Guidelines for conducting biofeedback-enhanced QoE studies in mulsemedia-enhanced virtual reality

Aleph Campos da Silveira
alephcampos@gmail.com
Federal University of Espirito Santos
Vitoria, Espirito Santo, Brazil

Roope Raisamo
roope.raisamo@tuni.fi
Tampere University
Tampere, Finland

Fotios Spyridonis
Fotios.Spyridonis@brunel.ac.uk
Brunel University London
London, United Kingdom

Alexandra Covaci
a.covaci@kent.ac.uk
University of Kent
Kent, United Kingdom

Gheorghita Ghinea
George.Ghinea@brunel.ac.uk
Brunel University London
London, United Kingdom

Celso Alberto Saibel Santos
saibel@inf.ufes.br
Federal University of Espirito Santos
Vitoria, Espirito Santo, Brazil

ABSTRACT

This paper presents the conclusions drawn from an ongoing experiment investigating the performance, usability, and impact of multisensory stimuli in the virtual environment, with a focus on proposing guidelines for conducting biofeedback-enhanced Quality of Experience (QoE) studies in multimedia-enhanced Virtual Reality (VR). The study evaluated various devices, including the Polar H10 and Grove GSR for biosignal measurement and an EEG device for brainwave analysis. Concerns related to participant comfort were highlighted, such as discomfort caused by tight electrodes and difficulties in achieving consistent contact with the scalp. Ergonomic issues with Head-Mounted Displays (HMDs) were also identified, emphasizing the need for comfortable and immersive experiences. The paper recommends addressing these concerns through inclusive design and user-friendly adaptation of devices. The findings emphasize the importance of integrated devices and user-friendly design to enhance QoE and facilitate the adoption of biofeedback technologies outside of the lab. By following the proposed guidelines, researchers and developers can improve the immersive experience and advance the field of biofeedback in VR environments.

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods**; **User studies**.

KEYWORDS

User Evaluation, Biofeedback, Guidelines, Quality of Experience, Mulsemedia

ACM Reference Format:
Aleph Campos da Silveira, Roope Raisamo, Fotios Spyridonis, Alexandra Covaci, Gheorghita Ghinea, and Celso Alberto Saibel Santos. 2023. Guidelines for conducting biofeedback-enhanced QoE studies in mulsemedia-enhanced virtual reality. In *Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia '23), October 23–27, 2023, Ribeirão Preto, Brazil.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3617023.3617029

1 INTRODUCTION

With the growing interest and development of Virtual Reality (VR) and Augmented Reality (AR) environments, the use of biosignals has proven to be a promising approach to improve human-machine interaction. Biosignals, in the context of this paper, are any signal measured and monitored from a human being, such as the electrical activity of the brain, the heart rate or skin resistance, which can be used to evaluate or control virtual devices or systems [38]. However, despite their potential, there are still some challenges and obstacles that limit the effective use of biosignals, mainly in VR and AR.

In recent years, there has been a notable shift in focus in the study of the senses in conjunction with other media devices. According to Calvert and Thesen [5], the adoption of a multisensory approach to human sensory perception has been influenced by technological advances and a deeper understanding of sensory neurophysiology. These technological advancements have occurred alongside an increase in knowledge about how our sensory systems work. It also has become evident that a comprehensive understanding of our perceptual systems requires consideration of how each sense is integrated with input from other sensory systems. As a result, the foundation for mulsemedia capabilities is established: integration of multiple sensory sources beyond audio and video to enhance the user's sense of presence [9, 15].

This shift in focus has driven the development of increasingly sophisticated and immersive VR and AR environments, which aim to offer richer and more engaging sensory experiences. However, the use of biosignals in such environments still faces some obstacles, such as the need for expensive and complex equipment for signal acquisition and processing [30, 39], difficulty in ensuring reliability and accuracy of collected data [16], and a lack of standardization and guidelines [1] for the use of biosignals in VR and AR. Additionally, there are important ethical issues related to the use of biosignals, such as privacy and security of collected personal data [11]. Such challenges require collaboration among professionals from different fields, such as engineers, neuroscientists, and experts in ethics and privacy, to overcome them and allow the use of biosignals in VR and AR environments to reach its full potential.

WebMedia '23, October 23–27, 2023, Ribeirão Preto, Brazil

ACM ISBN 979-8-4007-0908-1/23/10…$15.00
https://doi.org/10.1145/3617023.3617029

This paper explores the use of different biosignals as biofeedback to improve user experience (UX) evaluation. While the evaluations are currently ongoing, this paper mainly highlights the ergonomic challenges that arose during the study, specifically with the use of electroencephalograph (EEG), heart rate (HR), and galvanic skin response (GSR) for gathering biosignals, and Head Mounted Displays (HMD) and olfactory devices for rendering multisensory contents. While dealing with these obstacles, we took the opportunity to propose guidelines for those evaluations.

The paper is structured as follows: Section 2 presents the literature review; Section 3 introduces the design of the experiment used for UX evaluation which contains the methodology, as well as the devices used in the experimental setup; Section 4 discusses the preliminary results of our study and; finally, Section 5 introduces our guidelines for conducting biofeedback-enhanced QoE Evaluations, followed by conclusion in Section 6.

2 LITERATURE REVIEW ON USING BIOFEEDBACK TO ENHANCE QOE STUDIES

Human-Computer Interaction (HCI) research plays a crucial role in identifying, analyzing, and responding to user behavior [7]. Evaluating user interaction has been a central concern for HCI researchers, who have traditionally relied on methods that can disrupt the user experience or that the user may not recall accurately, particularly when assessing emotions during the interaction. As a result, researchers have focused on implicit methods to evaluate user interactions, such as biofeedback [23, 26, 32]. This method involves monitoring a person's body activity by using electrodes that attach to the scalp or the skin.

