ACTA IMEKO
ISSN: 2221-870X
June 2024, Volume 13, Number 2, 1 - 9

EEG measurements-based study for evaluating acoustic human perception: A pilot study

Silvia Angela Mansi[1], Reza Jamali[2], Gianmarco Battista[3], Marco Arnesano[1], Milena Martarelli[2], Paolo Chiariotti[4], Paolo Castellini[2]

[1] *Università Telematica eCampus, 22060 Novedrate (CO), Italy*
[2] *Department of Industrial Engineering and Mathematical Sciences, Università Politecnica delle Marche , 60131 Ancona, Italy*
[3] *Department of Engineering and Architecture, University of Parma, 43121 Parma, Italy*
[4] *Department of Mechanical Engineering, Politecnico di Milano, Milano, 20156, Italy*

ABSTRACT
Sound quality analysis and sound design are well-known human-centered strategies to evaluate the subjective perception of noise and to design machines and environments with pleasant and comfortable acoustic signatures. The subjective acoustic perception is conventionally measured by means of sound quality metrics determined through a correlation process with jury test results. The exploitation of electroencephalogram (EEG) measurements during the jury test for the registration of the brain activity in response to the acoustic stimuli presented to the jurors can allow us to estimate the jurors' perception directly from their physiological response. This study presents results from the application of an EEG wearable device to investigate changes in the EEG frequency domain at different acoustic stimuli. Forty-three participants were recruited, and the EEG signals were recorded using the wearable sensor. The analysis of power spectral densities (PSDs) was performed to investigate features correlated to acoustic sensation induced by audio stimuli. Statistically significant differences were found between the three audio stimuli. The results bring to the conclusion that wearable sensors could be used for EEG acquisition applied to acoustic perception evaluation.

Section: RESEARCH PAPER

Keywords: Sound quality; acoustic perception; electroencephalogram (EEG); wearable device

Citation: S. A. Mansi, R. Jamali, G. Battista, M. Arnesano, M. Martarelli, P. Chiariotti, P. Castellini, EEG measurements-based study for evaluating acoustic human perception: A pilot study, Acta IMEKO, vol. 13 (2024) no. 2, pp. 1-9. DOI: 10.21014/actaimeko.v13i2.1698

Section Editor: Francesco Lamonaca, University of Calabria, Italy

Received October 20, 2023; **In final form** April 3, 2024; **Published** June 2024

Funding: This project has received funding from the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Sklodowska-Curie Grant Agreement nº 858018.

Corresponding author: Silvia Angela Mansi, e-mail: silviaangela.mansi@uniecampus.it

1. INTRODUCTION

The assessment of sound quality plays a critical role across numerous contexts such as audio technology, product development, consumer electronics, telecommunications, automotive, entertainment, and the buildings sector. The acoustic indoor environmental quality (IEQ) of a building refers to the quality of the acoustic environment inside the building. It encompasses various factors related to sound and noise that can affect the comfort, well-being, and productivity of occupants [1], [2]. Acoustic IEQ is an important aspect of building design, particularly in spaces where speech intelligibility, privacy, and noise reduction are important. Positive acoustic conditions contribute to a more pleasant and functional space, while poor acoustic conditions can lead to discomfort, stress, and reduced performance. While objective sound quality metrics such as loudness, sharpness, roughness, and others serve as valuable insights into the technical facets of sound characterization, subjective evaluations by human listeners remain essential in capturing the perceptual and mental dimensions of sound quality [3], [4].

Jury testing (JT), also known as subjective sound quality evaluation, is one of the most diffuse approaches for assessing the perceived sound quality. The practice of JT consists of playing the set of sounds to be evaluated to a group of pre-instructed listeners, or better "jurors", who rank the sound by answering to pre-defined questions about the sounds, according to their own perspectives. However, the JT poses several challenges. On the one hand, subjective preferences are varied by individual preferences, cognitive biases, cultural background, environmental effects, and human perception. Furthermore, because the single juror might have different reaction to the same sound exposure, not all responses might be trustworthy. On the other hand, the test setup needs to be carefully planned in order to collect the required data regarding the sounds.

Recognizing these challenges, researchers and practitioners have devoted significant efforts to improve JT outcome in terms of reliability, accuracy, and repeatability of subjective evaluations while accounting for the complex nature of sound quality.

Previous research in this field has explored various paths for improving JT sound quality metrics. In [5], [6] it has been demonstrated that objective sound quality metrics have high correlation with the evaluation of discomfort given by participants to jury tests applied to the annoyance estimation of interior noise in helicopters and car cabins. Studies have investigated the integration of subjective and objective measures, combining perceptual evaluations with objective metrics derived from psychoacoustic models, acoustic analysis, and machine learning algorithms [7], [8]. By leveraging both human perception and technical measurements, researchers aim to establish a more holistic framework for sound quality assessment. Additionally, advancements have been made in the development of standardized test protocols and methodologies. These standardized protocols provide a consistent and controlled environment for subjective evaluations, facilitating comparisons between different products [9], [10]. Another helpful tool is the virtual reality. Robotham, et al. focused on sound quality assessment in Virtual Reality (VR) applications; this study compares different evaluation methods, including traditional jury testing, paired comparison tests, and rating scales. The authors emphasize the need for adapting sound quality evaluation approaches to the unique characteristics and challenges of VR environments [11].

Furthermore, advancements in statistical analysis techniques contributed to enhance the reliability and consistency of the results obtained from JT. Zhang, et al. demonstrated the importance of sample size determination and statistical modelling in reducing variability and enhancing the accuracy of JT outcomes [12].

The integration of wearable devices and physiological monitoring tools has also emerged as a promising approach for improving JT sound quality metrics. Wearable devices can capture real-time biometric data, such as heart rate, skin conductance, and brainwave activity, which can provide insights into listeners' physiological responses to audio stimuli. By correlating physiological measures with subjective ratings, researchers can gain a deeper understanding of the underlying mechanisms and physiological bases of sound quality perception [13].

Electroencephalography (EEG) refers to the process or technique of recording and analyzing the electrical activity of the brain through the use of an electroencephalogram [14]. Brainwaves are detected using electrodes placed on the scalp according to the 10-20 system [15] and they are classified according to their frequency content. Each brainwave class corresponds to a particular state of mind. In general, delta waves (0.1 Hz – 4 Hz) are associated with deep sleep, theta waves (4 Hz – 7.5 Hz) are related to consciousness sleep towards drowsiness, alpha waves (7.5 Hz – 12 Hz) are the prominent rhythm in relaxing and passive attention activities, beta (12 Hz – 30 Hz) is associated with active thinking and gamma waves (30 Hz – 45 Hz) are prominent during high mental activities. EEG has been largely adopted in the field of acoustic perception, given the possibility to monitor the human physiological response changes in real-time. Generally, the brain's electrical activity changes in response to the process of perception and cognition of environmental stimuli.

