Neurohumanities Lab: Physiological Signal Analysis Within an Educational Partially Immersive Environment

Rebeca Mellado-Reyes[1], Dr. Alexandro Ortiz[2], Dr. Carol S. Vélez-Saboyá[2], Dr. Jorge De-J. Lozoya-Santos[1], Dr. Mauricio A. Ramírez-Moreno[1], Dr. Manuel Cebral-Loureda[2*]

[1]Mechatronics Department, School of Engineering and Sciences, Tecnologico de Monterrey, Monterrey NL 64849, MX
[2]Department of Humanistic Studies, School of Humanities and Education, Tecnologico de Monterrey, Monterrey NL 64849, MX
Email: manuel.cebral@tec.mx

Abstract

The use of immersive virtual reality technologies in education has demonstrated an improvement in the learning process of professional level students; nevertheless, mostly limited to the fields of science and engineering. In this study, the Neurohumanities Lab was introduced as a semi-immersive space where, differences in electroencephalography and heart rate as physiological signals, plus the statistical results from the ITC-SOPI presence questionnaire, were analyzed. These results were compared to those of a traditional class. Supervised Machine Learning algorithms were tested, and the engagement ratio plus the power extraction in the gamma band were the most significant features, with 92.34% accuracy on average. Heart rate variations were related to changes in the state of presence, also observed by the questionnaire responses' results. Concluding that the Neurohumanities Lab has the potential to be a completely immersive environment, enhance the learning experience in the humanities area, and evaluate learning in an objective way.

Keywords: Emotions, Education, Humanities, Machine Learning, OpenBCI.

1. Introduction

Human sciences encompass all aspects of human society and culture, including language, literature, philosophy, law, politics, religion, art, history, and social psychology. The connection between specific domains of neuroscience and their counterparts within humanities has been previously emphasized. The partnership created between these fields will not only be mutually beneficial but also create a new dimension of significant and impactful scholarly activity [1]. However, the teaching of humanities has made slow progress in terms of educational innovation, especially when compared to other areas such as science and engineering [2].

Technologies that have been implemented with this purpose within the fields of engineering and science, include semi-immersive, immersive, and virtual reality spaces, which have been a popular option in response to their good results in improving learning [3]; but few studies have taken into account these tools in the humanities field. As a result, the Neurohumanities Laboratory (NH Lab) aims to implement a semi-immersive and interactive platform for humanities education. This platform will allow users to interact with the environment (the classroom) and have an impact on it in a way that aligns logically with the content being taught. For this purpose, the NH Lab includes a system that, with cameras and wearable biometric devices, can detect movements, emotion, and physiological and mental states, which can modify the environment.

Background

The study of emotions is fundamental to the understanding of human experience, how it is related to feelings, behaviors and thoughts. It is important to study how emotions influence cognition, learning, communication, and rational decision-making [4], [5]. Taking this into consideration, a machine learning, biometry-based, real-time emotion detection module based on Descartes passions is included as a fundamental element of the laboratory. It was trained with the DEAP Dataset [4] to be able to classify between six fundamental emotions: admiration, love, hate, desire, joy, and sadness [6], resulting in a 93% accuracy model [7].

The emotion classification process is based on the acquisition of electroencephalography (EEG), in this case, from an 8-electrode OpenBCI. Brain activity can be divided into frequency bands: Delta (0-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (12-30 Hz), and Gamma (30 Hz). Each band is linked to specific mental states, from sleep to high cognitive performance [8]. On the other hand, Heart Rate (HR), which is obtained from Inter-beat intervals (IBI) and blood volume pressure (BVP) through photoplethysmography (PPG) measures, are used to obtain the physiological response of each participant.

From this biometry acquisition module, the engagement index can be obtained quantitatively, which reflects a better immersion state and thus, an indicator of improved learning in virtual environments [9]. The ITC-Sense of Presence Inventory (ITC-SOPI) questionnaire [10] was used to evaluate this variable based on the participant's perception of presence, as it measures spatial presence of users' experience in a specific environment. The exploratory analysis of the data obtained from this study considers ecological validity, engagement, and behavioral effects related to an immersive technology.

Objectives

The NH Lab's objective, within the limitations of this study is to execute a 24-participant experimental protocol to analyze and compare biometric, physiological and subjective data from (12) participants within this space in comparison to those (12) of a traditional classroom setting. Secondarily, its aim is to test the emotion recognition feedback loop previously developed [7]; and use Machine Learning algorithms to identify EEG signals for better group (immersive vs traditional) differentiation.

