

# Evaluation of a headphones-fitted EEG system for the recording of auditory evoked potentials and mental workload assessment

Simon Ladouce, Max Pietzker, Dietrich Manzey, Frederic Dehais

### ▶ To cite this version:

Simon Ladouce, Max Pietzker, Dietrich Manzey, Frederic Dehais. Evaluation of a headphones-fitted EEG system for the recording of auditory evoked potentials and mental workload assessment. Behavioural Brain Research, 2024, 460, pp.114827. 10.1016/j.bbr.2023.114827 . hal-04369353

# HAL Id: hal-04369353 https://hal.science/hal-04369353

Submitted on 2 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Highlights

### Evaluation of a headphones-fitted EEG system for the recording of auditory evoked potentials and mental workload assessment

Simon Ladouce, Max Pietzker, Dietrich Manzey, Frederic Dehais

- The present findings demonstrate the feasibility of recording P300 ERP using a headphone-mounted mobile EEG system with sponge-based EEG sensors.
- The signal-to-noise ratio obtained at midline electrodes (placed over the connecting frame) was sufficient to extract average ERP features allowing to statistically discriminate workload conditions.
- In contrast, P300 ERP were not captured by sensors placed around-the-ear.
- The classification of mental workload based on the ERP response to target stimuli between condition pairs attained an accuracy ranging between 80 and 87%.

# Evaluation of a headphones-fitted EEG system for the recording of auditory evoked potentials and mental workload assessment

Simon Ladouce<sup>a</sup>, Max Pietzker<sup>b</sup>, Dietrich Manzey<sup>b</sup>, Frederic Dehais<sup>a,c</sup>

 <sup>a</sup>Human Factors and Neuroergonomics, ISAE-SUPAERO, 10 Av. Edouard Belin, Toulouse, 31400, Haute-Garonne, France
 <sup>b</sup>Department of Psychology and Ergonomics, Technical University Berlin, Straße des 17. Juni 135, 10623 Berlin, Berlin, 10623, Berlin, Germany
 <sup>c</sup>School of Biomedical Engineering, Science Health Systems, Drexel University, 3141 Chestnut St, Philadelphia, 19104, PA, United States

#### Abstract

Advancements in portable neuroimaging technologies open up new opportunities to gain insight into the neural dynamics and cognitive processes underlying day-to-day behaviors. In this study, we evaluated the relevance of a headphonemounted electroencephalogram (EEG) system for monitoring mental workload. The participants (N=12) were instructed to pay attention to auditory alarms presented sporadically while performing the Multi-Attribute Task Battery (MATB) whose difficulty was staged across three conditions to manipulate mental workload. The P300 Event-Related Potentials (ERP) elicited by the presentation of auditory alarms were used as probes of attentional resources available. The amplitude and latency of P300 ERPs were compared across experimental conditions. Our findings indicate that the P300 ERP component can be captured using a headphone-mounted EEG system. Moreover, neural responses to alarm could be used to classify mental workload with high accuracy (over 80%) at a single-trial level. Our analyses indicated that the signal-to-noise ratio acquired by the sponge-based sensors remained stable throughout the recordings. These results highlight the potential of portable neuroimaging technology for the development of neuroassistive applications while underscoring the current limitations and challenges associated with the integration of EEG sensors in everyday-life wearable technologies. Overall, our study contributes to the growing body of research exploring the feasibility and validity of wearable neuroimaging technologies for the study of human

Preprint submitted to Behavioural Brain Research

January 2, 2024

cognition and behavior in real-world settings.

Keywords: Wearable EEG, headphones, P300, Event-Related Potentials, Mental workload, sponge-based electrodesPACS: 0000, 11112000 MSC: 0000, 1111

## 1. Introduction

In recent years, a growing interest in studying human cognition and its neural correlates in unrestricted, real-life contexts has pushed neuroimaging research methods beyond laboratory walls [1, 2, 3]. The development of mobile brain and body imaging research methods (MoBI) 4, 5 aims to address longstanding critiques made toward exceedingly controlled experimental setups limiting the ecological validity of cognitive neuroscience research [6]. Indeed, it has been argued that the artificial nature of traditional lab-based experimental settings (presentation of abstract stimuli to which participants are instructed to respond by performing prototypical behaviors) places individuals in decontextualized situations and restricts their agency  $\boxed{\mathbf{7}}$ . Over the span of the last decade, the operationalization of this paradigm shift toward increasingly more ecologically valid research has prompted innovations in hardware (e.g., miniaturization of sensors) and software (i.e., data processing and analysis) solutions 8, 9 as well as conceptual adjustments (scalable research design 10) offering new opportunities for both basic and applied research 5, 11. Following the emergence of MoBI research methods, studies have shed light on brain dynamics underlying walking 12, 13, navigation 14 and allocation of attention 15, 16, 17 during real-world behavior. As of today, laboratories have shrunk to smartphone size 18, 19 and innovative MoBI appliances are being developed to further extend research horizons.

The development of MoBI methods is directly in line with the field of Neuroergonomics research (or the "study of the brain at work", [20]) which aims to understand the brain during unrestricted real-world tasks and in everyday life contexts" [21]. As an interdisciplinary field, it combines knowledge and methodologies from neuroscience, psychology, human factors, and computer sciences into one collective stream of research whose results hold the potential to benefit a wide range of domains such as safety [22]. workplace safety 23 and health 24, 25. A major topic in Neuroergonomics pertains to the characterization of neural correlates of mental effort during the execution of complex tasks and actions in natural environments 26, 27, 28, 29. Building such understanding of brain activity related to mental workload could be leveraged to inform monitoring systems that can dynamically adapt human-machine teaming to overcome cognitive limitations of the user. This ongoing research has been addressing the conceptual and theoretical aspects through an accumulation of empirical evidence using state-of-the-art laboratory methods. As aforementioned, the translation of these findings into real-world applications necessarily involves the adoption of minimally invasive MoBI systems. The capacity of such systems for recording highquality neurophysiological data without interfering with human operators' behaviors is a critical and challenging feature to attain. The meaning of noninvasiveness quality implies the optimization of different dimensions of user experience such as comfort and discreteness to ensure initial engagement with the technology but also long-term retention. An elegant solution has been proposed under the form of concealed electrodes placed inside the outer ear and on the concha 30, 31 and in-ear-recordings 32, 33, 34, 35. Rather than inside the ear, some authors achieved success in developing small, soft, curved electrode systems that can be integrated into the auricle 36. Others 37, 38 proposed a semi-circular electrode array placed around the ear and obtained viable signals for EEG analysis in both time and frequency domain <u>39</u> even under realistic settings [40, 41]. This approach allows comfortable and unobtrusive acquisition of high-quality EEG signals over prolonged periods. Although the grids can be used over several recording sessions, their setup requires the contribution of an external expert to affix the electrodes around the ear appropriately which is a time-consuming procedure 42. Another approach to acquiring brain activity during real-world behaviour inconspicuously can be found in fitting generic wearables (such as eyeglasses, caps, and headphones) with EEG sensors 43.

However, the use of these mobile devices poses new challenges in terms of handling the additional noise that inevitably contaminates EEG signals recorded during naturalistic behaviours (implying an increased proportion of motion-related artefacts) taking place in realistic settings (additional environmental sources of noise). Previous benchmarking studies have provided useful insight into the tradeoffs related to various technical characteristics of mobile EEG systems such as the electrode type used (e.g., wet versus dry) [44], minimal spatial coverage required to capture EEG features [45], or even head cap design [46]. The sensitivity of such devices to cognitive effects must be assessed under realistic conditions [39]. Empirical testing, benchmarking, and evaluation of new mobile EEG devices represent an important step from conventional and cumbersome systems to a fully portable and diverse toolbox of neuroimaging solutions [47]. As part of an iterative development process, it is therefore critical to identify and report limitations and shortcomings of such miniaturized EEG systems to inform the future designs of effective mobile neuroimaging methods [48].

The present study aims to evaluate the validity of a headphone-mounted EEG system for the recording of neural signatures of auditory attention This system allows for simultaneous EEG signal during complex tasks. acquisition and auditory stimulus presentation. Although not limited to it, the design of the assessed system appears particularly suited for consumer, clinical, and research applications aiming to record and interpret neural signatures of auditory processing. Indeed, the location of the electrodes around the ear (temporal recording sites) makes this device a particularly promising solution to record brain responses to auditory stimuli (referred to as auditory evoked potentials). Auditory evoked potentials comprise a complex of successive Event-Related Potential (ERP) components. Amongst these components, the well-documented P300 ERP is a positive deflection in the EEG around 300ms following the onset of visual and auditory stimuli. While earlier components may be evoked by both task-related and distracting sounds alike (reflecting early sensory processing), the P300 ERP is observed in response to surprising and or task-relevant stimuli. The P300 ERP component has consequently been linked to cognitive processes involved in higher-order functions such as attention and memory 49, 50, 51. A common paradigm used to evoke P300 ERPs is the auditory oddball paradigm, where participants are instructed to discriminate rare target sounds (oddballs) within a series of repeated and irrelevant frequent sounds. Time-locked EEG activity to the presentation of auditory stimuli reveals distinct and reliable P300-characteristics for target vs non-target stimuli 52, 53, 54, with strongest amplitudes typically observed over parietal and central areas 55. Recent studies have demonstrated that the P300 ERP could be effectively recorded using in-ear and aroundthe-ear EEG [38, 30, 56, 57]. The P300 ERP component has been extensively used as a proxy measure of cognitive workload in research protocols incorporating the P300 ERP elicitation paradigm 58, 59, 60, 61, 62. The amplitude of P300 ERP reflects the proportion of cognitive resources allocated to the processing of auditory stimuli from a limited pool of resources 58, 28, 62.