The use of these bio-electric signals of human body provides valuable insights into user attention and experience, like this cognitive and physical effects in various fields, especially in those that utilize multimedia technologies such as VR, AR, and mulsemedia. These technologies enable better evaluation of the impact on the user's body and interaction with the system. According to a recent study by Calvo-Morata et al. [6], electrocardiograms (ECG) and EEG are the most commonly used signals, primarily to assess user engagement, difficulty, and stress levels. However, the same authors also stated there are limitations to this approach, such as signal noise and the time required to calibrate devices during experiments, which make user testing more complex. To overcome these limitations, it is important to develop tools that allow experiments with large groups of users and to create open software and low-cost devices that enable more studies in this field. As the complexity of devices and the handling of data collected from them make it difficult to conduct experiments with actual users in real-world settings [6]. Calibration of the devices is often necessary, and signal noise and interference can render samples invalid.

The use of biosignals in large deployments remains a challenge. The review conducted by Calvo-Morata et al. [6] revealed that the majority of the studies in this matter were conducted with fewer than 30 users, and most experiments collecting parallel samples from only around 6 participants. As a result, studies often take weeks to collect sufficient data. Highly controlled rooms and environments are used to minimize these challenges. Additionally, ethical issues are addressed in many studies, but the treatment of user data remains an issue long after it is collected. It is important for studies to describe how user data is stored and secured, as well as the measures taken to ensure ethical exploitation of the data.

EEG is notorious know as being noise and hard to deal, and as stated by Cano et al. [7]. The complexity and cost of EEG using technologies have limited their use in HCI research in the past, but in recent years, commercial EEG devices and open-source alternatives have become available. However, the quality of the data collected and the sample size can vary between studies, as they are often experimental. Various neurophysiological measures, such as EEG, Functional Near-InfraRed spectroscopy (fNIRs), and Functional Magnetic Resonance Imaging (fMRI) have been used to assess the mental and emotional states of users through brain activity. However, the use of these methods can impact the effectiveness of measurements and increase costs. Most wearable devices are EEG-based and have been improved with less invasive hardware, including dry electrodes or water-based solutions, resulting in high signal quality and user comfort.

Despite the obstacles associated with using biosignals, the growth of pervasive and wearable devices that already gather a multitude of bio-electric measures from our bodies has opened up a range of new opportunities in the field of biofeedback. This said, biofeedback is a growing field of study that enables us to better understand how users interact with systems and devices, providing valuable insights into user behavior, attention, cognitive and physical effects, and emotional states. With the development of new tools and low-cost devices that allow experiments with large groups of users, the field of biofeedback is expected to expand further, enabling us to explore new possibilities and challenges in the domains of HCI and beyond.

To better comprehend how users engage with sound artifacts in 360° videos augmented with olfactory stimuli, we have gathered a selection of biosignal devices. As VR continues to expand, our goal is to enhance our comprehension of user experiences in these environment. In the upcoming section, we will outline the techniques and instruments employed in our research.

3 UX EVALUATION USING BIOSIGNALS AND METHODOLOGY

The methods employed for collecting HR and EDA data were established based on the frameworks proposed by Egan et al. [13] and Salgado et al. [34]. To ensure the ethical integrity of these evaluations, all procedures were granted approval by the Ethical Committee of Brunel University London review number 40020-LR-Oct/2022- 41826-3. To facilitate a comprehensive user experience (UX) assessment, a diverse group of participants was chosen, representing three distinct universities: Brunel University and Kent University in the United Kingdom, Tampere University in Finland, and ongoing evaluations at the Federal University of Espirito Santo (UFES) in Brazil. During the evaluations, participants were exposed to approximately 1-minute-long 360° videos within a controlled environment, as depicted in Figure 3. Throughout the evaluation process, meticulous monitoring of EEG, HR, and GSR signals was conducted for all participants.

In the evaluations of the 360° videos, three quarters of the participants were exposed to multisensory stimuli, specifically through the use of an olfactory device that diffused a pleasant lavender smell. This addition aimed to enhance the participants' sensory experience during the evaluations. Another one quarter were part of the control group and did not experienced the multisensory content. All the evaluation are being conducted with the same equipment and in a similar room between institutions.

Prior to the evaluation, participants were asked questions regarding any neurological or psychological disorders that could potentially impact their response to the videos. Subsequently, comprehensive information about the study, encompassing its objectives and procedures was provided to ensure participants' full understanding and informed consent.

To address the research question, three videos of comparable duration were selected for inclusion in the evaluation. The presentation order of these videos was randomized to ensure equal representation, ensuring that each video received evaluation from an equal number of participants.

The selection of videos was guided by the sequential order proposed by Comşa et al. [8], which categorizes them based on the level of coupling between audio and video. These videos encompass a spectrum of coupling levels, starting from low coupling progressing to mild and reaching high coupling.

To investigate how users perceive audio artifacts and desync in watching 360° video, we manipulated the videos' audio tracks by introducing delays and hastening effects ranging from 5 to 1 second in a similar way of the work of Brito et al. [4] for evaluating subtitles quality. The objective was to explore the impact of audio desync on user experience and determine whether the pleasant olfactory stimuli could have a beneficial effect in masking poor audiovisual quality.

Following each evaluation of the 360° video, participants were requested to complete two sets of questionnaires. The first set focused on assessing the participants' QoE, aiming to gain insights into their subjective perception and overall satisfaction with the video. The second questionnaire utilized the System Usability Scale (SUS), a standardized tool designed to evaluate the usability of systems and interfaces. By administering these questionnaires, we aimed to gather feedback from participants, enabling us to understand their experiences and gauge the experience of the 360° video.

Throughout the evaluation process, the EEG, HR, and GSR of each participant were monitored while the participant watched each one of the three videos. The physiological responses of the participant to each video, such as changes in brain activity, heart rate, and skin conductance were then collected and stored. The physiological data analysis will still be applied in order to to further explore the participants' responses to the three videos.

3.1 Devices Used

Devices used in the proposed experiment include the input/output collection devices as follows. Figure 3 illustrates a user using the HMD belts over the EMOTIV EPOC X. The Polar H10 is under the cloths and the GSR nodes attached to the hand.

