To obtain the MMN and P300 responses, the average standard cue waveform was subtracted from the average target cue waveform for each subject, with the standard events immediately before the oddballs being considered [31].

To statistically assess ASSR responses, the power of the 40 Hz frequency of interest is compared with the surrounding frequency bins, according to Equation 1 [19]:

$$40 \ Hz \ SNR = \frac{P \ (40 \ Hz)}{P_{\text{average}}(35 - 45 \ Hz)^*} \ , \ {}^*excluding \ 40 \ Hz \tag{1}$$

With the same ratio adapted for the SSVEP 10 Hz response as per Equation 2:

$$10 \ Hz \ SNR = \frac{P \ (10 \ Hz)}{P_{\text{average}}(5 - 15 \ Hz)^*} \ , \ {}^*excluding \ 10 \ Hz \tag{2}$$

A statistical f-test was also conducted on these ratios using the *anova1* function.

From the same reference, Equation 3 is used to calculate the power of the alpha block effect between the eyes closed and eyes open by averaging both same condition sections into one and considering the alpha frequencies defined between 8 Hz and 12 Hz.

$$R_{AM} = \frac{P_{\text{average}} \ (Alpha \ Band_{\text{Eyes Closed}})}{P_{\text{average}} \ (Alpha \ Band_{\text{Eyes Open}})} \ , \ Alpha \ Band \ (8 - 12 \ Hz) \tag{3}$$

The modulation significance was also assessed through a t-test using the *ttest* function.

ERP responses (AEP, VEP, and respective oddball counterparts) waveforms will be shown with highlighted components deemed significant through a t-test using the *ttest* function between the values of the averaged waveform of each subject, at each time point, and zero [32].

EOG saccade profiles will be quantified by the amplitude and polarity of each saccade type. For the EOG blinks, an amplitude factor between the averaged peak-to-peak amplitude of the intentional blinks over the regular blinks will be established [19].

Overall, a statistically significant p-value of 5% confidence (p < 0.05) was considered.

2.1.5. Alpha Block

As seen in Figure 7, a distinct effect on alpha wave blocking with the suppression of the 8 Hz to 12 Hz band when opening the eyes was achieved with the highest modulation of 6.1 dB at Oz, over the occipital lobe. Near the ear at T8, the modulation dropped to 3.7 dB.
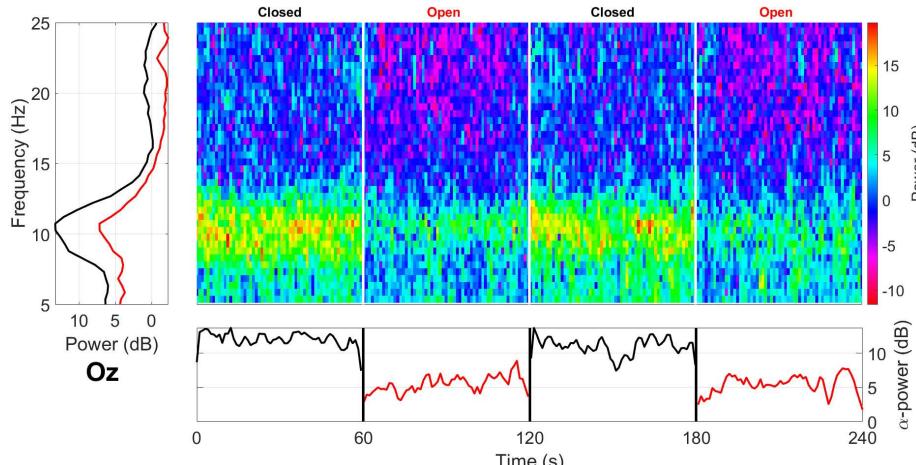

Figure 7. Alpha block grand average spectrogram over the control group at Oz (Cz referenced) - the bottom plot shows the alpha power mean power (8 Hz to 12 Hz) per section while the left vertical plot shows the frequency response between the two conditions.

2.1.6. ASSR

For the ASSR (Figure 8), peak frequency responses at the used 40 Hz modulation frequency of the stimulus were obtained, with the highest mean Signal-to-Noise Ratio (SNR) at P4 with 9.9 dB, lowering to a 4.7 dB significant response at T8.
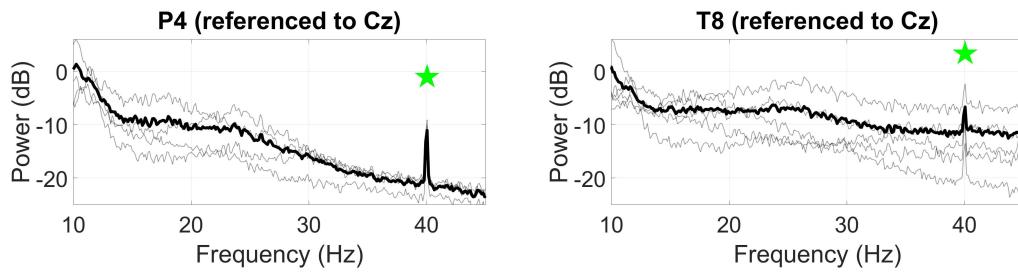

Figure 8. Grand average Auditory Steady-State Response (ASSR) (dB) to a 40 Hz frequency modulated stimuli, at **(Left)** P4 and **(Right)** T8. Statistically significant peaks are highlighted by the green star token, based on an f-test ($p < 0.05$).

2.1.7. SSVEP

The SSVEP expected peak frequency of 10 Hz response was better at the Oz location with a mean SNR of 11.0 dB, as seen in Figure 9, where the peaks for several harmonics of 10 Hz fundamental frequency are also present, up to the 8th harmonic. At T8, the ratio was lower, with a mean SNR of 7.5 db as we are further away from the visual cortex, with harmonic responses only distinguishable up to 40 Hz.
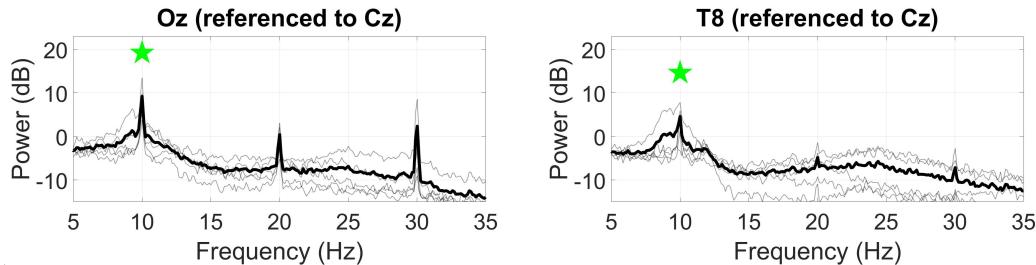

Figure 9. Grand average Steady-State Visual Evoked Potential (SSVEP) (dB) to a 10 Hz visual stimuli, at **(Left)** Oz and **(Right)** T8. Statistically significant peaks are highlighted by the green star token, based on an f-test ($p < 0.05$). Only the first harmonic was statistically evaluated.

2.1.8. AEP (N100)

The response to the auditory 1 kHz stimuli is shown in Figure 10, presenting three significant components, with the biggest being the expected deflection around 100 ms with a 5 µV amplitude corresponding to the N100 component (positive in this referencing), resulting in a peak-to-peak amplitude of 8.9 µV between the N1 and P2 components.
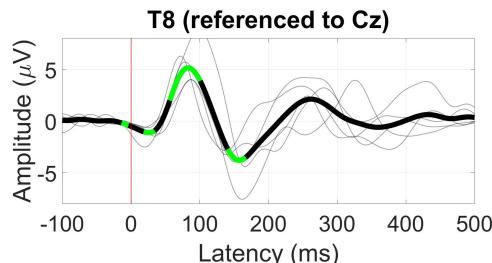

Figure 10. Grand average Auditory Evoked Potential (AEP) waveform at T8. Statistically significant segments are highlighted in green, based on a t-test ($p < 0.05$).

2.1.9. VEP

The obtained VEP waveform for the onset-offset pattern reversal paradigm is shown in Figure 11. At Oz, a significant component is present between 150 ms and 300 ms with a peak amplitude of 15 µV at 200 ms. In T8 near the ear, the same component isn't significant and, instead, shows two smaller positive deflections at 90 ms and 300 ms of 2 µV and 3 µV, respectively.
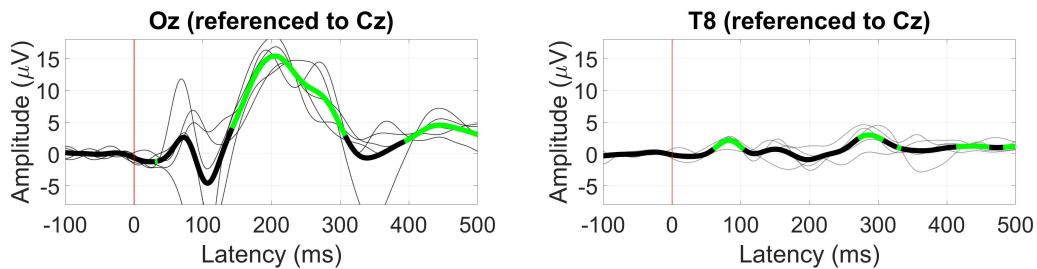

Figure 11. Grand average Visual Evoked Potential (VEP) waveform at **(Left)** Oz and **(Right)** T8. Statistically significant segments are highlighted in green, based on a t-test ($p < 0.05$).

2.1.10. AEP Oddball (MMN)

The MMN response was the strongest at electrode site Pz, shown in Figure 12. At 200 ms, the expected negative differentiation was observed between the oddball and standard stimuli with an amplitude of 2 µV, still present at T8, with the same amplitude.
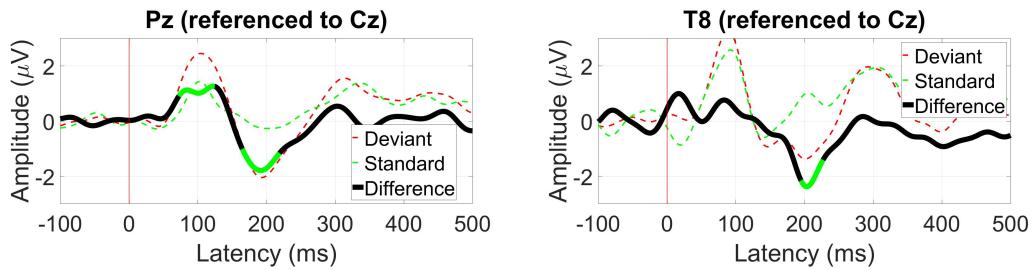

Figure 12. Grand average Mismatch Negativity (MMN) waveform at **(Left)** Pz and **(Right)** T8. Statistically significant segments are highlighted in green, based on a t-test ($p < 0.05$).

2.1.11. VEP Oddball (P300)

The P300 component characteristic of the subject's active response to an oddball stimulus was better observed at P4 between 280 ms with a 6 µV amplitude, lowering to 4 µV at T8 (Figure 13) with peak amplitude achieved at around 250 ms from the onset.
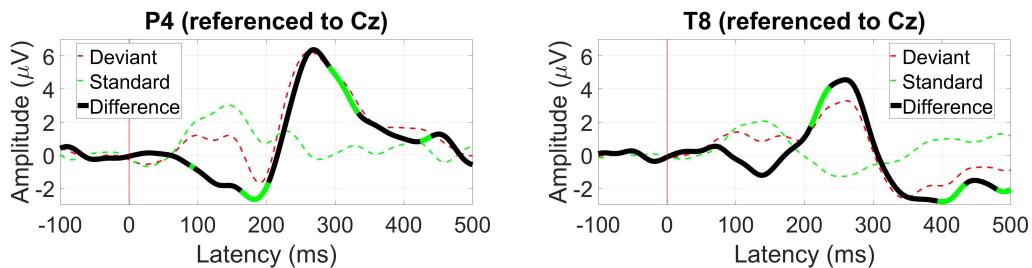

Figure 13. Grand average P300 waveform at **(Left)** P4 and **(Right)** T8. Statistically significant segments are highlighted in green, based on a t-test ($p < 0.05$).

2.1.12. EOG (Blinks and saccades)

While not cortical in origin, EOG responses regarding blinking amplitude and saccade profiles are also assessed as possible inputs for BCI applications. The response is naturally better at the scalp level at frontal electrodes like Fp1 and Fp2, near the eye muscles responsible for these movements. In our validation scalp setup, we considered F3 and F4 as approximate locations for Fp1 and Fp2.

At F3, in Figure 14, for an exemplary subject, the peak-to-peak amplitude difference between hard (8 mV) and regular (1.5 mV) can be observed, with a lower peak-to-peak amplitude but about the same ratio at T8.
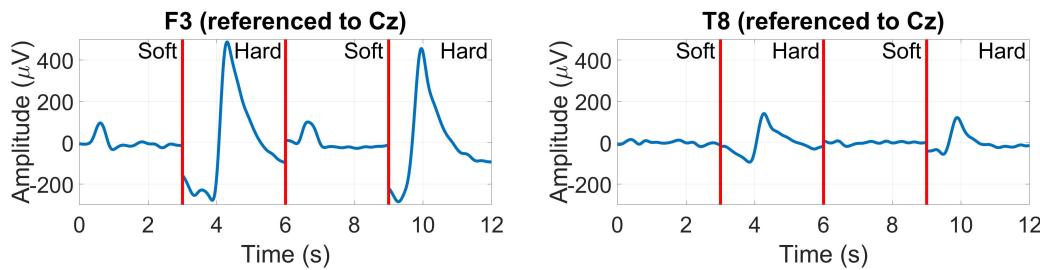

Figure 14. Soft/hard blink amplitudes, for an exemplary subject, at **(Left)** F3 and **(Right)** T8.

Figure 15 shows the saccade profiles in different directions. The saccades were assessed around 200 ms after stimulus onset, where the separation between directions is greater. Between T7 and T8, it is possible to see a 10 µV difference between the right (blue line) and left (green line) saccades, with their polarity inverting, depending on the side of the sensing electrode. The vertical-oriented saccades are less differentiated, with only around 2 µV amplitude separating their traces.
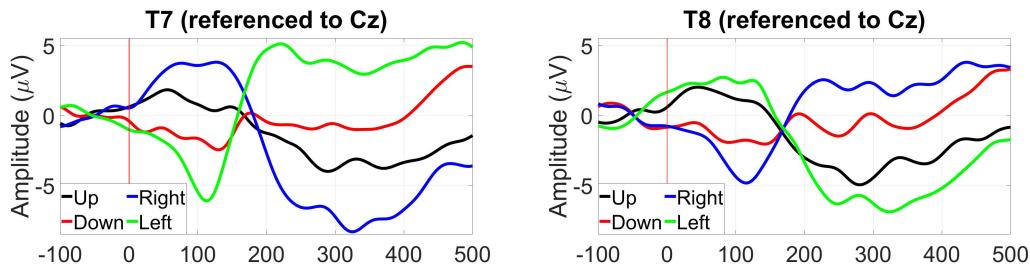

Figure 15. Grand average saccade profiles on the four cardinal directions, at **(Left)** T7 and **(Right)** T8.