Based on the limited pool of attentional resources theory [63, 64, 65], an increase in primary-task demands will require additional resources at the expense of resources allocated to a secondary task. This reallocation of attentional resources is reflected by a decrease of P300 ERP amplitude in response to stimuli related to the secondary task [66, 67, 68, 69].

The present study aims to assess the feasibility of recording the eventrelated potentials (ERPs) to auditory stimuli to measure cognitive workload using a headphones-fitted EEG device. Using a repeated measures experimental design, task difficulty will be staged across three conditions to manipulate mental workload. For this purpose, a classic single-task oddball counting condition will be contrasted with a dual-tasking and a multitasking condition. The tracking and two other perceptual-cognitive subtasks (system monitoring and resources management) from the multi-attribute task battery (MATB) [70] were selected to increase task demands. Based on previous findings, the P300 amplitude observed in the single-task condition is expected to be attenuated when participants are performing a secondary tracking task simultaneously (dual task) [66]. Furthermore, it is hypothesized that further increasing task difficulty by adding two additional subtasks (the multi tasking condition) would lead to further decrements in P300 ERP amplitude [71].

# 2. Material and Method

### 2.1. Participants

Twelve participants (three female; nine male) aged between 21 to 42 years old (M = 24.8, SD = 5.8) took part in the experiment. The call for participants was performed by advertising the study on the university's website and by placing posters around the campus. All participants had normal or corrected-to-normal vision, were free of sensory and motor impairments, were not following a psychoactive pharmacological treatment, and declared no history of neurological or psychiatric conditions. EEG data from one participant had to be dismissed due to large portions of missing data. The study was approved by the ethics committee of the University of Toulouse (CER approval number 2020-334) and conducted according to the principles expressed in the Declaration of Helsinki. Participation was voluntary and neither compensated by money nor participation hours. All participants

gave their written informed consent prior to the study.

### 2.2. Experimental Design

The study followed a within-subject design with task difficulty as an independent variable manipulated across three conditions (single, dual, and multitask conditions). The P300 ERPs elicited by the presentation of auditory oddball stimuli were recorded across all three conditions. The order of experimental conditions was counterbalanced across participants using a Latin square design to control for potential training and fatigue effects. The datasets of one participant had to be excluded from further analyses due to substantial portions of missing data throughout the recording.

### 2.3. Task and stimuli

In single-task experiments, manipulation of mental workload is commonly achieved through variations of working memory load (e.g., using N-back tasks [72]), task complexity (e.g., increasing sensory input by adding distracting information [73]) or overall task difficulty (e.g., altering simulated flight conditions [74]). An alternative approach lies in varying the number of simultaneous tasks to be performed at once [71]. This latter solution was preferred for the present study as it would allow assessing the validity of the Smartfones system for the recording of EEG signals as the number of simultaneous tasks increase, providing better insight into the capabilities of such a system for ecologically valid research.

#### 2.3.1. MATB Task

The Multi-Attribute Task Battery (MATB) was used as a primary task. The MATB provides a highly modular framework in which subtasks can be presented independently or simultaneously. Each subtask has been designed to target a range of sensory processes and to involve different cognitive functions [70]. The source code is open-access and provides the experimenter with full control over task parameters [75]. In its classic configuration, the MATB comprises a monitoring task, resource management, target tracking, and radio communication subtasks which are analogous to tasks performed by the cockpit crew during real flights. Its effective implementation alongside a secondary auditory oddball task has been demonstrated in previous research on mental workload [73, 59, 76]. Following the implementation described in [76], the radio communication subtask was removed from the environment to avoid auditive cluttering and contamination of the electrophysiological responses to auditory oddball stimuli. The task environment, as presented in the experiment, is shown in [1].

#### 2.3.2. Auditory Oddball Task

The auditory oddball paradigm in which infrequent target tones are presented amongst frequent non-target tones was used as a secondary task. The auditory oddball paradigm had already been used as a secondary task to provide proxy measures of mental workload in previous studies within the context of air traffic control 77, driving 78, and real-flight operations 22, 79. In all conditions, a randomized sequence of high (2000 Hz) and low (1000 Hz) pitch tones, sampled at 44100 Hz, was presented using Psycholobox 80 functions running on MATLAB (The MathWorks, Natick, USA) programming framework. The probability of occurrence was set to a ratio of 1:4 for target and nontarget tones, respectively. The stimulus duration was set to 100 milliseconds with an inter-trial interval of one second plus a random jitter (ranging from 0 to 1000ms) to counter habituation effects **81**. A total of 300 tones (60 targets and 240 non-target stimuli) were played in each condition to ensure a sufficient number of trials for ERP analyses and classification performed at single-trial level 82, 83. The participants were instructed to silently count the number of auditory oddball tones presented and report their count at the end of each condition.



Figure 1: Screenshot of the three experimental conditions: single-task (auditory oddball only), dual-task (oddball + tracking task), and multi-task (auditory oddball + MATB).

### 2.4. EEG recording

The Smartfones EEG system (mBrainTrain, Belgrade, Serbia) is equipped with a total of 11 sponge-based electrodes along with a reference and a ground electrode concealed within the earcups and the connecting frame of the device (see 2). All electrodes (L1, L2, L3, L4, R1, R2, R3, R4, C3, C4, Cz) were used for EEG signal acquisition and were online referenced to the integrated reference electrode located at the bottom of the left earcup at the left mastoid recording site 2. The signal was amplified (Smarting PRO amplifier, mBrainTrain) and recorded at a sampling rate of 500Hz. Prior to the experiment, the small sponges that connect the lowered electrodes with the scalp (referred to as sponge-based electrodes) were soaked in a saline solution for an hour. and were then inserted into the electrode shafts. The headphones were carefully placed to ensure participants' comfort and that the Cz-electrode was positioned in accordance with the international 10-20 layout [84]. The height of the earcups was then adjusted to make sure that both ears were covered by the speakers.



Figure 2: (A) Smarting Smartfones EEG device by mBrainTrain (Belgrade, Serbia). The headset is fitted with sponge-based electrodes. (B) Schematic representation of the Smartfones montage that includes 11 EEG recording sensors (green circles), a reference electrode positioned on the left mastoid recording site (blue circle), and the ground electrode placed on the right mastoid (black circle).

For each condition, event markers associated with the auditory stimuli presented were synchronized to the continuous EEG recordings using the LabStreamingLayer (LSL [85]) framework and stored under an xdf-file format.

#### 2.5. Procedure

Upon arrival, participants were briefed about the procedure, tasks, and equipment used for the experiment. The participants then gave their written informed consent by filling out a terms and conditions agreement form. Prior to undertaking the EEG setup procedure, subjects were first introduced to the auditory oddball task and the MATB controls through a tutorial lasting for about 15 minutes. The experimenter ensured that the participants could clearly hear both types of tones by adjusting the headphone's volume based on participant feedback. All the individual subtasks were trained individually before performing them simultaneously in the final stage of the training phase. At the end of the training, performance measures were displayed and reviewed by the experimenter to check if the instructions and controls had been understood and answer any questions the participants may have. The participants were then prepared for EEG data acquisition following a preparation protocol specific to the use of sponge-based electrodes. The skin around the participants' ears was cleaned using an alcohol wipe to remove dirt and sebum that would occlude skin conductance. Measurement of scalp midlines, from nasion to inion and left to right pre-auricular points, were carried out to identify the anchor point for the central electrode (Cz). Following the headphones setup, EEG signal quality and impedance were checked using the Smarting Streamer software (mBrainTrain, Belgrade, Serbia). Subjects were instructed to abstain from large or rapid movements during task performance and were given the opportunity to take short breaks in between conditions.

#### 2.6. EEG Processing and P300 ERP Feature Extraction

Data processing was performed with the EEGLAB toolbox [86] written in MATLAB (The MathWorks, Natick, USA). In the first step, the continuous data was downsampled to 250 Hz and subjected to a Finite Impulse Response (FIR) band-pass between 1 and 30 Hz with a filter order of 16500 and a -6dB cut-off. The filtered data was then epoched with a time window ranging from -200 ms to 1000 ms around experimental event markers. The

epoched data was baseline corrected (subtractive method) using the mean activity recorded over the prestimulus interval (200ms preceding stimulus onset). The ERP waveforms were computed by averaging single trials for each condition and stimulus type. The P300 ERP amplitude was extracted by averaging activity recorded between the 250 to 450 ms post-stimulus time window 87, 88. For the ERP analyses, epoched data containing aberrant activity (threshold of five standard deviations around the median voltage activity recorded across all epochs) were discarded. The following classification pipeline was applied to all single trials (without applying any trial rejection method) in order to have conditions similar to ERP online classification. The statistical analyses employed repeated measures ANOVA and paired-sample t-tests as appropriate to the factorial repeated-measures experimental design. To ensure that parametric analysis was appropriate we first confirmed that all data followed a normal Gaussian distribution. In addition, for all post-hoc t-tests, the Holm-Bonferroni correction for multiple comparisons was applied.

