2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) | 979-8-3503-2891-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISMAR-Adjunct60411.2023.00153

Multimodal Physiological Analysis of Cognition and the Correlations with Emotion in VR

Yang Shen
Beijing Normal University

Ming Li*
Beihang University

Yu Li
Beijing Normal University

Yang Gao
Beihang University

Junjun Pan†
Beihang University

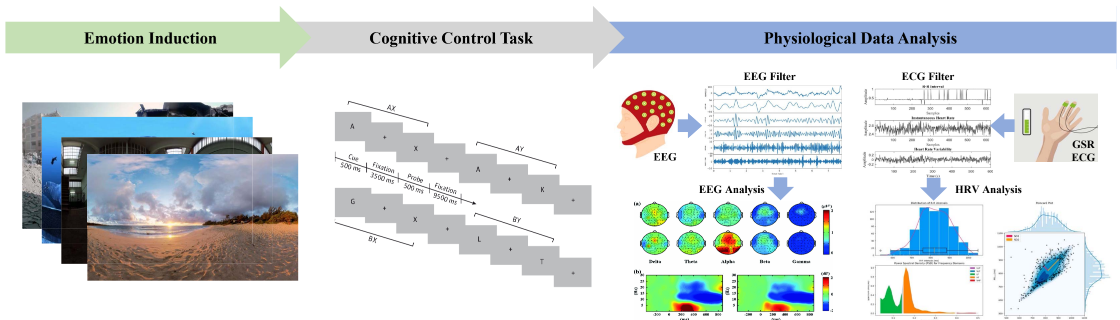

Figure 1: The experimental framework: emotion induction materials on the left, the AX-CPT cognitive task in the middle, and the analysis and processing of multimodal physiological data on the right.

Abstract

During VR cognitive rehabilitation, user performance in cognitive tasks is influenced by their emotional states. However, most current research primarily focuses on emotional valence, lacking a comprehensive examination of the impact of emotions on cognitive activities and their underlying physiological mechanisms based on an arousal-valence dimensional model. To address this gap, we have conducted an experiment in which participants are induced with four distinct emotions using VR videos, followed by cognitive tasks. Concurrently, we collect participants' EEG, ECG, and GSR signals. The results indicate that high arousal and high valence emotional states significantly improve user performance in cognitive tasks (shorter reaction times).

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Visualization—Visualization design and evaluation methods

1 Introduction

In VR-based neurorehabilitation, cognitive training tasks can enhance deficits in attention, perception, reasoning, and executive functions. During the rehabilitation process, cognitive task performance is influenced not only by the external environment but also by the individual's subjective feelings, particularly emotional states related to cognition. Prior research has shown that emotions affect task performance and responses, which vary among individuals [1]. As illustrated in Fig. 2, the impact of emotional arousal levels on individual performance generally follows Hebb's U-shaped curve model [2]. In this model, excessively high or low levels of emotional arousal tend to impede task performance, while a moderate level of emotional arousal stimulates users to perform optimally.

However, most current research primarily focuses on the impact of emotional valence on cognition, overlooking the role of emotional arousal. Mayer et al. [3] argue that cognitive load constrains information acquisition and integration. Both positive and negative emotions increase cognitive load and impair cognitive task performance. Consequently, a unified and comprehensive conclusion regarding the influence of arousal/valence emotions on cognition is still lacking.

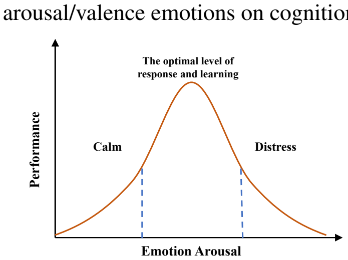

Figure 2: Schematic illustration of the association between emotion arousal levels and cognitive performance.

To this end, we design a VR-based emotion induction experiment using 8 VR videos from Stanford's public VR video datasets, YouTube, and Steam for emotion induction [4]. Following this, we conduct corresponding cognitive tasks (AX-Continuous Performance Test) to investigate the effects of emotions on cognitive activities. The performance of cognitive tasks is assessed based on test accuracy and reaction time measures. Furthermore, we collect electroencephalography (EEG), heart rate variability (HRV), and galvanic skin response (GSR) signals as objective data to evaluate their cognitive status. The overall framework of the experiment is illustrated in Fig. 1.

2 Materials and Methods

2.1 Experiment Apparatus and Participants

In this study, we use 1) an HTC Vive Pro HMD for displaying VR videos, featuring a resolution of 1200x1080 pixels per eye and a screen refresh rate of 90Hz, 2) an HP laptop computer equipped with an Intel Core i9 processor, GTX 3080Ti graphics card, and 32 GB RAM, 3) an ANT Neuro EEG signal acquisition system with 32 electrodes where all electrodes are referenced to CPz and grounded to the forehead. The sampling rate of raw EEG data is set to 512Hz. 4) a Shimmer3 GSR+ system for capturing HRV and GSR

*Co-first authors
†Corresponding authors: pan_junjun@buaa.edu.cn

2771-1110/23/$31.00 ©2023 IEEE
DOI 10.1109/ISMAR-Adjunct60411.2023.00153

signals. The GSR data are recorded at a sampling rate of 200Hz. We have recruited a total of 14 participants (7 male and 7 female), aged between 21 and 45 (M=31.2, SD=3.16). The study received approval from the local ethics committee.

2.2 Emotion Induction Material and Cognitive task

In accordance with the criteria used in prior studies, we select 8 VR videos to induce four emotional states based on Russell's arousal-valence emotion model: Low Arousal/Low Valence (LALV), Low Arousal/High Valence (LAHV), High Arousal/Low Valence (HALV), and High Arousal/High Valence (HAHV) [5]. The Self-Assessment Manikin (SAM) scale is employed to assess arousal, valence, and dominance [6].