Contributions

The main contribution of this work is the development of a semi-immersive environment that incorporates activities related to the understanding of a humanistic topic. Accordingly, it addresses the need for more technology-based education within this field to improve university students' learning experience in terms of engagement, personalization, and overall effectiveness. Also, a Machine Learning algorithm was designed as an evaluation method between control and experimental group, offering a numerical method of interpreting EEG signals in the context of Neurohumanities.

2. Methods and Methodology:

Data Collection

Before starting the experiment, 24 participants of ages 18 to 24 were: divided into experimental (7 men and 5 women) and control group (2 men and 10 women); asked about their dominant hand and age; and asked to answer two initial online questionnaires: General Health and 24-item version of the Trait Meta-Mood Scale (TMMS-24) for emotional intelligence. Secondly, participants were placed the Empatica E4 (Empatica, Milan, IT); and an UltraCortex Mark IV headset manufactured by OpenBCI, for which channels Fp1, Fp2, C3, C4, P7, P8, O1, and O2 were selected. The prior registered BVP at 64 Hz, electrodermal activity (EDA) at 4 Hz, body temperature at 1 Hz, and 3-axis acceleration data at 32 Hz. The latter registered EEG and acceleration data at 250 Hz. Thirdly, adjustments were made to register impedance values lower than 750kΩ (as suggested by the OpenBCI_GUI software). Then the participant was asked to blink five times and clench their jaw to ensure that the EEG signal was being recorded correctly. Fourthly, a 1-minute open eyes and 1-minute closed eyes resting-state EEG recording was conducted. The devices were to be worn during the whole experiment, which was made up of four 3-minute scenes, each consisting in different tasks (see more in Fig. 1). In the end, participants were asked to complete the ITC-SPOI questionnaire.

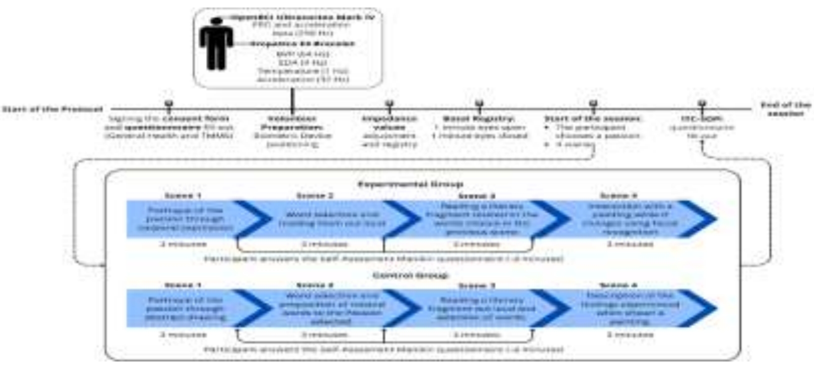

Figure 1: Experimental protocol diagram, including short description of each scene with modifications per group. Self-Assessment Manikin questionnaire plus EDA, temperature, and acceleration variables were not considered in this report.

This experimental protocol was approved by the institution's ethic committee "Comité Institucional de Ética en Investigación (CIEI) del Instituto Tecnológico y de Estudios Superiores de Monterrey" with an identification code: EHE-2023-03.

Experimental Setup

The physical space of the NH Lab consists of a 45 m^2 room, of which 20 m^2 are destined for the semi-immersive environment. It includes a Blackmagic Studio Camera 4K Plus (Blackmagic Design Pty. Ltd, Singapore) for the computer vision algorithms, a screen projector to display immersive visualizations, LED lights, and speakers (see Fig. 2-A). To run the whole experience, 4 computers were needed, plus a tablet given to answer the questionnaires. On the other hand, the experimental setup for the control group consisted of a small room with a table, chairs, printouts for each scene, color markers and pens (see Fig. 2-B). The codes for data recording, real-time feedback, and data sending to additional computers were made in Python. The data transfer between computers was facilitated through the OSC-Protocol. Additionally, an interface was implemented to synchronize different codes corresponding to each scene, managed either by the software Touch Designer for visual stimuli or Pure Data for audio and multimedia.

