# Estimation of Affective States in Virtual Reality Environments using EEG


Meghan Kumar
Virginia Commonwealth University
Richmond, Virginia, USA
kumarm2@vcu.edu

Connor Delaney
Virginia Commonwealth University
Richmond, Virginia, USA
delaneycm2@vcu.edu

Dean J. Krusienski
Virginia Commonwealth University
Richmond, Virginia, USA
djkrusienski@vcu.edu



## ABSTRACT

Recent developments in accessible virtual reality (VR) technologies have inspired research into increasing the user's sense of immersion via monitoring and feedback of physiological and cognitive states. This work presents a pilot study that models the user's affective state from electroencephalographic (EEG) measurements during observation of immersive VR videos. Participants passively viewed a series of short VR video clips and subjectively rated the level of valence, arousal, liking, and dominance. Separate regression and classification models were developed to estimate the subjective ratings for each category using EEG spectral features. The results indicate that both models are capable of capturing relevant EEG features for estimating the user's affective state ratings.



## KEYWORDS

virtual reality, electroencephalogram, self-assessment manikin, support vector machine

**ACM Reference Format:**

Meghan Kumar, Connor Delaney, and Dean J. Krusienski. 2022. Estimation of Affective States in Virtual Reality Environments using EEG. In *The15th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '22), June 29-July 1, 2022, Corfu, Greece.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3529190.3534738


## 1 INTRODUCTION

Through the incorporation of physiological measurements and feedback into virtual reality (VR) systems, it is possible to greatly enrich the user experience and functionality of VR applications [2, 4]. Recent brain-computer interface (BCI) studies have used electroencephalogram (EEG) sensors in conjunction with VR headsets to estimate human factors such as cognitive and affective states [7, 11, 13]. Such studies commonly detect and track power fluctuations in traditional EEG spectral bands (i.e., $\delta, \theta, \alpha, \beta, \gamma$) that correlate with modulation of particular cognitive and affective states, such as estimating cognitive workload during an $n$-back task [16].

The present pilot study extends the characterization of EEG features using the data set from [9] to develop and evaluate models that estimate subjective ratings of affective state from EEG recorded during passive observation of immersive VR videos. EEG data were collected from 10 participants during observation of 13 publicly available 360° VR videos [12]. Following each video, participants were asked to provide subjective ratings of their affective state based on the Russel Complex [14, 15]. The Russel Complex can be mapped into a two-dimensional space consisting of axes representing valence and arousal, respectively. For this experiment, the Russel Complex was extended to include dimensions representing liking and dominance [8]. Each axis is assigned a rating that ranges from one to nine. This extended Russel Complex was presented as a Self-Assessment Manikin (SAM) in the virtual environment for the participant to report his/her affective state for each video. The SAM utilizes a nonverbal pictorial assessment technique based on the dimensions of the Russel complex [3, 15].

EEG features were used to create a stepwise regression model to estimate the SAM ratings, as well as a support vector machine (SVM) model to classify the SAM ratings as above or below a threshold. The results indicate that both models are capable of estimating the user's affective state ratings based on EEG features, which will be informative for the development of future closed-loop affective-state BCIs.

## 2 METHODOLOGY

### 2.1 Participants and Experimental Setup

Ten healthy individuals (ages 19-34, mean 23.8) were recruited to participate in the experiment, which was approved by the Institutional Review Board of Virginia Commonwealth University. Each participant completed a screening process consisting of an informed consent form, demographic information form, and Motion Sickness Susceptibility Questionnaire (MSSQ) short form [5]. All participants satisfied the criteria by scoring a minimum of 19 on the MSSQ. Participants were provided guidelines to complete the SAM in terms of valence, arousal, dominance and liking prior to beginning the experimental task [3].

The HTC VIVE hardware system consists of a motion-tracked headset display and two "lighthouse" base stations that can provide 6 Degree of Freedom (6DoF) tracking. The VIVE wireless adapter was used in conjunction with the wireless EEG headset such that the participant was untethered and free to rotate while seated to view the 360° videos.