3.1.1 *Polar H10.*

Depicted in Figure 1(1), this is a chest strap heart rate monitoring device developed by Polar, a Finnish company specializing in wearable technology for fitness and sports [35]. The strap is worn around the chest and connects to a compatible mobile device or training equipment via Bluetooth or ANT+. This enables accurate real-time heart rate monitoring during exercise. The Polar H10 incorporates a medical-grade electrocardiogram (ECG) sensor, which filters noise and interference, providing reliable and accurate heart rate data in various conditions. It is compatible with a wide range of mobile training apps allowing users to monitor and track their exercise performance. Due to its high precision, robust construction, and wide compatibility, the Polar H10 is a popular choice for accurate heart rate monitoring and is commonly used in investigations such as this.

3.1.2 *Grove GSR.*

Illustrated in Figure 1(2), this is a sensory device that measures the electrical conductivity of the skin, which can be utilized as an indicator of emotional or stress responses in humans. It is commonly utilized in academic research [37], medical applications [12], and biofeedback devices [26]. The module comprises of two electrodes, typically placed on the fingers, that detect changes in the level of electrical activity in the skin when a person experiences an emotion or stress. The signal is then amplified and processed to generate a skin conductance reading. The Grove GSR module is regarded as simple to use and can be easily integrated into electronic projects using the standard Grove interface. It is compatible with various microcontroller platforms such as Arduino, Raspberry Pi, and others. Furthermore, numerous software and libraries are available that enable the integration of Grove GSR with various programming languages, making it a versatile tool for fields such as electronics, robotics, biofeedback, psychology, and more.

3.1.3 *EMOTIV EPOC X.*

Shown in Figure 1(3), this is a Brain-Computer Interface (EEG using) device that enables the capture of brain signals for translation into computer commands. It features a flexible and ergonomic design that allows for electrode adjustment, making it more comfortable and user-friendly than other EEG using devices. The device consists of electrodes that detect electrical currents generated by the brain and wirelessly transmit them to a computer. Specialized software then processes this data to extract useful information. This device is versatile and has numerous applications, including scientific research [32], gaming [27, 28], and education [41]. By detecting patterns of brain activity associated with mental states such as concentration, stress, and excitement, the device can personalize gaming and teaching experiences, as well as provide valuable insights into neuroscience.

3.1.4 *META 2.*

Shown in Figure 2(1), this is a VR headset developed by the META company. It allows users to visualize and interact with virtual objects in a real-world environment. The META 2 was designed to provide a more advanced AR experience than other headsets on the market, with a wide field of view, high resolution, and accurate head and hand tracking. It is used in a variety of applications such as product design, training, education, gaming, and entertainment. The META 2 can be used for academic research in a wide range of fields, especially those involving visualization and interaction with three-dimensional objects. Some examples of possible uses of the META 2 in academic research include: visualization of complex models [18], medical training [3], human-computer interaction studies [40], and psychology [29] and neuroscience [31] studies.

Figure 1: Input Devices: Biosignal Capture Devices.

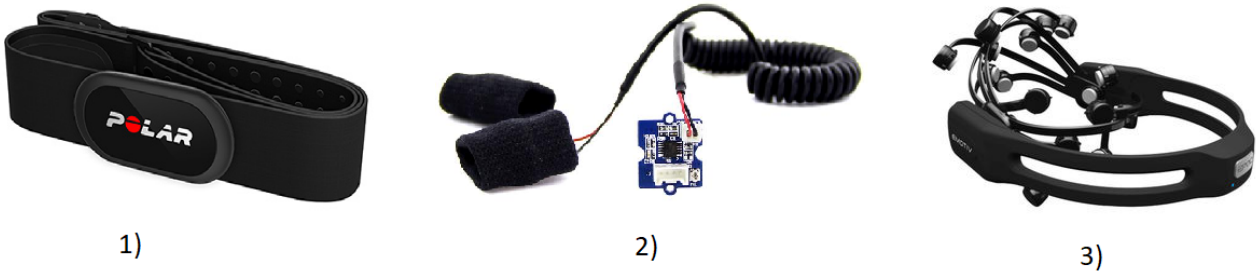


Figure 2: Output Devices: Immersive and Multi-sensory Devices.

3.1.5 ExHalia Sbi417. Depicted in Figure 2(2), this is a commercial device that uses air flow to provide (by default) one of four fragrances at a time [33]. According to Murray et al. [25], the SBi4 is "more reliable and more robust than other devices on the market" and the scents are more realistic. However, there are some considerations that researchers need to keep in mind when working with this olfactory display. According to Saleme et al. [33]:

- It's cartridges are made of perfumed polymer beads, which allow the scent to last less than other types of cartridges (for example, Dale Air Vortex18 employs alcohol-based fragrances soaked in cotton discs).
- The SBi4 connects to a USB port and allows the creation of codes to handle device activation. However, this allows the control of only a single fan at a time.

When used in conjunction with an HMD, the ExHalia olfactory device may be less efficient. One of the main reasons for this is that the device is in a fixed position, and moving the head breaks the line of sight between the device and the user's nose, which weakens the olfactory perception. This will be further discussed in the next section. Moreover, the position of the olfactory device may also impact its efficiency. Placing the device too close to the nose may result in overwhelming smells that could potentially cause discomfort to the user, while placing it too far may result in weaker smells that are difficult to notice. Additionally, regarding the development of olfactory devices for HMDs, the design of the device itself could also affect the efficiency, such as the type and size of the odor cartridges used, the method of odor diffusion, and the control of the intensity and duration of the smells.

4 PRELIMINARY RESULTS AND DISCUSSION

Although the experiment with the selected participants is still ongoing, we have gained experience and insights into the ergonomics of the tools and devices used in this study and allowed us to propose guidelines for future works.

Firstly, both the Polar H10 and Grove GSR devices demonstrated satisfactory performance with only minor adjustments required before each evaluation. While the Polar H10 may be considered more intrusive compared to the GSR and EEG capture device, as it needs to be placed beneath clothing close to the breasts and in direct contact with the skin, participants generally reported positive experiences. So far, the majority of participants found the device comfortable, with a small portion of individuals expressing concerns about the fit. For example, even after adjustments, one participant felt it was too tight, while the other found it to be too loose.

