

A Comparative Study of Mental States in 2D and 3D Virtual Environments Using EEG

Pinar Bilgin^{1,2}, Kat Agres^{2,3}, Neethu Robinson¹, Aung Aung Phyto Wai¹ and Cuntai Guan¹

Abstract—There is growing evidence that virtual reality (VR) can be used as an optional therapy method for stress relief and treatment of mental disorders such as a variety of anxiety disorders, depression and psychosis. However, more systematic studies to quantify the effects of VR for emotion elicitation and compare the effects of 3D and 2D environment for the same is necessary to design feasible and effective treatments. In this study, we design a cross-over experiment protocol comprising of two emotion eliciting environments relaxation and arousal - that is presented to the participant either in a 2D monitor or a 3D head-mounted (HMD) VR display. The EEG data is collected during the experiment and analyzed offline to classify emotion elicited by low and high arousal environments using SVM classification of band power features. A 10-fold cross validation is performed and classification accuracy for each subject is computed for 3D-VR and 2D-screen display. The average classification accuracies over subjects are obtained as 66.88% and 59.27% in 3D-VR and 2D-screen group respectively. The performance difference between 3D-VR and 2D screen is statistically significant ($p < 0.05$), indicating that 3D-VR generates more distinct EEG patterns associated with emotion elicitation. Further, significant differences in band powers from alpha, theta and beta bands are also observed between both groups. The results presented in this paper support the use of VR as an effective tool to study emotion elicitation and regulation, compared to a 2D screen.

I. INTRODUCTION

The growing number of stress related problems in our daily lives has caused an increase in research investigating how to help humans cope with stress and anxiety. Anxiety is one of the most common mental health problems, with a lifetime prevalence of slightly over 30%, and many undiagnosed cases [1], [2]. Cognitive behavior therapy (CBT) and pharmacotherapy are the current standard methods of treatment for anxiety, however these options are not accessible to most people, and as such, only a fraction of those diagnosed can receive sufficient treatment [2]. A more accessible option for treatment is relaxation and mindfulness training, which has been shown to have comparable effects to CBT in short term studies [2], [3]. By using affective Brain-Computer Interfaces (BCIs) with electroencephalogram (EEG) signals as a tool to automatically recognize, model and express emotions, researchers are able to create technological tools that are

able to interact with humans on an emotional level [4], [5], [6]. Through the use of affective EEG-BCI and virtual reality (VR) environments, we can create easily accessible technology-based therapy methods for anxiety disorders and stress relief.

VR technology has been successfully used with affective computing in treatment of mental disorders such as a variety of anxiety disorders, depression and psychosis [2], [7], [8]. Because of the high level of control over given stimuli, VR allows for precise implementation of therapeutic strategies and can assist with emotion, cognition and behavior assessment in an ecologically valid environment [9]. Mainly due to this and its wide availability and cost-effective nature, VR is preferred over in-vivo exposure therapy methods in the treatment of anxiety disorders including social anxiety disorder, phobias, and post-traumatic stress disorder (PTSD) [3], [9], [10], [8].

While several studies use VR for exposure therapy [2], [3], [7], [9], [10], [11], the effects of a 2D versus a 3D VR immersive environment are not widely studied and understood. The sense of presence during a 3D experience is purported to be higher than 2D [12], however, cognitive load has been observed to be higher in 2D tasks than 3D, and similar EEG results have been achieved through 2D and 3D tasks during less complex tasks. Lower cognitive load increases efficiency of learning, which suggests that 3D technologies should be advantageous in terms of cognitive load [13]. In terms of emotion elicitation, the comparison between 2D versus 3D has not been studied extensively. It has been observed that VR images create a stronger stimulus and higher beta power for dynamic videos, but the 2D screen produces greater alpha power, which may indicate higher levels of concentration [14].