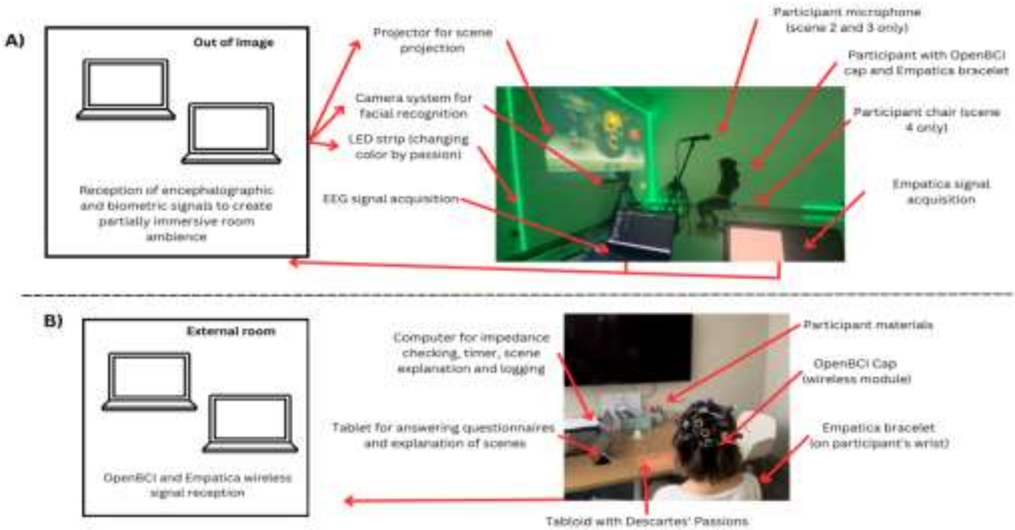

Figure 2: Experimental Setup: A) experimental group's semi-immersive setup during the 4th scene, where the color outputted from the LED strips is controlled by an Arduino UNO microcontroller. B) control group's area setup within a small, regular classroom, with less technological requirements.

Scene Descriptions

The experience is made up of 4 activities or scenes related to the understanding of a humanistic topic: Descartes's passions. The activities involve the participants choosing one of the passions (admiration, love, hate, desire, joy, and sadness) and base their experience on it. These scenes are included in Fig. 1.

In the first scene, participants represented the selected emotion by moving and interacting with the space. Based on the emotion and movement, the color of the lights, sound and visual elements changed. The real-time motion detection elements were coded with Python's library Mediapipe. For the control group, participants were instructed to draw the passion chosen as they perceived and felt it.

During the second scene, an Artificial Intelligence module was implemented to generate a list of words associated with the passion chosen, to portray them on the screen as a cloud of words. The participants received a standing microphone and were asked to say out loud the words that best represented the passion they chose. They could also propose words that were not included by the system, only if they considered it to be accurately descriptive. Python's library Whisper speech-to-text function was used to recognize words and store them during the experiment. For the control group, the participants were given a paper sheet with a cloud of words associated with the passion they chose and were asked to encircle the words or add new ones if needed.

For the third scene, a database of 60 literary pieces was assembled, from which an algorithm extracted short fragments related to the words previously chosen and displayed them on the screen one by one. The participant was asked to read the fragments out loud into the microphone in the most emphatic and comprehensible manner. The control group was provided with a sheet of paper where several fragments were written and were asked to read the fragments out loud and highlight the words that caught their attention. It must be noted that these paper sheets, as those from scene 2, were the same for each emotion, and did not vary for each participant, as in the experimental group.

During the last scene, the participant was asked to observe a painting with a skull in the center and objects on either side and explore the emotions it causes in them by interacting facially. The shape, color, and type of objects in the painting changed depending on the way they felt, and a skull mirrored their facial features, a feature enabled by Python's library Deep Face with face swapping the participant's face with a skull. The control group were shown the same painting but printed and were asked to describe on a sheet of paper the feelings experienced when observing that painting.

Data Processing

For the EEG data the pre-processing and Machine Learning processing made in MATLAB, the procedure illustrated in Fig. 3 was followed. For the initial cleaning, a MATLAB plug-in, PREP Pipeline, was used, followed by a 0.1-50 Hz bandpass filter. Finally, two artifact removal methods were employed: Artifact Subspace Reconstruction (ASR) to remove transient or large amplitude artifacts, and Independent Component Analysis (ICA) to differentiate and highlight artifacts to be later removed.