Following the participant screening, a visual calibration was performed using the HTC VIVE to correct for lens distance. The wireless 32-channel EEG cap was then placed on the participant's head and the EEG electrodes were filled with electrolyte gel. The electrode cap was covered with a protective plastic hair cap to protect the VR headset from the gel. The VR headset was placed over the EEG cap and the headset was tightened to comfortably fit the participant, as shown in Figure 1. Participants were then




*PETRA '22, June 29-July 1, 2022, Corfu, Greece*



ACM ISBN 978-1-4503-9631-8/22/06.
https://doi.org/10.1145/3529190.3534738

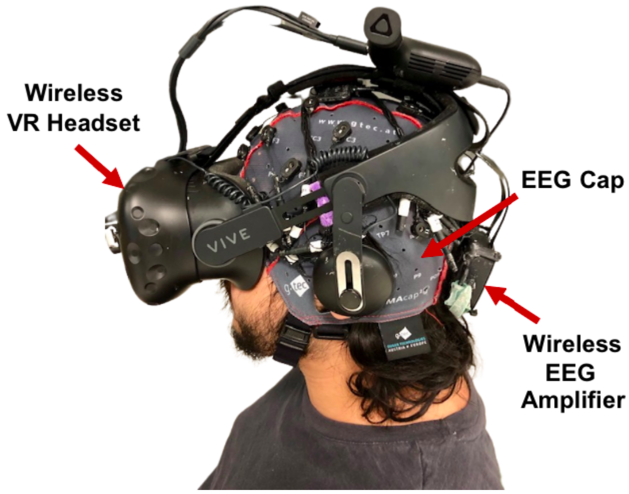


**Figure 1: Placement of wireless EEG cap and VR headset on participant [9].**

positioned approximately 1 meter from the recording computer in a seated position, in a swivel office chair.

### 2.2 Experimental Task

The stimuli are a series of twelve 360° videos viewed through the VR headset to induce various affective reactions. The videos were obtained from a public database which were used in a previous human emotion study [12]. The videos from the database were grouped into three categories based on the prior study's participant ratings: High-Valence-Low-Arousal (HVLA), High-Valence-High-Arousal (HVHA), and Low-Valence-Low-Arousal (LVLA). Each category consisted of 4 one-minute videos for a total of 12 video segments, and one neutral video to collect baseline EEG data at the beginning of the experiment.

At the start of the experiment, the participant viewed the neutral video and completed the SAM. The remaining 12 videos were then shown in a randomized order for each participant such that no two videos from the same affective category were not shown in sequence. For an individual video trial, the participant's task was to passively view a video, with the ability to rotate the chair and the headset to view the 360° environment from different perspectives. Immediately following the video, participants were prompted to complete the SAM using a scale of one to nine for each respective dimension of valence, arousal, dominance, and liking. The icons displayed for the SAM are shown in Figure 2 [3]. Each row represents a different dimension of affective state, and each column represents a different rating. The participant selects the ratings by directing their head toward the desired manikin and maintaining the cursor over the manikin for a 5-second dwell time, as indicated by a progress bar. A twenty-second interval between trials was used to collect inter-trial baseline EEG data. During this interval, participants were asked to keep their eyes open and remain still. The total duration of the experiment was kept to 25 min to reduce the likelihood of simulator sickness [6], which was not reported by any of the participants.