In this study, we compare the efficacy of 2D versus 3D environments for eliciting emotion in participants by examining the EEG signals induced in both conditions. We follow a cross-over design where participants are separated into two groups, and the first group does the 3D-VR task first and 2D-Screen task second, while second group follows in reverse order. The experiments involve two environments: the first is associated with a calm state (low-arousal, positive/neutral-valence), and the second is associated with a dynamic state (high-arousal, neutral-valence). The environments follow the Circumplex Model of Affect (CMA) and are created to induce relaxation and arousal, respectively [15]. CMA is a universally accepted model that places emotion on a two-dimensional scale of arousal-valence, with arousal varying from high to low, and valence varying from positive to nega-

¹ School of Computer Science and Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore pinar001@e.ntu.edu.sg, nrobinson@ntu.edu.sg, aungaung001@e.ntu.edu.sg, CTGuan@ntu.edu.sg

² Institute of High Performance Computing, A*STAR, 1 Fusionopolis Way, #16-38 Connexis North, Singapore kat_agres@ihpc.a-star.edu.sg

³ Yong Siew Toh Conservatory of Music, National University of Singapore, 3 Conservatory Drive, Singapore

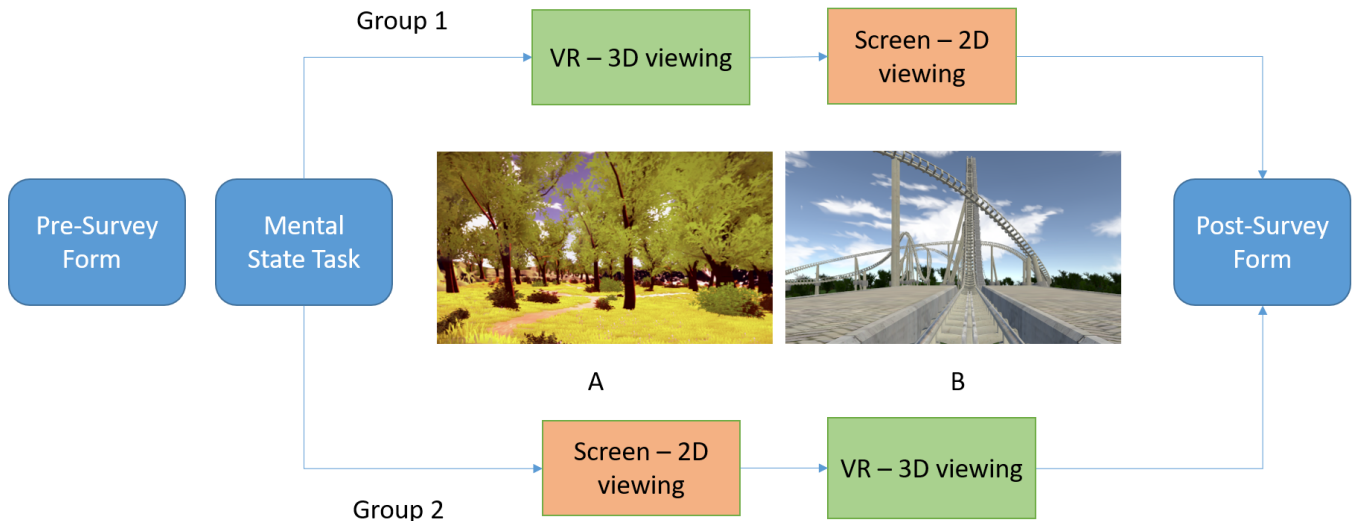


Fig. 1. Experiment Overview. The experiment involves two environments, a calm environment (A), and dynamic (arousal) environment (B), each 3 minutes in duration. The experiment is a total of 6 minutes in duration. The participants are separated into two groups (Group 1, Group 2).

tive. In our study, we classify users' EEG responses along the arousal dimension, because our experimental design relies mostly on neutral valence and evoking difference in arousal state.