#### 2.7. Classification

The classification pipeline was implemented using scikit-learn Python libraries for the general implementation of training and testing of classifiers. The Pyriemann library was used to perform the transformation of multivariate time series EEG data into covariance matrices, which allows the classification of the Riemannian geometry of symmetric positive definite matrices. First, the number of trials per class was balanced through the random selection of undersampled ensembles of trials. Then, the dimensionality of the multichannel EEG data was reduced using the unsupervised Xdawn method with a set of four spatial filters 89. By doing so, the Xdawn algorithm maximizes the difference between the ERP responses of a pair of classes (e.g., target versus non-target ERPs; single task target versus dual task target ERPs). Afterward, the spatially filtered signals were projected into the estimated signal subspace, yielding feature vectors that are subsequently input into a logistic regression model 90. For this purpose, we adopted the Riemannian geometry approaches 91, 92) that demonstrated unmatched performances within the frame of both single-trial ERP and passive BCI classification problems, as highlighted by a recent retrospective study 93. The classification pipeline used in this study can be formally described as follows:

A classifier was trained for each subject. For each class, a prototype response

P was obtained by averaging signals across trials. For each trial Xi, a template trial Ti was built using the concatenation of P and the trial Xi. The template trials Ti were converted into covariance matrices to capture the spatial structure of the signal [94]. The covariance matrices were then projected into their tangent space using the geometric mean of all covariance matrices as a reference point [90]. After this projection, each covariance matrix was represented by a vector upon which a logistic regression (without any regularization) was applied for classification. The performance was evaluated in terms of accuracy using 5-fold cross-validation. The folds were stratified to ensure that each of them contained the same number of trials for each class. The results for each class accuracy were computed and are reported under the form of confusion matrices in Figure 5 while the overall balanced accuracy of the classifier (averaged over folds) and the area under the ROC curve (ROC AUC) are reported under Table 1 and Table 2.

The assumption that chance level accuracy for a binary classification problem equals 50% only stands in theory for datasets with infinite number of samples 95. In the case of limited datasets with low number of samples per classes, which is a common scenario in neurophysiological recordings and applies to the present study, chance level needs to be adjusted to account for the variance in classification accuracies in small datasets. For this purpose, the statistical significance threshold of classification performance above chance level was determined with respect to a null distribution of classification accuracies computed through random permutations of class labels 95. For each single dataset, the original (unpermuted) classification accuracy was interpreted with respect to the distribution of classification performances obtained from the permutation of class labels of the exact same exact dataset repeated 99 times. The tails of the permutation distribution provide statistical significance boundaries for a given rate of false positive. As such, if the original classification accuracy is above the 95 or 99 percentiles of the empirical distribution then the classification performance is significant with  $\alpha = .05$ and  $\alpha = .01$ , respectively. The classification performances reaching statistical significance are flagged with asterisks in Tables 1 and 2.

#### 2.8. Signal quality over time

The root-mean-square (RMS) amplitude of the EEG signal is a general measure of the magnitude of the signal throughout the window, irrespective of frequency [88]. Windows containing high-amplitude artifacts are typically characterized by a substantially larger RMS amplitude than those with amplitudes in the normal EEG range. The relationship between the RMS amplitude of chronologically ordered epochs and time-on-task was then investigated through the computation of correlation coefficients.

## 3. Results

#### **3.1.** Task performance

The oddball task performance was analyzed to investigate the effect of the primary task load. For each condition, subjects were asked to report the number of oddball sounds presented over the last experimental condition. Reported values were subtracted from the correct number of oddballs within The performance measure was computed from absolute each condition. differences between the participant's reported count and the actual number of oddball stimuli presented. Auditory oddball task performance deteriorated as a function of task load as reflected by the increase in the absolute number as a function of task difficulty. While the number of errors was lowest in the single task condition (M = 1.1, SD = 2.4), the addition of the tracking task significantly increased the number of errors committed (M = 3.3, SD = 2.8). Task performance further decreased when participants performed the three MATB subtasks and the oddball task concurrently, resulting in the highest deviation from the correct oddball count (M = 5.8, SD = 3.2). A repeated measures ANOVA was conducted on task performance measures revealing a main effect of task difficulty on the oddball counting performance [F (2,20) = 17.87, p < .001,  $\eta_p^2 = .39$ ]. All pairwise comparisons between conditions reached statistical significance (p<.05) with Bonferroni-Holm corrections applied for multiple comparisons.

### **3.2.** Event-Related Potentials (ERP) analysis

#### 3.2.1. Establishing the presence of P300 ERP across recording sites

An initial analysis was dedicated to comparing signals recorded at earcup electrodes with the midline central electrodes to assess the benefits of integrated ear electrodes compared to standardized central locations. Although it is well established in the EEG literature that auditory-evoked potentials are most prominent over central and parietal areas [53], [54], [71], several studies have demonstrated the feasibility of recording far-field ERP using electrodes placed over temporal recording locations [96], [97]. Electrodes were clustered into three subgroups based on their location to provide a low-level factor for statistical analysis. As such the electrodes were spatially clustered into the following subgroups: left (L1, L2, L3, L4), central (C3, C4, Cz), and right (R1, R2, R3, R4).



Figure 3: Grand average (N=11) difference ERP waveforms for each condition (blue: single task, green: dual task, red: multi task) at each electrode location. The shaded area represents one standard error from the mean.

The grand average P300 ERP amplitude recorded across all conditions was subjected to a repeated measures ANOVA with stimulus type (target, non-target) and electrode site (left, central, right) as factors. The analysis revealed a main effect of stimulus type [F(1, 10) = 5.55, p < .05,  $\eta_p^2$  = .38]and an interaction effect with electrode site [F(2, 20) = 7.74, p < .01,  $\eta_p^2$ = .46]. Post-hoc paired-sample comparisons of stimulus type revealed that target sounds elicited higher amplitudes than non-target sounds with a mean difference of [t(10) = 2.36, p < .05]. Paired-sample t-tests were performed between P300 ERP amplitude elicited by target and non-target stimuli at left [t(10) = 1.58, p > .05)], right [t(10) = .65, p > .05], and central [t(10) =3.1, p < .05, d = .98] recording sites clusters, revealing that the P300 effect was only captured at midline central recording locations. The spatialization of the P300 ERP is clearly illustrated in Figure 3 where the largest P300 effects are observed over central electrodes in all conditions. These results indicate that only the central electrodes of the headphone-mounted EEG system could reliably capture the P300 effect.

Further inspection of ERPs recorded at the central electrode (Cz) for target and non-target trials revealed conventional waveform features like an early negative polarisation around 100 ms post-stimulus onset (N1 component) followed by a later positive change in voltage, representing the P300 response in target trials which diminishes almost fully for non-target trials 4. Across conditions, a continuous decrease in amplitude can be seen for target trials. This is also reflected in the difference waves, indicating an effect of additional task demands in dual and multitask conditions. Analogous dynamics are visible in the time-frequency domain 4, showing an event-related activation in slower frequency bands for target trials which correspond to expected effects **98**, **99**. Compared along conditions, the activation is inversely affected by increasing task load from single to dual and multitask recordings. Separate difference waves for conditions at all electrodes are shown in 3. Compared to the central electrodes, the loss in signal strength at ear electrodes is apparent but corresponds to reported findings in ear-EEG studies 100, 56. However, components are not as pronounced, waves fringe out, and condition effects seem arbitrary by visual inspection. The lack of strength also becomes visible in the spatial distribution of grand average P300 amplitude where a decrease in amplitude is only trackable in C3, C4, and Cz electrodes.

#### 3.2.2. Effect of task difficulty on P300 effect amplitude

As aforementioned, the performance of the secondary auditory oddball counting task deteriorated in dual and multitasking conditions, reflecting the increasing demands of the primary tasks. Likewise, the P300 response to secondary task sounds is expected to decrease with increasing task demands. Another repeated measure ANOVA was used to seek out the effects of task load on the P300 component among conditions. To reduce the number of factors and interaction terms, difference waves were used to extract the mean voltage within the predefined P300 time window (250 ms - 450 ms), releasing the factor for stimulus type. Two contrasts were defined a priori, testing for differences between single vs. dual, as well as dual vs. multitask conditions. Figure 3 shows the means of P300 amplitudes extracted from difference waves whereas only the central group of electrodes seems to capture the expected decrease in voltage across conditions. Only here, bars reflect the distinct gradation from single to multitask condition, although variation in the data is high.