2.2. Ear-EEG Phantom - Design and Validation

A physical test bench to evaluate ear-EEG devices was developed. The basis for the prototype was similar to the most common type of EEG phantoms: a mold that creates the proper shape and fit for use with an ear-EEG device and that can be filled with a conductive mixture.

In contrast to most common EEG phantoms, this work focused on developing a compact design solely focused on the ear anatomy instead of a complete head-shaped phantom. We also set out to create a modular ear section for the mold, allowing for testing different ear form factors without having to rebuild an entirely new mold, as well as having a way to secure the signal antennas in place during the setting process of the filling material.

The phantom was designed using the Fusion 360 Computer Aided Design (CAD) software and then exported as a *.stl* file for 3D printing. The base design is an 18 cm modular cylinder split into two parts along its length, with open ends. The bottom half has two longitudinal railings where the top fits, leaving two round holes centered in the middle for inserting antennas. The cylinder sides can then be closed with the lids. The top has two square openings for pouring of the conductive mixtures. A 0.1 mm gap was used for all fitting parts to account for the 3D printing resolution. The rendered prototype is shown in Figure 16, and the phantom prototype's full dimensions are presented in Appendix B.
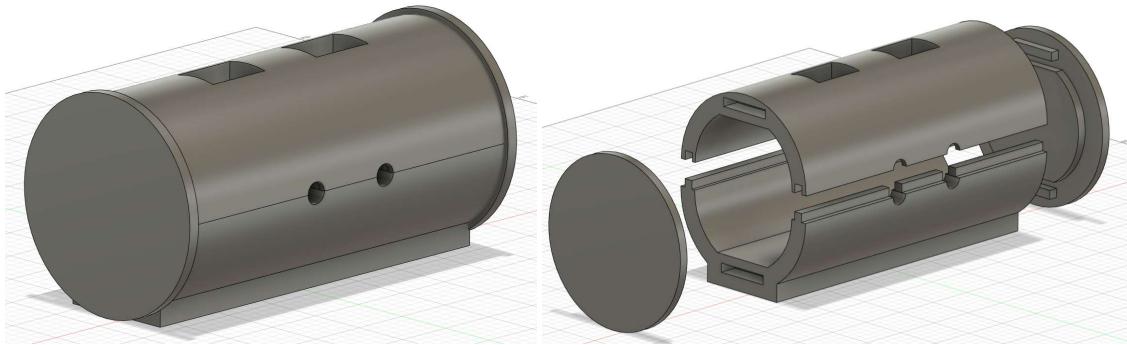
Figure 16. (Left) Closed render of the ear-EEG phantom. **(Right)** Exploded render of the ear-EEG phantom.

The inside of these lids can then accommodate the outer ear canal and concha shape to create an imprint on the material inside, where an ear-EEG device can be placed for testing. The current methodology used ear canal scans done professionally by a local audiologist, digitally scanned on-site, and delivered as a *.stl* file (as shown in Figure 17), which can be imported as a mesh to Fusion 360 and easily fused with the lid mesh by using the *Mesh Menu - Modify - Combine - Join Operation*.
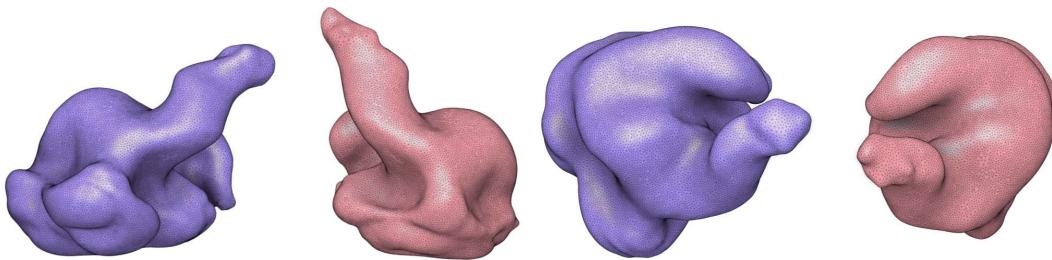

Figure 17. Outer ear scans (blue - left ear, red - right ear) from a subject, received as *.stl* files, shown from two perspectives, obtained by an expert audiologist.

Ear canal scans were centered on the lid's mesh with the concha cymba outline parallel to the mesh's top/bottom. Meshes intersected to a considerate adequate depth as long as the ear canal and concha structures were visible and protruding, similar to Figure 18.
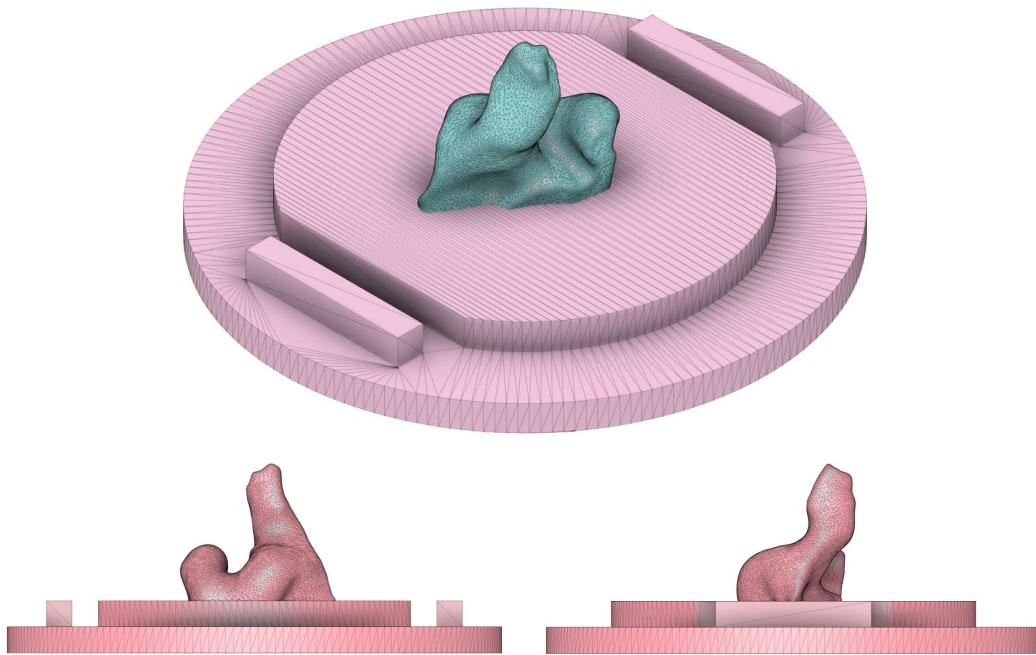
Figure 18. Example of a left ear scan being centered and orientated with the phantom's lid mesh and different views of the alignment and depth of the ear mesh and the lid mesh into a single rendered object.

The ear-EEG phantom prototype was 3D printed on a Prusa i3 MK3 3D printer using Polylactic Acid (PLA). PLA was chosen due to its ease of use, fast printing times, affordability, non-warping nature, and lack of post-processing requirements, making it an optimal choice for this design. The printing settings on the Prusa software were set to a *0.15mm QUALITY* printing resolution, using *Prusament PLA* filament with a default structural infill of 15%. The software automatically added necessary support structures *Everywhere* required. The combined printing time of the phantom body and lids was approximately 33 hours. To minimize difficult to remove supports, the cylinder halves were printed vertically. If the phantom's form-fitting factor needs to be changed, printing just another set of lids with the ear impression takes only 6 hours. Figure 19 shows the different modular parts after printing and support removal.


Figure 19. Disassembled ear-EEG phantom - bottom half (in yellow), top half (in white, similar to yellow), and two lids with a left and right ear imprint from one of the subjects.

2.2.1. Phantom assembly and bulk materials

To assemble the phantom, the antennas (pair of 3.5 mm AUX cables) were adjusted to fit the diameter of the cylinder holes by inserting one side of each cable inside PVC tubing and sealing it with duct tape. The cables are placed on the bottom half, then joined by the top and the lids with ear imprints on the side, holding the prototype together. Three strips of plumbing tape are used on the railing fittings, cables, and the inner circumference of the phantom lids to prevent leaks, as in Figure 20. The phantom can then be filled with a conductive substance.
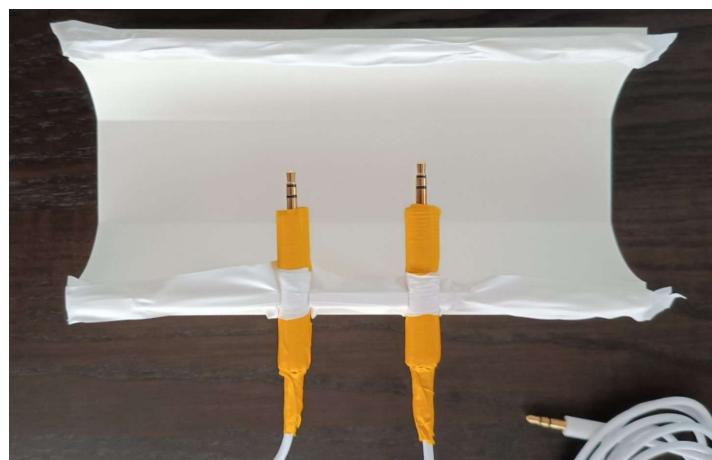
Figure 20. Ear-EEG phantom assembly - antennas and railing fittings were sealed with tape

Here, three materials were tested with our design: agar doped with salt, BG doped salt, and silicone doped with carbon fiber (CF), as a non-perishable option.

Agar and BG

Agar is a gelatinous vegan substance derived from certain types of seaweed, making a stable gel structure when cooled down after being boiled. An agar phantom can be created by mixing agar, water, and salt at the recommended weight percentages (% w/w) of 4% and 0.5%, of agar and salt, respectively, as per Equations 4, with A being the weight of agar, s the weight of salt and w the weight of water [10].

$$\begin{cases} \frac{A}{A+w+s} = 4.0\% \\ \frac{s}{A+w+s} = 0.5\% \end{cases} \Rightarrow \begin{cases} A = \frac{8}{191}w \\ s = \frac{1}{191}w \end{cases} \tag{4}$$

The phantom was overestimated to have a capacity of 700 ml (equivalent to 700 g of water). By substituting w for that value in the solved equations mentioned above, we get the necessary quantities of 30 g of agar (*Hoosier Hill Farm*, 10.20€ for 115 g, in Amazon) and 4 g of salt, rounded up to the nearest unit. The mixture was prepared as follows:

1. Boil 700 ml of regular tap water (or deionized water)
2. Add 30 g of agar slowly while stirring the mixture
3. Add 4 g of table salt while stirring until no granules are present - keep mixing while letting it cool down at room temperature for 10 minutes
4. Pour the mix into the assembled phantom through the top vents until the liquid reaches half the vent's height, let it sit in a refrigerator until it fully solidifies (minimum 2 hours, preferably overnight), which can be checked through the top

BG is a common material for EEG phantom making, usually made of cattle/pork gelatin. Various weight percentage combos of both gelatin and NaCl are shown to be an appropriate conductive medium [17]. Due to an accessible lesser amount of acquired gelatin bulk powder (*YouHerbIt* 240

Bloom, around 20€ for 350 g of powder, in Amazon), values within the middle range of the above study were used, aiming for a 15% weight percentage on gelatin and 8% weight percentage of NaCl, with *BG* meaning the weight of gelatin powder instead of agar, which when solved wield the Equations shown in 5.

$$\begin{cases} \frac{BG}{BG+w+s} = 15.0\% \\ \frac{s}{BG+w+s} = 8\% \end{cases} \Rightarrow \begin{cases} BG = \frac{15}{77}w \\ s = \frac{8}{77}w \end{cases} \tag{5}$$

The same steps as for agar were followed to prepare the gelatin solution, with 140 g of gelatin powder mixed with 70 g of salt in 700 ml of hot (not boiling) water. Unlike agar, gelatin is better mixed at a lower temperature to prevent the formation of air pockets.

Both agar and BG ear-EEG phantoms are shown in Figure 21.
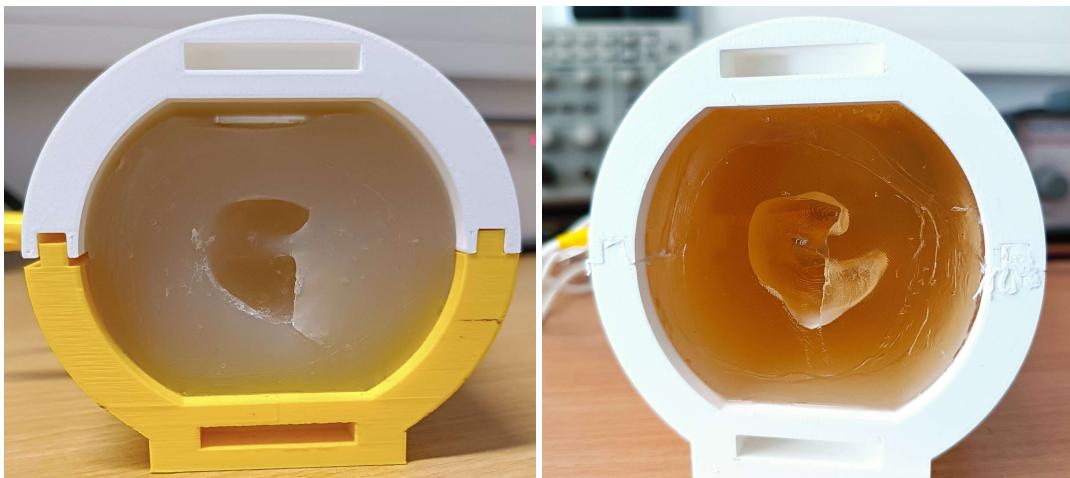
Figure 21. (Left) Agar ear-EEG phantom. **(Right)** Ballistic Gelatin (BG) ear-EEG phantom.

Silicone doped with carbon fiber

A non-perishable phantom that can maintain signal integrity indefinitely was also envisioned, as the aforementioned organic materials tend to deteriorate over time.

Conductive fillers have been tested to make silicone rubbers conductive, with carbon fibers found to be the most effective even at a 0.5% weight percentage measuring a 2 kohm resistance. Increasing to 1% further lowered resistance to 200 ohm, and higher percentages gave similar resistance, indicating that the percolation threshold (filler concentration at what a conductive network is fully established) was reached [33].