The increased sense of presence and lower cognitive load associated with 3D environments is expected to lead to higher levels of stimulation and more accurate emotion classification than the 2D screen, especially for the dynamic environment. It has been observed in previous research that there is a statistically significant difference between alpha and beta bands for 2D and 3D displays [14]. There is also evidence of higher variance in the theta band for 2D compared with 3D setups, and a similar effect for alpha band activity for 2D compared with 3D learning tasks [13]. Higher theta band power was also observed in the 3D setup for a navigation task [12]. To the authors' knowledge, there is no existing research that compares the ability of 2D and 3D display methods for eliciting emotion in viewers. In this study, we use EEG frequency bands as features to test the accuracy of the 2D-Screen versus the 3D-VR conditions in emotion classification, and observe and compare the effects of 2D versus 3D display on evoking emotion and the frequency bands that have been affected for each display. We collect empirical data from the participants to test the two conditions. Due to greater sense of presence, especially during the dynamic environment, we expect that alpha and beta band activity will be higher during 3D-VR task than those in 2D-Screen environment, as alpha and beta band are highly correlated with arousal state.

The rest of the paper is organized as follows. In section II, we explain the experiment protocol, the environment designs and the signal processing steps taken for emotion classification. We then explain our classification results and discuss what they mean in section III, and give our conclusions in section IV.

II. RESEARCH METHODOLOGY

This is a preliminary study that attempts to differentiate the levels of immersiveness and emotion elicitation between 3D-VR and 2D-screen displays using EEG signals.

A. EXPERIMENT PROTOCOL

For this experiment, a flat 17-inch Razer Blade Pro laptop screen was used for the 2D display, and an OSVR HDK2 VR HMD was used for the 3D-VR display. The EEG signals were collected using a Muse headband with 4 channels (TP9, FP1, FP2, TP10) and the sampling rate was 256 Hz. The Muse headband has been used in several research studies such as [16], [17], [18] in wearable BCI systems to study attention. For this preliminary experiment, 10 participants were recruited (2 female and 8 male participants between the ages of 21-35). To decrease any sense of imbalance and potential cybersickness induced by the VR, an armchair was used for the experiments. Participants filled out a survey form before and after the experiment, to indicate their levels of alertness, emotional state, as well as physical and mental fatigue. The experiment description and rationale was given in the consent form and verbally explained to each participant before they began the experiment, and participants were instructed to sit in a comfortable position and remain still during the task. This research was approved by the Institutional Review Board of Nanyang Technological University, Singapore.

The experiment was separated into two parts in total, as calm environment (relaxation) and dynamic environment (arousal). The same experiment process was used for 2D-Screen and 3D-VR for each subject. To eliminate the effect of order, a cross-over design is used in which the subjects were separated into two groups of 5, and the order of conditions was counterbalanced across participants. The first group viewed the VR display first and the Screen display second, and the second group viewed the Screen display first