The GSR device demonstrated overall satisfactory performance and low intrusiveness, requiring minor adjustments and calibration prior to each session, although the resistor had to be manually adjusted with the use of a screw driver until the serial output was 512 before being attached to the fingers. Besides this minor inconvenience, no critical complains were highlighted. However, some discomfort was reported, expressing that the two electrodes can be too tight on some user fingers. Additionally, in long exposures,

Figure 3: A research participant using the devices for biosignals capture. The Grove GSR attached to the fingers, the EMOTIV EPOC X under the Meta 2. The polar H10 cannot be seeing under the shirt.

it was observed that the fabric surrounding the fingers caused excessive sweating in some participants, resulting in a decrease in skin resistance, leading to high skin conductance that might be no related to their emotional level. To address this issue, the decision was made to halt the experiment, remove and reinsert the GSR device after extended periods, in order to maintain the accuracy and reliability of the collected data. However, this approach can result in extended evaluation times and may potentially impact the authenticity of subjective user responses. These observations highlight the importance of ensuring participant comfort and maintaining suitable environmental conditions during data collection, as variations in ambient temperatures can lead to inaccurate biosignal readings.

Capturing biosignals with the EEG device proved to be the most challenging aspect of the study. The loose claws of the EMOTIV EPOC X were prone to displacement due to motion, presenting a particular issue when using the device in conjunction with an HMD and 360° videos that encourage head movement. The constant contact requirement between the nodes and the scalp also posed difficulties, especially when participants had an abundance of hair. This aspect of the device made it particularly challenging to achieve consistent contact and obtain accurate readings. Also, the EEG device exhibited limitations in accommodating different hair types and haircuts. Participants with long hair had to endure longer evaluation time due to adjustments, as the presence of excessive hair made it difficult to use the device effectively. The longer exposure to evaluation may lead to tiredness, boredom and exhaustion. This limitation may lead to discrimination against certain hairstyles, potentially impacting the inclusivity and usability of the EEG device.

Also, since we are dealing with 360° videos, the use of EEG and HDM display is susceptible to noisy data. Besides motion leading to artifacts, as stated by Mikhail et al. [23], sweat can lead to changes in the impedance of electrodes that are used to record brain activity, which can also result in noisy or corrupted data. EMOTIV EPOC X allows to position each electrode in accordance to the Electrode locations on the scalp labelled using AtlasLabel in FSL and the Harvard-Oxford Cortical Structural atlas recording [36], as seeing in the Figure 6. These various types of noise, as seen in the horizontal axes (AF3, C1, O1, etc.) in the Figure 4, can make processing EEG challenging, especially in real-time environments where there is no control over the environment or the subject. Notice the top right corner the difference of quality signal by the EmotivPRO Software between a good quality signal (100%) in Figure 5 contrasting with a bad quality one (42%) in Figure 4.

This is a well-known issue, as mentioned by Cano et al. [7], and there are some alternatives available, albeit at a high cost. One such alternative is the Galea HMD [17], a hardware and software platform that merges next-generation biometrics with mixed reality. This device integrates EEG, EMG, EDA, PPG, and eye-tracking into a single headset. It can also be integrated into existing AR and VR head-mounted displays and includes SDKs for bringing rich and tightly time-locked biometric data into 3D development engines, 3rd-party applications, and all common programming languages.

Regarding the output devices used, it is important to address some common complaints regarding the META 2 (and it is present in most HMD, as stated by Mehrfard et al. [22]). One frequently reported issue is the cumbersome nature of the device, which can cause fatigue and discomfort during prolonged sessions. This feedback was especially notable among participants who had to wear the HMD for an average duration of over 30 minutes, leading to tiredness and discomfort.

The ExHalia Sbi4 study yielded mixed results concerning the perception of olfactory stimuli. So far, participants perceived the smells as weak, prompting further examination of this issue. We believe that the primary contributing factor is the line of sight disruption between the user and the fixed stream of olfactory stimuli due to the HMD and 360 allowing head movement. This phenomenon has been previously discussed in studies [20] which propose alternative solutions, such as inhaled devices directly attached to the HMD, delivering smells directly to the user's nose. Another factor impacting the weak perception of smells is the potential overstimulation of sensations caused by the 360° experience, as the user is immersed in a peripheral surround, attending primarily to both visual and auditory stimuli, while keeping the olfactory stimuli as secondary.

Statistical data analysis for the questionnaires and the impact of multisensory stimuli on the captured biosignals is currently in progress and will be presented in a forthcoming paper.

Figure 4: Example of noise detected in the EMOTIV Pro signal during head movement. Signal Quality at 42%.