Earlier research results of human brain physiology have shown that brain activities occur simultaneously with external stimulation (visual and acoustic stimuli) [16], [17]. In the acoustic field, for example, Trimmel [18] noticed that noise exposure can change the activity of the central nervous system, and the degree of impact is related to sound types and Sound Pressure Level (SPL). Other studies found a correlation between EEG power spectral densities (PSDs) and individual acoustic perception. some of these investigated the relationship between emotional state and EEG response induced by acoustic stimuli.

Schmidt and Trainor [19] presented happy and sad musical excerpts and found decreased alpha power at left frontal electrodes during happy music, whereas sad music was associated with a more pronounced alpha power decrease at the right frontal leads. Kabuto et al. [20] analysed the PSDs changes induced by pleasant music, showing that alpha-power is high under the psychosomatic state with "pleasant and calm" feelings, and that its increase is related to relaxation.

J. L. Walker [21] reported a correlation between self-reports of paying little attention to the music and high theta- and high delta-wave production.

Nevertheless, none of the above-mentioned studies used wearable device to investigate brain responses to audio stimuli. The recent development of non-invasive wearable sensing technologies highlights the potential of monitoring specific human physiological signals throughout daily life, having more reliable signals that can be correlated to human preference and perception, rather than potentially biased surveys-only data sources. Wearable EEG devices have found various applications across different fields due to their portability and non-invasive nature. They have been used in developing Brain-Computer Interfaces (BCIs), allowing individuals to control electronic devices or computers directly with their brain signals. This technology has applications in assistive technologies for people with disabilities and may enable hands-free control [22]. EEG wearables have been demonstrated to be capable of assessing cognitive performance, attention levels, and mental workload in real-time [23]. Studies revealed their applicability also in monitoring sleep stages, and identifying sleep-related disorders [24]. Another field of application includes monitoring stress levels and emotional responses, useful in various settings such as workplace stress management programs and research on emotional well-being [25]. In general, EEG portable devices offer several advantages: they are low cost, simple to be used, their comfortable design allows to reduce the time of application, and more importantly, they considerably attenuate the obstructiveness of measurements, making the experimental time not unpleasant for the participants. However, they are strongly prone to collect environmental noise and artifacts due to subject movements (e.g., eyeblink, muscular artifacts), which means that acquired data need to be processed before they became reliable for acoustic perception measurements.

The aim of the presented study is to evaluate the applicability of wearable sensor in the field of acoustic perception estimation when targeting sounds and noises in automotive applications. The experimental protocol, the signal processing procedure, and the statistical analysis are illustrated together with results from the measurement campaign performed in a controlled environment. Results demonstrate the feasibility of the proposed approach, that could be used as a pilot investigation towards JT application.

The paper is organised in five sections. Section 2 presents the measurement equipment used, the JT protocol and the audio signals submitted to the jurors. Section 3 describes the data processing methodology and the statistical analysis while Section 4 reports the results achieved. In Section 5 and 6 the discussion of results and the conclusions are drawn.

2. MATERIALS AND METHODS

2.1. Measurement equipment

2.1.1. EEG device

A brain sensing headband, Muse 2 (Interexon Inc.) [26], was used for recording EEG signals. Researchers provided evidence that Muse is an effective portable tool for continuous recording EEG data [27], [28], appliable outside its designed functionality (meditation and training device). The EEG signals were obtained from 4 electrodes. The two electrodes are on the forehead (left and right of the reference: AF7, AF8, silver made) and one electrode above each ear (TP9 and TP10, conductive silicone - rubber). Three reference electrodes (FPz - CMS/DRL) are placed in the middle between the two input electrodes on the forehead, Figure 1. EEG data were recorded using the Muse application [29] paired with a smartphone via Bluetooth Low Energy (BLE) at 256 Hz sampling frequency.

2.1.2. Earphones

Over-ear headphones are not compatible with the usage of Muse 2 EEG headband because they would cause bad adhesion of the electrodes. Therefore, a compact earphone, the Sony MDR-E9LP, were utilized (Figure 2).

2.2. Experimental procedure

To estimate the effective sample size of subjects for the experimental campaign, the G*Power3 statistical test was performed [30]. Table 1 shows the priori analysis setting.

The obtained sample size was 36 and 38 for the parametric and non-parametric tests, respectively.

The experimental campaign accounted for 43 subjects, 21 females and 22 males. The listeners were between 19 to 61 years old; but the 80 percent of the sample was in the range of 28 ± 7 years old.

The experiments took place in the acoustic laboratory of "Università Politecnica delle Marche", Italy, in September 2022.

More specifically, the subject, once entered the test room, was instructed about the test (where to sit, how to wear physiological monitoring devices, and when to answer the questionnaire in Google Form accessible via QR code). All subjects signed the informed consent. Each test session lasted 6 min; the subjects were exposed to three different audio signals; each audio signal was preceded by a minute of silence. After each sounds exposure the subjects filled out a questionnaire to indicate their acoustic perception about the audio signals they listened to, Figure 3. EEG acquisition was done continuously. Subjects were instructed to stay as quiet as possible during the 6 minutes of physiological data collection, just sitting and relaxing to reduce measurement artifacts due to body movements, Figure 2. Finally, it is worth noting that recruited subjects were instructed not to smoke, not to perform any intense physical effort, not to eat or drink coffee or any other exciting beverage since at least 1 hour before their scheduled tests. All the tests took place in the morning from 10 AM to 1 PM and in the afternoon, in the time slot from 2:30 PM to 6:30 PM. The questionnaire the subjects were asked to fill aimed at describing the subjects' acoustic perception of each sound in terms of annoying/pleasant,

Table 1. CAPTION.

Parameters	Values
Required power level (1 - β)	0.95
Prespecified significant level α	0.05
Effect size f	0.25
Number of measurements	3
Number of groups	1
Statistical test	repeated measures - within factor ANOVA and test Wilcoxon signed-rank test

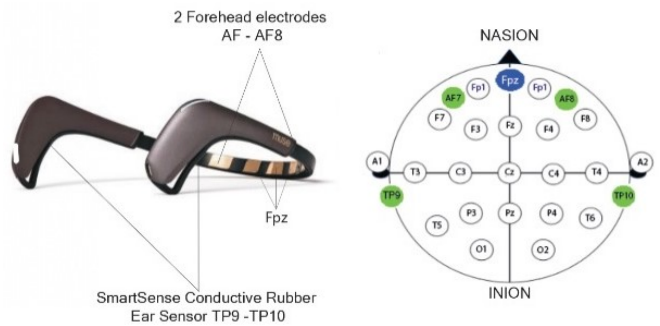

Figure 1. Left: MUSE 2 headband sensors overview. Right: Top-down view of the EEG electrode positions on the subject's head.