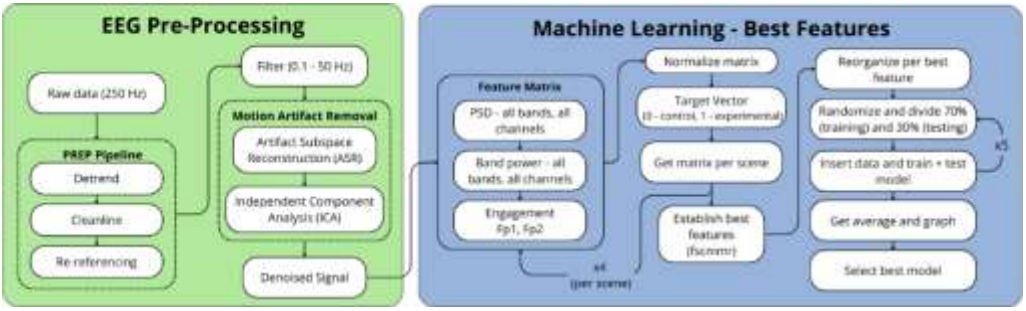

Fig. 3: EEG Data pre-processing and Machine Learning processing flow diagram.

The cleaned signal is then processed to obtain the total power and the Power Density Spectrum (PSD), calculated with the Welch method, from all frequency bands in the 8 electrodes. Also, the engagement index was obtained from the average of frontal electrodes, using the formula: Beta/Theta+Alpha. The 82 features that resulted from this processing step were selected due to their relationship with presence and engagement in an immersive state [11]. Subsequently, these features were placed in an 83-column matrix, where the last column represented the target vector identifying the sample as class 0 (control) or class 1 (experimental).

The Maximum Relevance Minimum Redundancy (MRMR) algorithm was applied to the created matrix to sort the features by the relevance weights for a better class distinction. From this step, the 10 best features per scene were obtained for model testing. Afterward, the new matrix created was randomly divided (70%-training, 30%-testing) and looped with three algorithms: decision tree and linear and quadratic discrimination. Overfitting was avoided with 5-fold cross-validations, which consisted of randomizing the training and testing matrices, and obtaining the mean accuracy of the algorithm. To evaluate the number of features needed for the highest accuracy, the model was tested starting at 2 features until reaching the 82, ranking them in descending order. On the other hand, the HR was calculated from the BVP data obtained with the Empatica at 64 Hz by finding its peaks and the timespan between each.

3. Results and Discussion:

After performing the feature extraction process, the results are divided into the different classification algorithms and the features used for each. The 10 best features obtained from each scene through the MRMR algorithm are presented in Table 1. The results highlight that the engagement of one or both of the frontal electrodes is a differentiator for class distinction, as it was included in this selection. The level of engagement was also plotted for each scene and each group, as seen in Fig. 4. The results vary with each scene and each group, but overall were higher for the experimental or immersive group, excepting scene 1. This means that the level of engagement depends on the type of activity and the presence of the participant, which is also related to the level of learning: the higher level of engagement, the higher the level of learning [9]. The fact that in the first scene, the engagement is lower for the experimental group can be an indication that the activity was not as stimulating.

Table 1: Selection of the 10 best features per scene after being sorted by the MRMR algorithm.

Scene 1	Scene 2
FP2_beta	C3_alpha
fp1_engagement	O2_delta_psd
O1_delta	O2_gamma
C4_delta_psd	fp2_engagement
FP2_gamma	C3_delta
P8_theta_psd	FP1_gamma_psd
P7_delta	P7_theta
FP1_gamma_psd	P8_beta_psd
C4_alpha	O2_theta
O2_gamma_psd	C4_delta_psd
Scene 3	**Scene 4**
FP2_beta	P8_delta
C3_gamma_psd	fp2_engagement
fp1_engagement	FP1_gamma
P8_delta_psd	O1 _theta _psd
O1_alpha_psd	P7_alpha
FP1_alpha	P8_beta
FP2_theta	P8_theta_psd
P8_theta_psd	FP2_alpha_psd
O1_delta	O2_beta
FP2_alpha	C4_gamma

Likewise, when limiting the number of features to 10, the decision tree algorithm outperformed the other two with similar accuracy percentages, as with an undefined number of features. Upon examining the results of linear and quadratic discrimination methods, the accuracy values remain low even with 10 features. This means that the model will predict with a mean accuracy of 92.34% with 10 features or 93.71% with more than 60 features (Table 2), whether a participant is learning within a traditional or a semi-immersive environment.