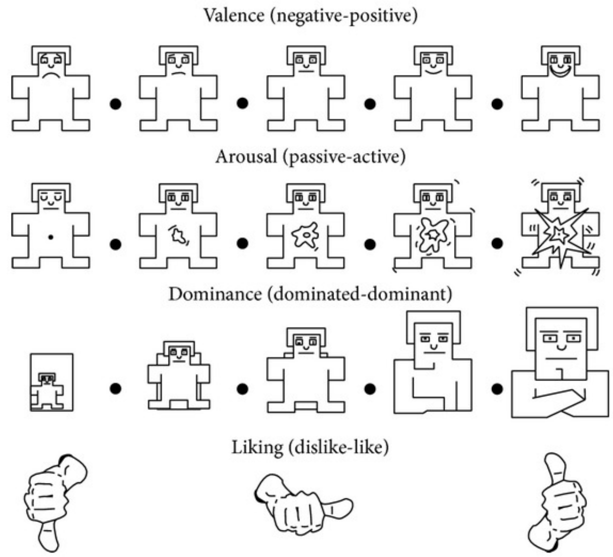


**Figure 2: Self-assessment manikin (SAM) for affective states [8].**

### 2.3 Data Collection

EEG data were collected using a 32-channel wireless bio-signal amplifier (g.Nautilus, Guger Technologies), with electrode positions based on the International 10-20 system as shown in Figure 3. The recordings were grounded to location AF3 and referenced to the right earlobe. The sampling rate was 250 Hz with a bandpass filter from 0.1-100 Hz and a notch filter between 58-62 Hz. Additionally, the position and acceleration data from the VR headset were recorded for analysis of movement artifacts. Communication between the VR environment (developed in Unity [17]) and the EEG recording was performed via Lab Streaming Layer [10] and recorded using the Lab Recorder application.

## 3 DATA ANALYSIS

### 3.1 Artifact Suppression

Because the participants were allowed to move their heads and swivel in the chair to explore the VR environments, EEG artifact suppression methods were implemented to reduce the influence of movement and other potential artifacts. The EEG data were bandpass filtered between 4-47 Hz, according to the frequency bands used for the subsequent feature extraction. The EEG and corresponding headset movement data were segmented into 12 one-minute trials for each participant. In order to correct for unrelated stimulus variation in power, the baseline EEG signal from the 20 second inter-trial intervals was used. Second Order Blind Identification (SOBI) Independent Component Analysis (ICA) [1] was applied to parse the EEG data into statistically independent components. The Spearman correlation coefficient was computed between each individual component and each axis of the headset acceleration data. Independent components with a Spearman correlation coefficient outside two standard deviations from the mean

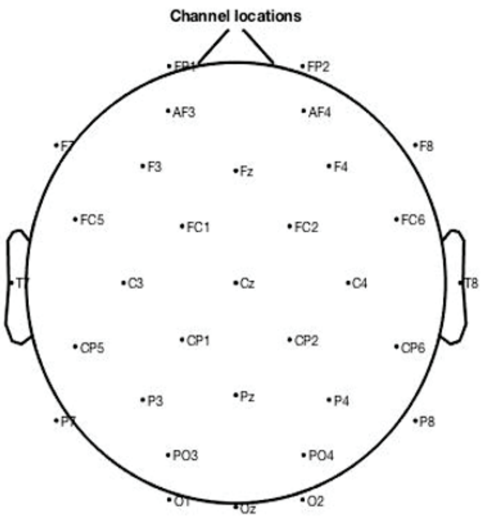


**Figure 3: The EEG electrode locations based on the International 10-20 system.**

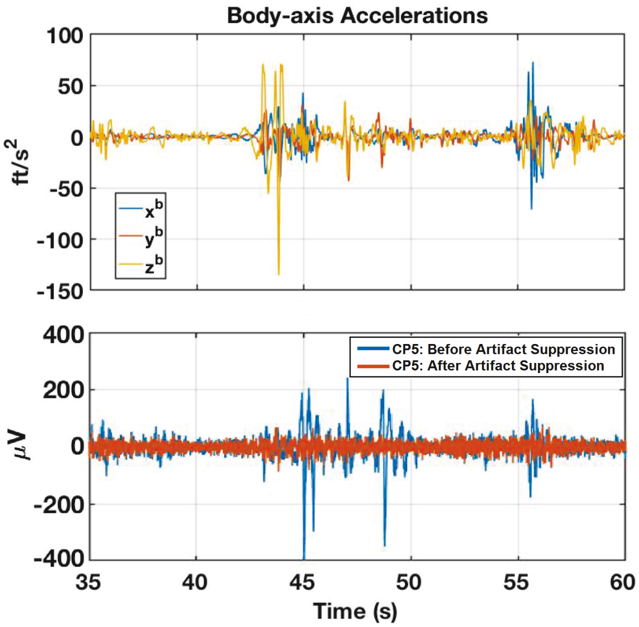


**Figure 4: Example of accelerometer data and corresponding artifact suppression. Top panel: 3-axis accelerometer traces during a representative video trial. Bottom panel: corresponding EEG traces from a representative channel (CP5) before and after artifact suppression.**

correlation coefficient were selected for removal, and the EEG signal was then reconstructed using retained components. Example accelerometer data and corresponding artifact suppression for a representative EEG channel are provided in Figure 4. It can be observed that larger fluctuations in the accelerometer data temporally correspond with larger fluctuations in the EEG recording, which are effectively suppressed using the proposed approach.