A Do-It-Yourself (DIY) approach in doping the silicone was followed [34], the acquired materials being:

- Chopped carbon fibers, 3 mm in length, 30€ for 500 g, on Amazon
- Two-part A/B system platinum curable silicone, mixing ratio of 1:1, 23€ for around 630 ml, on Amazon (two were acquired for the phantom)

The silicone was estimated to have a density of 1.11 g/cm^3. Various weight of carbon fiber percentages were tried according to Equation 6 to establish a percolation threshold:

$$\frac{CF}{CF+S} = X\% \ (where \ X = 0.5, 1.0, 1.5, 2.0 \ ...) \tag{6}$$

With *CF* being the weight of carbon fibers and *S* the weight of silicone. It was observed that when the weight percentage of CF exceeded 1%, the mixture became too dense and unsuitable for molding with the current mixing methods. Therefore, only 0.5% and 1.0% percentages were considered and prepared according to Equations 7:

$$\begin{cases} CF = \frac{S}{199} \ , \ for \ X = 0.5 \\ CF = \frac{S}{99} \ , \ for \ X = 1.0 \end{cases} \tag{7}$$

This non-perishable phantom was ultimately made for a 1% percentage, with 8 g of carbon fibers (rounded to nearest unit) being used for a total of 700 ml of silicone, following the next steps [34]:

1. Measure 8 g of carbon fibers into a disposable cup (use a mask/gloves when handling carbon fibers)
2. Wet the carbon fibers with a small amount of rubbing alcohol, spread them around, and let almost entirely evaporate (to release strands of hair that surround the carbon fibers)
3. Add the carbon fibers to 350 ml of part A silicone and mix thoroughly (an electric mixer with a wider spatula attachment was used) until the mix presents a grey/blueish tint
4. Add 350 ml of part B silicone and keep mixing for up to 25 minutes to the same tint
5. Pour into the phantom, equally through each vent, and let cure for 6 hours

To characterize the conductivity profiles of our chosen materials, fine strips of each material (30mm x 10mm x 2mm) were prepared specifically for conductivity testing. Both ends of each sample were coated with conductive silver ink and standard two-wire method was employed to measure the resistance across the sample. Conductivity of the samples was then calculated according to Equation 8.

$$\begin{cases} \rho = R\frac{A}{L} \\ \sigma = \frac{1}{\rho} \end{cases} \Rightarrow \sigma = \frac{L}{R \times A} \tag{8}$$

Where ρ is the material resistivity, R is the measured resistance, A is the area of the cross-section of the sample, and L is the sample's length. σ is the material's conductivity as the inverse of resistivity. . The reference conductivity values for each material are listed in Table 1.

Table 1. Electrical conductivity, in Siemens per meter, for samples of agar, BG and silicone doped with carbon fiber (CF) (1%) as the proposed materials for the ear-EEG phantom.

	Conductivity [S/m]
Agar	0.309
BG	0.918
CF (1%)	14.035

While the sample preparation for the carbon-doped silicone samples was successful, the mixing method employed was not suitable for a homogeneous distribution of carbon fibers in the for the larger phantom mold. This resulted in the creation of conductive (darker) and non-conductive (translucent, low or no carbon fibers) areas as perceived in Figure 22. Therefore this material was discarded from further testing.
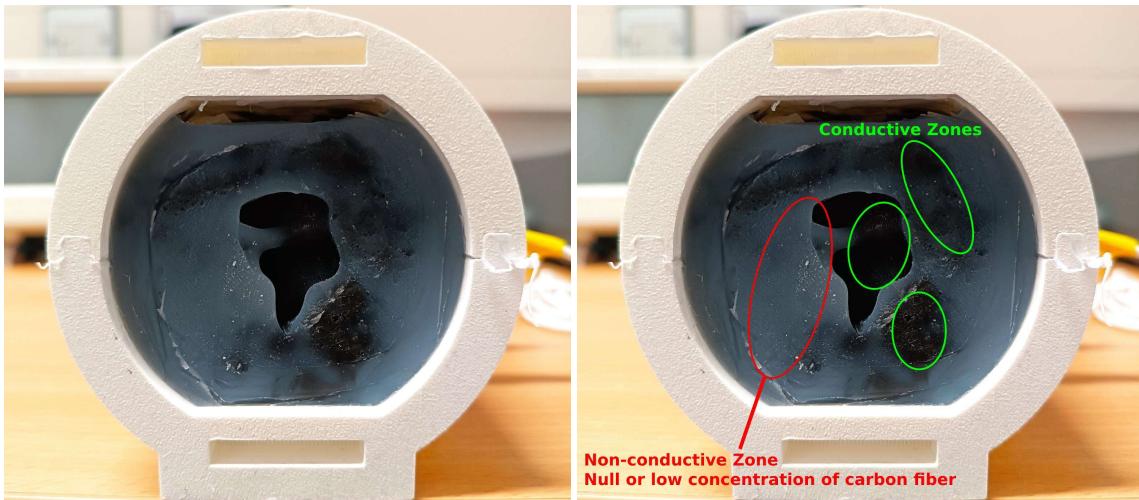
Figure 22. Silicone doped with carbon fibers ear-EEG phantom - the lack of conductive homogeneity is highlighted on the right, with conductive and non-conductive zones being visible.

2.2.2. Ear-EEG phantom testing protocol and setup

The agar and BG phantoms were tested over eight days, alternating daily for a total of four sessions each (i.e., "Day 4" is the last testing day per phantom), and refrigerated while not in use (4ºC). All phantoms were made using the ear scans of the same subject, for this reason, the same pair of ear-EEG sensors were used for all phantom testing.

The phantom was first weighed without the side lids and with the AUX cables stacked on top. The antennas were connected to a signal generator, simulating a 10 Hz (alpha band) 100 mV square wave. The oscilloscope probe was jammed in the side of the phantom to measure the signal through the material alone and test for the material's integrity. Then, both earpieces were inserted, and electrode contact impedances were measured using an analog impedance meter (D175, Digitimer), capable of measurement values up to 50 kOhm. The impedance meter's reference electrode was connected to electrode Ex1. The oscilloscope probe was then connected directly to each electrode to visualize the simulated signal, which allowed to test if any electrode wasn't functioning or had a diminished signal amplitude compared to the one measured through the material. To simulate an ear-EEG acquisition, the earbuds were then connected to an OpenBCI Cyton amplifier (8 channels, 250 Hz sampling rate, 24 bits ADC) and connected per the recommended settings EEG acquisitions [35,36]. Ex1 was assigned to the bottom SRB2 pin (reference), and Ex2 was assigned to the bottom BIAS pin (noise-canceling), *a priori*. The remaining electrodes were connected in ascending order to the bottom N1P to N6P analog input pins (Ex3 is N1P, and Ex8 is N6P, on the board). Impedance values were also rechecked through the OpenBCI GUI. Then 20 s recordings were obtained through the amplifier on each ear, with a gain of 24: a recording with the simulated wave and an acquisition with no signal to assess the recorded noise floor at each electrode, both offline filtered between 0.3 Hz and 100 Hz.

The protocol was repeated after applying conductive paste (Ten20) to each ear electrode. The testing setup is illustrated in Figure 23.

Noise floor Root Mean Square (RMS) values were obtained with the *rms* function and the alpha simulation SNR (at 1% confidence) was calculated as per Equation 2.
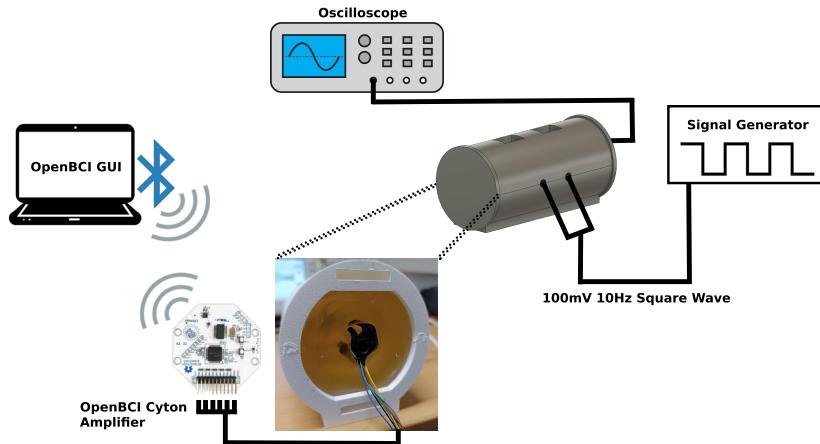
Figure 23. Schematic testing setup of the proposed ear-EEG phantom.

2.2.3. Phantom integrity and durability

When measuring the signal through the phantom, it is important to track the usability of this tool over time. This was achieved by monitoring the net weight of each phantom and measuring the input signal through the medium.

Table 2 shows that during a week-long testing period (by Day 4), the agar and BG phantoms lost no more than 10g in weight (with a higher percentage loss on agar). The reason for this weight loss can be attributed to the evaporation of water in both phantoms.

Table 2. Measured mass (g) of each ear-EEG phantom over the testing days, on agar and BG.

	Day 1	Day 2	Day 3	Day 4
Agar	855	851	850	845
BG	963	959	958	956

Table 3. shows the measurement of signal integrity passing through the phantom over a one week period. The agar phantom had the more significant amplitude loss, from the initial 100 mV, only achieving 44 mV on day one, declining to 36 mV by the last testing day. The BG phantom started at 80 mV and lost half its amplitude throughout testing due to the phantom material drying up. The phantoms suffered extensive pilot testing before the first day of quantified measurements and were kept in suboptimal refrigerating conditions.

Table 3. Measured signal amplitude (mV) at the sides of each ear-EEG phantom, directly through the material, as a measure of signal integrity, on agar and BG.

	Day 1	Day 2	Day 3	Day 4
Agar	44	40	40	36
BG	80	52	52	40

2.2.4. Electrode impedance

Figure 24 shows an example of electrode contact impedances (measured via the OpenBCI Cyton board and validated via the analog impedance meter) for wet and dry electrode conditions with the agar phantom's second day of recordings. Impedance measurements show a clear improvement across electrodes when applying conductive paste to the electrodes, reliably bringing the impedance under 10 kOhm, which is acceptable for EEG data collection (the same effect happened for BG measurements, with overall even lower values than agar). From this data we can see that the Ex3, EL7, and EL8 electrodes had the most difficulty achieving suitable impedance for dry electrode conditions, with impedance values over 50 kOhm, indicative of no or very poor contact on the electrodes. This effect was consistent across testing sessions (suggesting ER4 as a better channel than ER3 for dry recordings, at

least on this specific subject). However, applying conductive paste brings these electrodes' impedance down to usable equivalent levels.
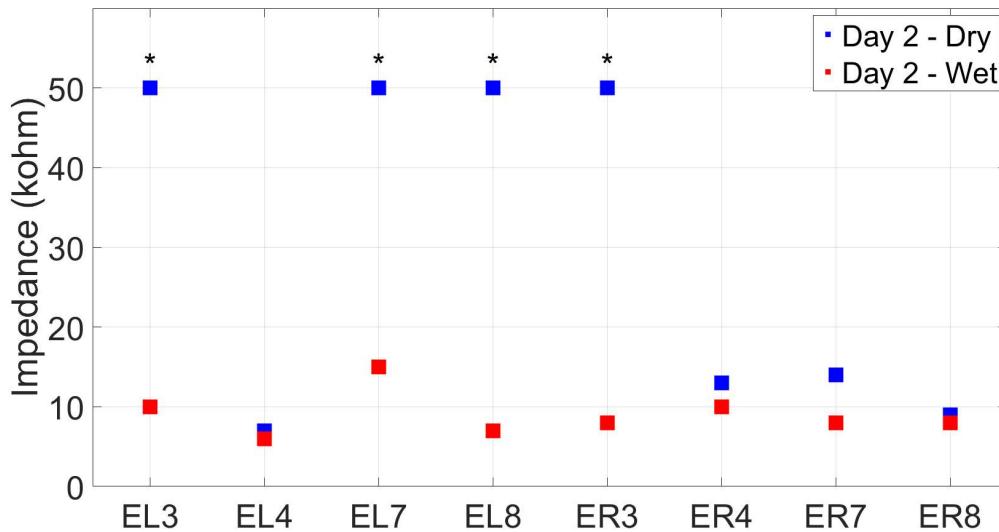

Figure 24. Contact impedance measurements (kOhm) measured on an agar ear-EEG phantom, in the second day of testing, for wet and dry electrode conditions. Values marked with an * surpassed the value of 50 in the graph, for the respective unit.

2.2.5. Noise floor measurements

Like impedance, noise floor measurements (without the input simulated signal) greatly benefit from the usage of paste at each electrode site, bringing the values across all channels close or under 1 µVrms across the board, as seen in Figure 25 for the second day of testing (first for noise measurements) on an agar phantom. These observations indicate that the electrode material and shape are adequate for wet electrode recordings. Comparing the left ear to the right ear results in the same table on a dry setting, we can predict that the left ear is more prone to have noisier channels (with RMS values in the order of hundreds of µV), when compared to the right side electrodes, suggesting a better performance when assessing the right earpiece on this subject.
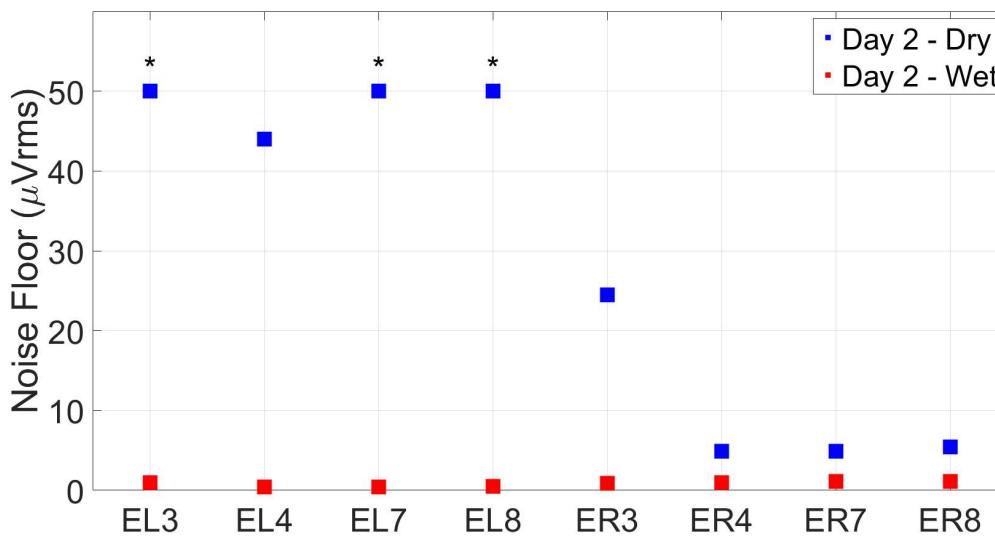
Figure 25. Noise floor measurements (µVrms) measured on an agar ear-EEG phantom, in the second day of testing, for wet and dry electrode conditions. Values marked with an * surpassed the value of 50 in the graph, for the respective unit.