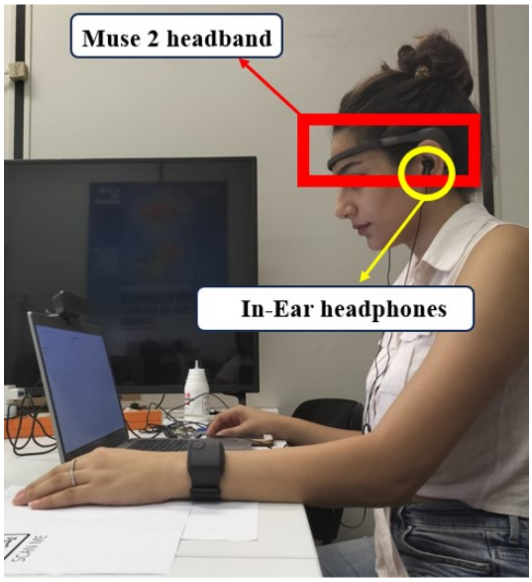

Figure 2. Measurements of EEG signals with Muse 2 headband during test session.

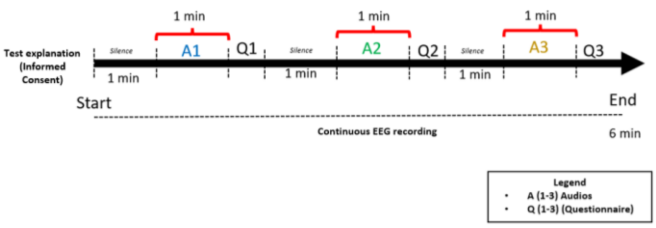

Figure 3. JT protocol submitted to each juror.

relaxing/stressful, quiet/loud sensations on a 5-point Likert scale ranging from 1, meaning annoying/relaxing/quiet, up to 5, meaning pleasant/stressful/loud. In this evaluation scale, the sense of neutrality was set at level 3 [31].

2.3. Auditory stimuli characteristics

In this subsection, we provide a detailed description of the acoustic stimuli utilized in our experiment. The acoustic description begins with a fundamental audio spectral metrics, which serve as key introduction indicators in our analysis.

2.3.1. Definition of audio spectral metrics

Several audio spectral metrics, commonly adopted in the field of audio analysis, are used here to provide an objective description of the spectral characteristics of sound signals, making them valuable tools for audio signal analysis and processing [32]. For time varying signals the metrics are computed from the spectrogram. The spectrograms are computed for time chunks lasting 0.4 s and overlapped of 90%. A Hanning time window is applied to reduce the leakage. The resulting spectrograms are depicted in Figure 4 for the frequency range from 40 Hz to 10 kHz, which corresponds to the range used for audio metrics computation. For each time chunk, the audio power spectrum (PS) is denoted by $P(b)$, where b represents the discrete frequency bins in the range of analysis.

Spectral Centroid

The Spectral Centroid (CNT) represents the center of mass of the power spectrum on the frequency axis. It indicates where the bulk of the acoustic energy is concentrated in the frequency domain. A higher CNT value typically corresponds to a brighter sound, while a lower value indicates a darker or bass-heavy sound. The CNT indicator is calculated as follows:

$$CNT = \mu_b = \frac{\sum_f b \cdot P(b)}{\sum_f P(b)} .$$

Spectral Spread

The Spectral Spread (SPR) is the standard deviation of the PS around the spectral CNT. This represents an indication of the bandwidth of the spectrum.

$$SPR = \sigma_b = \sqrt{\frac{\sum_f \left(b - \mu_f\right)^2 P(b)}{\sum_f P(b)}} .$$

Spectral Skewness

The Spectral Skewness (SKW) quantifies the asymmetry of the spectral distribution around the CNT. It measures whether the distribution is skewed to the left or right with respect to the centroid. Positive values indicate a dominance of low frequencies, while negative values a dominance of high frequencies. It is calculated as:

$$SKW = \frac{\sum_f (b - \mu_b)^3 \cdot P(b)}{(\sigma_b)^3 \sum_b P(b)} .$$

Spectral Kurtosis

The Spectral Kurtosis (KUR) is used to assess the "peakiness" of the spectral distribution. It actually measures the flatness, or non-Gaussianity, of the spectrum around its CNT and is calculated as follows:

$$KUR = \frac{\sum_b (b - \mu_b)^4 P(b)}{(\sigma_b)^4 \sum_b P(b)} .$$

Spectral Entropy

The Spectral Entropy (ENT) measures the degree of disorder in the spectral distribution of a signal. The more frequencies contribute to the full spectrum, also considering their amplitude, the higher the ENT. For example, a pure tone has low Spectral ENT, while white noise PS has higher ENT. However, pink noise PS has less ENT of white noise. It is calculated using Shannon's ENT formula:

$$ENT = -\frac{1}{\log N_b} \sum_b P(b) \cdot \log\left(P(b)\right) .$$

Spectral Flux

The Spectral Flux (FLX) is a measure of the spectrum changes over time. It is computed using two consecutive spectra in the spectrogram, denoted respectively by $P(b,t)$ and $P(b,t-1)$:

$$FLX(t) = \left(\sum_f \left(P(b,t) - P(b,t-1) \right)^2 \right)^{\frac{1}{2}} .$$

2.3.2. Spectral metrics of the auditory stimuli provided to subjects

In this study, three distinct sounds were used to stimulate emotions in listeners. These sounds were chosen to be significantly different from each other. Figure 4 depicts the spectrograms corresponding to the three audio stimuli, while Table 2 reports the spectral metrics associated with each of the three sounds. In the table, the average of each metric together

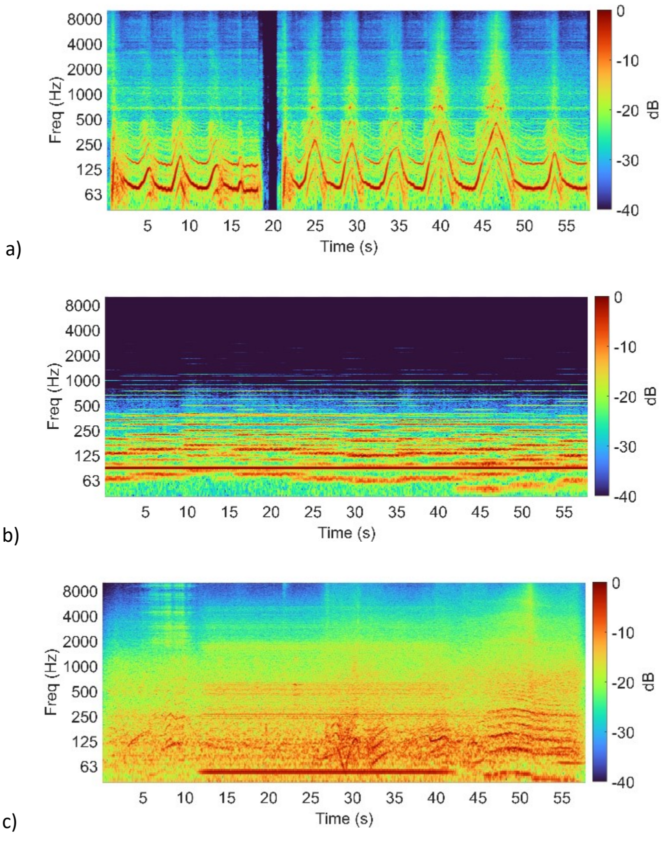

Figure 4. Spectrograms illustrating the spectral content of three sound stimuli over time. (a) engine noise, (b) soothing music, and (c) road noise.