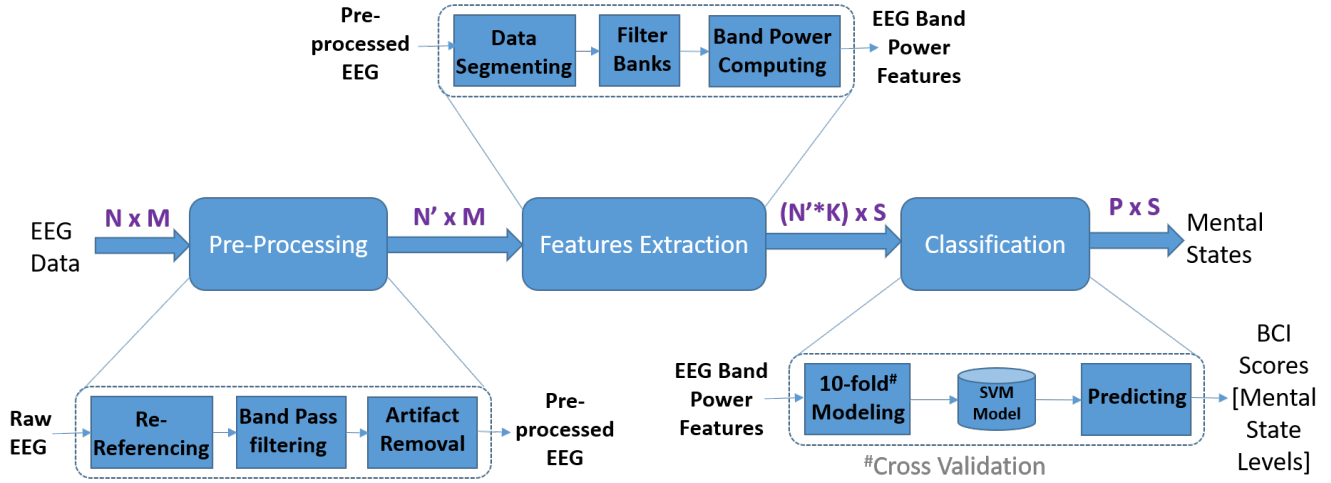


Fig. 2. Block diagram of Signal Processing steps. EEG data is pre-processed using bandpass filtering and artifact removal, and 6 frequency bands are extracted for each channel. Data is then fed into an SVM classification algorithm to calculate the accuracies and BCI scores (mental states). N - # channels from Muse (4), M - # EEG data samples, N' - # pre-processed channels (2), K - # bands (6), S - # data segment, P - # BCI Scores presenting mental states (1).

and the VR display second. There was a rest period of at least 30 minutes between tasks. The order of experiment for both displays were the same, starting with 3 minutes of a calm environment including a brief nature-based mindfulness task, and 3 minutes of a dynamic environment involving a first-person point-of-view roller coaster ride, totaling 6 minutes. Both of the environments were created using Unity® video game software. More information about the two environments is provided below:

1) *Calm Environment*: The calm environment, as seen in Figure 1(A), consists of a first-person point of view walk through a sunny forest. This scene was selected based on related studies of anxiety relaxation techniques used in meditation therapy [2], [7]. These studies indicate that scenes with warm, bright, nature-based environments and low illumination (as not to bother participants eyes), as well as mindfulness instructional cues and calming music with neutral valence, are most effective in calming participants [2].

2) *Dynamic Environment*: A roller coaster ride was selected as the dynamic environment, as seen in Figure 1(B). Previous studies have shown that this scene elicits high-arousal effects with a high sense of presence in the VR display [15], [19]. Except for track noise and wind audio, there were no music or audio cues during the roller coaster task, so as not to affect the valence state of the participants.

B. SIGNAL PROCESSING

The collected Muse EEG data were processed using MATLAB®. Figure 2 shows all the steps taken to process the data. Those signal processing steps are as follows:

1) *Pre-processing*: The raw EEG data from Muse were re-referenced by conversion from four channels (TP9, FP1, FP2, TP10) to two channels $FP1/\frac{(TP9+TP10)}{2}$, and $FP2/\frac{(TP9+TP10)}{2}$. Data were then separated into two classes, Calm and Arousal, and pre-processed with a Butterworth bandpass filter along

0.3-45 Hz. The artifact removal is done by subtracting smoothed references from the EEG. The smoothed references are obtained using moving averaging with multiple moving windows.

2) *Feature Extraction*: The pre-processed data was segmented into epochs with a 200ms shift between the epochs. We performed that analysis for epoch lengths, 1s, 2s, 3s, and 4s, and computed accuracy in each. The delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), low beta (12-18 Hz), high beta (18-30 Hz) and gamma (30-45 Hz) frequency band powers were extracted as features for each channel, totaling in 12 features and an average of 357 trials for calm class and 356 trials for arousal class for 1 second epoch length.

3) *Classification*: LIBSVM [20] was used for classification of the two classes using an SVM-SVR model with radial basis function, epsilon-SVR of cost 5, and gamma 0.5, using within subject 10-fold cross validation.