2.2.6. Alpha wave simulation

When simulating a known signal into the phantom through the two antennas, the input frequency (10 Hz - within the alpha band) was successfully recorded at all ear electrode sites for the first day of testing in agar and BG. As expected, the modulation SNR value was higher when measured in the ear canal electrodes like ER8 (further away from the within-ear reference at ER1) than, for example, ER3 or ER4, which are closer to the reference.

In Figure 26, the resulting alpha modulation spectra at ER8 are present, with clear peaks at the 10 Hz and 30 Hz odd harmonic frequencies that compose the square wave. The noise floor aspect of the spectra resembles what would be expected from an EEG recording, although that was not always the case in the following days of testing when the agar or BG phantom noise floor readings started to increase. Nonetheless, the phantom antenna signal delivery and materials proved efficient in simulating a given signal, and the present methods regarding this aspect were positively validated, with our prototype and employed methods show feasibility for this type of testing.
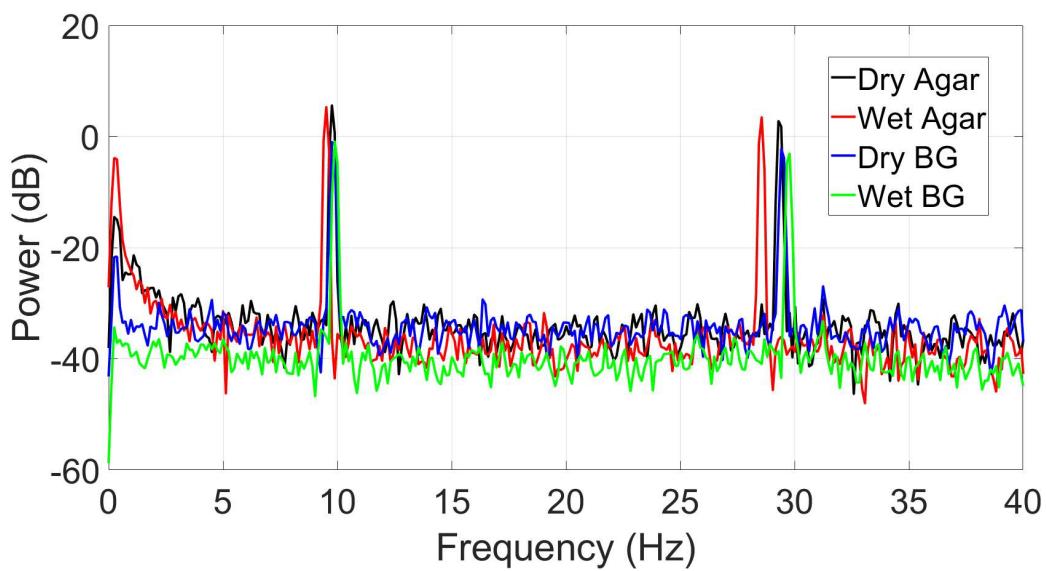
Figure 26. Ear-EEG alpha wave (10 Hz input) simulation frequency response (dB), for agar and BG, in dry and wet electrode settings, at the ER8 electrode.

3. Results: Toolkit Use Case: Validation of an ear-EEG sensor

3.1. Tested ear-EEG devices and setup

The tested ear-EEG devices utilized were fabricated by a third-party company using a process based on the unique *.stl* of each individual subject, resulting in personalized ear-EEG buds for every participant (5 subjects total, age avg. 36.2 ± 10.6 std.). Each electrode on the devices is associated with a color-coded wire, and also numbered for easier referencing, according to Figure 27.
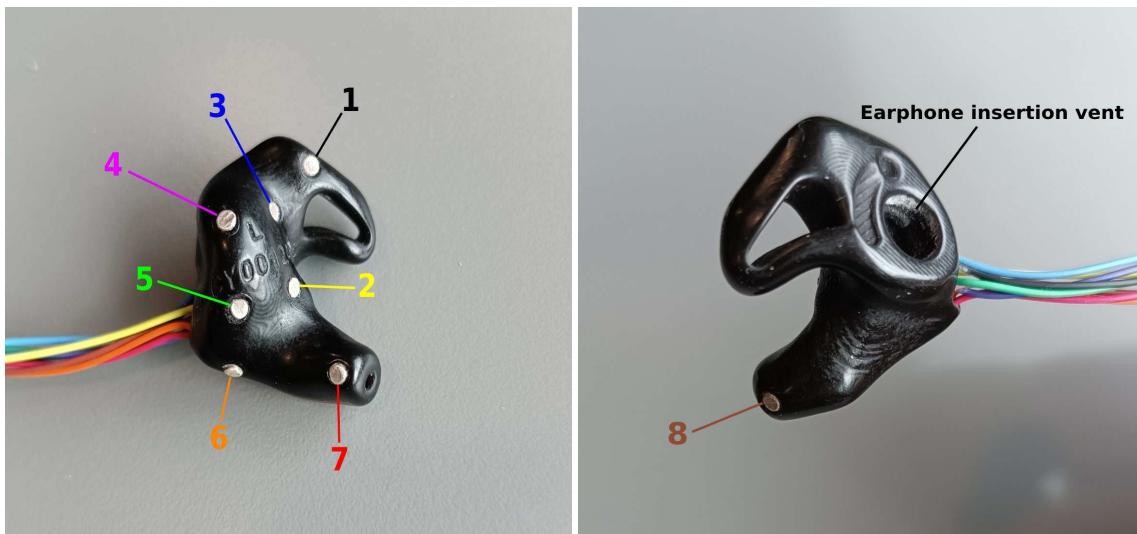
Figure 27. (Left) Internal side of the tested earbuds. **(Right)** External side of the tested earbuds.

The earbuds have eight 2mm Ag/AgCl electrodes on the surface. One is located in the concha cymba area (1), four around the concha cavum (2-5), one near the antitragus (6), and the last two inside the ear canal (7 and 8), facing the posterior and anterior directions, respectively. The electrodes have wires welded to them that exit near electrode 6. On the external side, the earbud boasts an earphone insertion vent for sound delivery.

The ear-EEG data were acquired under wet and dry ear electrode conditions, simultaneously with conventional EEG from the scalp, using the same amplifier system as for the validation results, to provide a direct reference with an EEG gold standard. However, since this amplifier only supports 24 data channels outside the CMS and DRL, four scalp channels (F3, F4, P3, and P4) from the control group were disconnected to make room for all ear-EEG electrodes on the amplifier connector. In Figure 28, the ear-EEG recording layout is illustrated, with the full referencing scheme used in this work.
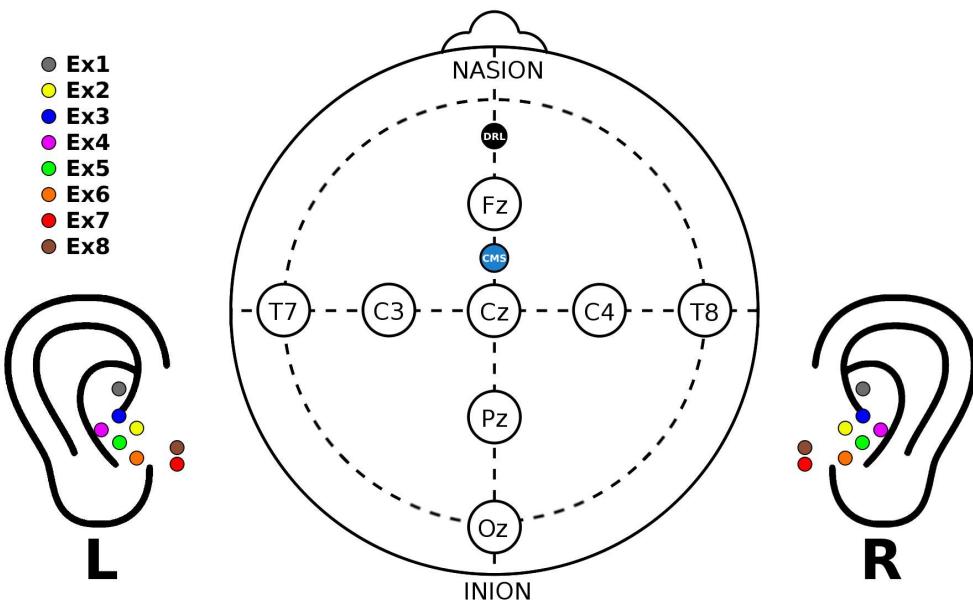
Figure 28. Ear-EEG scalp and ears acquisition configuration showing the approximate position of ear electrodes, with colors and numbers matching those in Figure 27 - referencing for ear electrodes is provided in the legend, for example, "Ex1" is used to label ear electrode 1 (black), where "x" is replaced by **L** or **R** for left or right ear, respectively.

The earphones were strapped to the participant's back and shoulders. The ear-EEG wiring was taped to the cap with skin-safe tape and connected to the amplifier, as seen in Figure 29.
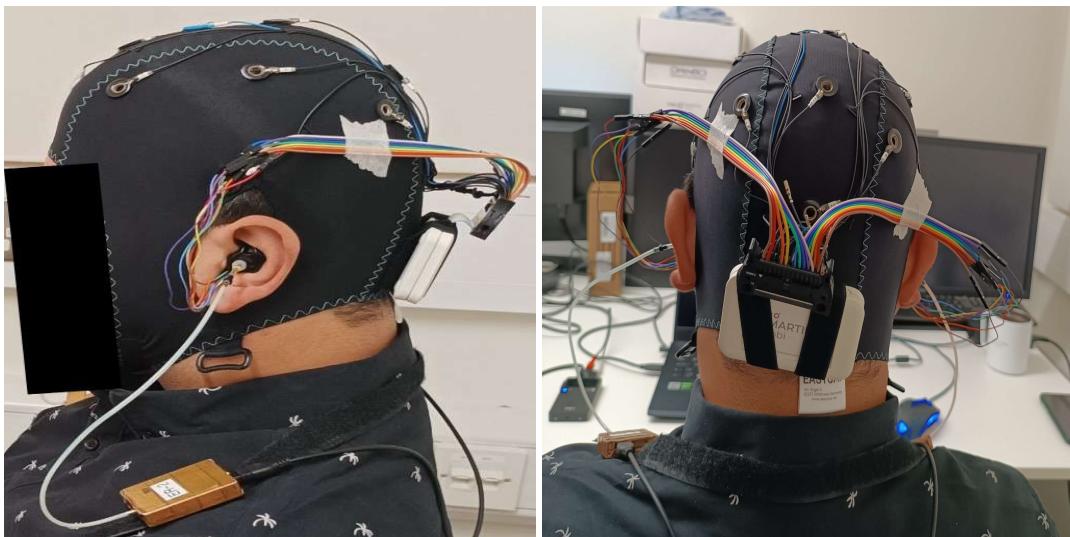
Figure 29. (Left) Ear-EEG head setup - side view. **(Right)** Ear-EEG head setup - posterior view.

A small amount of conductive electrode paste (Ten20) was used for wet electrode ear recordings. The subjects inserted the earpieces themselves at their comfort. Ear electrode impedance was generally too high or immeasurable by the Streamer, so it won't be assessed. The same test battery was utilized.

For the ear-EEG recordings, the lowest volume setting from the validation group was adopted for all subjects.

The same processing and statistical analysis was applied, with the ear-EEG recordings being referenced to the following, for different assessment of ear-EEG configurations:

- **Cz**, as a standard scalp reference
- **T8**, as a scalp but closer to the ear reference
- **ER3**, as a within-ear (assessing ER8, for example) and between-ears reference (assessing EL8, for instance)

3.2. EaR-P Lab Applied to Ear-EEG Study

This section presents the result obtained per paradigm of EaR-P Lab for wet ear electrode conditions, showing this tool being applied in the assessment of ear-EEG devices.

3.2.1. Alpha block

Figure 30 shows the grand average alpha block power for different ear-EEG references. The modulation was of 3.7 dB for the scalp Cz reference, about 3 dB lower than the control group value at Oz. Referencing to T8, closer to the ear, the modulation decreases to 3.1 dB and 2 dB for the within-ear reference at ER3, with a lower value for the between-ears assessment.
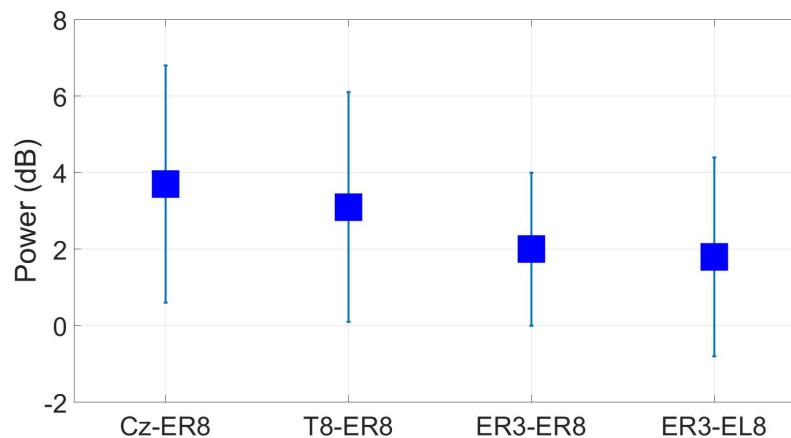
Figure 30. Grande average alpha block modulation (dB) for the wet ear-EEG recordings for the different referencing setups. Omitted results are not significant based on a t-test ($p < 0.05$).

3.2.2. ASSR

The ASSR SNR at ER8 (Figure 31) for the Cz reference has a value of 7.9 dB, a decrease of 1.5 dB over the best response in the control group. Comparing the T8 and ER3 references, we can see the SNR is 1.2 dB higher for the in-ear ER3 reference over the temporal site.
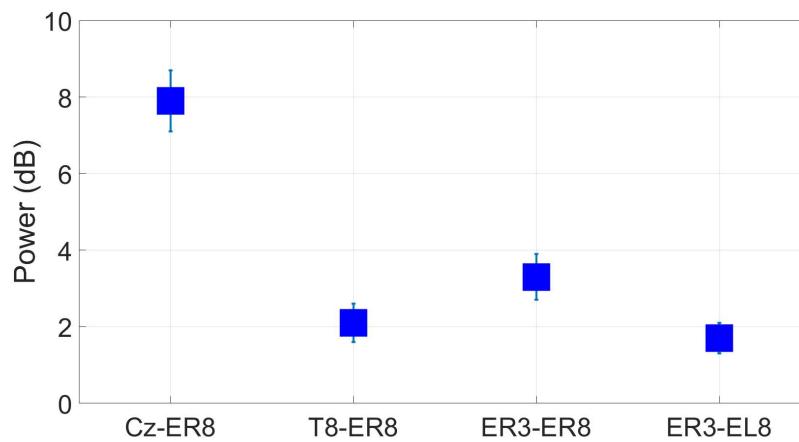
Figure 31. Grand average ASSR (dB) to a 40 Hz frequency modulated auditory stimuli for the wet ear-EEG recordings for the different referencing setups. Omitted results are not significant based on an f-test ($p < 0.05$).