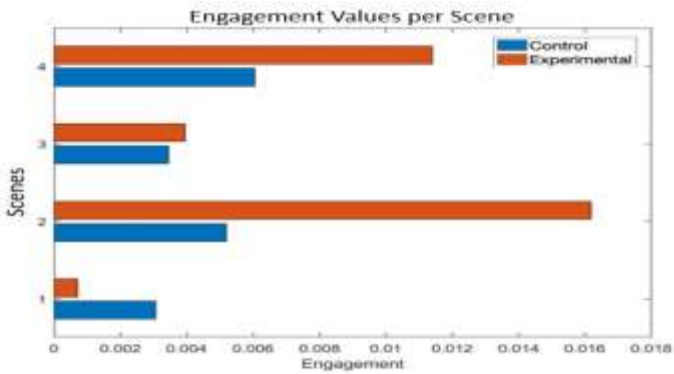

Figure 4: Normalized values of engagement across the four scenes within control and immersive groups

Table 2: Number of features with the highest average accuracy for the Machine Learning models, plus highest average accuracy when using the best 10 features per algorithm.

Scene	Decision Tree	Linear Discrimination	Quadratic Discrimination	Decision Tree (10 features)	Linear Discrimination (10 features)	Quadratic Discrimination (10 features)
1	76 features 94.77%	9 features 57.07%	64 features 81.93%	92.80%	56.03%	62.57%

2	80 features 92.79%	8 features 52.60%	79 features 82.93%	91.82%	49.15%	52.05%
3	81 features 93.80%	34 features 55.58%	80 features 69.23%	92.27%	53.01%	54.69%
4	65 features 93.48%	32 features 56.86%	82 features 63.61%	92.44%	55.33%	56.95%

On the other hand, the HR, that was calculated from the BVP increases considerably more for the control group in comparison to the experimental group, as seen in Figure 5. It should be noted that, at no time, the heart rate is lower than the baseline value.

Figure 5: Average heart rate for each scene in Beats per minute.

Nevertheless, studies show that the HR is indirectly proportional to the presence of participants: the lower the HR, the greater the presence [11]. This difference is mostly noted in the experimental groups during scenes 1, 2 and 4, meaning that participants experience greater presence when expressing bodily a passion, mentioning words out loud, or interacting with a live painting, in comparison to creating an abstract drawing, writing words, or analyzing a painting, respectively.

Nevertheless, the results from scene 3 contradict those of the engagement: the control group showed higher presence than their counterparts. Therefore, reading aloud the literary quotes and selecting words make participants feel more present than reading literary quotes in the experimental group without word selection. This demonstrates that engagement is not necessarily directly related to presence.

Finally, the ITC-SOPI questionnaire shows the average of the results in four categories to compare different aspects of the participants' experience, which is part of the methodology of the questionnaire [8]. Figure 6 shows the results. As they seem very similar, the ANOVA test was used to test the significance of the data obtained. The p-level obtained for spatial presence, engagement, naturalness, and negative effects are as follows: 0.011, 0.135, 0.005, and 0.344 respectively.

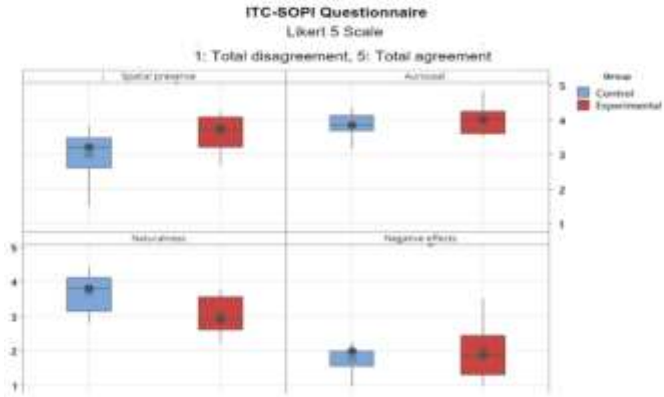

Figure 6: Results of the ITC-SOPI questionnaire comparing the different categories with both groups used (control and experimental), where the dot indicates: media and the cross: arithmetic mean shown in Minitab.

These results indicate that engagement and negative effect categories do not show significant results to be able to make a comparison between both groups. Nevertheless, the results obtained in the categories of spatial presence and ecological validity, or naturalness show significant differences. A higher value in spatial presence is related to the quality of immersion and presence in the experiment, while a higher naturalness indicates how real the experimentation is or how it is assimilated to real life [6].