Figure 4: (A) Grand average (N=11) Event-Related Potentials waveforms recorded at central electrode Cz in response to target (top) and non-target (bottom) auditory stimuli grouped by task difficulty (single, dual, and multi task conditions). The shaded area represents one standard error of the mean. The dotted square marks the time window (250 to 450ms) used for P300 ERP features extraction. (B) The difference waveform between target and non-target ERP extends to the event-related spectral perturbation (ERSP) plots highlighting a prototypical increase in low frequency (delta and theta) activity over the P300 ERP time period. (C) The spatial distribution of grand average P300 ERP component amplitude recorded over sponge-based electrodes for each condition.

The two-way repeated measures ANOVA, with task type (single, dual, multi) and recording site (left, central, right) as factors, revealed a significant main effect of recording sites  $[F(2, 20) = 7.18, p = .001, \eta_p^2 = .52)]$ . The P300 amplitude was found to be highest at central electrodes (M = 3.31, SD = 4.2, surpassing both left-side (M = 1.2, SD = 2.1, t(10) = 3.42, p < .01, d = .79, and right-side (M = 0.7, SD = 2.5, t(10) = 4.193, p < .01, d = .97) sensors clusters. There was also an interaction effect between the task type and recording site (F(4,40) = 4.19, p < .01,  $\eta_p^2$  = .32). Further repeated measures of one-way ANOVA with task as a factor were carried out separately for each recording cluster. Only the central recording sites showed a main effect of task on P300 amplitude (F(2,18) = 7.05, p < .01,  $\eta_p^2 = .44$ ), whereas left (F(2,20) = .58, p > .56,  $\eta_p^2$  = .06) and right (F(2,20) = .01, p > .05,  $\eta_p^2 < .01$ ) electrodes were not sensitive to the task type effect. Further post hoc paired-samples t-tests revealed that P300 ERP amplitude recorded over central electrode sites was higher for the single task (M = 5.5, SD = 4)than for the multi (M = 1.98, SD = 4.3, t(10) = 3.69, p < .01, d = .92) task but only marginally higher than for the dual task (M = 3.5, SD = 2.9, t(10))= 2.42, p > .05, d = .60).

#### 3.3. Classification of auditory processing

A first analysis aimed to assess whether single-trial responses to task-related (target) and distractor stimuli could be effectively identified across different workload conditions using the headphones-fitted EEG system. For this purpose, the classifier was trained on single-trial ERP responses to target and nontarget stimuli for each workload condition (single-dual, single-multi, dualmulti task). Table 1 presents the balanced accuracy and ROC AUC achieved by the Xdawn classifier to distinguish between ERP responses elicited by target and non-target stimuli across each condition. Statistical significance thresholds of classification accuracy beyond chance level were computed following a permutation-based approach [95] (detailed in the classification methods section). Across subjects and conditions, the classification performance significance boundaries were 67.8% and 71.2% for  $\alpha = .05$  and .01, respectively. The reported classification accuracies above statistical significance boundaries are flagged (\*p < .05, \*\*p < .01). The top row of Figure 5 present the confusion matrices for the classification of target versus non-target ERP responses for each of the three workload conditions.

Table 1: Subject-level classification performance of the classifier evaluated through balanced accuracy (in %) and ROC AUC score to distinguish EEG responses to target and non-target stimuli across single, dual, and multi-task conditions. The feature extraction was based on the Riemannian geometry of spatial covariance matrices whose discriminability across pairs of classes was enhanced by the Xdawn algorithm. Classification performances above statistical significance boundaries for chance level decoding are flagged (\* $\alpha = .05$ , \*\* $\alpha = .01$ ).

	Single		Dual		Multi	
	Accuracy	ROC AUC	Accuracy	ROC AUC	Accuracy	ROC AUC
P1	80.7**	.90	65.7	.82	50.7	.65
P2	73.4**	.79	67.8	.63	71.3**	.64
P3	$73.8^{**}$	.77	67.4	.82	66.4	.78
P4	$73.6^{**}$	.75	62.1	.79	52.7	.58
P5	66.3	.70	$68.1^{*}$	.71	51.7	.65
P6	79.3**	.80	46.2	.56	57.9	.76
P7	$70^{*}$	.69	65.9	.80	57.9	.64
P8	59.3	.46	60.7	.62	53.8	.58
P9	65.5	.73	57.2	.59	49.3	.45
P10	60	.46	50.3	.51	53.8	.55
P11	54.3	.54	47.6	.50	52.1	.60
Mean	68.7	.69	59.9	.67	56.2	.62
$\mathbf{SD}$	8.5	.14	8.4	.12	6.9	.09

### 3.4. Classification of workload

The classifier was trained on single-trial ERP responses to target stimuli pooled from the three couple combinations of workload conditions (singledual,single-multi, dual-multi task). Table 2 presents the balanced accuracy and ROC AUC achieved by the classifier to distinguish between ERP responses of the contrasted conditions across participants. Statistical significance thresholds of classification accuracy beyond chance level were computed following a permutation-based approach [95] as described in the classification methods section. Across subjects and conditions, the classification performance significance boundaries were 63.5% and 68.3% for  $\alpha = .05$  and .01, respectively. The reported classification accuracies above statistical significance boundaries are flagged (\*p < .05, \*\*p < .01). The bottom row of Figure 5 present the confusion matrices for the binary classification of workload for each of the

#### three conditions couples.

Table 2: Subject-level classification performance of the classifier evaluated through balanced accuracy (in %) and ROC AUC score to distinguish the level of mental workload across paired conditions. The feature extraction was based on the Riemannian geometry of spatial covariance matrices whose discriminability across pairs of classes was enhanced by the Xdawn algorithm. Classification performances above statistical significance boundaries for chance level decoding are flagged (\* $\alpha = .05$ , \*\* $\alpha = .01$ ).

	Single-Dual		Single-Multi		Dual-Multi	
	Accuracy	ROC AUC	Accuracy	ROC AUC	Accuracy	ROC AUC
P1	69.3**	.82	83.6**	.94	85**	.88
P2	74.1**	.79	86.8**	.88	$64.3^{*}$	.77
P3	85.2**	.93	75**	.86	74**	.79
P4	88.6**	.98	86.4**	.90	69.7**	.81
P5	$76.6^{**}$	.74	84.4**	.91	69.6**	.81
P6	$67^{*}$	.75	87.1**	.86	81.4**	.89
P7	$93.3^{**}$	.98	91.4**	.95	$93.5^{**}$	.98
P8	94.8**	.99	86.9**	.96	90**	.96
P9	$69.7^{**}$	.80	93.2**	.89	90**	.98
P10	95.2**	.99	$98.6^{**}$	.98	89.7**	.95
P11	90**	.97	91.4**	.97	80**	.89
Mean	82.1	.89	87.7	.92	80.6	.88
$\mathbf{SD}$	11.0	.1	6.1	.04	10	.08



Classification of target versus non-target ERP at single-trial level

Figure 5: Confusion matrices for the Tangent Space Classifiers trained with Riemannian geometry spatial covariance matrices of ERP data enhanced with the Xdawn spatial filter. Top row: classification of target versus non-target single-trials ERP for each workload condition (single, dual, and multi task). Bottom row: Classification of mental workload based on single-trial ERP responses to target stimuli.

### 3.5. Analysis of Signal-to-Noise Ratio (SNR) over time

The RMS of EEG signals recorded over all electrodes was computed for every consecutive epoch in each condition. There was no effect of experimental condition on the average EEG signal RMS (F(2,20) = .078, p > .05). Moreover, no correlation was found between signal RMS and time on task (r(265) = .04, p > .05), suggesting neither improvement nor degradation of EEG signal quality over time.



Figure 6: Distribution of signal quality across participants throughout the study. The signal quality was measured as the Root Mean Square (RMS) of single-trial EEG signals averaged over all electrodes.

## 4. Discussion

The desire to observe truly enacted cognition under unrestricted conditions motivates the development of innovative solutions for the recording of brain and body dynamics that can be taken outside of the laboratory. Over the last decade, a multiplicity of portable neuroimaging hardware has emerged providing a range of tailored solutions for specific use cases. The necessary trade-off made during the design of such miniaturized systems between user experience (comfort and discreetness) and signal quality is a crucial aspect to consider. The design of the headphones-fitted EEG system evaluated in the present study is promising for the quick and discreet recording of auditory evoked potentials in ecologically valid settings. This design also includes sponge-based electrodes which are quick to set up and don't require the application of sticky gel on participants' scalp. The system comprises a much lower number of sensors (11 channels) compared to research-grade EEG systems (commonly ranging from 32 to 128 channels). This lower electrode density precludes the application of effective processing methods relying on a larger and more distributed head coverage. The present study did not compare the signals captured by the headphones-fitted EEG system to a research-grade EEG. Therefore it remains to be assessed whether such a system could provide recordings on par with higher-density wet-EEG systems in terms of signal quality. Nevertheless, the present study characterizes the advantages and pitfalls of the evaluated system which is critical to inform the development of novel mobile EEG solutions. These insights can then be leveraged to inform the design of future devices but also provide researchers with objective information upon which they can base their choice of research methods [48]. This study aimed to determine if a headphones-fitted spongebased EEG device could be applied to mental workload assessment in multitasking contexts. Moreover, we also wanted to evaluate if this minimalistic design and its concealed ear electrodes would be suitable for the capture of auditory evoked potentials. An auditory oddball task was implemented as a single task and extended by two additional conditions of increased task load. Under single-task performance, the type of sound (target and non-target) showed a significant effect on auditory evoked P300 amplitudes that was observed only at central electrodes, replicating well-documented effects of higher and stronger amplitude in response to target sounds over central areas 52, 53, 54. The amplitude of the P300 ERP recorded using the headphones-fitted EEG system was comparable to those commonly reported in studies using researchgrade wet-EEG systems 101 and around-the-ear systems 39.