## 3.2 Feature Extraction

The power spectral density of the EEG signal for each trial and baseline period were computed using Welch's Method on 5-second non-overlapping windows. The baseline power was subtracted from trial power, yielding the change of power relative to pre-stimulus period. The net changes of power spectral density (PSD) were averaged over 4 frequency bands $\theta$-band (4-7 Hz), $\alpha$-band (8-13 Hz), $\beta$-band (14-29 Hz), and $\gamma$-band (30-47 Hz) and used as the feature vector for each observation window.

## 3.3 Affective State Estimation Models

*3.3.1 Stepwise Linear Regression Models.* A stepwise linear regression model was constructed to estimate the affective state rating as a function of the EEG features from each trial. Because of the high-dimensional feature space, a stepwise regression routine was utilized to identify statistically significant model terms. The pool of candidate explanatory variables consisted of the power bands for each electrode, resulting in 128 potential terms (32 channels x 4 power bands). Using the *stepwisefit* function in MATLAB R2022a, the process begins with a forward selection procedure where no terms are initially in the model structure. At each step, the p-value and coefficient of determination ($R^2$) are evaluated to determine whether a term should remain in the model. The criteria requires that each term retained in the model has a p-value less than 0.05 and increases the $R^2$ by a minimum of 0.5%.

Four-fold cross validation was applied for evaluation. For a given fold, 216 segments (12 trials x 18 segments) were utilized for model training and 72 segments (12 trials x 6 segments) for model testing. To assess the impact of artifact suppression on model performance, models were developed using EEG spectral data before and after artifact suppression was applied. Additionally, a model was developed using exclusively headset acceleration data without EEG.

*3.3.2 Binary Support Vector Machine Models.* Practical performance metrics such as the minimum level of fidelity or accuracy needed for reliably estimating affective state from EEG have yet to be defined. For example, while the capability to precisely predict rating levels over a continuum may be desirable, this level of resolution might not be necessary for a particular application. Alternatively, it may be more useful to simply distinguish between low and high levels of an affective state.

To explore this scenario, the SAM ratings were thresholded to binary low and high scores, which were modeled using a nonlinear Support Vector Machine (SVM) classifier with a Gaussian kernel (using the default hyperparameters of the *fitcsvm* function in MATLAB R2022a). The SAM ratings (1-9) were transformed into binary ratings (0 or 1) using a threshold determined by the mean rating for each individual participant and affective state category, respectively. Using the relevant EEG features determined by the stepwise regression process, an SVM model was trained for each participant

**Table 1: The means of the participant-wise intercorrelations between the scales of valence, arousal, liking, dominance across 10 participants. Significant correlations ($p \leq 0.05$) using Fisher's method are indicated by *.**

| | Valence | Arousal | Dominance | Liking |
|---|---|---|---|---|
| **Valence** | 1 | 0.46* | 0.12* | 0.83* |
| **Arousal** | - | 1 | 0.58* | 0.50* |
| **Dominance** | - | - | 1 | 0.31* |
| **Liking** | - | - | - | 1 |

and affective dimension, respectively, using the four-fold cross validation approach described in the previous section. Similar to the regression models, SVM models were created using EEG spectral bands before and after artifact suppression, and using exclusively headset acceleration data without EEG.

# 4 RESULTS

## 4.1 Affective State Ratings

Figure 5 shows the distribution of ratings for the three categories of videos: HVLA, HVHA, LVLA. Videos in the HVHA category had a median rating of 8 for valence, arousal, and liking, and 6 for dominance. For the LVLA video ratings, valence, arousal, and liking had significantly lower medians of 3, 4, and 4.5, respectively. For the LVLA category, dominance had the highest median rating of 7. Videos categorized as HVLA medians of 8, 6, 4, and 8 for valence, arousal, dominance, and liking, respectively.