3.2.3. SSVEP

Comparing the results from Figure 32, all measured SNR results were lower when compared to the control group SNR values of 11 dB and 7.5 dB at Oz and T8, respectively. When looking at scalp references (Cz and T8), the SNR at ER8 holds up at 6.4 dB and 5.9 dB. However, the within-ear and between-ears referencing was systematically shown as not significant, even for other electrodes.
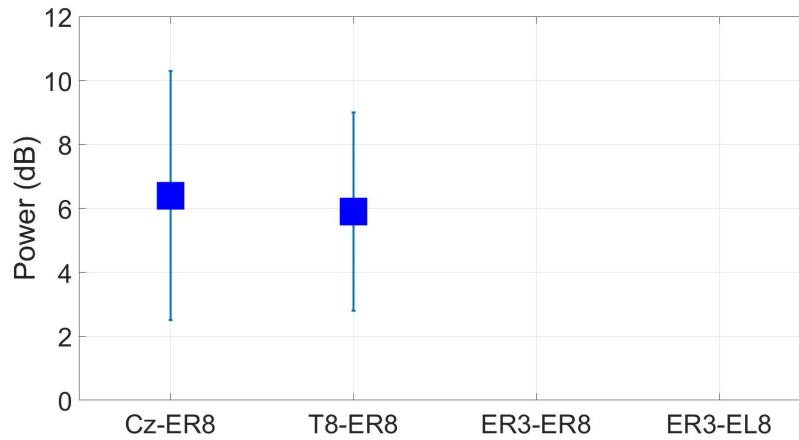
Figure 32. Grand average SSVEP (dB) to a 10 Hz visual stimuli for the wet ear-EEG recordings for the different referencing setups. Omitted results are not significant based on an f-test ($p < 0.05$).

3.2.4. AEP (N100)

When referenced to Cz, the ER8 electrode acquired the N100 component (albeit inverted) at an amplitude of around 6 µV for wet electrode ear-EEG recordings. Referenced to T8, the significant component was diminished, with a peak amplitude of around 2 µV, as per Figure 33.
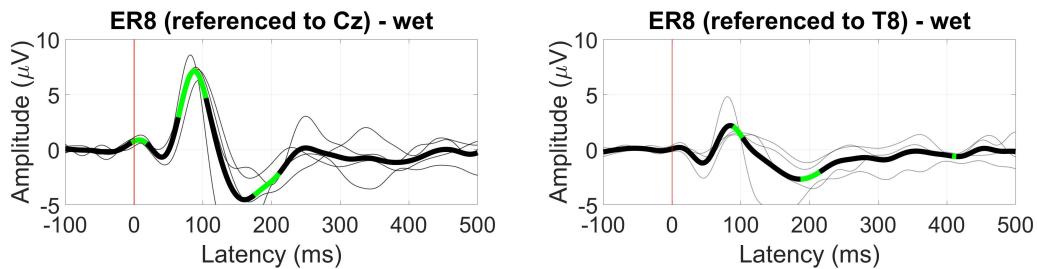
Figure 33. Grand average wet ear-EEG AEP waveform at ER8 for a **(Left)** Cz and a **(Right)** T8 reference. Statistically significant segments are highlighted in green, based on a t-test (p < 0.05).

For the ear referenced assessments (Figure 34), only the between-ears configuration measuring at EL8 showed a significant N100 component with a 1 µV amplitude.
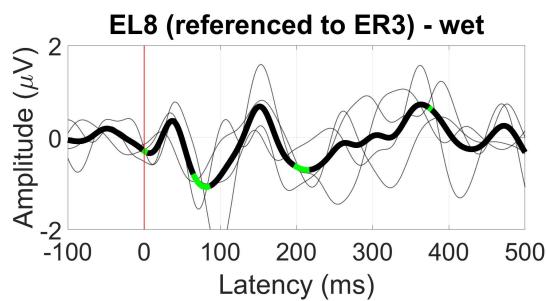

Figure 34. Grand average wet ear-EEG AEP waveform at EL8 for an ER3 reference. Statistically significant segments are highlighted in green, based on a t-test (p < 0.05).

3.2.5. VEP

While the VEP waveform seen at Oz for the control group is not maintained near the ear, the resulting waveform is maintained from the Cz to the T8 referencing for ear-EEG, with a 2 µV drop in amplitude for the latter, as seen in Figure 35.
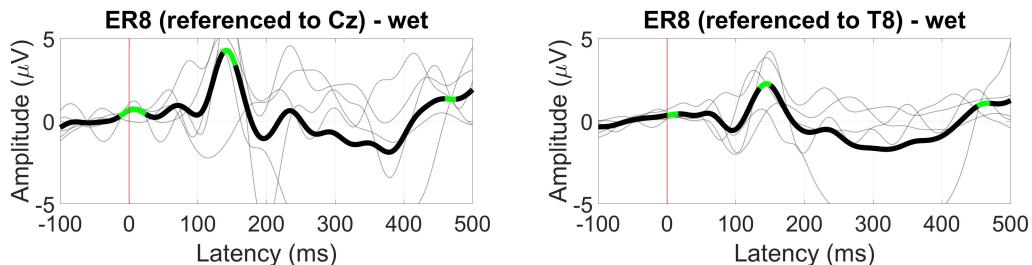
Figure 35. Grand average wet ear-EEG VEP waveform at ER8 for a **(Left)** Cz and a **(Right)** T8 reference. Statistically significant segments are highlighted in green, based on a t-test (p < 0.05).

For the ear ER3 reference, only the in-between ears configuration presented statistical significance in the waveform with a later 200ms component being deemed significant at 1µV.
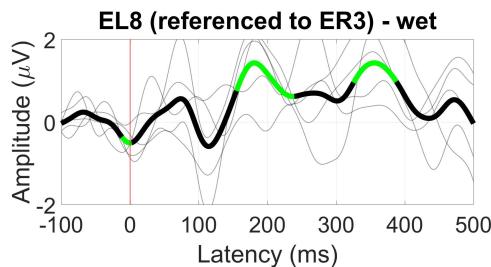
Figure 36. Grand average wet ear-EEG VEP waveform at EL8 for an ER3 reference. Statistically significant segments are highlighted in green, based on a t-test (p < 0.05).

3.2.6. AEP Oddball (MMN)
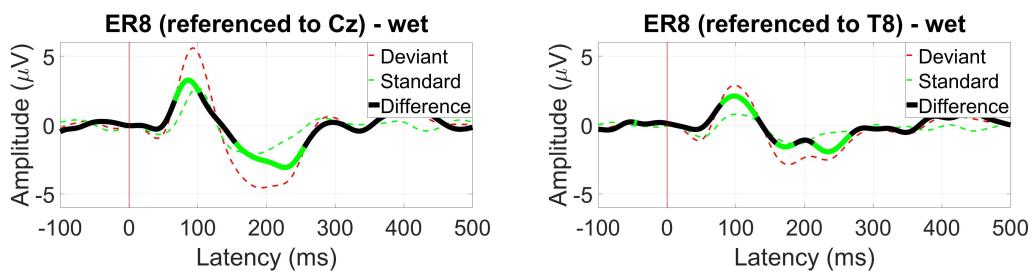
Figure 37. Grand average wet ear-EEG MMN waveform at ER8 for a **(Left)** Cz and a **(Right)** T8 reference. Statistically significant segments are highlighted in green, based on a t-test (p < 0.05).

For within-ear and between-ears referencing, no significant components resembling the MMN response were found.

3.2.7. VEP Oddball (P300)

Measured from the scalp to the ear, the greatest observed difference between the oddball and standard stimuli shows a negative deflection around the 350 ms latency, for Cz (5 µV to 6 µV amplitude) and T8 (1.5 to 2 µV amplitude) referencing as per Figure 38, suggesting a later onset for this component than on the scalp.
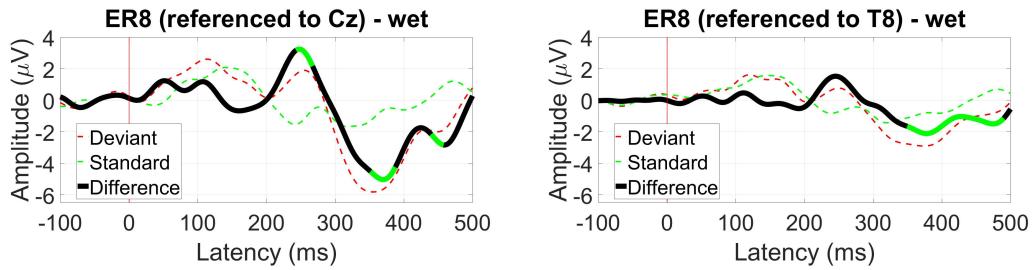
Figure 38. Grand average wet ear-EEG P300 waveform at ER8 for a **(Left)** Cz and a **(Right)** T8 reference. Statistically significant segments are highlighted in green, based on a t-test (p < 0.05).

Ear-EEG referenced P300 waveforms didn't show significant components at expected latencies.

3.2.8. EOG (Blinks and saccades)

Concerning EOG blink ratios, from Figure 39, the ER3 reference was about 1.5 times higher than scalp references near the ear, with diminished results for the between-ears configuration at EL8. The best ratio altogether was obtained for the within-ear ear-EEG reference, at a ratio of 3.
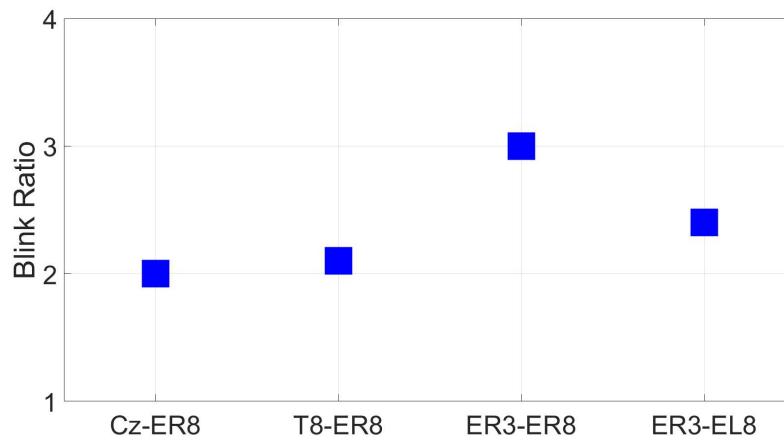
Figure 39. Grand average ear-EEG soft/hard blink ratio, for the different ear-EEG reference configurations.

Measuring at the ER8, saccade profiles were maintained for the different directions, with greater amplitudes for the Cz reference but better differentiation when referencing to T8, with a 10µV difference between up and down saccades, per Figure 40.
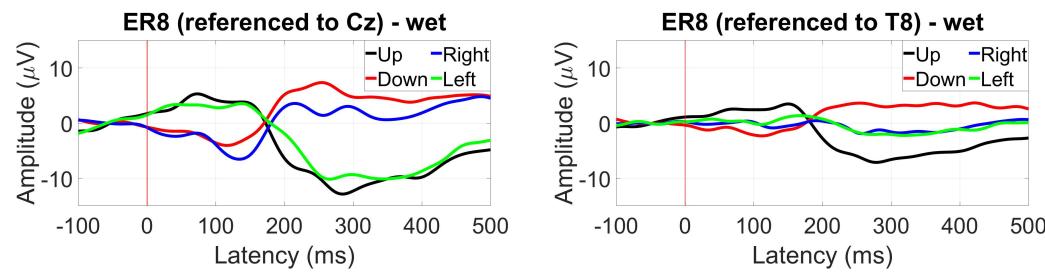
**Figure 40.** Grand average wet ear-EEG saccade profiles on the four cardinal directions at ER8 for a **(Left)** Cz and a **(Right)** T8 reference.

For within referencing, the vertical saccade plane presents a better differentiation of 3 to 4µV, while recording across the ears shows the best results in amplitude and difference between right and left saccades (the vertical saccades present the same profile), as seen in Figure 41.
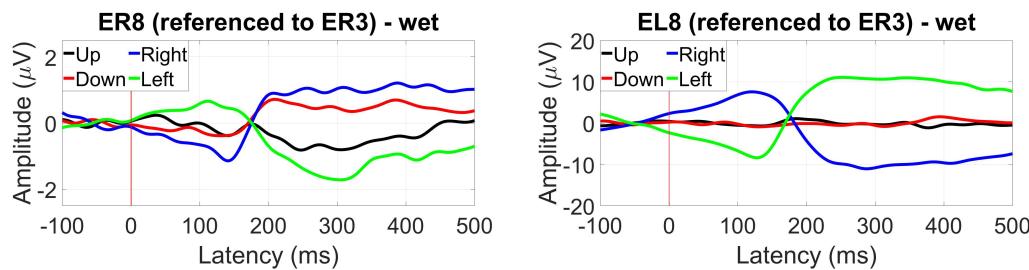

Figure 41. Grand average wet ear-EEG saccade profiles on the four cardinal directions at ER8 for a **(Left)** within-ear and a **(Right)** between-ears reference.

3.3. Reassesment of dry ear-EEG ASSR data

From the ear-EEG phantom results, it is suggested that the ER4 electrode has better connectivity to the ear than ER3, which was previously selected as a reference, for the live ear-EEG recordings. Figure 42 shows a direct comparison of the ASSR dry data (most viable ear-based paradigm) reanalyzed with ER4 as a reference for this single subject that was used for molding the phantom. The suggested electrode ER4 from the phantom improved the SNR at ER8 (and the other electrodes) by about 2.3 dB,

exemplifying how the phantom can predict better usability of ear devices without the need to conduct testing on a live subject.
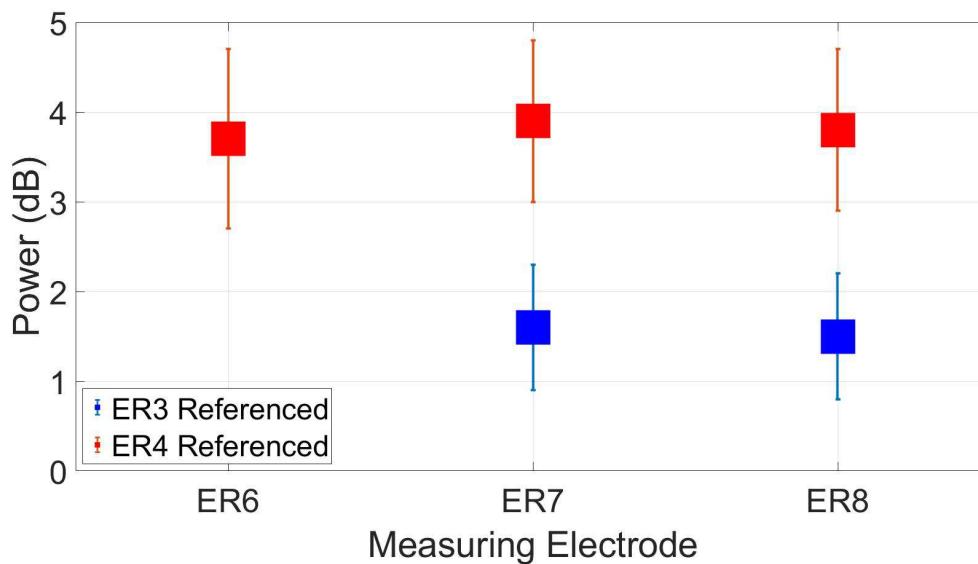

Figure 42. Reassessment of dry electrodes ear-EEG ASSR data for the subject utilised to make the ear-EEG phantom. Data was rereferenced from ER3 (original reference) to ER4 (better electrode proposed by the phantom) resulting in a dB increase for this response.