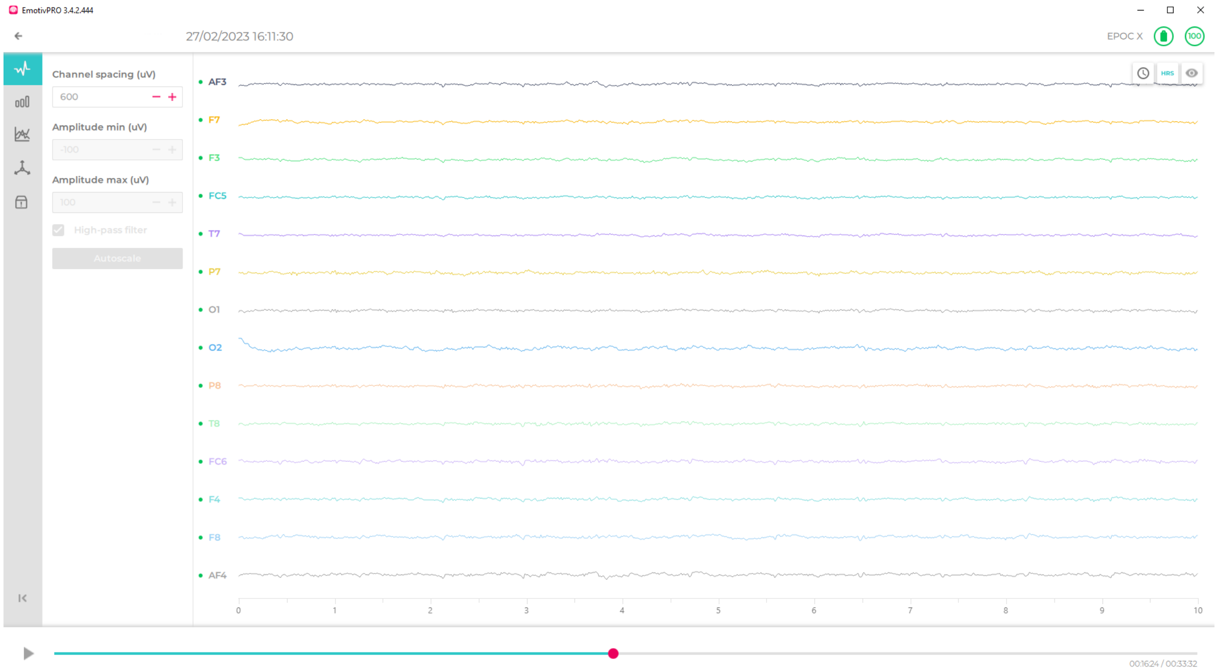
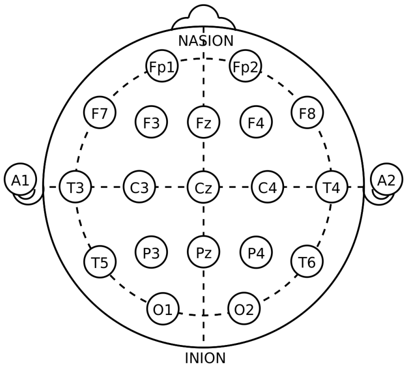


Figure 5: Clean and stable signal amplitude compared to noise result of head motion. Signal Quality at 100%.

5 PROPOSED GUIDELINES

Based on the information provided while running the evaluations so far and on the guidelines of Murray et al. [24] and the heuristics of Machado Neto and Pimentel [21], we propose the following novel guidelines for or conducting biofeedback-enhanced QoE studies in mulsemedia:

Figure 6: International 10-20 system for EEG (electroencephalography)



5.1 Regarding Participants Pool

A. Determine an appropriate sample size for the study, considering statistical power and generalization of results.

B. Aim for a diverse participant pool that represents different demographics (e.g., age, gender, cultural background) to ensure the findings are more inclusive and applicable to a broader population.

C. Obtain informed consent from participants, clearly explaining the purpose, procedures, and potential risks or benefits of the study.

D. Follow ethical guidelines and regulations related to data privacy, confidentiality, and participant well-being.

E. Provide participants with sufficient training and familiarization sessions with the biofeedback devices and the VR environment before the actual study.

F. Ensure participants are comfortable and confident in using the devices and interacting with the VR content to minimize potential learning or adaptation effects during the study, also to guarantee an easy input from the participant [21].

G. Participants should not be affected by cold, flu or fever since it can lead to altered biosignals, have good dental and overall body hygiene since it can alter skin related signals [24].

H. Participants should avoid consuming caffeine or any stimulants before the evaluation since it can lead alteration of heart rate and skin resistance.

5.2 Regarding Participant Comfort and Device Placement

A. Prioritize participant comfort when selecting and placing biofeedback devices. Consider the intrusiveness and potential discomfort associated with each device.

B. Provide options for device adjustments to accommodate individual preferences and physical characteristics, such as ensuring proper fit for devices to avoid tightness or looseness.

C. Address issues related to discomfort and fatigue associated with wearing HMDs for prolonged periods. Consider the weight and design of the devices to minimize discomfort during longer sessions.

D. Recognize differences in user experience between different types of evaluations and tailor the session duration accordingly to mitigate discomfort.

E. Address challenges related to motion and displacement of EEG device nodes by exploring alternative device options with improved stability, specifically designed for use with HMDs and motion-intensive activities.

F. Consider the limitations of EEG devices in accommodating different hair types and haircuts. Ensure inclusivity and usability by selecting devices that can effectively accommodate participants with long hair.

G. Address the evaluation time for the participants and make clear the the adjustment of the devices can take a bigger portion of the experiment length.

5.3 Regarding Bio-signal Quality

A. Regularly monitor and evaluate the signal quality of each biofeedback device during data collection. Take steps to mitigate any signal artifacts, noise, or interference.

B. Employ standardized calibration procedures for each biofeedback device used in the experiment. This ensures consistency across participants and sessions.

C. Seek to adopt environmental awareness devices than can mitigate inaccuracies led by different environments.

D. Implement signal filtering techniques or algorithms to improve the quality of monitored biosignals, particularly in real-time environments where noise can be challenging to control. This can be done during pre-processing and post-processing of the data.

E. Maintain appropriate environmental conditions during data collection to ensure accurate biosignal readings. Pay attention to ambient temperature and humidity levels, as they can affect the accuracy of biosignal data.

F. Mitigate sweating-related issues by monitoring participants for excessive sweating caused by the devices.

G. For longer exposures, remove the device after extended periods for participant rest [24], cleaning and position readjustment to maintain accuracy.

H. Address noise artifacts caused by excessive sweat during data processing by exploring alternative hardware options which integrates multiple biometric sensors into a single device.

5.4 Regarding Multi-sensory Stimuli

A. Assess the impact of multisensory stimuli in virtual environments and exploring alternative solutions, such as, in case of olfactory stimuli, inhaled devices directly attached to the HMDs.

B. Investigate the potential over stimulation of other senses, such as visual and auditory stimuli, in the presence of multisensory stimuli.

C. Understand the balance and hierarchy of sensory experiences to optimize the impact of multisensory stimuli on the overall QoE.

D. Ensure proper synchronization between different sensory stimuli (visual, auditory, olfactory, etc.) to create a coherent and realistic multisensory experience.
E. Use precise timing mechanisms to synchronize the presentation of different stimuli, minimizing any perceptual desync that may affect the overall QoE.
F. Investigate cross-modal effects between sensory modalities and how they influence the perception and evaluation of the multisensory experience.
G. Explore potential interactions and synergies between different sensory modalities that may enhance or diminish the overall QoE.
H. Assess the contextual relevance of each sensory modality within the virtual environment and its contribution to the overall narrative or task at hand.
I. For olfactory stimuli, participants and researchers should not have used any perfume, deodorants or aftershave before the experiment [24].