4. Conclusion

Within the field of humanities, the learning process and evaluation methods that come along cannot be entirely objective, as emotions and subjective variables come along with it, and alter the information presented. With this study, it was found that there is indeed a way in which the educational and learning process of university students in humanities can be evaluated. The physiological states, based on EEG and HR data, can be an objective method to evaluate the learning process of a student, through presence and engagement within a partially immersive environment. The association can be made between these results and the learning and understanding of a humanistic topic. However, the field of study is broad and there are still limitations in the research that must be taken into account for future analysis. These limitations include the type of data recollected, and the analysis made in this specific study, as more data was recollected but not analyzed in this work, such as the Self-Assessment Manikin questionnaires, the most relevant features, and detected emotions during the experiments. Therefore, this paper is only the first part of the NH Lab's results.

In addition, the supervised Machine Learning model showed a high accuracy with 10 features, enough to demonstrate that there is evidently a difference to be noted between traditional and technology-using educational settings. Consequently, all these findings can provide a better understanding of the implications of teaching methods in the humanities and the effectiveness of the use of immersive technologies in learning. In the future, it is expected that the use of this first designed space can be scaled to a completely immersive version with different humanities-related tasks and feedback methods, and more complex and accurate real-time classification models. Also, regarding the classification model, a future implementation can be adapted to real-time measurements to obtain more information during the learning process of students and provide assessments in the context of emotional intelligence or mental health conditions.

Acknowledgement:

We Due to the limitation on the number of authors allowed in this journal, we could not include Alejandro Villarreal-Villarreal, Mei Li L. Cham-Perez, Rebeca Romo-De León, Verónica A. Elizondo-Villegas, Alejandro Galindo for their contributions to this work.

Funding Statement:

This research was supported by Tecnologico de Monterrey and CONAHCYT Mexico, grant numbers # E061 - EHE-GI02 - D-T3 – E and CF-2023-G-583.

Conflict of interest: The authors declare that there is no conflict of interest.

WORKS CITED

T. J. Carew and M. Ramaswami, "The Neurohumanities: An Emerging Partnership for Exploring the Human Experience", Neuron, vol. 108, núm. 4, pp. 590–593, nov. 2020, doi: https://10.1016/j.neuron.2020.10.019.

M. Bond, K. Buntins, S. Bedenlier, O. Zawacki-Richter, y M. Kerres, "Mapping research in student engagement and educational technology in higher education: a systematic evidence map", Int. J. Educ. Technol. High. Educ., vol. 17, num. 1, p. 2, dic. 2020, doi: https://10.1186/s41239-019-0176-8

J. Hutson and T. Olsen, "Digital Humanities and Virtual Reality: A Review of Theories and Best Practices for Art History", Int. J. Technol. Educ., vol. 4, n´um. 3, pp. 491–500, jul. 2021, doi: https://10.46328/ijte.150

S. Koelstra et al., "DEAP: A Database for Emotion Analysis; Using Physiological Signals," IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18–31, Jan. 2012, doi: https://doi.org/10.1109/t-affc.2011.15

M. Mikhail, K. E. Ayat, J. A. Coan, y J. J. B. Allen, "Using minimal number of electrodes for emotion detection using brain signals produced from a new elicitation technique", Int. J. Auton. Adapt. Commun. Syst., vol. 6, n´um. 1, p. 80, 2013, doi: https://10.1504/IJAACS.2013.050696

R. Descartes, Las pasiones del alma, vol. 290. Edaf, 2010.

M. A. Blanco-Ríos et al., "Real-time EEG-based emotion recognition for neurohumanities: perspectives from principal component analysis and tree-based algorithms," Frontiers in human neuroscience, vol. 18, Mar. 2024, doi: https://doi.org/10.3389/fnhum.2024.1319574 .

Md. R. Islam et al., "Emotion Recognition From EEG Signal Focusing on Deep Learning and Shallow Learning Techniques", IEEE Access, vol. 9, pp. 94601–94624, 2021, doi: https://10.1109/ACCESS.2021.3091487

C. Ochs y A. Sonderegger, "The Interplay Between Presence and Learning", Front. Virtual Real., vol. 3, p. 742509, feb. 2022, doi: https://10.3389/frvir.2022.742509

J. Lessiter, J. Freeman, E. Keogh, y J. Davidoff, "A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory", Presence Teleoperators Virtual Environ., vol. 10, n´um. 3, pp. 282–297, jun. 2001, doi: https://10.1162/105474601300343612

S. Grassini y K. Laumann, "Questionnaire Measures and Physiological Correlates of Presence: A Systematic Review", Front. Psychol., vol. 11, p. 349, mar. 2020, doi: https://10.3389/fpsyg.2020.00349