III. RESULTS & DISCUSSION

For this study, the classification accuracies for the calm and arousal states are compared for the 3D-VR and 2D-Screen tasks (Table I). Different window lengths are tested (Table II) to observe the effect of epoch length in classification accuracy. The results of accuracy for calm and arousal classes are also compared for VR and Screen (Table III). The frequency band powers are plotted from EEG data to observe differences in level of stimulation between VR and Screen tasks. The accuracies for each frequency band is calculated using the same SVM algorithm, and results are compared (Table IV).

A. Classification Accuracy

The classification accuracy obtained for each subject with 2D and 3D environment in demonstrated in Fig.3, Table I and Fig. 4. The average accuracy over subjects, 66.88% for 3D is significantly higher than for 2D, 59.27%. Fig.3

TABLE I
AVERAGE ACCURACY (%) PER PARTICIPANT

	3D-VR	2D-Screen
s1	59.28	60
s2	83.28	56.14
s3	71.57	50.70
s4	63.14	70.71
s5	60.57	53.94
s6	62.71	60.56
s7	53.85	58.73
s8	72.28	62.25
s9	70	58.71
s10	72.14	61

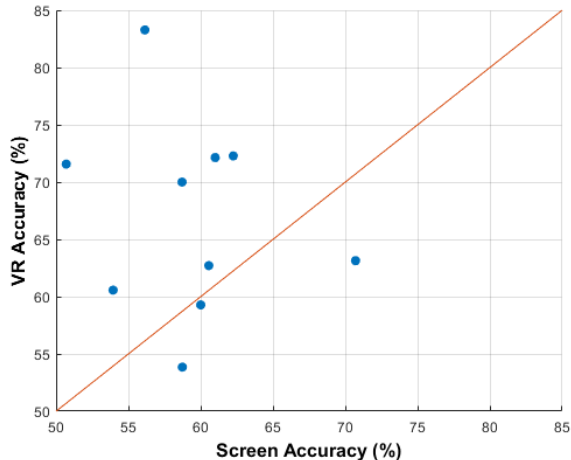


Fig. 3. 3D-VR versus 2D-Screen Average Accuracy (%) per Participant

indicates that for majority of subjects the 3D environment is more effective in evoking distinct emotions. The Receiver Operating Characteristic (ROC) curve of 3D versus 2D in Fig. 4 shows us that the sensitivity for VR is higher than Screen, indicating better results. The findings support the hypothesis that 3D-VR is more effective in evoking emotion than 2D-Screen. With a larger sample size, we should be able to see a more homogeneous comparison, because outlier subjects would not have as large an effect on the overall results. It can be observed that a few subjects demonstrated higher accuracy in 2D environment. This might be due to the quality of the VR headset and the effect of cybersickness for some subjects.

We tested different window lengths to observe its effect in overall accuracy. The results for each window length are similar with little variance within the 3D-VR and 2D-Screen tasks. The average accuracy across all participants for 3D-VR is greater than the accuracy for 2D-Screen in each window length, as shown in Table II, with a statistically significant difference between the two tasks in 1 second window length ($p < 0.05$).

Table III shows the differences between classes for 3D-VR and 2D-Screen for each subject. It was hypothesized that the dynamic environment would have a greater effect during the 3D-VR condition than the 2D-Screen condition, and indeed, this can be observed in the average true prediction

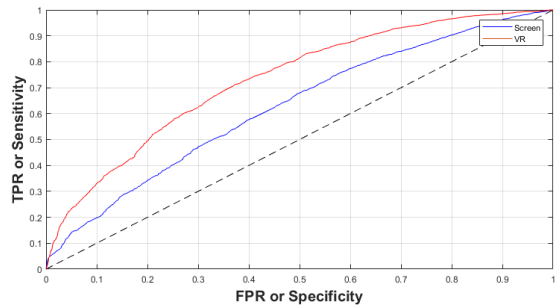


Fig. 4. 3D-VR versus 2D-Screen ROC curve, TPR: True Positive Rate, FPR: False Positive Rate

TABLE II
AVERAGE ACCURACY FOR ALL PARTICIPANTS.