4. Discussion

Ear-level EEG devices are an active field of research and development with potential to disrupt the way in which brain activity can be factored into daily life. We can expect the increase of new applications, form factors and sensor technology being developed as the field grows. Currently, there is scientific evidence to support the use of ear-EEG [21,22] and we set out to provide a set of tools that could benefit further testing and characterization of ear-EEG devices. Our approach was to develop a validation toolkit that would allow the characterization of ear-EEG devices from hardware to neural signal acquisition.

We developed an EEG acquisition application based on nine commonly utilized EEG paradigms in the field. Our validation results show that this implementation of EaR-P Lab can indeed be used to elicit the expected neural responses from scalp EEG recordings. When evaluating these responses from the perspective a near-ear electrode location like T8, we observed that the VEP responses were the most affected, indicating that ear-EEG may have a limitation on the type of measurable paradigms.

We also developed an EEG phantom suitable for evaluating ear-EEG sensors by combining 3D scans of a participant's ear impressions and 3D printing a mould to be filled with conductive material. This allowed us to assess ear-EEG devices in terms of their electrode contact impedance, measured noise floor, and acquisition of a known signal. To the best of our knowledge this is the first EEG Phantom dedicated for ear-level devices, as traditional head phantoms neglect the structures of the ear needed for this evaluation [8].

Our phantom development highlighted salt-doped agar as the best material substrate for this purpose, with conductivity values equivalent to those reported in literature for whole brain anatomy at 0.33 S/m [37,38]. While salt-doped ballistic gelatin and carbon fibre doped silicone were also investigated as potential material substrates, there were notable limitations with these options.

Ballistic gelatin showed a drastic change on signal transmission integrity across a week of testing. This could be related to the storage conditions, highlighting the susceptibility of the phantom to the effect of time and the environment, which could hinder the repeatability of results when evaluating ear sensors. While agar is also an organic and perishable material, it showed a more stable material

performance despite being kept under the same storage conditions as the BG phantom. To address the susceptibility to degradation over time, we investigated the use a synthetic non-perishable material composition for the phantom by doping platinum cured silicone with carbon fibres. This approach, however, did not scale from the prepared samples for conductivity testing to the larger ear-EEG phantom.

The mechanism by which carbon fibres turn silicone into a conductive medium is different to that of salt-doped agar or BG. In the latter, salt fully dissolves in the medium creating a homogeneous conductive substrate, while in the former, carbon fibres disperse in the silicone creating conductive paths through heterogeneously mixed fibres. When the samples were small strips, this dispersion of carbon fibres was effective throughout the volume, however when a larger volume was used, the stirring method did not achieve dispersion of fibres over the full volume. This led to patches of non-conductive silicone, clearly visible in the final product. A different mixing approach should be defined in the future to ensure that the CF phantom is conductive throughout. It should be noted however that the conductivity of this synthetic substrate is orders of magnitude larger than that of real anatomical structure and the agar/BG substrates.

We piloted the use of these tools through the evaluation of a third-party ear-EEG sensors in a small feasibility study. Our results showed that the ear-EEG sensors were functional and EEG responses were successfully recorded when using a scalp reference (Cz and T8). However, in-ear or cross-ear references only showed significant EEG responses for steady state paradigms and alpha blocking. Further investigation should be done on the reference choice for these ERP paradigms, as it has been shown that this is crucial for obtaining characteristic ERP responses [39] . However, this was not the scope of this feasibility study. Importantly, we could see clear EOG responses within-ear and cross ear reference configuration, promising for further ear-based BCI applications.

We have also shown how the data obtained with the ear-EEG phantom, without the need of a testing subject, shows feasibility on selecting optimal electrode locations in the ear and can be tied with the actual ear-EEG recordings of the same tested subject to improve the SNR of the ASSR.

Next steps with this tool include a better executed testing protocol and the assessment of a generic ear canal for the testing and comparison of different earpiece designs. Our ear-EEG phantom methodology should also be considered in the future for the controlled study of how real-world factors, like gait, affect ear-EEG [40].

The presented ear-EEG validation Toolkit is available to the scientific community, via the Open Science Framework (OSF) repository named: "Brain Wearables: Validation toolkit for Ear Level EEG sensors" (https://osf.io/2dxs4/). All future modifications to the Ear-P Lab and the ear-EEG phantom generated by the authors will also be updated in the mentioned repository.

Author Contributions: Conceptualization, A.L.V. and M.J.C.; methodology, G.C. and A.L.V.; design, G.C..; validation, G.C.; formal analysis, G.C.; investigation, G.C. and A.L.V; writing—original draft preparation, G.C.; writing—review and editing, G.C., A.L.V. and M.J.C; visualization, G.C., A.L.V. and M.J.C; supervision, A.L.V. All authors have read and agreed to the published version of the manuscript

Institutional Review Board Statement: Ethical review and approval were waived for this study as the data presented were acquired as a proof of function and exploratory basis for the development of the toolkit and not for the purpose of understanding human normal or abnormal function.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The presented ear-EEG validation Toolkit and the validation data is available to the scientific community, via the Open Science Framework (OSF) repository named: "Brain Wearables: Validation toolkit for Ear Level EEG sensors" (https://osf.io/2dxs4/). All future modifications to the Ear-P Lab and the ear-EEG phantom generated by the authors will also be updated in the mentioned repository.

Conflicts of Interest: During the development of this project and at the time of submission of this manuscript M.J.C. is has been employed at Segotia, the third-party company who provided the ear-EEG sensors for this study.

Abbreviations

The following abbreviations are used in this manuscript:

EEG   Electroencephalography
BCI   Brain-Computer Interface
ERP   Event-Related Potential
ASSR   Auditory Steady-State Response
SSVEP   Steady-State Visual Evoked Potential
AEP   Auditory Evoked Potential
VEP   Visual Evoked Potential
MMN   Mismatch Negativity
EOG   Electro-Oculography
ISI   Interstimulus Interval
BG   Ballistic Gelatin
CF   Carbon Fibers

Appendix A. EaR-P Lab Paradigms and Settings

This appendix section outlines the testing sequences initialized by each button in the Main Menu of EaR-P Lab.

General Recording

This test paradigm displays a white cross centered on the screen (Figure A1), for the selected duration of time. This option was implemented to record resting state EEG or any custom test not readily provided by default in EaR-P Lab.
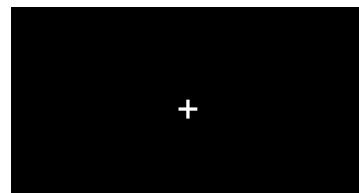
Figure A1. Focus cross utilized by the "General Recording", "ASSR" and "Alpha Block" paradigms

Auditory Steady-State Response (ASSR)

This test paradigm displays the same focus as Figure A1, while playing an amplitude modulated sound file for an ASSR test. The duration of the experiment is the same as the duration of the sound file.

Steady-State Visual Evoked Potential (SSVEP)

This paradigm consists of a flickering target like the one presented in in Figure A2. The target consists of a radial black and white pattern centered on a black background, with a diameter that is equal to the screen height. The flicker frequency and the duration of the test can be adjusted in the "Settings" menu.
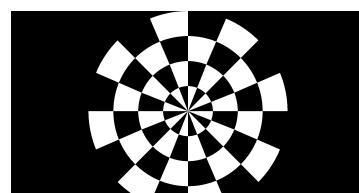

Figure A2. Target used for the "SSVEP" experiment in **EaR-P Lab**

Alpha Block

This paradigm starts with onscreen instructions for the subject to press "SPACE" and close the eyes. After a specified amount of time, there is a transition phase where an artificially generated voice will warn the subject to open its eyes and look at the cross for the same time as before. These instructions repeat one more time so in total there are two phases with eyes closed and two with eyes open, that are alternated between them. At the beginning of each phase, the specific markers seen in Figure A3 are also sent through LSL.
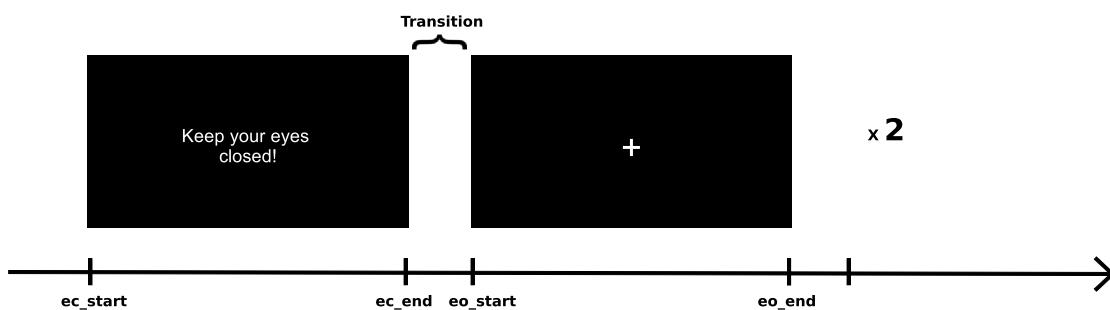
Figure A3. "Alpha Block" paradigm functioning, with the markers that are sent at the start and end of each phase

Auditory Evoked Potential (AEP)

In this paradigm a sequence of discreet auditory stimuli is played. The stimuli can be self generated tones of specific frequency and duration via PsychoPy or custom made sound files added to the EaR-P Lab source folder. The stimuli are separated by an adjustable ISI, and each time the stimuli start the "aep"marker is sent to LabRecorder. In order to keep the subject engaged (to help avoid drowsiness while the test is running), the *Inscapes* animation is also played in the background, as shown in Figure A4. This animation was originally crafted to maintain children "engaged while minimizing the certain aspects of cognitive processing" [41]. The original sound was removed from the animation, and it will loop back to the beginning until the test is finished.
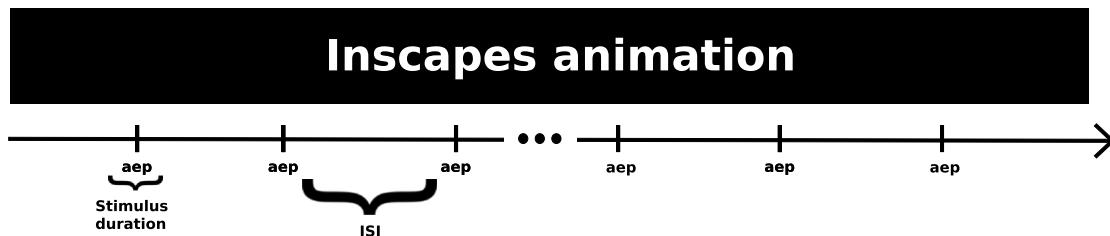

Figure A4. "AEP" sequence featured in **EaR-P Lab**

Visual Evoked Potential (VEP)

This paradigm is similar to its auditory counterpart, with the default option of a shape being displayed on screen for a specified amount of time and ISI. Different versions of VEP are possible to acquire through this paradigm. Flash VEP can be elicited where the screen is black and flashes white for a specified duration. Two patterned stimuli, a checkerboard and a dartboard (as shown in Figure A5) can be used for an onset-offset VEP or pattern reversal VEP (if the chosen ISI is set to zero). A "vep" marker is sent every time the on screen stimuli changes to a new one.

32 of 38
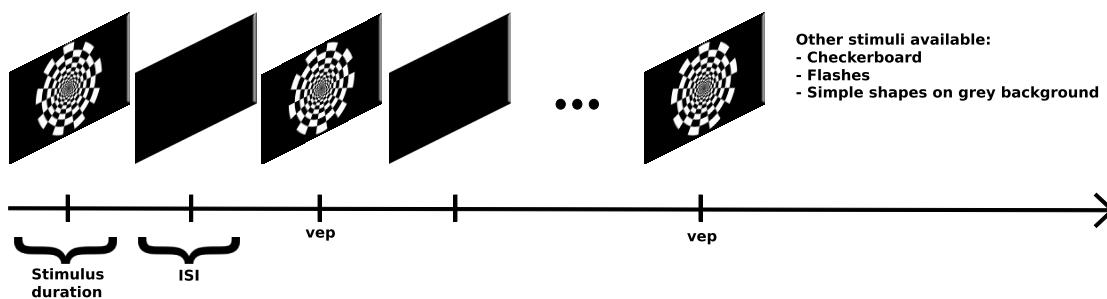
Figure A5. "VEP" button paradigm structure and options

Oddball Paradigms

These paradigms are identical for visual and auditory modalities. The stimulus sequence starts with 20 standard stimuli followed by a randomly generated sequence of standard and target events, as depicted in Figure A6. The oddball paradigm has a fixed, non-modifiable, 80-20 configuration (i.e., the target stimuli appears at random with a probability of 20%). For the auditory version of the oddballs, similar to the regular "AEP" sequence, the *Inscapes* movie is played in the background. Additionally, two rules are also implemented during these paradigms, these being that there can't be a succession of two targets in a row (i.e., there is always at least one standard stimuli before and after each oddball) and that in the case there is a part of the sequence made up of eight standard events in a row (excluding the first twenty events), the next event is compulsorily a target stimulus. Markers for the "standard" and "target" as well as a marker for a "response" in case of an active task (e.g., P300) are sent accordingly.
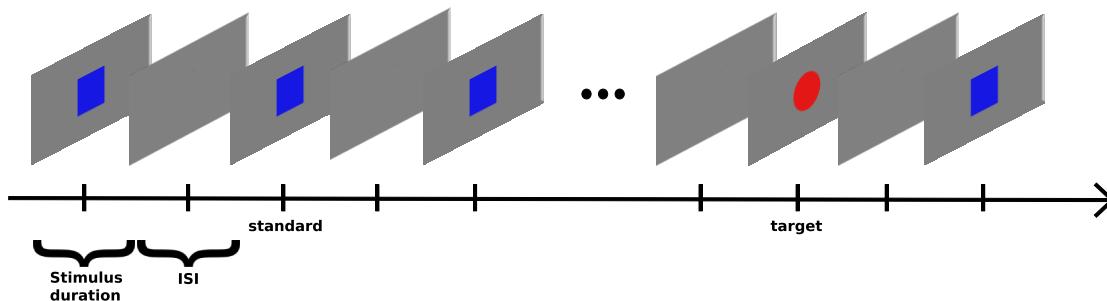
Figure A6. "Oddball" type paradigms sequence - example of visual oddballs

Electro-Oculography (EOG)

This paradigm is split into two parts: eye blinks and saccades. During the eye blinks phase, the subject is prompted to do a total of four eye blinks, alternated between a normal or "soft" blink and an intentional or "hard" blink. For each blink, the subject has a three second window that starts once "SPACE" is pressed, according to the instructions. As shown in Figure A3, the "soft_blink_start", "soft_blink_end", "hard_blink_start" and "hard_blink_end" static markers are used to limit each blink window.