Table 2. Spectral metrics of the sounds. Time averages and extreme values of the metrics are indicated.

Audio	Description	CNT (Hz)	SPR (Hz)
A1	Engine noise	210 [88, 1106]	330 [134, 2100]
A2	Music	126 [100, 168]	70 [38, 110]
A3	Road noise	402 [239, 1031]	686 [363, 2100]
		SKW	**KUR**
A1	Engine noise	22.5 [2.8, 36.1]	775 [12, 1633]
A2	Music	2.2 [0.8, 4.5]	10 [3, 32]
A3	Road noise	5.2 [2.9, 7.7]	46 [12, 97]
		ENT	**FLX**
A1	Engine noise	0.33 [0.15, 0.73]	1.95e-3 [8.06e-5, 7.12e-3]
A2	Music	0.27 [0.19, 0.33]	6.93e-4 [1.42e-4, 1.95e-3]
A3	Road noise	0.57 [0.47, 0.71]	5.84e-4 [4.01e-5, 2.30e-3]

with the ranges assumed by the metric itself is given. The average is calculated over time for each metric and it represents the barycentre of the values assumed by the metric in the whole audio duration. The comparison of these objective values among the three sounds makes it possible to draw an acoustic profile for each of them and to highlight the differences.

Audio 1 (A1) is the sound an internal combustion engine of a sports car, which should elicit the idea of power, mechanical precision, and intensity. This sound has a CNT of 210 Hz and a SPR of 330 Hz. The high KUR value (775) confirms the prevalence of tones in the sound, while the positive SKW (22.5) indicates that lower tones are predominant. The ENT is moderate (0.33), while the FLX is the highest compared to the other two sounds.

Audio 2 (A2) consists of soothing music, which evokes feelings of tranquillity, relaxation, and harmony. This sound exhibits a CNT value of 126 Hz and a SPR of 70 Hz, indicating a lower frequency concentration. The SKW is almost neutral (2.2) and the KUR (10) is the lowest among the three sounds, therefore, there are not really dominating peaks in the spectrum.
Low ENT and FLX values, respectively 0.27 and 0.00069, suggest a relatively simple and stationary distribution of frequencies.

Audio 3 (A3) represents road traffic noise, which connotes a noisy and stressful environment. This sound is characterized by the highest CNT and SPT values, respectively of 538 Hz and 686 Hz., suggesting a relevant content in a broader frequency range. The SKW (5.2) and the KUR (46) are moderate compared to the Sound No. 1. This sound has the highest ENT (0.57), thus meaning that many frequencies contribute to the sound, on the average. The FLX is the lowest one (0.00058) due to the presence of constant background noise.

3. DATA ANALYSIS

3.1. EEG pre-processing and features extraction

Muse 2 headband sensor has an on-board Digital Signal Processing (DSP) module which pre-processes raw data using a noise filter (0.1 – 45 bandpass and notch at 60 Hz), eye blink and jaw clench artifact removal and Frequency Domain analysis (window of 256 samples, with a step of 22) giving the access to the PSD of five frequency bands Delta (0.5 – 4 Hz), Theta (4 – 8 Hz), Alpha (8 – 12 Hz), Beta (12 – 30 Hz) and Gamma (> 30 Hz) [33].

For each one of the five frequency bands, across the 4 input channels (TP9, AF7, AF8, and TP10), the following features were computed:

- *Relative power*:
  $\sum_0^i B \, / \, \sum_0^i T$ (1)
  where B is the power of signal in a specific frequency band T is the total power
- *Frontal Asymmetry (FA)* and *Temporal Asymmetry (TA)* using the formulas (2) and (3):
  FA: B_{AF8} - B_{AF7} (2)
  TA: B_{TP10} - B_{TP9} (3)
- *Band ratio* in each channel:
  alpha/beta . (4)

From now on, the capitalized terms Delta, Theta, Alpha, Beta, and Gamma will be used to denote the absolute power of the signal in each frequency band, along with the reference electrode. The term "Relative" accompanied by the corresponding band name will be employed to refer to relative power within the respective bands.

3.2. Statistical analysis

3.2.1. Statistics on EEG data

The relative frequency of individual responses about personal perception of sounds were computed. A Python custom code was developed to perform statistical analysis to establish the correlation level between individual acoustic perception and EEG features changes under the three audios stimuli. The outlier's identification and removal were performed using the Z-score method (cut-off value equal 2). According to the Guide to the expression of uncertainty in measurement (GUM) [34], the mean ($\bar{f}$) and the standard uncertainty of the mean (u_f) of each EEG features were calculated. The Shapiro test [35], and Bartlett test [36] were performed to verify the assumption of normality and the homogeneity of variance, respectively. Across the entire group of features only the (delta_AF8) revealed a non-gaussian distribution. Two types of statistical methods were used to compare means of EEG features across variables based on repeated observations in different conditions, the repeated measures ANOVA model [37] in case of satisfaction of both normality and homogeneity of data, and the Kruskal-Wallies non parametric test [38] on contrary. Finally, the EEG features ability to discriminate between A1-A2, A1-A3, and A2-A3 was investigated with a post-hoc analysis using Dwass-Steel-Critchlow-Fligner pairwise comparison test [39].

3.2.2. Correlation between EEG features and audio metrics

To identify a possible relationship between EEG features and audio metrics, the Pearson's correlation coefficient is calculated using MATLAB "*corrcoef*" function. Only the statistically significant EEG features, obtained from the post-hoc analysis were used to compute the correlation with audio metrics. The time-averaged audio metrics of Table 2 are used for all subjects.

4. RESULTS AND DISCUSSION

4.1. Questionnaire for the evaluation of the subjective perception

The statistical analysis of the questionnaire results is summarized in Figure 5. The results of relative frequency distribution of subjective acoustic perception revealed that A1 was perceived more annoying than pleasant (39% of subjects'

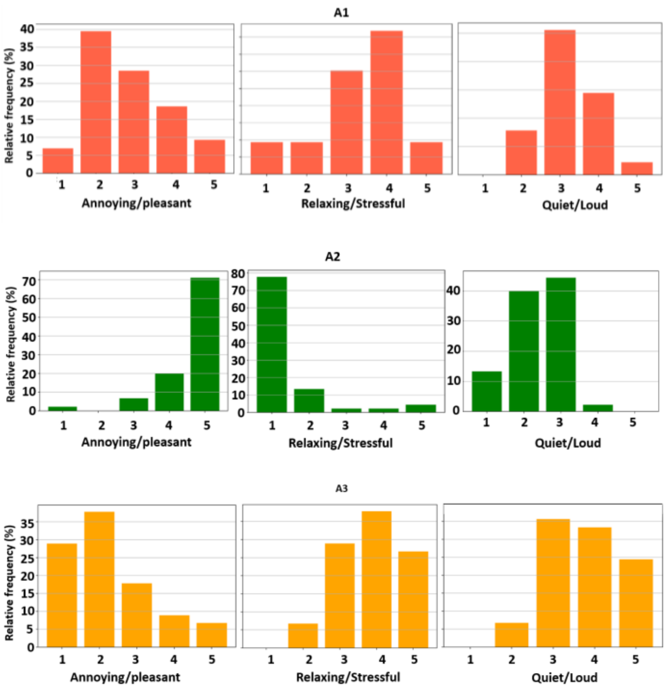

Figure 5. Frequency distribution in percentage (%) of individual acoustic perception, and illustrating the average ratings for each dimension.

score equal 2), more stressful than relaxing (43% of responses for score 4), and more than 50% of subjects perceived A1 neutral in terms of quiet/loud perception. For the A2 sound, more than 70% of subjects gave a score equal to 5 (pleasant); this sound was perceived mostly relaxing (78% for a score equal to 1) and quiet/loud almost neutral (score 2 – 3). The results related to A3 showed that the perception of 37%, and 28% (score 2 and 1 respectively) of subjects was annoying, it was also perceived stressful, only the 6% of subjects gave a score minor than 3 (relaxing), in terms of quiet/loud most of the responses were in the range from 3 to 5 (loud).