The results aligned with the affective state categorizations as established in [12]. HVHA exhibited the highest median for valence and arousal; supporting the videos viewed in the category were pleasing and rated above a 5 for each state due to positive emotion. Videos in the LVLA category had the lowest median for valence, arousal, and liking, indicating videos intended to elicit negative emotion, videos were rated as unpleasant and generally disliked. However, dominance had a median rating above 5 suggesting the videos from the LVLA category elicited a consistent impact on the participants. HVLA consisted of videos which were meant to present a calm and relaxing environment to the participant. The median value valence and liking is the same for the category suggesting these videos were generally liked and considered insignificant due to the low median value of 4 dominance.

The mean intercorrelations of the four affective states across 10 participants are presented in Table 1. All intercorrelations were found to be significant, with largest inter correlation observed between liking and valence ($\rho$= 0.83). Correlations ranging from $\rho$=0.12 to $\rho$=0.58 were observed between dominance/liking arousal/liking, arousal/dominance, valence/arousal, and valence/dominance with valence/dominance having the lowest intercorrelation ($\rho$=0.12). The large intercorrelation observed between liking and valence is as expected due to positive/negative emotions (valence) being associated with liking/dislike.

## 4.2 Stepwise Linear Regression Models

Figure 6 shows boxplots of $R^2$ values for the the stepwise regression models based on four-fold cross validation. The median $R^2$ value

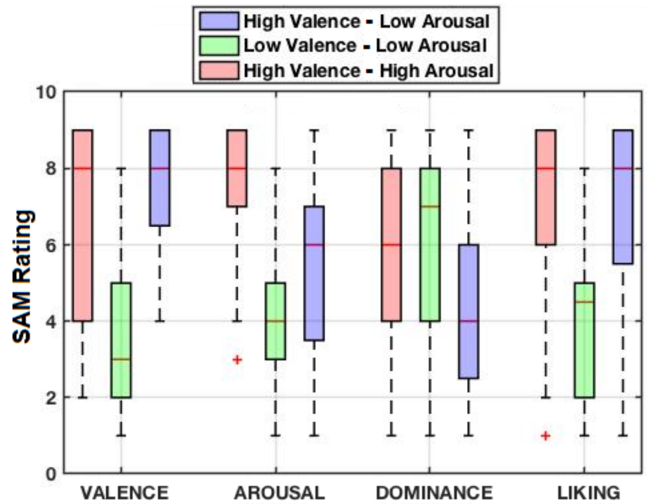


**Figure 5: Boxplot of participant SAM ratings for video categories**

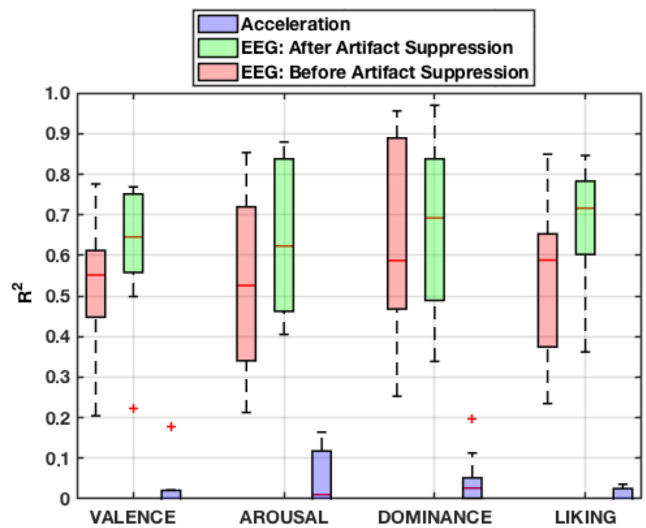


**Figure 6: Boxplot across participants of $R^2$ between SAM rating and stepwise regression model output.**

is between 0.5-0.6 for all affective states for the EEG data before artifact suppression. Artifact suppression resulted in approximately a 10% increase in median $R^2$, indicating the movement artifacts were likely masking relevant EEG features. This is further supported by the low $R^2$ between the acceleration data alone and the SAM ratings, providing evidence that any residual contamination from movement artifacts are not significantly contributing to the predictive power of the EEG recordings.