These guidelines aim to improve the overall methodology, device selection, participant comfort, and data accuracy in future biofeedback-enhanced QoE evaluations conducted in multisensory-enhanced VR and AR environments.

6 CONCLUSION

In conclusion, the ongoing experiment has provided insights into the performance, usability, and impact of multisensory stimuli in the virtual environment and helped us propose the guidelines for conducting biofeedback-enhanced QoE studies in multimedia-enhanced VR.

Regarding the devices, both the Polar H10 and Grove GSR demonstrated satisfactory performance with minor adjustments and calibration before each evaluation. Some discomfort using the GSR was stated due to the tightness of the electrodes on their fingers, highlighting the need to prioritize participant comfort in biosignal device design and usage.

The challenges posed by the EEG device, including the loose claws and difficulty achieving consistent contact with the scalp, highlight the guidelines emphasis also on addressing participant comfort and maintaining appropriate environmental conditions during data collection. The limitations in accommodating different hair types and haircuts further emphasize the need for inclusive design and consideration of diverse user populations in biofeedback studies conducted within VR environments. The concomitant use of EEG with 360° VR presents an ongoing challenge due to the movement of users' heads induced by the content presented in HMD, which results in node displacement and introduces noise. Further studies are needed to address this issue and develop effective strategies to overcome this obstacle.

Concerns regarding the META 2 and other HMDs being cumbersome and causing fatigue and discomfort were the basis for recommendation to address ergonomic issues in HMD design to ensure a comfortable and immersive experience for users.

The perception of olfactory stimuli in the virtual environment will be further investigated in needed to better understand their impact on enhancing the immersive experience and overall QoE in VR and AR applications.

Another important point of discussion is the multimodal data fusion and analysis. Although it is an active area of research [14, 19] how this is undertaken in the case of multiple sensory modalities and mulsemedia data including, for instance, EEG, EDA or ECG signals, as well as eye-tracking data is not well understood. Whilst one may employ methods from other non-mulsemedia domains, more work is needed to validate appropriate methods for mulsemedia data fusion and analysis - which obviously goes beyond the application of classic statistical unimodal approaches - in order to draw meaningful conclusions. In a mulsemedia context this is even more poignant given masking and cross-modal effects [2, 10].

The findings presented in this study further underscore the importance of an integrated device that combines all necessary components to enhance and measure the QoE. Such integration would significantly reduce the time-consuming preparation required for evaluations. In line with our objective of exploring the practical applications of these devices in everyday use, it becomes crucial to prioritize the adaptation of these devices to be more user-friendly. This not only improves the efficiency of data collection but also contributes to a more positive user experience, facilitating the adoption of biofeedback technologies in real-world "outside the lab" scenarios. Moreover, the pursuit of user-friendliness should extend beyond the integration of devices. Attention should be given to the design and functionality of individual components, ensuring they are intuitive, comfortable, and adaptable to diverse user needs.

Furthermore, it is important to note that the conduction of these experiments may require a significant amount of time and be quite time-consuming. The process of preparing participants, calibrating devices and collecting data can be lengthy and intricate. Researchers should allocate sufficient resources and plan accordingly to account for the time required to conduct these biofeedback-enhanced QoE studies in VR and AR.

In summary, these preliminary findings not only provide insights into the performance and impact of the devices and multisensory stimuli used but also highlight the importance of guidelines for conducting biofeedback-enhanced QoE studies in multimedia-enhanced VR. By following these guidelines, we propose that researchers and developers ensure participant comfort, accurate biosignal capture, appropriate evaluation design, and careful implementation of multisensory stimuli, in order of enhancing the immersive experience and advancing the field of biofeedback in VR and AR environments.

ACKNOWLEDGMENTS

This study was financed partly by the Coordination for the Improvement of Higher Education Personnel (CAPES, Brazil) – Finance Code 88887.570688/2020-00 (Programa de Pós-Graduação Sanduíche no Exterior), Coordination for the Improvement of Higher Education Personnel (CAPES, Brazil) – Finance Code 88881.689984/2022-01 (Programa de Pós-Graduação Sanduíche no Exterior), National Council for Scientific and Technological Development (CNPQ, Brazil) – Finance Code 307718/2020-4 and Fundação de Amparo à Pesquisa e Inovação do Espírito Santo (FAPES, Brazil) – Finance Code 2021-GL60J

REFERENCES

[1] Mohamed Abdelazez, Sreeraman Rajan, and Adrian D.C. Chan. 2021. Automated Biosignal Quality Analysis of Electrocardiograms. *IEEE Instrumentation & Measurement Magazine* 24, 2 (2021), 37–44. https://doi.org/10.1109/MIM.2021.9400951

[2] Oluwakemi A. Ademoye, Niall Murray, Gabriel-Miro Muntean, and Gheorghita Ghinea. 2016. Audio Masking Effect on Inter-Component Skews in Olfaction-Enhanced Multimedia Presentations. *ACM Trans. Multimedia Comput. Commun. Appl.* 12, 4, Article 51 (aug 2016), 14 pages. https://doi.org/10.1145/2957753

[3] Tayebeh Baniasadi, Seyed Mohammad Ayyoubzadeh, and Niloofar Mohammadzadeh. 2020. Challenges and practical considerations in applying virtual reality in medical education and treatment. *Oman medical journal* 35, 3 (2020), e125.

[4] Jessica Oliveira Brito, Celso A. S. Santos, Rodrigo Laiola Guimarães, and Thiago Felipe Corrêa Borges. 2019. Toward Understanding the Quality of Subtitle Synchronization to Improve the Viewer Experience. In *Proceedings of the 25th Brazillian Symposium on Multimedia and the Web* (Rio de Janeiro, Brazil) *(WebMedia '19)*. Association for Computing Machinery, New York, NY, USA, 209–216. https://doi.org/10.1145/3323503.3349565

[5] Gemma A. Calvert and Thomas Thesen. 2004. Multisensory integration: methodological approaches and emerging principles in the human brain. *Journal of Physiology-Paris* 98, 1 (2004), 191–205. https://doi.org/10.1016/j.jphysparis.2004.03.018 Representation of 3-D Space Using Different Senses In Different Species.