4.2. Correlation between EEG features and subjective perception

In general, EEG measurements showed a correlation with acoustic perception in terms of increase or decrease of power of brain waves. Higher value of Delta in right frontal and temporal electrodes (TP10, AF8), and Theta TP10 were found under A1 exposure compared to the others, as reported in Table 3. As reported in see Table 4 and Figure 6, the post-hoc analysis revealed a statistical difference between both A1-A2, and A1-A3 for Delta TP10, Delta AF8, and Theta TP10 (significant values were accepted for *p-value* < 0.05 - significant level *alpha* = 0,05-). Relative Alpha TP10 showed higher values in A2 condition. An increase of TA Alpha was found under A3 audio exposure. TA showed a statistical difference in contrasts A2-A3, and A1-A3. Slow-waves delta and theta in the right temporal site showed a correlation with score on the "annoying" scale in A1 and A3. Delta waves are the lowest recorded brain waves in human being. They are associated with the deepest level of relaxation and restorative, healing sleep. Delta activity has been assessed to be also prominent in some emotional experience. J. L. Walker [40] found a correlation between high-delta and high-theta production and "unpleasant" music. The results of study showed also a correlation between high delta activity and "pay little attention", supporting the hypothesis that predominant delta activity is modulated by annoying sound perception. The results about theta increase induced by annoying sound is in line with

Table 3. Repeated measured ANOVA - Kruskal-Walli's test results. Significance level *alpha* = 0.05.

EEG features	$(\bar{f} \pm u_f)$ A1	$(\bar{f} \pm u_f)$ A2	$(\bar{f} \pm u_f)$ A3	H-statistic	p-value
Delta TP9	0.6 ± 0.07	0.5 ± 0.08	0.5 ± 0.06	7.5	**0.001**
Delta AF7	0.6 ± 0.06	0.5 ± 0.09	0.5 ± 0.08	5.1	**0.009**
Delta AF8	0.7 ± 0.06	0.6 ± 0.08	0.6 ± 0.09	10.6	**0.005**
Delta TP10	0.7 ± 0.05	0.5 ± 0.07	0.5 ± 0.06	16.7	**< 0.001**
Theta AF7	0.2 ± 0.07	0.2 ± 0.06	0.2 ± 0.06	3.3	**0.04**
Theta AF8	0.3 ± 0.07	0.3 ± 0.07	0.2 ± 0.07	5.6	**0.005**
Theta TP10	0.4 ± 0.06	0.3 ± 0.06	0.3 ± 0.04	11.1	**< 0.001**
Alpha AF7	0.4 ± 0.05	0.3 ± 0.05	0.3 ± 0.04	4.9	**0.01**
Alpha AF8	0.4 ± 0.05	0.4 ± 0.06	0.3 ± 0.06	5.5	**0.006**
Beta TP10	0.4 ± 0.03	0.4 ± 0.04	0.4 ± 0.04	4.5	**0.01**
Gamma TP9	0.1 ± 0.06	0.03 ± 0.05	0.5 ± 0.06	3.3	**0.04**
Gamma TP10	0.1 ± 0.04	0.05 ± 0.05	0.03 ± 0.04	4.8	**0.01**
Relative Delta TP10	0.3 ± 0.02	0.3 ± 0.02	0.3 ± 0.04	3.6	**0.03**
Relative Alpha TP10	0.3 ± 0.02	0.3 ± 0.02	0.3 ± 0.02	11.9	**< 0.001**
Relative Gamma TP10	0.04 ± 0.02	0.03 ± 0.02	0.01 ± 0.02	4.1	**0.02**
TA Theta	-0.04 ± 0.03	-0.02 ± 0.03	0.01 ± 0.02	3.4	**0.04**
TA Alpha	-0.06 ± 0.03	-0.003 ± 0.03	0.05 ± 0.03	8.3	**0.001**

Table 4. Dwass_Seel_Critchlow-Fligner pairwise comparison results. Significance level *alpha* = 0.05. (* stands for *p-value* < 0.001).

EEG features	A1-A2	A1-A3 p-values	A2-A3
Delta TP9	0.1	0.1	0.9
Delta AF7	0.1	0.2	0.9
Delta AF8	**0.01**	**0.01**	0.9
Delta TP10	**0.004**	**0.001***	0.9
Theta AF7	0.3	0.2	0.9
Theta AF8	0.3	0.1	0.7
Theta TP10	**0.04**	**0.01**	0.9
Alpha AF7	0.5	0.2	0.8
Alpha AF8	0.7	0.2	0.6
Beta TP10	0.8	0.5	0.5
Gamma TP9	0.3	0.2	0.9
Gamma TP10	0.4	0.9	0.8
Relative Delta TP10	0.2	0.1	0.3
Relative Alpha TP10	**0.01**	0.2	0.9
Relative Gamma TP10	0.7	0.2	0.7
TA Theta	0.6	0.1	0.3
TA Alpha	0.8	**0.01**	**0.03**

Zheng-Guang Li et al. [41] work, they found that theta waves increased with subjective annoyance under noise exposure.

Concerning alpha activity, a correlation between high attentiveness to sounds, pleasant/relaxing state and the increase of alpha waves has been demonstrated. An emotionally significant stimulus automatically attracts attention [42] and it is, therefore, conceivable that due to the pleasant emotions induced in the course of the consonant pieces, participants listened more attentively to the pleasant-sounding excerpts. The results of the present work are in line with [43], [44].

