

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Identifying Mind Wandering Episodes during Virtual Cognitive Stimulation Therapy through Gaze Estimation from Videos

Permalink

<https://escholarship.org/uc/item/4qq6q35b>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Teng, Xiaoru

Li, Yong

Wong, Gloria

et al.

Publication Date

2026

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Identifying Mind Wandering Episodes during Virtual Cognitive Stimulation

Therapy through Gaze Estimation from Videos

Xiaoru Teng

u3306585@connect.hku.hk

Department of Psychology, University of Hong Kong

Yong Li

101013946@seu.edu.cn

School of Computer Science and Engineering, Southeast University

Gloria H.Y. Wong

hoiyangloria.wong@reading.ac.uk

School of Psychology and Clinical Language Sciences, University of Reading

Antoni B. Chan

abchan@cityu.edu.hk

Department of Computer Science, City University of Hong Kong

Janet H. Hsiao

jhsiao@ust.hk

Division of Social Science, Department of Computer Science & Engineering, Hong Kong University of Science and Technology

Abstract

Recent studies suggest that eye movements may be used to monitor task-specific mind wandering (MW) episodes during online learning. We examined whether we could replicate these lab-based findings in real-life virtual cognitive stimulation therapy (vCST) sessions using eye movements estimated from video by machine learning methods without eye trackers. We found that lower joint attention was a reliable indicator of MW for tasks involving well-defined strategies. For tasks involving well-learned visual routines, in contrast to previous studies, higher rather than lower eye movement consistency was associated with MW. This may be because real-life vCST sessions involved less structured discussions with more varied task demands, where lower eye movement consistency may reflect active engagement. Our results suggest the feasibility of using machine learning to monitor eye movements from video for MW detection, and raise the issue of generalizability from lab research to real-life scenarios due to potential differences in task demands.

Acknowledgements

We are grateful to Research Grant Council of Hong Kong (CRF #C7055-21G).