The next part of this paradigm is a "follow-the-dot" test, shown in Figure A7. Here, a red dot starting in the center of the screen disappears and reappears at one of the four main cardinal directions for the subject to follow to test for eye saccade amplitudes and profiles. The dot jumps a distance equal to half the screen height for each saccade, is held in that position for a specified duration, and returns to the middle during the ISI period. The appropriate marker stating "top", "right", "left" or "bottom" is sent when the dot moves to the respective location. The static markers "follow_start" and "follow_end" also limit this part of the coded "EOG" paradigm.
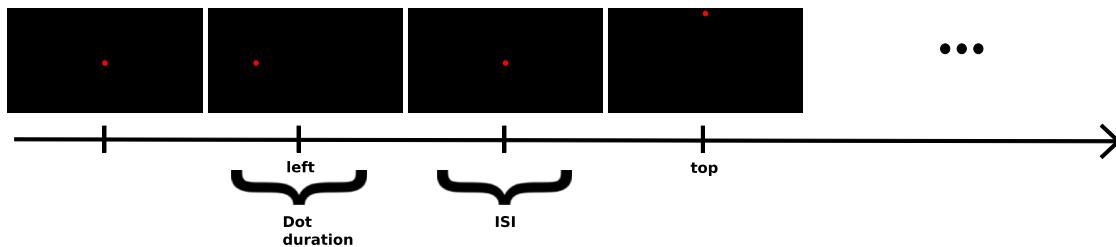
Figure A7. "Follow-the-dot" phase of the "EOG" paradigm in **EaR-P Lab**

Settings Menu

The "Settings" button opens up a secondary window (Figure A8) with different boxes and options that control what happens and is shown in each paradigm mentioned above. These settings range from the type of stimuli, frequency of the SSVEP, testing durations, total trial number for the AEP/VEP paradigms, how many oddballs until the oddball type tasks are complete, and how many repetitions for each saccade in the EOG test. Additionally, there is some information about how to proceed to add more stimuli to EaR-P Lab, a disclaimer about the ISI for auditory paradigms, and a button to automatically measure the angle of the saccades in the EOG task dependent on the dimensions of the stimuli presentation monitor used and the subject's distance to said screen.
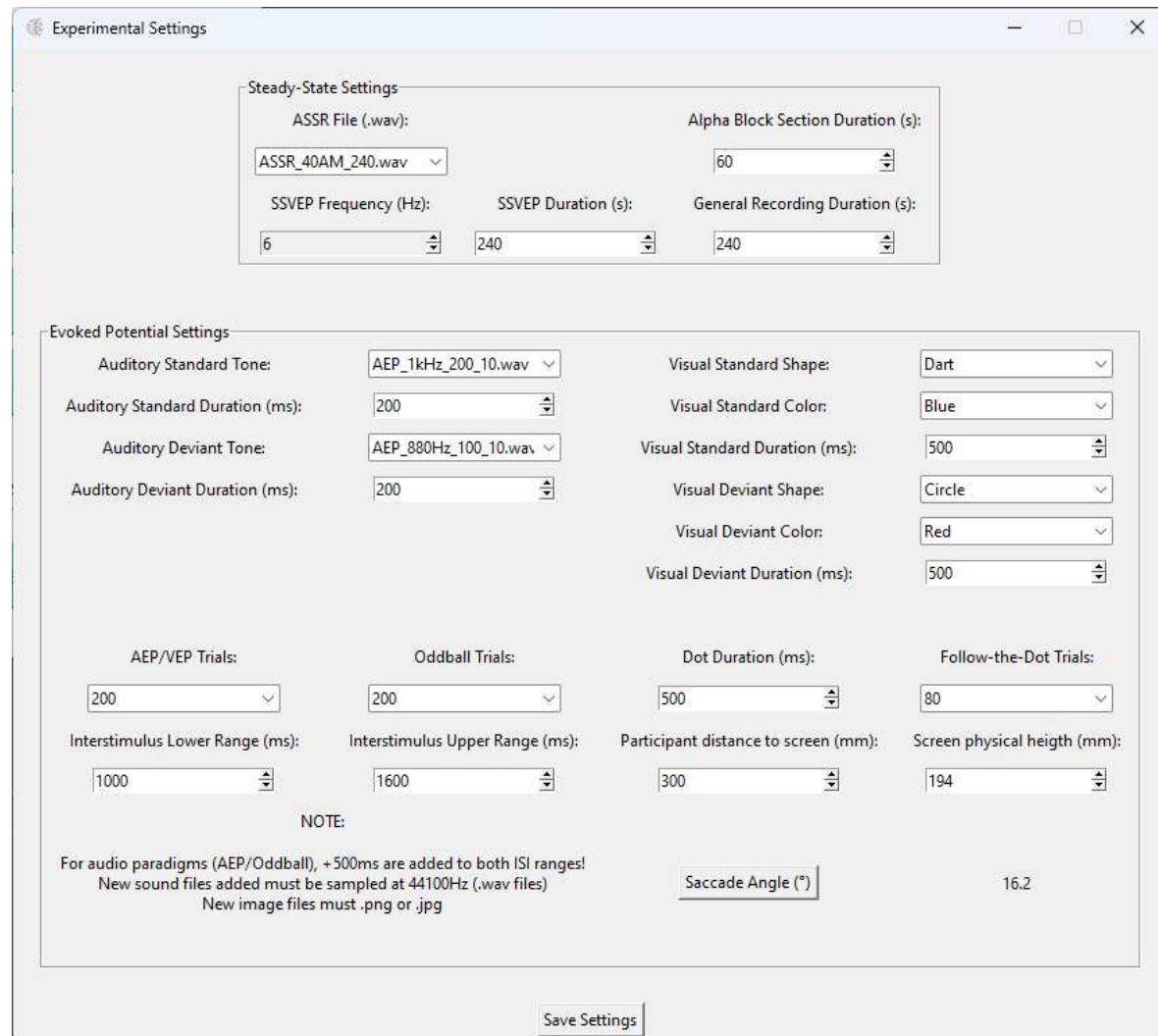
Figure A8. EaR-P Lab - Settings Menu

Markers Menu

Like the "Settings" menu, here a secondary window (Figure A9) is presented with text boxes about each paradigm so that the user can set its delimiting markers for each experiment. This feature is useful, for instance, to record the same type of paradigm within the same *.xdf* file and to help with accurate data parsing. At the bottom of the window, there is information about the static (meaning non-changeable in the GUI) markers used by EaR-P Lab, as shown in the above paragraphs.
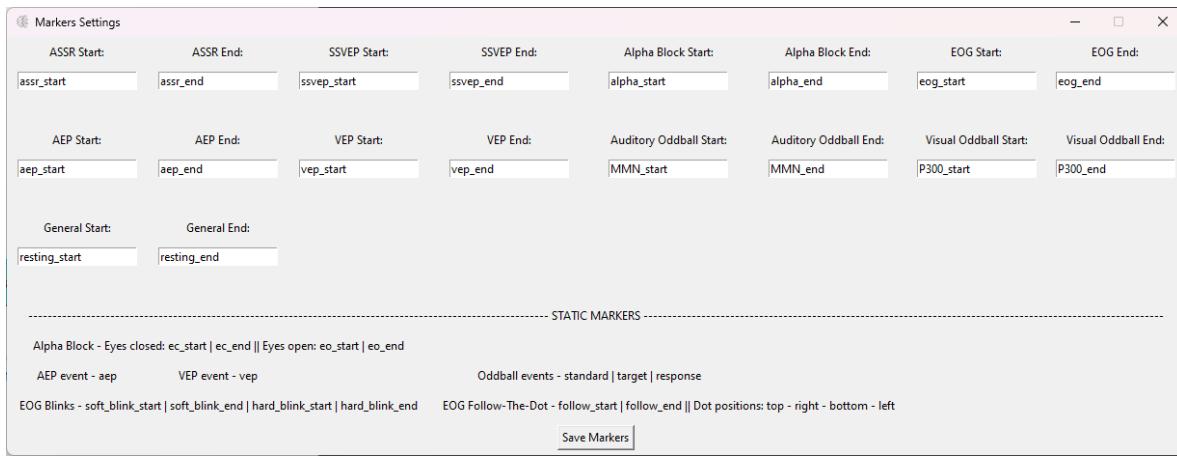
Figure A9. EaR-P Lab - Markers Menu

Appendix B. Ear-EEG Phantom Dimensions

This appendix section will illustrated the detailed dimensions and design drawings of the Ear-EEG Phantom.
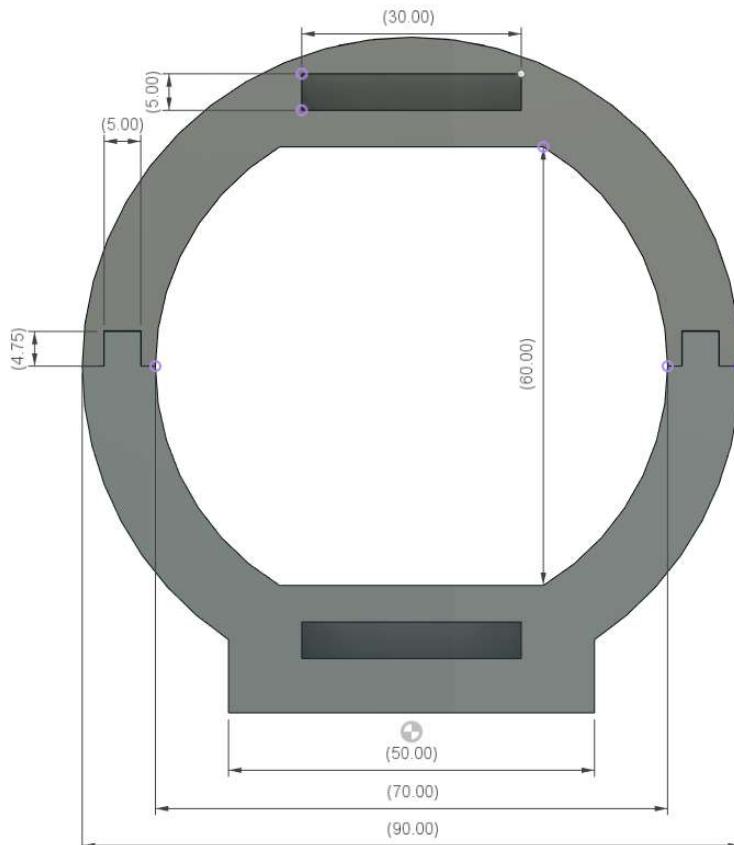