Statistical difference was also found between A2 and A3 audios exposure for TA Alpha, showing a higher value when subjects were exposed to the A3 sounds, suggesting a greeter activation of the right hemisphere of the brain during audio stimuli processing. Interhemispheric asymmetry is explained by

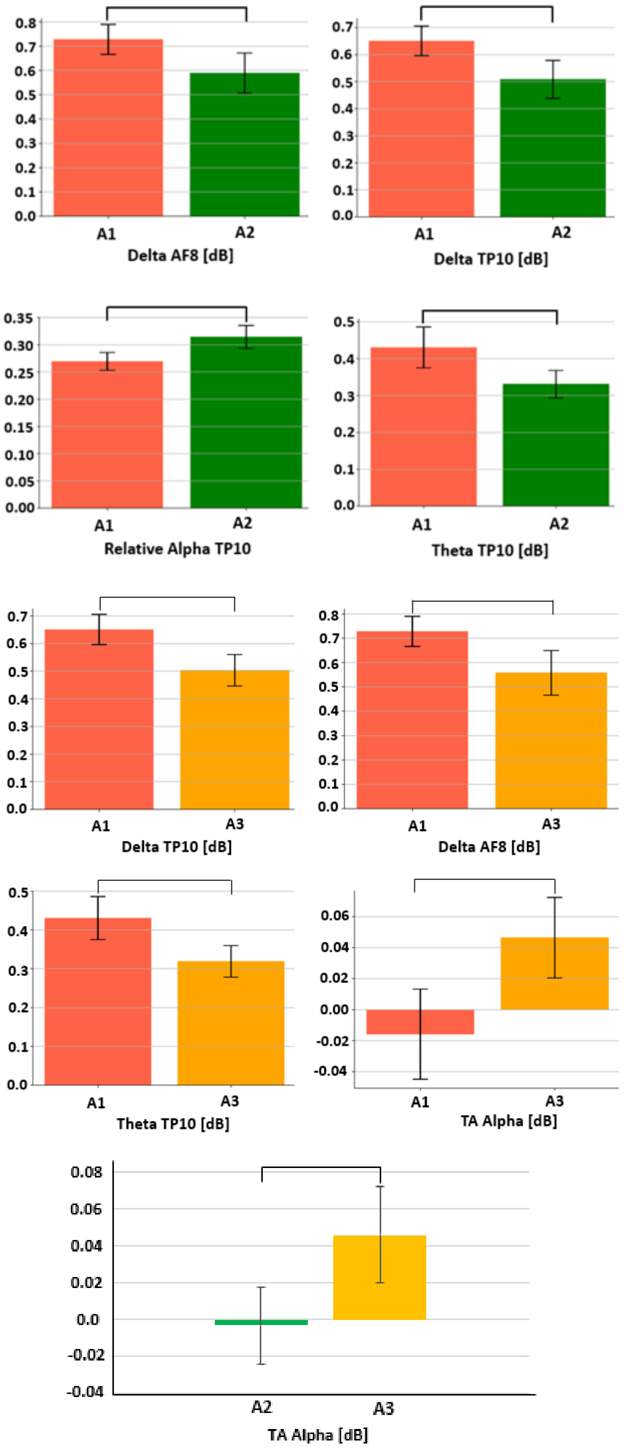

Figure 6. Histograms of Dwass_Seel_Critchlow-Fligner pairwise comparison. The $\bar{f}$ and u_f of EEG statistically significant by contrast A1-A2, A1-A3, A2-A3.

the study of functional lateralization in subcortical and cortical auditory circuits. However, none of the hemispheres can be said to process distinct aspects of audio inputs with absolute dominance. This is consistent with the core tenets of the notion of system dynamic localization of functions, which postulates that the entire brain participates in the realization of each function through the collaboration of distributed neuron ensembles[45]. The higher value of TA alpha under A3 exposure is supported by M. J. Jafari et al. [46], they found an increase of alpha bands when the participants were exposed to background noise.

Table 5. Pearson's correlation coefficients and p-values between audio metrics and EEG features. Significance level *alpha* = 0.05.

Audio Metrics	Delta AF8	Delta TP10	Theta TP10	Relative Alpha TP10	TA Alpha
			(R-value)		
CNT	+0.11	-0.08	-0.09	-0.01	**0.25**
SPR	-0.07	-0.04	-0.05	-0.05	**0.23**
SKW	**0.26**	**0.32**	**0.28**	**-0.27**	-0.16
KUR	**0.27**	**0.33**	**0.28**	**-0.27**	**-0.18**
ENT	-0.14	-0.12	-0.12	0.02	**0.27**
FLX	**0.27**	**0.33**	**0.29**	**-0.26**	**-0.20**
			(p-value)		
CNT	0.2287	0.3605	0.3123	0.8845	**0.0040**
SPR	0.4145	0.6615	0.5493	0.5844	**0.0082**
SKW	**0.0027**	**0.0002**	**0.0014**	**0.0019**	0.0732
KUR	**0.0020**	**0.0002**	**0.0011**	**0.0023**	**0.0413**
ENT	0.1185	0.1730	0.1634	0.8060	**0.0022**
FLX	**0.0017**	**0.0002**	**0.0010**	**0.0032**	**0.0203**

4.3. Correlation between EEG features and audio metrics

The correlation analysis, described in Section 3.2.2, between EEG features and the audio spectral metrics are reported in Table 5. Even though the results revealed a moderate linear relationship, the corresponding p-values confirm the significance of these correlations. All the audio metrics used in this analysis show a statistically significant correlation with at least one EEG feature. These results are consistent with other studies in this field. The highest correlation occurs between EEG metrics and spectral SKW, spectral kurtosis and FLX. This is confirmed from previous studies reported in literature [47] where it has been demonstrated that spectral SKW and kurtosis are widely correlated with perceived affective quality of soundscape and specifically with traffic noise.

5. CONCLUSION

The presented research aimed at evaluating individual acoustic perception to audio stimuli using a wearable device. A laboratory-based experimental protocol was conducted to expose voluntaries subjects (43 in total) to three different audios.

EEG data collection was pursued through commercially available wearable device. The results partially confirmed previous researcher outcomes in the field of acoustic perception to audio stimuli, highlighting the reliability of EEG measurements through wearable sensor.

The results showed moderate to substantial correlations between subjective reactions to audios and EEG indicators taken during audio listening tests. The present study found significant differences between the three audio stimuli in terms of increase and decrease of specific EEG features. Considering the relative frequency distribution of subjective acoustic perception collected via questionnaire, A1 was perceived annoying, stressful and with neutral loudness, A2 was perceived by almost subjects relaxing

and pleasant, A3 was categorized as annoying and stressful sounds. Those subjective perception found support into statistical differences of power of brain waves. Delta power (in TP10 and AF8 channels) and Theta TP10 showed higher value when subjects were exposed to A1 suggesting the hypothesis that those brain waves are modulated by annoying perception of audio stimuli, relative alpha had higher value under pleasant sound exposure (A2), TA in alpha frequency band was higher under A3 audio exposure (annoying/stressful) suggesting right lateralization of brain in processing sound content and a relation with negative emotion due to the presence of background noise in A3.

The present work is a pilot study for understanding the applicability of EEG measures for acoustic perception evaluation and it has demonstrated the potential of EEG measurements to increase objectivity in auditory perception. This methodology can be applied across various fields, for instance, in the evaluation of jurors' involvement in the acoustic jury tests and judicial domain, as demonstrated in this paper. Other application scenarios can be emotion recognition under acoustic stimuli or those where the subject cooperation is limited, such as cases involving degenerative diseases.