The oddball counting performance was found to be significantly affected by the complexity of the primary task. Indeed, counting errors increased as a function of primary-task demands, with the least number of errors observed in the single task condition and the most errors for the multi task condition. The oddball task performance was a reliable behavioural marker reflecting the scarcity of resources available for auditory processing during the simultaneous performance of multiple cognitive-motor tasks. It could be argued that this juxtaposition of the auditory oddball paradigm contaminates the behavior and cognitive experiences investigated. Moreover, using such a paradigm as a probe to extract measures of cognitive resources may be impractical in applied field research that may span over hours [43], [15], [102]. Indeed, continuous counting and reporting are simply impossible, again making the case for wearable and discreet EEG devices that would allow capturing cognitive markers of mental workload reliably during the performance of natural behaviors. In line with the behavioural results, the ERP analysis revealed a modulation of the P300 amplitude as a function of task load. This effect was only observed over central recording sites which is consistent with previous findings highlighting a spatial distribution of the P300 effect most prominent over midline central and parietal recording sites 22, 58, 59, 60, 61, 62, 103. Further post hoc analyses revealed that this effect was mainly driven by differences between the single and multi-task conditions in terms of their averaged P300 ERP amplitude, whereas no significant difference was found between the dual condition and any of the other conditions. This is a particularly underwhelming result as the characterization of gradual changes in mental workload represents one of the main objectives of the present research. One could argue that these P300 amplitude differences could be attributed to higher noise resulting from increased participants' motion to handle the joystick and electrode displacement in higher load conditions. However, the analysis of signal-to-noise ratio (SNR) over time revealed neither improvement nor degradation of the EEG signal over time nor any significant difference between the different experimental conditions. However, the P300 ERP effect was captured neither by the left nor the right sensor clusters. This absence of clear P300 ERP response is in contrast to previous studies showcasing the feasibility of capturing ERP components using around-the-ears EEG montages [38, 104, 32]. It is worth noting that the aforementioned studies utilized a different electrode layout comprising wet electrodes placed around-the-ear of the participants which were firmly held in place with adhesive bands. The higher adherence of such apparatus may explain the higher SNR as a direct consequence of an overall lower and stable impedance 15. It is also unclear whether the absence of the P300 ERP response at left and right earcup recording locations should be attributed to the sponge-based electrodes' limitations to capture far-field potentials, electrode displacements during the recordings, or if analogous electrode sites in a conventional system would have also failed to record the effects. To address this uncertainty, it would be valuable to benchmark the headphones-fitted EEG system against a reference research-grade EEG system. Lastly, an inspection of ERP waveforms recorded over left earcup sensors revealed modulations of late event-related potentials. While the investigation of such responses is beyond the scope of the present study, their modulation may effectively be leveraged to better assess mental workload. It is indeed possible that lateral electrodes, and in this case sensors fitted in the left earcup, although not sensitive to P300 ERP modulations may however be more adequate to capture later and more distal signals such as components of the Late Positive Complex (LPC, [105, 106]) reflecting higher-order cognitive processes [107].

Another aspect of our analysis focused on performing single-trial classification of event-related brain responses. To accomplish this, we implemented a signal processing pipeline using state-of-the-art methods for ERP classification, which achieved a mean accuracy of around 61% across all task load conditions. Important differences were observed across workload conditions in terms of classification performance of ERP response to target versus non-target stimuli. Furthermore, only the higher classification accuracy achieved in the single task condition was found to be above chance level, and this only for a subset of participants. These results highlight the inconsistency in classification of ERP evoked by different types of auditory stimuli. It should also be noted that this performance is substantially lower than the accuracy of 70% in the distinction of target and non-target neural responses achieved through the use of around-the-ear EEG arrays 43. This low performance is even more surprising considering the contribution of the additional three central electrodes of the headphones-fitted EEG system. Indeed, the proximity of these central electrodes to the midline parietal recording locations where the P300 ERP effect is most prominent should provide higher SNR than the around-the-ear montage which only allows the capture of far-field P300 ERP signals. It appears that these central electrodes do not compensate for an overall lower SNR, which we can only speculatively attribute to the spongebased design of the sensors fitted in the headphones in the absence of a proper comparison with a wet-EEG system. Furthermore, it is worth noting that we were able to consistently classify the different workload conditions across participants with mean accuracies ranging from 80% to 87% using P300 ERP features. The accuracy scores are on par with previous studies that utilized research-grade electrodes and reported accuracies ranging from 70 to 90%**108**, **76**. Such results hold promising implications for the implementation of unobtrusive neurotechnology to monitor mental workload and auditory attention performance in real-world settings 43, 103, 40.

However, it is important to consider further hardware developments to enhance the practicality of such systems in everyday life scenarios. Mounting the smartphones with external assistance is preferable as the rigid frame hinders easy adjustment such as parting hair to ensure good adherence of the electrodes with the scalp or the eventual rehydration of the sponges to maintain signal quality over time as the sensors dry. Moreover, it should be noted that even a slight movement of the headphones of the system would cause a shift and introduce noise across all electrode sites simultaneously. Although the headphones could be tightened to adhere more closely to participants' heads, such an adjustment came at the price of wearing comfort, as reported verbally by the participants at the occasion of an informal debrief carried out at the end of the experiment. This important trade-off between user comfort and signal quality is dependent on whether the headphones were tightly or loosely fitted is a critical aspect to improve in future renditions of headphones-fitted EEG devices. A potential solution to improve both user experience and signal quality could be to adopt a hybrid hardware design including quick-to-set-up sponge-based sensors in the frame of the headphones along with around-the-ear wet EEG sensors. It is also important to note that, although the headphones-fitted EEG system allows for recording of brain activity during whole body movements thanks to the lightweight hardware and unterhered connection between the amplifier and recording unit, the system was however used in stationary settings within the frame of the present study. While portability represents an important advantage of the system, the validity of measures under whole-body motion remains to be investigated.

Another important aspect of the present work that requires further consideration concerns the experimental oddball paradigm. In the present study, an active auditory oddball paradigm was used during which participants were instructed to silently count the number of target stimuli presented over the whole duration of an experimental condition. While such an active oddball paradigm evokes strong neural responses that can be leveraged to probe attentional processes and be used as a proxy measure of cognitive resources available, they however introduce additional workload. It can be argued that increasing mental workload by the addition of the counting oddball task defeats the purpose of monitoring mental workload as it places the agents in artificial conditions. A potential solution to this issue could be found in using a passive variant of the auditory oddball task instead of active counting. This task would not involve the distributed network of brain regions involved in context updating and maintenance of counting information in working memory. The use of passive oddball has been demonstrated as an effective approach to distinguish different levels of mental workload 28. Future studies should investigate the use of such a passive oddball paradigm to capture neural markers of mental workload. Indeed, opting for passive paradigms would

only yield minimal contamination of the actual primary task (or ensemble of tasks) on which an operator focuses within the frame of a natural working environment.

# 5. Conclusion and perspectives

The present results indicate that a headphones-fitted EEG system with wet sponge-based sensors can effectively capture neural signatures of auditory processing using electrodes spread over the connecting frame. However, the electrodes fitted in both earcups did not capture the P300 ERP. The temporal recording locations could therefore not be leveraged to assess mental workload nor distinguish types of auditory stimuli based on the modulation of P300 ERP. Although interesting in terms of its unobtrusive and inconspicuous design, the device was reportedly described as uncomfortable to wear over time by the participants of this study. This limitation renders the assessed system unsuitable for long recordings. The signal quality recorded by the sponge-based sensors fitted on the connecting frames allowed for the classification of mental workload with high accuracy. The development of wearable and inconspicuous EEG devices is instrumental in the diversification and extension of the mobile-brain imaging toolbox. Provided that future iterations of headphones-fitted EEG systems improve upon the aforementioned user experience (wearing comfort) along with signal quality pitfalls reported at earcup sensors, this type of wearable EEG device would offer an elegant and tailored solution to the investigation of auditory processing in real-world settings.

# 6. ACKNOWLEDGMENT

This work was funded by AID (Powerbrain project) and ANITI (Chair for Neuroadaptive Technology).

# References

 S. Makeig, K. Gramann, T.-P. Jung, T. Sejnowski, H. Poizner, Linking brain, mind and behavior, International Journal of Psychophysiology 73 (2) (2009) 95–100. doi:10.1016/j.ijpsycho.2008.11.008.