Figure A10. Proposed ear-EEG phantom prototype dimensions - Side View
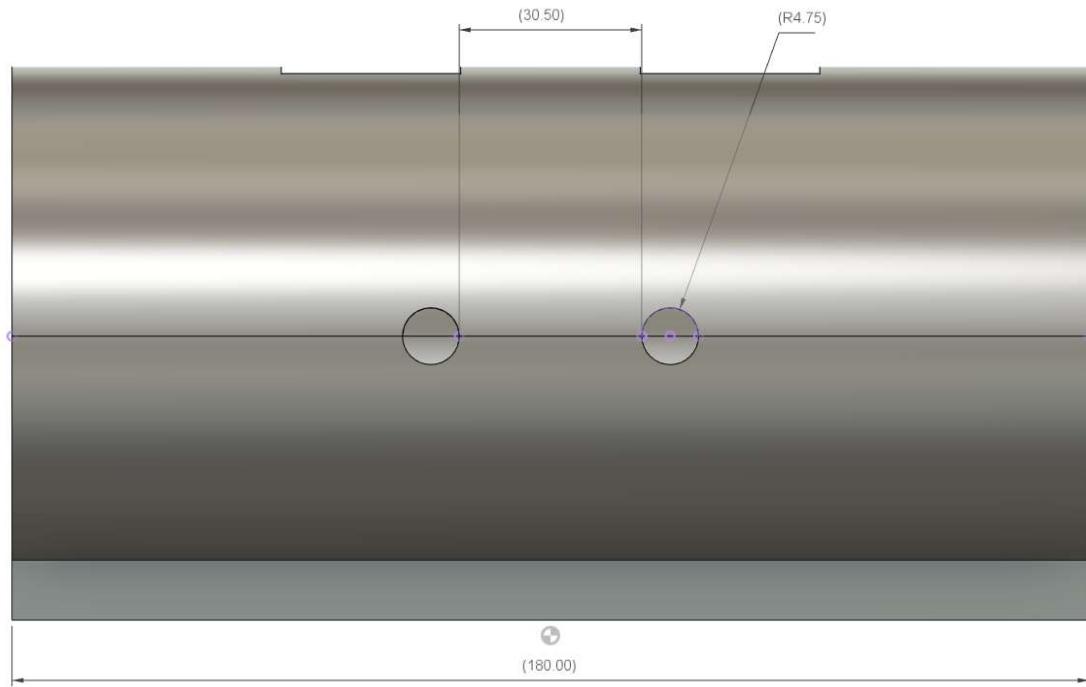
**Figure A11.** Proposed ear-EEG phantom prototype dimensions - Front View
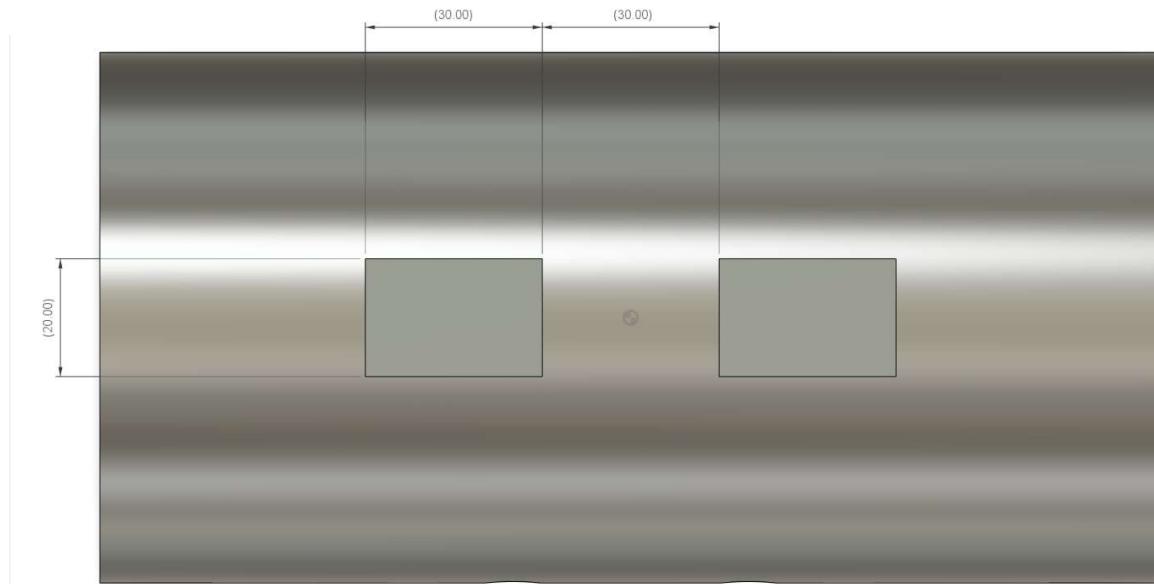
**Figure A12.** Proposed ear-EEG phantom prototype dimensions - Top View
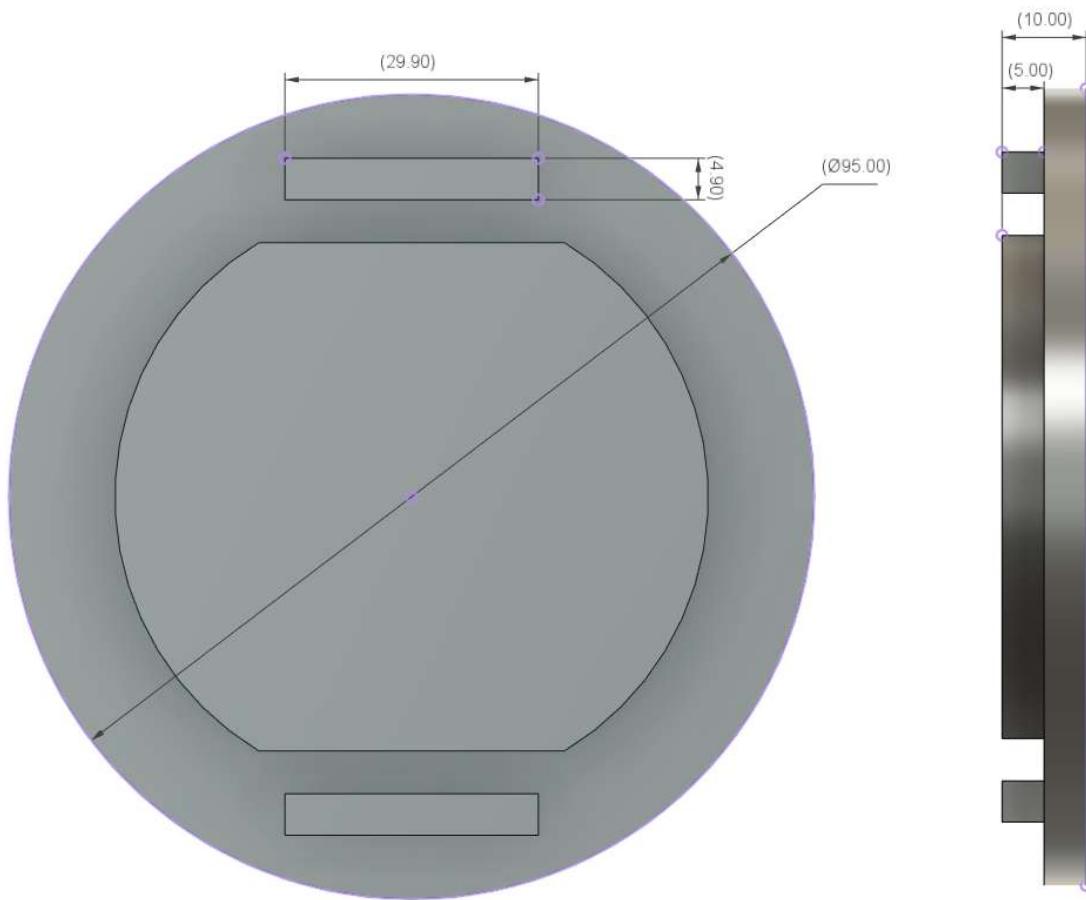
Figure A13. Proposed ear-EEG phantom prototype dimensions - Lids

References

1. Casson, A.; Yates, D.; Smith, S.; Duncan, J.; Rodriguez-Villegas, E. Wearable Electroencephalography. *IEEE Engineering in Medicine and Biology Magazine* **2010**, *29*, 44–56. doi:10.1109/MEMB.2010.936545.

2. Casson, A.J. Wearable EEG and beyond. *Biomedical Engineering Letters* **2019**, *9*, 53–71. doi:10.1007/s13534-018-00093-6.

3. Soufineyestani, M.; Dowling, D.; Khan, A. Electroencephalography (EEG) Technology Applications and Available Devices. *Applied Sciences* **2020**, *10*, 7453. doi:10.3390/app10217453.

4. Looney, D.; Kidmose, P.; Park, C.; Ungstrup, M.; Rank, M.; Rosenkranz, K.; Mandic, D. The In-the-Ear Recording Concept: User-Centered and Wearable Brain Monitoring. *IEEE Pulse* **2012**, *3*, 32–42. doi:10.1109/MPUL.2012.2216717.

5. Wyckoff, S.N.; Sherlin, L.H.; Ford, N.L.; Dalke, D. Validation of a wireless dry electrode system for electroencephalography. *Journal of NeuroEngineering and Rehabilitation* **2015**, *12*, 95. doi:10.1186/s12984-015-0089-2.

6. What are imaging phantoms? https://www.nist.gov/physics/what-are-imaging-phantoms. Available online. Accessed: 11-12-2023.

7. Pet-CT phantom. https://www.mirion.com/products/medical/nuclear-medicine-instrumentation/quality-assurance/phantoms/pet-ct-phantom. Available online. Accessed: 11-12-2023.

8. Hairston, W.D.; Slipher, G.A.; Yu, A.B. Ballistic gelatin as a putative substrate for EEG phantom devices, 2016. doi:10.48550/ARXIV.1609.07691.

9. Richer, N.; Downey, R.J.; Nordin, A.D.; Hairston, W.D.; Ferris, D.P. Adding neck muscle activity to a head phantom device to validate mobile EEG muscle and motion artifact removal. 2019 9th International IEEE/EMBS Conference on Neural Engineering (NER), 2019, pp. 275–278. doi:10.1109/NER.2019.8716959.

10. Chowdhury, M.E.; Khandakar, A.; Hossain, B.; Alzoubi, K. Effects of the phantom shape on the gradient artefact of electroencephalography (EEG) data in simultaneous EEG-fMRI. *Applied Sciences (Switzerland)* **2018**, *8*. doi:10.3390/app8101969.

11. Audette, W.E.; Allen, L.V. Design and Demonstration of a Head Phantom for Testing of Electroencephalography (EEG) Equipment Hearing Protection View project DPOAE Level and Phase Mapping View project. doi:10.13140/RG.2.2.12078.25920.

12. Wood, S.; Martins, T.; Ibrahim, T.S. How to design and construct a 3D-printed human head phantom. *Journal of 3D Printing in Medicine* **2019**, *3*, 119–125. doi:10.2217/3dp-2019-0016.

13. Marathe, A.R.; Ries, A.J.; McDowell, K. Sliding HDCA: Single-trial eeg classification to overcome and quantify temporal variability. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **2014**, *22*, 201–211. doi:10.1109/TNSRE.2014.2304884.

14. Leahy, R.; Mosher, J.; Spencer, M.; Huang, M.; Lewine, J. A study of dipole localization accuracy for MEG and EEG using a human skull phantom. *Electroencephalography and Clinical Neurophysiology* **1998**, *107*, 159–173. doi:10.1016/S0013-4694(98)00057-1.

15. Nordin, A.D.; Hairston, W.D.; Ferris, D.P. Dual-electrode motion artifact cancellation for mobile electroencephalography. *Journal of Neural Engineering* **2018**, *15*. doi:10.1088/1741-2552/aad7d7.

16. Kuratko, D.; Lacik, J.; Koudelka, V.; Vejmola, C.; Wojcik, D.K.; Raida, Z. Forward Model of Rat Electroencephalogram: Comparative Study of Numerical Simulations With Measurements on Rat Head Phantoms. *IEEE Access* **2022**, *10*, 92023–92035. doi:10.1109/ACCESS.2022.3202206.

17. Owda, A.Y.; Casson, A.J. Investigating Gelatine Based Head Phantoms for Electroencephalography Compared to Electrical and Ex Vivo Porcine Skin Models. *IEEE Access* **2021**, *9*, 96722–96738. doi:10.1109/ACCESS.2021.3095220.

18. Velcescu, A.; Lindley, A.; Cursio, C.; Krachunov, S.; Beach, C.; Brown, C.A.; Jones, A.K.; Casson, A.J. Flexible 3D-printed EEG electrodes. *Sensors (Switzerland)* **2019**, *19*. doi:10.3390/s19071650.

19. Kaveh, R.; Doong, J.; Zhou, A.; Schwendeman, C.; Gopalan, K.; Burghardt, F.L.; Arias, A.C.; Maharbiz, M.M.; Muller, R. Wireless User-Generic Ear EEG. *IEEE Transactions on Biomedical Circuits and Systems* **2020**, *14*, 727–737. doi:10.1109/TBCAS.2020.3001265.

20. Kappel, S.L.; Rank, M.L.; Toft, H.O.; Andersen, M.; Kidmose, P. Dry-Contact Electrode Ear-EEG. *IEEE Transactions on Biomedical Engineering* **2019**, *66*, 150–158. doi:10.1109/TBME.2018.2835778.

21. Kaongoen, N.; Choi, J.; Choi, J.W.; Kwon, H.; Hwang, C.; Hwang, G.; Kim, B.H.; Jo, S. The future of wearable EEG: a review of ear-EEG technology and its applications. *Journal of Neural Engineering* **2023**, *20*, 051002. doi:10.1088/1741-2552/acfcda.

22. Röddiger, T.; Clarke, C.; Breitling, P.; Schneegans, T.; Zhao, H.; Gellersen, H.; Beigl, M. Sensing with Earables: A Systematic Literature Review and Taxonomy of Phenomena. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2022**, *6*. doi:10.1145/3550314.

23. Peirce, J.; Gray, J.R.; Simpson, S.; ad Richard Höchenberger, M.M.; Sogo, H.; Kastman, E.; Lindeløv, J.K. PsychoPy2: Experiments in behavior made easy. *Behaviour Research Methods* **2019**, pp. 195–203. doi:https://doi.org/10.3758/s13428-018-01193-y.

24. Kothe, C.; Medine, D.; Boulay, C.; Grivich, M.; Stenner, T. Lab Streaming Layer open source repository. https://github.com/sccn/labstreaminglayer.

25. Razavi, M.; Yamauchi, T.; Janfaza, V.; Leontyev, A.; Longmire-Monford, S.; Orr, J. Multimodal-Multisensory Experiments. *Preprints* **2020**. doi:10.20944/preprints202008.0614.v1.

26. Event triggering and data synchronization with mobile EEG fully mobile EEG devices. https://mbraintrain.com/event-triggering-and-data-synchronization-with-mobile-eeg/. Available online. Accessed: 11-12-2023.

27. Setting up precise sound stimulation with psychopy. https://mbraintrain.com/how-to-set-up-precise-sound-stimulation-with-psychopy-and-pylsl. Available online. Accessed: 11-12-2023.

28. Bridges, D.; Pitiot, A.; MacAskill, M.R.; Peirce, J.W. The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ* **2020**, *8*, e9414. doi:10.7717/peerj.9414.

29. Can PsychoPy deliver millisecond precision. https://www.psychopy.org/general/timing/millisecondPrecision. Available online. Accessed: 11-12-2023.

30. Mobile EEG - smarting mobi. https://mbraintrain.com/smarting-mobi/. Available online. Accessed: 11-12-2023.

31. Bishop, D.V.M.; Hardiman, M.J. Measurement of mismatch negativity in individuals: A study using single-trial analysis. *Psychophysiology* **2010**, *47*, 697–705. doi:10.1111/j.1469-8986.2009.00970.x.

32. Goverdovsky, V.; Looney, D.; Kidmose, P.; Mandic, D.P. In-Ear EEG From Viscoelastic Generic Earpieces: Robust and Unobtrusive 24/7 Monitoring. *IEEE Sensors Journal* **2016**, *16*, 271–277. doi:10.1109/JSEN.2015.2471183.

33. Saleem, A.; Frormann, L.; Soever, A. Fabrication of extrinsically conductive silicone rubbers with high elasticity and analysis of their mechanical and electrical characteristics. *Polymers* **2010**, *2*, 200–210. doi:10.3390/polym2030200.

34. SILC circuits: High performance conductive silicone. https://www.instructables.com/Silc-Circuits-High-Performance-Conductive-Silicone/. Available online. Accessed: 11-12-2023.

35. OpenBCI Cyton. https://shop.openbci.com/products/cyton-biosensing-board-8-channel. Available online. Accessed: 11-12-2023.

36. OpenBCI EEG Setup. https://docs.openbci.com/GettingStarted/Biosensing-Setups/EEGSetup/. Available online. Accessed: 11-12-2023.

37. McCann, H.; Pisano, G.; Beltrachini, L. Variation in reported human head tissue electrical conductivity values. *Brain Topography* **2019**, *32*, 825–858. doi:10.1007/s10548-019-00710-2.

38. Koessler, L.; Colnat-Coulbois, S.; Cecchin, T.; Hofmanis, J.; Dmochowski, J.P.; Norcia, A.M.; Maillard, L.G. In-vivo measurements of human brain tissue conductivity using focal electrical current injection through intracerebral multicontact electrodes. *Human Brain Mapping* **2016**, *38*, 974–986. doi:10.1002/hbm.23431.

39. Mikkelsen, K.B.; Kappel, S.L.; Mandic, D.P.; Kidmose, P. EEG Recorded from the Ear: Characterizing the Ear-EEG Method. *Frontiers in Neuroscience* **2015**, *9*. doi:10.3389/fnins.2015.00438.

40. Oliveira, A.S.; Schlink, B.R.; Hairston, W.D.; König, P.; Ferris, D.P. Induction and separation of motion artifacts in EEG data using a mobile phantom head device. *Journal of Neural Engineering* **2016**, *13*. doi:10.1088/1741-2560/13/3/036014.

41. Vanderwal, T.; Kelly, C.; Eilbott, J.; Mayes, L.C.; Castellanos, F.X. Inscapes : A movie paradigm to improve compliance in functional magnetic resonance imaging. *NeuroImage* **2015**, *122*, 222–232. doi:10.1016/j.neuroimage.2015.07.069.