## 4.3 Binary SVM Models

Figure 7 shows results of affective state estimation using the binary SVM models. Similar to stepwise regression, the model performance is generally consistent across the four affective states. However, the

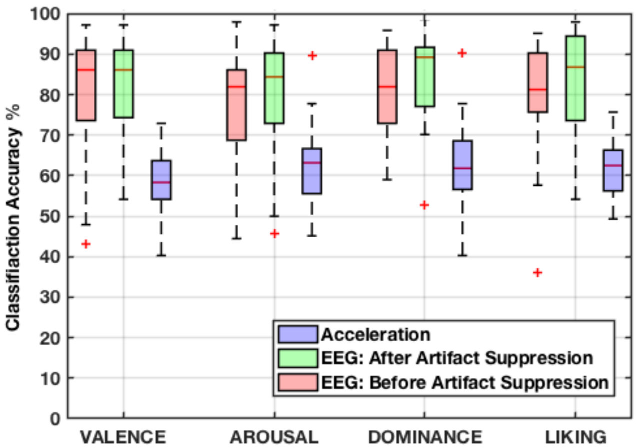


**Figure 7: Boxplot across participants of classification accuracy generated by the binary support vector machine (SVM) model for predicting high vs. low SAM ratings.**

favorable effects of artifact suppression are less prominent. The median performance after artifact suppression increases between 2-8% for arousal, dominance, liking, while valence remains unchanged.

The models using headset acceleration have median performances between 58-65%. The performance discrepancies between the linear stepwise models and the nonlinear SVM models using the accelerometer data suggests there is likely a nonlinear mapping of the accelerometer data that can provide predictive power. However, the model performance using EEG features remains superior to acceleration data alone, indicating that the EEG features provide additional predictive power beyond any residual artifacts related to the movement recordings.

## 5 DISCUSSION

The results of this pilot study indicate that subjective self-assessment ratings of a participant's affective state can reliably be estimated via EEG using regression and classification models. The regression models with artifact suppression yielded median $R^2$ values across each affective state ranging from 0.62-0.71. By simplifying the problem to estimate thresholded high versus low affective state ratings, a binary SVM classifier produced a classification accuracy above 85% for all affective states.

Although the computational cost for the proposed signal processing pipeline was not formally quantified in this study, the spectral analysis, artifact suppression, and affective state estimation models can be implemented with minimal latency for real-time neurofeedback in VR applications. However, depending on the application, affective state fluctuations or transitions may not occur rapidly and may not require immediate user feedback. Thus, acceptable processing latencies will depend on the application scenario and expected dynamics of the targeted affective states and user feedback.

While the relevant EEG features for this paradigm were examined in [9], further analysis and a larger participant pool is needed to thoroughly explore the relative feature contributions to the present models. Although the employed artifact suppression methods appear to effectively suppress head movement artifacts in EEG, no method exists that can definitively separate EEG and electromyogram (EMG) or movement artifacts occupying the same spectral frequency bands [16]. Although not collected and analyzed as part of the present study, it would be informative to capture video and/or EMG of facial expressions that vary with affective state and likely impact the EEG recordings. Ultimately, the combination of EEG, EMG, and movement information would likely improve modeling results, which may be more realistic and practical for end-user applications. Due to the subjective nature of the task and varying style and content of the available 360° videos, further analysis needs to be performed on the impact of the specific video stimuli (e.g., motion, audio, realism, optical flow, etc.) on the EEG recordings.

## 6 CONCLUSION

This pilot study demonstrates that it is possible to reliably estimate a user's affective state ratings using regression and classification models of spectral measurements obtained via EEG recordings. These models were designed to evaluate respective performance for two extremes of feedback implementations, continuous and binary, since different applications may require different degrees of granularity for the estimates. As with any EEG study involving movements and task engagement, it is challenging to definitively isolate the contributions of brain and muscle activity on the performance and the results must be interpreted accordingly.

The affective state estimation modeling techniques explored in this study could be implemented to provide closed-loop biofeedback to allow users to alter the virtual reality interactions according to their affective state. This has the potential to increase system functionality and enhance the user's experience. However, such a closed-loop scenario introduces various challenges as appropriate feedback and altered affective state dynamics must be investigated in detail. This pilot study provides important insights toward closing the loop with affective state estimation and feedback for VR applications.

## ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (1944389). The authors would like to acknowledge Pedram Zanganeh Soroush and Yusuke Yamani for assisting with the development and testing of the experimental paradigm.