- [2] S. Ladouce, D. I. Donaldson, P. A. Dudchenko, M. Ietswaart, Understanding minds in real-world environments: Toward a mobile cognition approach, Frontiers in Human Neuroscience 10 (2017) 694.
- [3] M. Rosenkranz, T. Cetin, V. N. Uslar, M. G. Bleichner, Investigating the attentional focus to workplace-related soundscapes in a complex audio-visual-motor task using eeg, Frontiers in Neuroergonomics 3 (2023) 38.
- [4] K. Gramann, J. T. Gwin, D. P. Ferris, K. Oie, T.-P. Jung, C.-T. Lin, L. D. Liao, S. Makeig, Cognition in action: Imaging brain/body dynamics in mobile humans, Reviews in the Neurosciences 22 (6) (2011) 593–608. doi:10.1515/RNS.2011.047.
- [5] E. Jungnickel, L. Gehrke, M. Klug, K. Gramann, Chapter 10—mobi—mobile brain/body imaging, in: H. Ayaz, F. Dehais (Eds.), Neuroergonomics, Academic Press, 2019, pp. 59–63. doi:10.1016/B978-0-12-811926-6.00010-5.
- [6] P. De Sanctis, T. Solis-Escalante, M. Seeber, J. Wagner, D. P. Ferris, K. Gramann, Time to move: Brain dynamics underlying natural action and cognition, European Journal of Neuroscience 54 (12) (2021) 8075– 8080. doi:10.1111/ejn.15562.
- S. G. Shamay-Tsoory, A. Mendelsohn, Real-life neuroscience: An ecological approach to brain and behavior research, Perspectives on Psychological Science 14 (5) (2019) 841–859. doi:10.1177/1745691619856350.
- [8] P. M. R. Reis, F. Hebenstreit, F. Gabsteiger, V. von Tscharner, M. Lochmann, Methodological aspects of EEG and body dynamics measurements during motion, Frontiers in Human Neuroscience 8 (156) (2014). doi:10.3389/fnhum.2014.00156. URL http://journal.frontiersin.org/article/10.3389/fnhum. 2014.00156/abstract
- [9] M. Klug, S. Jeung, A. Wunderlich, L. Gehrke, J. Protzak, Z. Djebbara, A. Argubi-Wollesen, B. Wollesen, K. Gramann, The BeMoBIL Pipeline for automated analyses of multimodal mobile brain and body imaging data, bioRxiv (2022).

- [10] F. J. Parada, Understanding Natural Cognition in Everyday Settings: 3 Pressing Challenges, Frontiers in Human Neuroscience 12 (September) (2018) 1–5. doi:10.3389/fnhum.2018.00386.
- [11] A. K. Engel, A. Maye, M. Kurthen, P. König, Where's the action? the pragmatic turn in cognitive science, Trends in Cognitive Sciences 17 (5) (2013) 202–209. doi:10.1016/j.tics.2013.03.006.
- [12] S. Pizzamiglio, U. Naeem, H. Abdalla, D. L. Turner, Neural correlates of single- and dual-task walking in the real world, Frontiers in Human Neuroscience 11 (2017) 460. doi:10.3389/fnhum.2017.00460.
- [13] M. Mustile, D. Kourtis, S. Ladouce, G. Learmonth, M. Edwards, D. Donaldson, M. Ietswaart, Mobile EEG reveals functionally dissociable dynamic processes supporting real-world ambulatory obstacle avoidance: Evidence for early proactive control, European Journal of Neuroscience (2021). doi:10.1111/ejn.15120.
- [14] A. Wunderlich, K. Gramann, Eye movement-related brain potentials during assisted navigation in real-world environments, European Journal of Neuroscience (2021). doi:10.1111/ejn.15095.
- [15] D. Hölle, J. Meekes, M. G. Bleichner, Mobile ear-eeg to study auditory attention in everyday life, Behavior Research Methods 53 (5) (2021) 2025–2036. doi:10.3758/s13428-021-01538-0.
- [16] S. Ladouce, D. I. Donaldson, P. A. Dudchenko, M. Ietswaart, Mobile eeg identifies the re-allocation of attention during real-world activity, Scientific Reports 9 (1) (2019) 15851.
- [17] S. Ladouce, M. Mustile, M. Ietswaart, F. Dehais, Capturing Cognitive Events Embedded in the Real World Using Mobile Electroencephalography and Eye-Tracking, Journal of Cognitive Neuroscience (2022). doi:10.1162/jocn\_a\_01903.
- [18] A. Stopczynski, C. Stahlhut, J. E. Larsen, M. K. Petersen, L. K. Hansen, The smartphone brain scanner: A portable real-time neuroimaging system, PLOS ONE 9 (2) (2014) e86733. doi:10.1371/ journal.pone.0086733.

- [19] A. Stopczynski, C. Stahlhut, M. K. Petersen, J. E. Larsen, C. F. Jensen, M. G. Ivanova, T. S. Andersen, L. K. Hansen, Smartphones as pocketable labs: Visions for mobile brain imaging and neurofeedback, International Journal of Psychophysiology 91 (1) (2014) 54–66. doi: 10.1016/j.ijpsycho.2013.08.007.
- [20] R. Mehta, R. Parasuraman, Neuroergonomics: A review of applications to physical and cognitive work, Frontiers in Human Neuroscience 7 (2013). doi:10.3389/fnhum.2013.00889.
- [21] F. Dehais, W. Karwowski, H. Ayaz, Brain at work and in everyday life as the next frontier: Grand field challenges for neuroergonomics, Frontiers in Neuroergonomics 1 (2020) 583733. doi:10.3389/fnrgo. 2020.583733.
- [22] F. Dehais, A. Dupres, S. Blum, N. Drougard, S. Scannella, R. Roy, F. Lotte, Monitoring pilot's mental workload using erps and spectral power with a six-dry-electrode eeg system in real flight conditions, Sensors 19 (6) (2019) 1324. doi:10.3390/s19061324.
- [23] E. Wascher, H. Heppner, S. Hoffmann, Towards the measurement of event-related EEG activity in real-life working environments, International Journal of Psychophysiology 91 (1) (2014) 3–9. doi: 10.1016/j.ijpsycho.2013.10.006.
- [24] C. Kranczioch, C. Zich, I. Schierholz, A. Sterr, Mobile eeg and its potential to promote the theory and application of imagery-based motor rehabilitation, International Journal of Psychophysiology 91 (1) (2014) 10–15.
- [25] N. Spychala, S. Debener, E. Bongartz, H. H. O. M"uller, J. D. Thorne, A. Philipsen, N. Braun, Exploring self-paced embodiable neurofeedback for post-stroke motor rehabilitation, Frontiers in Human Neuroscience 13 (2020) 461. doi:10.3389/fnhum.2019.00461.
- [26] F. Dehais, A. Lafont, R. Roy, S. Fairclough, A neuroergonomics approach to mental workload, engagement and human performance, Frontiers in neuroscience 14 (2020) 268.

- [27] A. F. Kramer, Physiological metrics of mental workload: A review of recent progress, in: Multiple-task performance, CRC Press, 1991, pp. 279–328.
- [28] M. Miller, J. Rietschel, C. McDonald, B. Hatfield, A novel approach to the physiological measurement of mental workload, International Journal of Psychophysiology 80 (1) (2011) 75–78. doi:10.1016/j. ijpsycho.2011.02.003.
- [29] G. F. Wilson, F. T. Eggemeier, Psychophysiological assessment of workload in multi-task environments, in: Multiple-task performance, CRC Press, 1991.
- [30] D. Looney, C. Park, P. Kidmose, M. Rank, M. Ungstrup, K. Rosenkranz, D. Mandic, An in-the-ear platform for recording electroencephalogram, in: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2011, pp. 6882–6885. doi:10.1109/IEMBS.2011.6091733.
- [31] H. Manabe, M. Fukumoto, T. Yagi, Conductive rubber electrodes for earphone-based eye gesture input interface, in: Proceedings of the 2013 International Symposium on Wearable Computers, 2013, pp. 33–40.
- [32] O. Valentin, G. Viallet, A. Delnavaz, G. Cretot-Richert, M. Ducharme, H. Monsarat-Chanon, J. Voix, Custom-fitted in-and around-the-ear sensors for unobtrusive and on-the-go eeg acquisitions: Development and validation, Sensors 21 (9) (2021) 2953.
- [33] D. Looney, P. Kidmose, C. Park, M. Ungstrup, M. Rank, K. Rosenkranz, D. Mandic, The in-the-ear recording concept: Usercentered and wearable brain monitoring, IEEE Pulse 3 (6) (2012) 32– 42. doi:10.1109/MPUL.2012.2216717.
- [34] J. H. Lee, S. M. Lee, H. J. Byeon, J. S. Hong, K. S. Park, S.-H. Lee, Cnt/pdms-based canal-typed ear electrodes for inconspicuous eeg recording, Journal of neural engineering 11 (4) (2014) 046014.
- [35] M. G. Bleichner, M. Lundbeck, M. Selisky, F. Minow, M. Jäger, R. Emkes, S. Debener, M. De Vos, Exploring miniaturized eeg electrodes for brain-computer interfaces. an eeg you do not see?, Physiological Reports 3 (4) (2015). doi:10.14814/phy2.12362.