Window Length	Screen-2D (%) / std*	VR-3D (%) / std*
1 second	59.27 / 8.8	66.88 / 11.67
2 second	60.26 / 10.26	66.97 / 13.99
3 second	60.31 / 12.16	65.45 / 14.46
4 second	56.43 / 13.87	65.18 / 15.67

rate in Table III. The 2D-Screen and 3D-VR for the calm environment was expected to yield similar effects, however, the results show that the true prediction rate of 3D-VR was much higher than the true prediction rate of 2D-Screen for the calm class. We obtained the significance for Table III using the Wilcoxon paired t-test method where in all tests, hypothesis is measurement one greater than measurement two. We use Wilcoxon paired t-test method to determine the statistical significance with a non normal distribution for the difference between 2D and 3D for each class. We find that there is a significant difference between 2D and 3D for the calm class with ($p < 0.05$), while the difference between 2D and 3D for the dynamic class is not statistically significant.

B. Performance Across Frequency Bands

The spectral power for every band is calculated for both VR and Screen. Comparing Fig. 5 and Fig. 6, it can be observed that there are variations in band power values between 2D and 3D groups. The power in the delta, theta, alpha and beta bands is higher for 3D-VR than the 2D-Screen condition for majority of the subjects. This is in line with the findings in literature [12], [13], [14], [21] that alpha and beta bands are correlated with arousal states, while the theta band is associated with relaxation/meditation states [13].

As shown in Table IV, all of the frequency bands yielded higher accuracy results for VR than Screen. The SVM results were very similar for the Theta, High Beta, and Gamma bands. The differences were statistically significant only for the Delta band ($p < 0.05$) and High Beta band ($p < 0.05$) according to Wilcoxon paired t-test method. Averages of accuracies per subject yield higher results for VR for Delta, Low Beta and Alpha band, however, difference in accuracies for Low Beta and Alpha band is not statistically significant. Our results support the hypothesis that 3D-VR induces stronger emotion in both calm and dynamic environments

TABLE III
CALM VERSUS AROUSAL AVERAGE ACCURACY (%) PER PARTICIPANT FOR VR/SCREEN

	VR - Calm	Screen - Calm	VR - Arousal	Screen - Arousal
s1	49.71	54.57	66.86	61.71
s2	79.71	57.71	86.28	51.14
s3	73.43	45.55	68.57	43.14
s4	62.57	68	60.28	72.86
s5	57.14	47.5	60.86	55.43
s6	58.28	58.61	66	59.14
s7	43.43	41.66	62.57	74.28
s8	69.71	70.83	74.28	52.28
s9	73.43	50	66.28	64.28
s10	72	57.14	71.71	62.28571429
Average	63.95	55.16	68.37	59.66

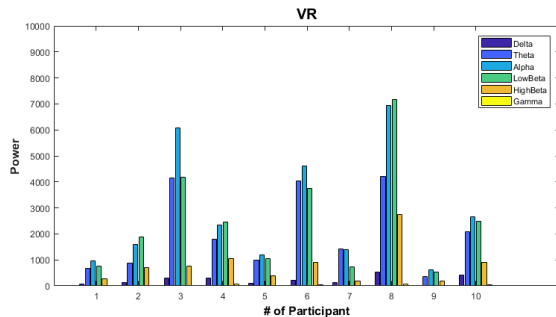


Fig. 5. VR-Frequency band powers per Participant

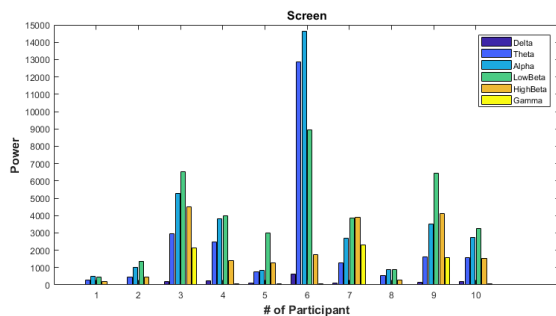


Fig. 6. Screen-Frequency band powers per Participant

than 2D-Screen, and that VR yields greater discrimination between the Calm and Arousal classes.