Following emotion induction, we administer a classic psychological cognitive control task called AX-CPT, for which we develop and utilize a VR version. This paradigm assesses participants' ability to perform tasks without losing attention. The task consists of a total of 200 trials. Accuracy is calculated as the number of correct trials divided by 200. Ultimately, we record the task accuracy and average reaction time.

2.3 Multimodal Physiological Data

For neurophysiological data, we employ EEGLAB for EEG preprocessing. Specifically, the signals are then resampled to 128 Hz and re-referenced to facilitate faster feature extraction. A bandpass filter from 4 to 47 Hz is applied to the signals. Visual inspection is conducted to remove any abnormal signals with amplitudes exceeding ±100 (μv). Lastly, independent component analysis (ICA) is used to remove artifacts. Based on the processed data, we utilize the Welch method in Python-MNE to calculate the power spectral density (PSD) features for each electrode in the θ (4-7 Hz), α (7-13 Hz), β (14-29 Hz), and γ (30-47 Hz) bands.

For HRV analysis, we use global or local thresholding for QRS detection. Detected NN intervals are corrected through algorithms and visual inspection. NN intervals are interpolated linearly at 4 Hz. Time-domain features include SDNN, RMSSD, PNN50. Frequency-domain features encompass LF power(0.04–0.15 Hz), HF power(0.15–0.4 Hz), and LF/HF.

For GSR signal, extracted features include mean skin resistance, mean of derivatives, proportion of negative derivatives, count of local minima, average rising time, spectral power in 0-2.4 Hz, zero-crossing rate of skin conductance slow response (0-0.2 Hz), and zero-crossing rate of skin conductance very slow response (0-0.08 Hz).

2.4 Experimental Procedures

All participants willingly took part in the study and provided written informed consent before the experiment. The experiment commenced with the display of a fixation cross on the VR screen. Subsequently, participants randomly viewed one of the 8 VR videos. Then filled out the SAM emotion questionnaire. Following that, the AX-CPT cognitive task was administered. The aforementioned process was carried out a total of 8 times, and each participants watched all 8 videos. Each VR video had an approximate duration of 3min30s, while the cognitive task lasted around 5min. The entire experiment lasted for an hour and a half.

3 Results

In this study, the Wilcoxon signed-rank test is used to analyze the differences in accuracy and reaction time of the cognitive task across different emotional states. The significance level for all experiments is set at 0.05. The results, as shown in Table 1, indicate that there are no significant differences in the accuracy of the cognitive task among the four emotional states. However, the reaction time for the HAHV emotional state is significantly lower than the other three emotional states (p = 0.028).

Table 1: The accuracy and reaction time of the cognitive task across different emotional states (* = $p < 0.05$).

Emotion	Reaction Time (s)	Accuracy
HAHV	26.449*	97.67
HALV	27.511	98.17
LAHV	29.183	97.83
LALV	29.648	98.16

The results from the physiological data are shown in Table 2, there is a significant decrease in SDNN (p = 0.017) and a significant increase in LF power (p = 0.014) during the HAHV emotional state compared to the other three emotional states, indicating activation of the sympathetic nervous system. Moreover, there is a significant decrease in α band power (p = 0.023) in the frontal lobe, which aligns with previous research suggesting an inverse association between α band power in the frontal lobe and cognitive load. In the HAHV emotional state, participants experienced higher cognitive load, leading to a reduction in α power.

Table 2: Physiological features during AX-CPT task in different emotional states (* = $p < 0.05$).

Emotion	SDNN	LF power	α
HAHV	26.78*	73.89*	0.34*
HALV	30.45	68.23	0.42
LAHV	32.37	60.67	0.56
LALV	40.54	64.53	0.63

4 Conclusion

This study aimed to investigate the impact of different emotional states on cognitive activities, drawing upon Russell's arousal-valence model of emotion. The results demonstrate that experiencing a high arousal-high valence emotional state significantly improves performance in cognitive tasks, accompanied by the activation of the sympathetic nervous system. Additionally, there is a significant decrease in α band power in the frontal lobe. These physiological markers can serve as reliable indicators for assessing cognitive activities and play a crucial role in guiding the adaptive adjustment of VR rehabilitation systems, thereby enhancing their effectiveness in promoting cognitive improvements.

Acknowledgments

This work was supported by the Research Program Funds of the Collaborative Innovation Center of Assessment toward Basic Education Quality at Beijing Normal University: 2022-01-003-BZK01& 2022-01-131-BZK01; the Fundamental Research Funds for the Central Universities:2021NTSS33

References

[1] J. L. Plass and S. Kalyuga, "Four ways of considering emotion in cognitive load theory," *Educational Psychology Review*, vol. 31, pp. 339–359, 2019.

[2] D. O. Hebb, "Drives and the cns (conceptual nervous system)." *Psychological review*, vol. 62, no. 4, p. 243, 1955.

[3] R. E. Mayer, *The Cambridge handbook of multimedia learning*. Cambridge university press, 2005.

[4] B. J. Li, J. N. Bailenson, A. Pines, W. J. Greenleaf, and L. M. Williams, "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*, vol. 8, p. 2116, 2017.

[5] M. Li, J. Pan, Y. Gao, Y. Shen, F. Luo, J. Dai, A. Hao, and H. Qin, "Neurophysiological and subjective analysis of vr emotion induction paradigm," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 11, pp. 3832–3842, 2022.

[6] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.