- [36] J. J. Norton, D. S. Lee, J. W. Lee, W. Lee, O. Kwon, P. Won, S.-Y. Jung, H. Cheng, J.-W. Jeong, A. Akce, et al., Soft, curved electrode systems capable of integration on the auricle as a persistent brain-computer interface, Proceedings of the National Academy of Sciences 112 (13) (2015) 3920–3925.
- [37] N. Pham, T. Kim, F. M. Thayer, A. Nguyen, T. Vu, Earable–an ear-worn biosignal sensing platform for cognitive state monitoring and human-computer interaction, in: Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services, 2019, pp. 685–686.
- [38] M. G. Bleichner, S. Debener, Concealed, unobtrusive ear-centered eeg acquisition: Ceegrids for transparent eeg, Frontiers in Human Neuroscience 11 (2017). doi:10.3389/fnhum.2017.00163. URL https://www.frontiersin.org/article/10.3389/fnhum. 2017.00163
- [39] M. Pacharra, S. Debener, E. Wascher, Concealed around-the-ear eeg captures cognitive processing in a visual simon task, Frontiers in Human Neuroscience 11 (2017). doi:10.3389/fnhum.2017.00290.
- [40] B. Somon, Y. Giebeler, L. Darmet, F. Dehais, Benchmarking ceegrid and solid gel-based electrodes to classify inattentional deafness in a flight simulator, Frontiers in Neuroergonomics 2 (2022) 802486.
- [41] S. Getzmann, J. E. Reiser, M. Karthaus, G. Rudinger, E. Wascher, Measuring correlates of mental workload during simulated driving using ceegrid electrodes: a test-retest reliability analysis, Frontiers in neuroergonomics 2 (2021) 729197.
- [42] D. Hölle, M. G. Bleichner, Recording Brain Activity with Ear-Electroencephalography, Journal of Visualized Experiments (193) (2023) 1-12. doi:10.3791/64897. URL https://www.biorxiv.org/content/10.1101/2022. 10.07.511290v1{%}OAhttps://www.jove.com/t/64897/ recording-brain-activity-with-ear-electroencephalography
- [43] S. Debener, R. Emkes, M. De Vos, M. Bleichner, Unobtrusive ambulatory eeg using a smartphone and flexible printed electrodes

around the ear, Scientific Reports 5 (1) (2015) 16743. doi:10.1038/ srep16743.

- [44] G. Di Flumeri, P. Aric'o, G. Borghini, N. Sciaraffa, A. Di Florio, F. Babiloni, The dry revolution: Evaluation of three different eeg dry electrode types in terms of signal spectral features, mental states classification and usability, Sensors 19 (6) (2019) 1365. doi:10.3390/ s19061365.
- [45] T. M. Lau, J. T. Gwin, D. P. Ferris, How many electrodes are really needed for eeg-based mobile brain imaging?, Journal of Behavioral and Brain Science 02 (03) (2012) 387–393. doi:10.4236/jbbs.2012. 23044.
- [46] A. von Lühmann, S. R. Soekadar, B. Blankertz, K.-R. Müller, Headgear for mobile neurotechnology: Looking into alternatives for EEG and NIRS probes, GBCIC (2017).
- [47] C. He, Y.-Y. Chen, C.-R. Phang, C. Stevenson, I.-P. Chen, T.-P. Jung, L.-W. Ko, Diversity and suitability of the state-of-the-art wearable and wireless eeg systems review, IEEE Journal of Biomedical and Health Informatics (2023) 1–14doi:10.1109/JBHI.2023.3239053.
- [48] G. Niso, E. Romero, J. T. Moreau, A. Araujo, L. R. Krol, Wireless EEG: A survey of systems and studies, NeuroImage 269 (December 2022) (2023). doi:10.1016/j.neuroimage.2022.119774.
- [49] B. R. Dunn, D. A. Dunn, M. Languis, D. Andrews, The relation of erp components to complex memory processing, Brain and Cognition 36 (3) (1998) 355–376. doi:10.1006/brcg.1998.0998.
- [50] R. N"a"at"anen, Selective attention and evoked potentials in humans—a critical review, Biological Psychology 2 (4) (1975) 237–307.
   doi:10.1016/0301-0511(75)90038-1.
- [51] J. Polich, Updating p300: an integrative theory of p3a and p3b, Clinical Neurophysiology 118 (10) (2007) 2128–2148. doi:10.1016/j.clinph.
   [2007.04.019].
- [52] R. J. Barry, S. Kirkaikul, D. Hodder, Eeg alpha activity and the erp to target stimuli in an auditory oddball paradigm, International Journal of

Psychophysiology 39 (1) (2000) 39–50. doi:10.1016/S0167-8760(00) 00114-8.

- [53] D. d. M. Didon'e, M. V. Garcia, S. J. Oppitz, T. F. F. da Silva, S. N. dos Santos, R. S. Bruno, V. A. V. dos Santos Filha, P. L. C'oser, Auditory evoked potential p300 in adults: Reference values, Einstein (S ao Paulo) 14 (2016) 208–212. doi:10.1590/S1679-45082016A03586.
- [54] J. Polich, Frequency, intensity, and duration as determinants of p300 from auditory stimuli, Journal of Clinical Neurophysiology 6 (3) (1989) 277–286. doi:10.1097/00004691-198907000-00003.
- [55] H. d. O. Simões, A. C. F. Frizzo, S. Zanchetta, M. Hyppolito, A. C. M. B. Reis, Variables in p300 recording: Task type and electrode position, CoDAS 28 (2016) 355–361. doi:10.1590/2317-1782/20162015189.
- [56] A. Meiser, M. Bleichner, Ear-eeg compares well to cap-eeg in recording auditory erps: A quantification of signal loss, Journal of Neural Engineering (2022). doi:10.1088/1741-2552/ac5fcb.
- [57] B. Mirkovic, M. G. Bleichner, M. De Vos, S. Debener, Target speaker detection with concealed eeg around the ear, Frontiers in Neuroscience 10 (2016). doi:10.3389/fnins.2016.00349.
- [58] C. L. Baldwin, J. T. Coyne, Dissosiable aspects of mental workload: Examinations of the p300 erp component and performance assessment, Psychologia 48 (2) (2005) 102–119. doi:10.2117/psysoc.2005.102.
- [59] U. Ghani, N. Signal, I. K. Niazi, D. Taylor, Erp based measures of cognitive workload: A review, Neuroscience & Biobehavioral Reviews 118 (2020) 18–26. doi:10.1016/j.neubiorev.2020.07.020.
- [60] K. Natani, F. E. Gomer, Electrocortical activity and operator workload: A comparison of changes in the electroencephalogram and in event-related potentials, Tech. rep., MCDONNELL DOUGLAS ASTRONAUTICS CO ST LOUIS MO (1981). URL https://apps.dtic.mil/sti/citations/ADA237188
- [61] P. Ullsperger, G. Freude, U. Erdmann, Auditory probe sensitivity to mental workload changes - an event-related potential study,

International Journal of Psychophysiology 40 (3) (2001) 201–209. doi: 10.1016/S0167-8760(00)00188-4.

- [62] C. Wickens, Application of event-related brain potential research to problems in human factors, Basic Issues and Applications (1990) 301– 309.
- [63] C. D. Wickens, The structure of attentional resources, Attention and Performance VIII 8 (1980) 239–257.
- [64] C. D. Wickens, Multiple resources and performance prediction, Theoretical Issues in Ergonomics Science 3 (2) (2002) 159–177. doi: 10.1080/14639220210123806.
- [65] C. D. Wickens, Multiple resources and mental workload, Human Factors: The Journal of the Human Factors and Ergonomics Society 50 (3) (2008) 449–455. doi:10.1518/001872008X288394.
- [66] J. B. Isreal, G. L. Chesney, C. D. Wickens, E. Donchin, P300 and tracking difficulty: evidence for multiple resources in dual-task performance, Psychophysiology 17 (3) (1980) 259–273. doi:10.1111/ j.1469-8986.1980.tb00146.x.
- [67] E. J. Sirevaag, A. F. Kramer, M. G. H. Coles, E. Donchin, Resource reciprocity: An event-related brain potentials analysis, Acta Psychologica 70 (1) (1989) 77–97. doi:10.1016/0001-6918(89)
   [90061-9].
- [68] C. Wickens, A. Kramer, L. Vanasse, E. Donchin, The reciprocity of primary and secondary task resources: Evidence from the p300 component of the erp, Science 221 (1983) 1080–1082.
- [69] C. Wickens, A. Kramer, L. Vanasse, E. Donchin, Performance of concurrent tasks: A psychophysiological analysis of the reciprocity of information-processing resources, Science 221 (4615) (1983) 1080–1082. doi:10.1126/science.6879207.
- [70] J. R. Comstock, R. Arnegard, et al., The multi-attribute task battery for human operator workload and strategic behavior research (1992).