Although we provide promising pilot results, this study has a number of limitations. The study will be extended with a larger sample size, using different classification methods to compare results. Since the Muse headband has a limited number of electrodes, future study will use research grade EEG amplifier. The response from pre/post-survey forms will be used in future studies involving more subjects to support

TABLE IV
ACCURACIES (%) ACROSS EEG FREQUENCY BANDS

	Delta	Theta	Alpha	LBeta*	HBeta*	Gamma
VR	62.27	54.33	57.43	58.53	52.81	56.83
Screen	54.50	53.63	53.06	52.82	51.71	54.56
p-value	0.032	0.652	0.116	0.042	0.539	0.561

*Low Beta, * High Beta

the research findings with subjective characteristics. Also, we propose that the experimental paradigm could be extended as a large-scale EEG-BCI online (real-time) study with closed-loop emotion classification incorporating biofeedback, and using longer exposure times in both environments.

IV. CONCLUSIONS

In this preliminary study, we collected EEG data from 10 participants and compared the results for 2D-Screen display and 3D-VR display. We used an SVM algorithm for classification, and identified that the overall accuracies for VR were higher for each window length and frequency band. Due to the small number of participants, however, the power of statistics is small, as individual subjects perform differently across both tasks. For future studies, we should include more participants and longer exposure times to obtain more homogeneous and informative results. Overall, our study supports the claim that VR indeed evokes stronger emotion in participants than flat 2D screen, and can be used as an effective tool in future emotion elicitation and regulation studies.

REFERENCES

- [1] P. Lindner, A. Miloff, E. Zetterlund, L. Reuterskild, G. Andersson, and P. Carlbring, "Attitudes toward and familiarity with virtual reality therapy among practicing cognitive behavior therapists: A cross-sectional survey study in the era of consumer vr platforms," *Frontiers in Psychology*, vol. 10, p. 176, 2019.
- [2] J. Tarrant, J. Viczko, and H. Cope, "Virtual reality for anxiety reduction demonstrated by quantitative eeg: A pilot study," *Frontiers in Psychology*, vol. 9, p. 1280, 2018.
- [3] I. L. Kampmann, P. M. Emmelkamp, D. Hartanto, W.-P. Brinkman, B. J. Zijlstra, and N. Morina, "Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial," *Behaviour Research and Therapy*, vol. 77, pp. 147 – 156, 2016.
- [4] M. Soleymani, M. Pantic, and T. Pun, "Multimodal emotion recognition in response to videos," *IEEE Transactions on Affective Computing*, vol. 3, pp. 211–223, April 2012.
- [5] Y.-J. Liu, M. Yu, G. Zhao, J. Song, Y. Ge, and Y. Shi, "Real-time movie-induced discrete emotion recognition from eeg signals," *IEEE Transactions on Affective Computing*, vol. PP, pp. 1–1, 01 2017.
- [6] R. Sitaram, T. Ros, L. Stoeckel, S. Haller, F. Scharnowski, J. Lewis-Peacock, N. Weiskopf, M. L. Blefari, M. Rana, E. Oblak, N. Birbaumer, and J. Sulzer, "Closed-loop brain training: the science of neurofeedback," *Nature Reviews Neuroscience*, vol. 18, pp. 86–100, 2017.
- [7] D. Freeman, S. Reeve, A. Robinson, A. Ehlers, D. Clark, B. Spanlang, and M. Slater, "Virtual reality in the assessment, understanding, and treatment of mental health disorders," *Psychological Medicine*, vol. 47, no. 14, p. 23932400, 2017.