[6] Antonio Calvo-Morata, Manuel Freire, Iván Martínez-Ortiz, and Baltasar Fernández-Manjón. 2022. Scoping review of bioelectrical signals uses in videogames for evaluation purposes. *IEEE Access* (2022).

[7] Sandra Cano, Jonathan Soto, Laura Acosta, Victor M. Peñeñory, and Fernando Moreira. 2022. Using Brain -Computer Interface to evaluate the User eXperience in interactive systems. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* (May 2022), 1–9. https://doi.org/10.1080/21681163.2022.2072398

[8] Ioan-Sorin Comşa, Estêvão Bissoli Saleme, Alexandra Covaci, Gebremariam Mesfin Assres, Ramona Trestian, Celso A. S. Santos, and Gheorghiţă Ghinea. 2020. Do I Smell Coffee? The Tale of a 360° Mulsemedia Experience. *IEEE MultiMedia* 27, 1 (2020), 27–36. https://doi.org/10.1109/MMUL.2019.2954405

[9] Alexandra Covaci, Estêvão B Saleme, Gebremariam Mesfin, Ioan-Sorin Comsa, Ramona Trestian, Celso AS Santos, and George Ghinea. 2022. Multisensory 360 videos under varying resolution levels enhance presence. *IEEE Transactions on Visualization and Computer Graphics* 29, 4 (2022), 2093–2101.

[10] Alexandra Covaci, Estêvão Bissoli Saleme, Gebremariam Mesfin, Nadia Hussain, Elahe Kani-Zabihi, and Gheorghita Ghinea. 2020. How Do We Experience Crossmodal Correspondent Mulsemedia Content? *IEEE Transactions on Multimedia* 22, 5 (2020), 1249–1258. https://doi.org/10.1109/TMM.2019.2941274

[11] Sebastião Baümberg Tavares da Silva. 2021. *A Framework for Supporting Privacy in the Computation of Biosignals*. Ph.D. Dissertation. Universidade de Lisboa (Portugal).

[12] Tarak Das, Camellia Mitra, Hritwika Paul, and Sahil Banerjee. 2022. Study the Changes of Bioelectrical Skin Impedance of Human body Associated with Different Physiological Parameters. *Journal homepage: www. ijrpr. com ISSN* 2582 (2022), 7421.

[13] Darragh Egan, Sean Brennan, John Barrett, Yuansong Qiao, Christian Timmerer, and Niall Murray. 2016. An evaluation of Heart Rate and ElectroDermal Activity as an objective QoE evaluation method for immersive virtual reality environments. In *2016 eighth international conference on quality of multimedia experience (QoMEX)*. IEEE, 1–6.

[14] Jing Gao, Peng Li, Zhikui Chen, and Jianing Zhang. 2020. A Survey on Deep Learning for Multimodal Data Fusion. *Neural Computation* 32, 5 (05 2020), 829–864. https://doi.org/10.1162/neco_a_01273 arXiv:https://direct.mit.edu/neco/article-pdf/32/5/829/1865303/neco_a_01273.pdf

[15] George Ghinea, Frederic Andres, and Stephen Gulliver. 2012. Multiple Sensorial Media Advances and Applications: New Developments in. (2012).

[16] Giorgos Giannakakis, Dimitris Grigoriadis, Katerina Giannakaki, Olympia Simantiraki, Alexandros Roniotis, and Manolis Tsiknakis. 2019. Review on psychological stress detection using biosignals. *IEEE Transactions on Affective Computing* 13, 1 (2019), 440–460.

[17] Kunal Gupta, Yuewei Zhang, Tamil Selvan Gunasekaran, Prasanth Sasikumar, Nanditha Krishna, Yun Suen Pai, and Mark Billinghurst. 2023. VRdoGraphy: An Empathic VR Photography Experience. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 1013–1014.

[18] Marc Janeras, Joan Roca, Josep A Gili, Oriol Pedraza, Gerald Magnusson, M Amparo Núñez-Andrés, and Kathryn Franklin. 2022. Using Mixed Reality for the Visualization and Dissemination of Complex 3D Models in Geosciences—Application to the Montserrat Massif (Spain). *Geosciences* 12, 10 (2022), 370.

[19] Dana Lahat, Tülay Adali, and Christian Jutten. 2015. Multimodal Data Fusion: An Overview of Methods, Challenges, and Prospects. *Proc. IEEE* 103, 9 (2015), 1449–1477. https://doi.org/10.1109/JPROC.2015.2460697

[20] Yiming Liu, Chun Ki Yiu, Zhao Zhao, Wooyoung Park, Rui Shi, Xingcan Huang, Yuyang Zeng, Kuan Wang, Tsz Hung Wong, Shengxin Jia, Jingkun Zhou, Zhan Gao, Ling Zhao, Kuanming Yao, Jian Li, Chuanlu Sha, Yuyu Gao, Guangyao Zhao, Ya Huang, Dengfeng Li, Qinglei Guo, Yuhang Li, and Xinge Yu. 2023. Soft, miniaturized, wireless olfactory interface for virtual reality. *Nature Communications* 14, 1 (May 2023). https://doi.org/10.1038/s41467-023-37678-4

[21] Olibário Machado Neto and Maria da Graça Pimentel. 2013. Heuristics for the Assessment of Interfaces of Mobile Devices. In *Proceedings of the 19th Brazilian Symposium on Multimedia and the Web* (Salvador, Brazil) *(WebMedia '13)*. Association for Computing Machinery, New York, NY, USA, 93–96. https://doi.org/10.1145/2526188.2526237

[22] Arian Mehrfard, Javad Fotouhi, Giacomo Taylor, Tess Forster, Nassir Navab, and Bernhard Fuerst. 2019. A comparative analysis of virtual reality head-mounted display systems. *arXiv preprint arXiv:1912.02913* (2019).