- [71] I. Solís-Marcos, K. Kircher, Event-related potentials as indices of mental workload while using an in-vehicle information system, Cognition, Technology Work 21 (1) (2019) 55–67. doi:10.1007/ s10111-018-0485-z.
- [72] M. Causse, E. Fabre, L. Giraudet, M. Gonzalez, V. Peysakhovich, Eeg/erp as a measure of mental workload in a simple piloting task, Procedia Manufacturing 3 (2015) 5230-5236. doi:10.1016/j.promfg. 2015.07.594.
- [73] B. Z. Allison, J. Polich, Workload assessment of computer gaming using a single-stimulus event-related potential paradigm, Biological Psychology 77 (3) (2008) 277–283. doi:10.1016/j.biopsycho.2007.
   [10.014].
- [74] A. F. Kramer, E. J. Sirevaag, R. Braune, A psychophysiological assessment of operator workload during simulated flight missions, Human Factors: The Journal of the Human Factors and Ergonomics Society 29 (2) (1987) 145–160.
- [75] Y. Santiago-Espada, R. Myer, K. Latorella, J. Comstock Jr, The multiattribute task battery ii (matb-ii) software for human performance and workload research: A user's guide (2011).
- [76] R. N. Roy, S. Bonnet, S. Charbonnier, A. Campagne, Efficient workload classification based on ignored auditory probes: A proof of concept, Frontiers in Human Neuroscience 10 (2016) 519. doi:10.3389/fnhum.
   [2016.00519]
- [77] L. Giraudet, J.-P. Imbert, M. Bérenger, S. Tremblay, M. Causse, The neuroergonomic evaluation of human machine interface design in air traffic control using behavioral and eeg/erp measures, Behavioural Brain Research 294 (2015) 246-253. doi:10.1016/j.bbr.2015.07.
  [041]
- [78] Y. Takeda, K. Inoue, M. Kimura, T. Sato, C. Nagai, Electrophysiological assessment of driving pleasure and difficulty using a task-irrelevant probe technique, Biological psychology 120 (2016) 137-141. doi:10.1016/j.biopsycho.2016.09.009.

- [79] F. Dehais, I. Rida, R. N. Roy, J. Iversen, T. Mullen, D. Callan, A pbci to predict attentional error before it happens in real flight conditions, in: 2019 ieee international conference on systems, man and cybernetics (smc), IEEE, 2019, pp. 4155–4160.
- [80] M. Kleiner, D. Brainard, D. Pelli, A. Ingling, R. Murray, C. Broussard, What's new in psychoolbox-3, Perception 36 (14) (2007) 1–16.
- [81] D. Str"uber, J. Polich, P300 and slow wave from oddball and singlestimulus visual tasks: Inter-stimulus interval effects, International Journal of Psychophysiology 45 (3) (2002) 187–196. doi:10.1016/ S0167-8760(02)00071-5.
- [82] M. A. Boudewyn, S. J. Luck, J. L. Farrens, E. S. Kappenman, How many trials does it take to get a significant erp effect? it depends, Psychophysiology 55 (6) (2018) e13049. doi:10.1111/psyp.13049.
- [83] S. W. Smith, The Scientist and Engineer's Guide to Digital Signal Processing, 1st Edition, California Technical Pub, 1997.
- [84] H. H. Jasper, The ten-twenty electrode system of the intenational federation, Electroenceph Clin Neutro Physiol 10 (1958) 367–380.
- [85] C. Kothe, et al., Lab streaming layer (lsl) (2014).
- [86] A. Delorme, S. Makeig, Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis, Journal of Neuroscience Methods 134 (1) (2004) 9–21. doi:10.1016/ j.jneumeth.2003.10.009.
- [87] M. Congedo, The analysis of event-related potentials, in: Computational EEG Analysis, Springer Singapore, 2018, pp. 55– 82. doi:10.1007/978-981-13-0908-3\_4.
- [88] E. S. Kappenman, S. J. Luck, The effects of electrode impedance on data quality and statistical significance in erp recordings, Psychophysiology 47 (5) (2011) 888–904.
- [89] B. Rivet, A. Souloumiac, V. Attina, G. Gibert, xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain-Computer Interface, IEEE Transactions on Biomedical Engineering 56 (8) (2009) 2035–2043. doi:10.1109/TBME.2009.2012869.

- [90] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, Multiclass braincomputer interface classification by Riemannian geometry, IEEE Transactions on Biomedical Engineering 59 (4) (2012) 920–928. doi: 10.1109/TBME.2011.2172210.
- [91] A. Appriou, A. Cichocki, F. Lotte, Modern machine-learning algorithms: for classifying cognitive and affective states from electroencephalography signals, IEEE Systems, Man, and Cybernetics Magazine 6 (3) (2020) 29–38.
- [92] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, F. Yger, A review of classification algorithms for eeg-based brain-computer interfaces: a 10 year update, Journal of neural engineering 15 (3) (2018) 031005.
- [93] R. N. Roy, M. F. Hinss, L. Darmet, S. Ladouce, E. S. Jahanpour, B. Somon, X. Xu, N. Drougard, F. Dehais, F. Lotte, Retrospective on the first passive brain-computer interface competition on cross-session workload estimation, Frontiers in Neuroergonomics 3 (2022) 838342.
- [94] A. Barachant, M. Congedo, A Plug&Play P300 BCI Using Information Geometry (2014) 1–9arXiv:1409.0107.
   URL http://arxiv.org/abs/1409.0107
- [95] E. Combrisson, K. Jerbi, Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy, Journal of Neuroscience Methods 250 (2015) 126-136. doi:10.1016/j.jneumeth.2015.01.010. URL http://dx.doi.org/10.1016/j.jneumeth.2015.01.010
- [96] A. Meiser, F. Tadel, S. Debener, M. G. Bleichner, The Sensitivity of Ear-EEG: Evaluating the Source-Sensor Relationship Using Forward Modeling, Brain Topography 33 (6) (2020) 665–676. doi:10.1007/ s10548-020-00793-2.
   URL https://doi.org/10.1007/s10548-020-00793-2
- [97] M. G. Bleichner, B. Mirkovic, S. Debener, Identifying auditory attention with ear-EEG: cEEGrid versus high-density cap-EEG comparison, Journal of Neural Engineering 13 (6) (2016) 066004. doi:10.1088/1741-2560/13/6/066004.

URL http://stacks.iop.org/1741-2552/13/i=6/a=066004?key= crossref.3cead442258228d397ef0718763deb8d

- [98] Y. Ke, T. Jiang, S. Liu, Y. Cao, X. Jiao, J. Jiang, D. Ming, Crosstask consistency of electroencephalography-based mental workload indicators: Comparisons between power spectral density and taskirrelevant auditory event-related potentials, Frontiers in Neuroscience 15 (2021) 703139. doi:10.3389/fnins.2021.703139.
- [99] M. J. Sugg, J. Polich, P300 from auditory stimuli: Intensity and frequency effects, Biological psychology 41 (3) (1995) 255-269. doi: 10.1016/0301-0511(95)05136-8.
- [100] M. G. Bleichner, B. Mirkovic, S. Debener, Identifying auditory attention with ear-eeg: Ceegrid versus high-density cap-eeg comparison, Journal of Neural Engineering 13 (6) (2016) 066004. doi:10.1088/1741-2560/13/6/066004.
- [101] C. Gonsalvez, J. Polich, P300 amplitude is determined by targetto-target interval., Psychophysiology 39 (3) (2002) 388–396. doi: 10.1017/S0048577201393137.
- [102] A. Sterr, J. K. Ebajemito, K. B. Mikkelsen, M. A. Bonmati-Carrion, N. Santhi, C. della Monica, L. Grainger, G. Atzori, V. Revell, S. Debener, D.-J. Dijk, M. DeVos, Sleep eeg derived from behindthe-ear electrodes (ceegrid) compared to standard polysomnography: A proof of concept study, Frontiers in Human Neuroscience 12 (2018). doi:10.3389/fnhum.2018.00452.
- [103] F. Dehais, S. Ladouce, J. J. T. Tresols, L. Darmet, D. Callan, Mental workload classification during simulated flight operations based on cardiac and neural dynamics recorded using the muse 2 low-cost system, in: 14th International Conference on Applied Human Factors and Ergonomics (AHFE 2023), 2023.
- [104] N. S. J. Jacobsen, S. Blum, K. Witt, S. Debener, A walk in the park? characterizing gait-related artifacts in mobile eeg recordings, European Journal of Neuroscience 54 (12) (2021) 8421–8440. doi:10.1111/ejn. 14965.

- [105] K. A. Brookhuis, G. Mulder, L. J. Mulder, A. B. Gloerich, H. J. Van Dellen, J. J. Van Der Meere, H. Ellermann, Late positive components and stimulus evaluation time, Biological Psychology (1981). doi:10. 1016/0301-0511(81)90030-2.
- [106] S. SUTTON, D. S. RUCHKIN, The Late Positive Complex: Advances and New Problems, Annals of the New York Academy of Sciences (1984). doi:10.1111/j.1749-6632.1984.tb23520.x.
- [107] H. Yang, G. Laforge, B. Stojanoski, E. S. Nichols, K. McRae, S. Köhler, Late positive complex in event-related potentials tracks memory signals when they are decision relevant, Scientific Reports (2019). doi:10. 1038/s41598-019-45880-y.
- [108] A.-M. Brouwer, M. A. Hogervorst, J. B. Van Erp, T. Heffelaar, P. H. Zimmerman, R. Oostenveld, Estimating workload using eeg spectral power and erps in the n-back task, Journal of neural engineering 9 (4) (2012) 045008.