Wearable EEG measures in jury tests have both advantages and limitations. Understanding these can help determine the suitability and potential of using EEG technology in this context.

In terms of advantages, it is important to underline the fact that wearable EEG devices are non-invasive and can be relatively comfortable to wear, minimizing interference with jurors' normal activities. EEG can provide biometric data that may be useful for assessing juror engagement, stress levels, or emotional reactions during trials. EEG provides objective data on brain activity, which can complement subjective assessments of juror behaviour and decision-making.

Nevertheless, wearable EEG devices typically have limited spatial resolution compared to traditional EEG setups, which can impact the precision of data collected. EEG measurements can be affected by external factors such as noise and electromagnetic interference, potentially leading to data inaccuracies. Individual differences in baseline brain activity and electrode placement can introduce variability in EEG data, making it challenging to establish consistent benchmarks. An important future step will be to improve the quality of EEG signals acquired from wearable devices with a limited number of electrodes. This can be achieved by means of optimization of signal processing for data cleaning and artifact removal. Future works should also investigate methodologies for EEG and audio feature extraction and correlation, for example, more advanced statistical techniques and machine learning.

Another issue is the choice of a compatible earphone and EEG wearable device which became crucial in term of comfortable wearability for jurors, to ensure that they do not interfere with jury members' ability to concentrate or deliberate effectively. Also, these aspects have room for further improvements and can be object of future works.

Implementing EEG technology in the context of jury tests requires careful consideration of these limitations to addressing associated challenges. Furthermore, additional research would be useful in the study for continuously refining the use of EEG technology in jury tests based on feedback, research findings, and technological advancements.

ACKNOWLEDGEMENT

This project has received funding from the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Sklodowska-Curie Grant Agreement nº 858018. Additionally, the authors also acknowledge the support of the young researcher PRIN project NEXT.COM (20172FSCH4_002), initiative from the Italian Ministry of Research.

REFERENCES

[1] G. Battista, S. Serroni, M. Martarelli, M. Arnesano, G.M. Revel, Innovative measurements for Indoor Environmental Quality (IEQ) assessment in residential buildings, 2022 IEEE Int. Work. Metrol. Living Environ. MetroLivEn 2022 - Proc. (2022), pp. 170–173.
DOI: 10.1109/metrolivenv54405.2022.9826982

[2] G. Chinazzo, R.K. Andersen, E. Azar, V.M. Barthelmes, C. Becchio, (+ 31 more authors), Quality criteria for multi-domain studies in the indoor environment: Critical review towards research guidelines and recommendations, Build. Environ. 226 (2022).
DOI: 10.1016/j.buildenv.2022.109719

[3] V. Giannella, C. Colangeli, J. Cuenca, R. Citarella, M. Barbarino, Experimental/Numerical Acoustic Assessment of Aircraft Seat Headrests Based on Electrospun Mats, Appl. Sci. 2021, Vol. 11, Page 6400. 11 (2021) 6400.
DOI: 10.3390/app11146400

[4] I. Dimino, C. Colangeli, J. Cuenca, P. Vitiello, M. Barbarino, Active Noise Control for Aircraft Cabin Seats, Appl. Sci. 2022, Vol. 12, Page 5610. 12 (2022) 5610.
DOI: 10.3390/APP12115610

[5] L. Delcor, E. Parizet, J. Ganivet-Ouzeneau, J. Caillet, Model of sound and vibration discomfort in helicopter cabins, Appl. Acoust. 195 (2022) 108847.
DOI: 10.1016/ j.apacoust.2022.108847

[6] C. J. D. S. Brizon, E. Bauzer Medeiros, Combining subjective and objective assessments to improve acoustic comfort evaluation of motor cars, Appl. Acoust. 73 (2012), pp. 913–920.
DOI: 10.1016/ j.apacoust.2012.03.013

[7] V. der Auweraer, Neural Network Models For The Subjective And Objective Assessment Of A Propeller Aircraft Interior Sound Quality, (n.d.).

[8] R. Jamali, G. Battista, M. Martarelli, P. Chiariotti, D. S. Kunte, C. Colangeli, P. Castellini, Objective-Subjective Sound Quality Correlation Performance Comparison of Genetic Algorithm Based Regression Models and Neural Network Based Approach, J. Phys. Conf. Ser. 2041 (2021) 012015.
DOI: 10.1088/1742-6596/2041/1/012015.

[9] M. Münder, C. C. Carbon, Howl, whirr, and whistle: The perception of electric powertrain noise and its importance for perceived quality in electrified vehicles, Appl. Acoust. 185 (2022) 108412.
DOI: 10.1016/j.apacoust.2021.108412

[10] Z. Xiaojuan, L. Yan, Z. Changbin, Sound Quality Subjective Evaluation Analysis of Noise Inside High-speed Trains, (2012).
DOI: 10.2991/mems.2012.25

[11] T. Robotham, O. S. Rummukainen, M. Kurz, M. Eckert, E. A. P. Habets, Comparing Direct and Indirect Methods of Audio Quality Evaluation in Virtual Reality Scenes of Varying Complexity, IEEE Trans. Vis. Comput. Graph. 28 (2022), pp. 2091–2101.
DOI: 10.1109/tvcg.2022.3150491

[12] E. Zhang, L. Hou, C. Shen, Y. Shi, Y. Zhang, Sound quality prediction of vehicle interior noise and mathematical modeling using a back propagation neural network (BPNN) based on particle swarm optimization (PSO), Meas. Sci. Technol. 27 (2015) 015801.
DOI: 10.1088/0957-0233/27/1/015801

[13] N. Otto, S. Amman, C. Eaton, S. Lake, Guidelines for Jury Evaluations of Automotive Sounds, (1999).

[14] J. S. Kumar, P. Bhuvaneswari, Analysis of Electroencephalography (EEG) Signals and Its Categorization–A Study, Procedia Eng. 38 (2012), pp. 2525–2536.
DOI: 10.1016/j.proeng.2012.06.298

[15] U. Herwig, P. Satrapi, C. Schönfeldt-Lecuona, Using the International 10-20 EEG System for Positioning of Transcranial Magnetic Stimulation, Brain Topogr. 16 (2003), pp. 95–99.
DOI: 10.1023/B:BRAT.0000006333.93597.9D

[16] W. G. Walter, V. J. Dovey, H. Shipton, Analysis of the electrical response of the human cortex to photic stimulation, Nature. 158 (1946), pp. 540–541.
DOI: 10.1038/158540A0

[17] A. Compston, The Berger rhythm: potential changes from the occipital lobes in man, Brain. 133 (2010) 3–6.
DOI: 10.1093/brain/awp324

[18] M. Trimmel, M. Kundi, G. Binder, E. Groll-Knapp, M. Haider, Combined effects of mental load and background noise on CNS activity indicated by brain DC potentials, Environ. Int. 22 (1996), pp. 83–92.
DOI: 10.1016/0160-4120(95)00106-9

[19] L.A. Schmidt, L.J. Trainor, Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions, 15 (2010), pp. 487–500.
DOI: 10.1080/02699930126048