- [8] X. B. Lin, T.-S. Lee, Y. B. Cheung, J. Ling, S. H. Poon, L. Lim, H. H. Zhang, Z. Y. Chin, C. Wang, R. Krishnan, and C. Guan, "Exposure therapy with personalized real-time arousal detection and feedback to alleviate social anxiety symptoms in an analogue adult sample: Pilot proof-of-concept randomized controlled trial," *JMIR Mental Health*, vol. 6(6), 2019.
- [9] L. R. Valmaggia, L. Latif, M. J. Kempton, and M. Rus-Calafell, "Virtual reality in the psychological treatment for mental health problems: An systematic review of recent evidence," *Psychiatry Research*, vol. 236, pp. 189 – 195, 2016.
- [10] E. Carl, A. T. Stein, A. Levihn-Coon, J. R. Pogue, B. Rothbaum, P. Emmelkamp, G. J. Asmundson, P. Carlbring, and M. B. Powers, "Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials," *Journal of Anxiety Disorders*, vol. 61, pp. 27 – 36, 2019. Virtual reality applications for the anxiety disorders.
- [11] P. Lindner, A. Miloff, W. Hamilton, L. Reuterskild, G. Andersson, M. B. Powers, and P. Carlbring, "Creating state of the art, next-generation virtual reality exposure therapies for anxiety disorders using consumer hardware platforms: design considerations and future directions," *Cognitive Behaviour Therapy*, vol. 46, no. 5, pp. 404–420, 2017. PMID: 28270059.
- [12] S. M. Slobounov, W. Ray, B. Johnson, E. Slobounov, and K. M. Newell, "Modulation of cortical activity in 2d versus 3d virtual reality environments: An eeg study," *International Journal of Psychophysiology*, vol. 95, no. 3, pp. 254 – 260, 2015.
- [13] A. Dan and M. Reiner, "Eeg-based cognitive load of processing events in 3d virtual worlds is lower than processing events in 2d displays," *International Journal of Psychophysiology*, vol. 122, pp. 75 – 84, 2017. Neural Patterns of Learning, Cognitive Enhancement and Affect.
- [14] S. H. Kweon, H. J. Kweon, S.-j. Kim, X. Li, X. Liu, and H. L. Kweon, "A brain wave research on vr (virtual reality) usage: Comparison between vr and 2d video in eeg measurement," in *Advances in Human Factors and Systems Interaction* (I. L. Nunes, ed.), (Cham), pp. 194–203, Springer International Publishing, 2018.
- [15] J. Posner, J. A Russell, and B. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Development and psychopathology*, vol. 17, pp. 715–34, 02 2005.
- [16] J. Amores, R. Richer, N. Zhao, P. Maes, and B. M. Eskofier, "Promoting relaxation using virtual reality, olfactory interfaces and wearable eeg," pp. 98–101, March 2018.
- [17] B. Byrom, M. McCarthy, P. Schueler, and W. Muehlhausen, "Brain monitoring devices in neuroscience clinical research: The potential of remote monitoring using sensors, wearables, and mobile devices," *Clinical Pharmacology & Therapeutics*, vol. 104, no. 1, pp. 59–71.
- [18] G. Wiechert, M. Triff, Z. Liu, Z. Yin, S. Zhao, Z. Zhong, R. Zhaou, and P. Lingras, "Identifying users and activities with cognitive signal processing from a wearable headband," pp. 129–136, Aug 2016.
- [19] T. Baumgartner, D. Speck, D. Wettstein, O. Masnari, G. Beeli, and L. Incke, "Feeling present in arousing virtual reality worlds: prefrontal brain regions differentially orchestrate presence experience in adults and children," *Frontiers in Human Neuroscience*, vol. 2, p. 8, 2008.
- [20] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [21] J. Marn-Morales, J. L. Higuera-Trujillo, A. Greco, J. Guixeres, C. Llinares, E. P. Scilingo, M. Alcaiz, and G. Valenza, "Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors," *Scientific Reports*, vol. 8, 2018.