## REFERENCES

[1] Adel Belouchrani, Karim Abed-Meraim, Jean-François Cardoso, and Éric Moulines. 1997. A blind source separation technique using second-order statistics. *IEEE Trans. Signal Process.* 45 (1997), 434–444.

[2] Jim Blascovich, Jack Loomis, Andrew C. Beall, Kimberly R. Swinth, Crystal L. Hoyt, and Jeremy N. Bailenson. 2002. TARGET ARTICLE: Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology. *Psychological Inquiry* 13, 2 (2002), 103–124. https://doi.org/10.1207/s15327965pli1302_01

[3] Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 1 (1994), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

[4] Raymundo Cassani, Marc-Antoine Moinnereau, Liviu Ivanescu, Olivier Rosanne, and Tiago Falk. 2020. Neural Interface Instrumented Virtual Reality Headsets: Toward Next-Generation Immersive Applications. *IEEE Systems, Man, and Cybernetics Magazine* 6 (07 2020), 20–28. https://doi.org/10.1109/MSMC.2019.2953627

[5] John F. Golding. 2006. Predicting individual differences in motion sickness susceptibility by questionnaire. *Personality and Individual Differences* 41, 2 (2006), 237–248. https://doi.org/10.1016/j.paid.2006.01.012

[6] Robert S. Kennedy, Kay M. Stanney, and William P. Dunlap. 2000. Duration and Exposure to Virtual Environments: Sickness Curves During and Across Sessions. *Presence: Teleoperators and Virtual Environments* 9, 5 (10 2000), 463–472. https://doi.org/10.1162/105474600566952

[7] Bojan Kerous, Filip Skola, and Fotis Liarokapis. 2018. EEG-based BCI and video games: a progress report. *Virtual Reality* 22, 2 (2018), 119–135.

[8] S. Koelstra, C. Muhl, M. Soleymani, Jong-Seok Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. 2012. DEAP: A Database for Emotion Analysis Using Physiological Signals. *IEEE Transactions on Affective Computing* 3, 1 (2012), 18–31. https://doi.org/10.1109/t-affc.2011.15

[9] Meghan Kumar, Connor Delaney, Pedram Zanganeh Soroush, Yusuke Yamani, and Dean J. Krusienski. 2021. Characterization of Affective States in Virtual Reality Environments using EEG. In *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 2689–2693. https://doi.org/10.1109/SMC52423.2021.9658916

[10] Lab Streaming Layer. 2022. https://github.com/sccn/labstreaminglayer.

[11] Anatole Lécuyer, Fabien Lotte, Richard B Reilly, Robert Leeb, Michitaka Hirose, and Mel Slater. 2008. Brain-computer interfaces, virtual reality, and videogames. *Computer* 41, 10 (2008), 66–72.

[12] Benjamin J. Li, Jeremy N. Bailenson, Adam Pines, Walter J. Greenleaf, and Leanne M. Williams. 2017. A Public Database of Immersive VR Videos with Corresponding Ratings of Arousal, Valence, and Correlations between Head Movements and Self Report Measures. *Frontiers in Psychology* 8 (2017). https://doi.org/10.3389/fpsyg.2017.02116

[13] Fabien Lotte, Josef Faller, Christoph Guger, Yann Renard, Gert Pfurtscheller, Anatole Lécuyer, and Robert Leeb. 2012. Combining BCI with virtual reality: towards new applications and improved BCI. In *Towards practical brain-computer interfaces*. Springer, 197–220.

[14] Robert Plutchik. 1988. The Nature of Emotions: Clinical Implications. *Emotions and Psychopathology* (1988), 1–20. https://doi.org/10.1007/978-1-4757-1987-1_1

[15] James A. Russell. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology* 39, 6 (1980), 1161–1178. https://doi.org/10.1037/h0077714

[16] Christoph Tremmel, Christian Herff, Tetsuya Sato, Krzysztof Rechowicz, Yusuke Yamani, and Dean J. Krusienski. 2019. Estimating Cognitive Workload in an Interactive Virtual Reality Environment Using EEG. *Frontiers in Human Neuroscience* 13 (2019). https://doi.org/10.3389/fnhum.2019.00401

[17] Unity Real-time Development Platform. 2022. https://unity.com/.