[20] M. Kabuto, T. Kageyama, H. Nitta, EEG power spectrum changes due to listening to pleasant music and their relation to relaxation effects, Nihon Eiseigaku Zasshi. 48 (1993), pp. 807–818.
DOI: 10.1265/jjh.48.807

[21] J. L. Walker, Subjective reactions to music and brainwave rhythms, Physiol. Psychol. 5 (1977), pp. 483–489.
DOI: 10.3758/BF03337859/metrics

[22] J. Zhang, J. Li, Z. Huang, D. Huang, H. Yu, Z. Li, Recent Progress in Wearable Brain–Computer Interface (BCI) Devices Based on Electroencephalogram (EEG) for Medical Applications: A Review, Heal. Data Sci. 3 (2023).
DOI: 10.34133/hds.0096

[23] B.A. Hickey, T. Chalmers, P. Newton, C.T. Lin, D. Sibbritt, C.S. McLachlan, R. Clifton-Bligh, J. Morley, S. Lal, Smart Devices and Wearable Technologies to Detect and Monitor Mental Health Conditions and Stress: A Systematic Review, Sensors (Basel). 21 (2021).
DOI: 10.3390/S21103461

[24] S.A. Imtiaz, A Systematic Review of Sensing Technologies for Wearable Sleep Staging, Sensors (Basel). 21 (2021), pp. 1–21.
DOI: 10.3390/S21051562

[25] J.W. Ahn, Y. Ku, H.C. Kim, A Novel Wearable EEG and ECG Recording System for Stress Assessment, Sensors (Basel). 19 (2019).
DOI: 10.3390/S19091991

[26] O.E. Krigolson, C.C. Williams, A. Norton, C.D. Hassall, F.L. Colino, Choosing MUSE: Validation of a Low-Cost, Portable EEG System for ERP Research, Front. InNeuroscience. 11 (2017).
DOI: 10.3389/fnins.2017.00109

[27] M. Abujelala, A. Sharma, C. Abellanoza, F. Makedon, Brain-EE: Brain enjoyment evaluation using commercial EEG headband, ACM Int. Conf. Proceeding Ser. 29-June-2016 (2016).
DOI: 10.1145/2910674.2910691

[28] G. Wiechert, M. Triff, Z. Liu, Z. Yin, S. Zhao, Z. Zhong, R. Zhaou, P. Lingras, Identifying users and activities with cognitive signal processing from a wearable headband, Proc. 2016 IEEE 15th Int. Conf. Cogn. Informatics Cogn. Comput. ICCI*CC 2016. (2017), pp. 129–136.
DOI: 10.1109/ICCI-CC.2016.7862025

[29] Mind Monitor website, Online [Accessed 8 May 2024] https://mind-monitor.com

[30] F. Faul, E. Erdfelder, A.G. Lang, A. Buchner, G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences, Behav. Res. Methods. 39 (2007), pp. 175–191.
DOI: 10.3758/BF03193146

[31] A technique for the measurement of attitudes., (n.d.). Online [Accessed 8 May 2024]
https://psycnet.apa.org/record/1933-01885-001

[32] Spectral Descriptors - MATLAB & Simulink - MathWorks Italia, (n.d.). Online [Accessed 8 May 2024]
https://it.mathworks.com/help/audio/ug/spectral-descriptors.html

[33] M. Ramires Bonfim, A. da Costa Lamas, Analysis of Electrical Neural Data Using MUSE and EEGLAB, (n.d.).
DOI: 10.1007/978-3-031-31007-2_10.

[34] BIPM - JCGM, Evaluation of measurement data - Guide to the expression of uncertainty in measurement, Int. Organ. Stand. Geneva ISBN. 50 (2008) 134. Online [Accessed 8 May 2024] http://www.bipm.org/en/publications/guides/gum.html

[35] M. Brzezinski, The Chen-Shapiro test for normality, Stata J. 12 (2012), pp. 368–374.
DOI: 10.1177/1536867X1201200302

[36] H. Arsham, M. Lovric, Bartlett's Test, Int. Encycl. Stat. Sci. (2011), pp. 87–88.
DOI: 10.1007/978-3-642-04898-2_132

[37] P. A. Games, Alternative Analyses of Repeated-Measure Designs by ANOVA and MANOVA, Stat. Methods Longitud. Res. (1990), pp. 81–121.
DOI: 10.1016/B978-0-12-724960-5.50008-1

[38] E. Ostertagová, O. Ostertag, J. Kováč, Methodology and application of the Kruskal-Wallis test, Appl. Mech. Mater. 611 (2014) 115–120.
DOI: 10.4028/www.scientific.net/amm.611.115

[39] D. E. Critchlow, M. A. Fligner, On distribution-free multiple comparisons in the one-way analysis of variance.
DOI: 10.1080/03610929108830487

[40] J. L. Walker, Subjective reactions to music and brainwave rhythms, Physiol. Psychol. 5 (1977) 483–489. Online [Accessed 8 May 2024] https://link.springer.com/content/pdf/10.3758/BF03337859.pdf

[41] Z. G. Li, G. Q. Di, L. Jia, Relationship between Electroencephalogram variation and subjective annoyance under noise exposure, Appl. Acoust. 75 (2014) 37–42.
DOI: 10.1016/j.apacoust.2013.06.011

[42] M. M. Bradley, M. Codispoti, B. N. Cuthbert, P. J. Lang, Emotion and Motivation I: Defensive and Appetitive Reactions in Picture Processing, Emotion. 1 (2001) 276–298.
DOI: 10.1037/1528-3542.1.3.276

[43] S. Nishifuji, I. Miyahara, Amplitude and Phase Responses of Brain Alpha Wave to Unpleasant and Pleasant Sounds, Transactions of the Society of Instrument and Control Engineers 44(10), pp. :776-778
DOI: 10.9746/ve.sicetr1965.44.776

[44] N. Hurless, A. Mekic, S. Peña, E. Humphries, H. Gentry, D. F. Nichols, Music genre preference and tempo alter alpha and beta waves in human non-musicians, Impuls. Prem. Undergrad. Neurosci. J. (2013). Online [Accessed 8 May 2024]
https://assets.pubpub.org/m2h1swki/41643404198032.pdf

[45] S. F. Vaitulevich, E. A. Petropavlovskaya, L. B. Shestopalova, N. I. Nikitin, Functional Interhemispheric Asymmetry of Human Brain and Audition, Hum. Physiol. 2019 452. 45 (2019), pp. 202–212.
DOI: 10.1134/S0362119719020129

[46] M. J. Jafari, R. Khosrowabadi, S. Khodakarim, F. Mohammadian, The Effect of Noise Exposure on Cognitive Performance and Brain Activity Patterns, Open Access Maced. J. Med. Sci. 7 (2019), pp. 2924–2931.
DOI: 10.3889/oamjms.2019.742

[47] J. K. A. Tan, Y. Hasegawa, S. K. Lau, S. K. Tang, The effects of visual landscape and traffic type on soundscape perception in high-rise residential estates of an urban city, Appl. Acoust. 189 (2022) 108580.
DOI: 10.1016/j.apacoust.2021.108580