[23] Mina Mikhail, Khaled El-Ayat, Rana El Kaliouby, James Coan, and John J. B. Allen. 2010. Emotion Detection Using Noisy EEG Data. In *Proceedings of the 1st Augmented Human International Conference* (Megève, France) *(AH '10)*. Association for Computing Machinery, New York, NY, USA, Article 7, 7 pages. https://doi.org/10.1145/1785455.1785462

[24] Niall Murray, Oluwakemi A. Ademoye, Gheorghita Ghinea, and Gabriel-Miro Muntean. 2017. A Tutorial for Olfaction-Based Multisensorial Media Application Design and Evaluation. *ACM Comput. Surv.* 50, 5, Article 67 (sep 2017), 30 pages. https://doi.org/10.1145/3108243

[25] Niall Murray, Brian Lee, Yuansong Qiao, and Gabriel-Miro Muntean. 2014. Multiple-Scent Enhanced Multimedia Synchronization. *ACM Trans. Multimedia Comput. Commun. Appl.* 11, 1s, Article 12 (oct 2014), 28 pages. https://doi.org/10.1145/2637293

[26] Diogo Kionori Cândido Nishikawa, Roberta Pereira Brandão, and Vânia Paula de Almeida Neris. 2020. Um estudo empírico sobre reações emocionais de usuários na interação com interfaces web pautadas na Gestalt. In *Anais do XI Workshop sobre Aspectos da Interação Humano-Computador para a Web Social*. SBC, 25–32.

[27] Pratheep Kumar Paranthaman, Nikesh Bajaj, Nicholas Solovey, and David Jennings. 2021. Comparative Evaluation of the EEG Performance Metrics and Player Ratings on the Virtual Reality Games. In *2021 IEEE Conference on Games (CoG)*. IEEE, 1–8.

[28] Szczepan Paszkiel and Szczepan Paszkiel. 2020. Using BCI and VR technology in neurogaming. *Analysis and Classification of EEG Signals for Brain–Computer Interfaces* (2020), 93–99.

[29] Jiali Qian, Daniel J McDonough, and Zan Gao. 2020. The effectiveness of virtual reality exercise on individual's physiological, psychological and rehabilitative outcomes: a systematic review. *International journal of environmental research and public health* 17, 11 (2020), 4133.

[30] Plínio M.S. Ramos, Caio B.S. Maior, Márcio C. Moura, and Isis D. Lins. 2022. Automatic drowsiness detection for safety-critical operations using ensemble models and EEG signals. *Process Safety and Environmental Protection* 164 (2022), 566–581. https://doi.org/10.1016/j.psep.2022.06.039

[31] Giuseppe Riva, Brenda K Wiederhold, and Fabrizia Mantovani. 2019. Neuroscience of virtual reality: from virtual exposure to embodied medicine. *Cyberpsychology, behavior, and social networking* 22, 1 (2019), 82–96.

[32] Theerat Saichoo, Poonpong Boonbrahm, and Yunyong Punsawad. 2021. Facial-Machine interface-based virtual Reality Wheelchair control using EEG artifacts of Emotiv neuroheadset. In *2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. IEEE, 781–784.

[33] Estêvão B. Saleme, Alexandra Covaci, Gebremariam Mesfin, Celso A. S. Santos, and Gheorghita Ghinea. 2019. Mulsemedia DIY: A Survey of Devices and a Tutorial for Building Your Own Mulsemedia Environment. *ACM Comput. Surv.* 52, 3, Article 58 (jun 2019), 29 pages. https://doi.org/10.1145/3319853

[34] Débora Pereira Salgado, Felipe Roque Martins, Thiago Braga Rodrigues, Conor Keighrey, Ronan Flynn, Eduardo Lázaro Martins Naves, and Niall Murray. 2018. A QoE assessment method based on EDA, heart rate and EEG of a virtual reality assistive technology system. In *Proceedings of the 9th ACM Multimedia Systems Conference*. 517–520.

[35] Marcelle Schaffarczyk, Bruce Rogers, Rüdiger Reer, and Thomas Gronwald. 2022. Validity of the polar H10 sensor for heart rate variability analysis during resting state and incremental exercise in recreational men and women. *Sensors* 22, 17 (2022), 6536.

[36] Catriona L. Scrivener and Arran T. Reader. 2022. Variability of EEG electrode positions and their underlying brain regions: visualizing gel artifacts from a simultaneous EEG-fMRI dataset. *Brain and Behavior* 12, 2 (Jan. 2022). https://doi.org/10.1002/brb3.2476

[37] Jungryul Seo, Teemu H Laine, and Kyung-Ah Sohn. 2019. An exploration of machine learning methods for robust boredom classification using EEG and GSR data. *Sensors* 19, 20 (2019), 4561.

[38] Tanuja Subba and Tejbanta Singh Chingtham. 2022. A Review on Types of Machine Learning Techniques for Biosignal Evaluation for Human Computer Interaction. *Advanced Computational Paradigms and Hybrid Intelligent Computing: Proceedings of ICACCP 2021* (2022), 457–466.

[39] Moritz Tacke, Katharina Janson, Katharina Vill, Florian Heinen, Lucia Gerstl, Karl Reiter, and Ingo Borggraefe. 2022. Effects of a reduction of the number

of electrodes in the EEG montage on the number of identified seizure patterns. *Scientific Reports* 12, 1 (2022), 1–7.

[40] Tijana Vuletic, Alex Duffy, Laura Hay, Chris McTeague, Gerard Campbell, and Madeleine Grealy. 2019. Systematic literature review of hand gestures used in human computer interaction interfaces. *International Journal of Human-Computer Studies* 129 (2019), 74–94.

[41] Efy Yosrita, Yaya Heryadi, Lili Ayu Wulandhari, and Widodo Budiharto. 2019. EEG Based Identification of Words on Exam Models with Yes-No Answers for Students with Visual Impairments. In *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*. IEEE, 1–5.