



## OPEN ACCESS

EDITED AND REVIEWED BY  
Roberto Therón,  
University of Salamanca, Spain

\*CORRESPONDENCE  
Iulia Lefter  
✉ I.Lefter@tudelft.nl

RECEIVED 05 April 2024  
ACCEPTED 09 April 2024  
PUBLISHED 30 April 2024

CITATION  
Lefter I, Rook L and Chaspari T (2024)  
Editorial: Multimodal interaction technologies  
for mental well-being.  
*Front. Comput. Sci.* 6:1412727.  
doi: 10.3389/fcomp.2024.1412727

COPYRIGHT  
© 2024 Lefter, Rook and Chaspari. This is an  
open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](#). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that the  
original publication in this journal is cited, in  
accordance with accepted academic practice.  
No use, distribution or reproduction is  
permitted which does not comply with these  
terms.

# Editorial: Multimodal interaction technologies for mental well-being

Iulia Lefter<sup>1\*</sup>, Laurens Rook<sup>1</sup> and Theodora Chaspari<sup>2</sup>

<sup>1</sup>Faculty of Technology, Policy and Management, Delft University of Technology, Delft, Netherlands,  
<sup>2</sup>Faculty of Computer Science, Institute of Cognitive Science, University of Colorado Boulder, Boulder,  
CO, United States

## KEYWORDS

mental wellbeing, multimodal interaction, multimodal sensing, affective computing, virtual agents, e-health, responsible AI

## Editorial on the Research Topic

### Multimodal interaction technologies for mental well-being

## 1 Introduction

The world is witnessing a dramatic increase in people suffering from mental disorders and disorders of the brain—the majority of which put a heavy burden on the quality of life of individuals, families, institutions and society as a whole (Wittchen et al., 2011). Research shows that mental wellbeing of patients, caretakers and healthcare professionals generally suffers from the complicated nature of the mental disorder, the length and intensity of the treatment, the possibility of relapse, and long waiting lists that prevent timely reception of adequate care (Holmes et al., 2014; Roefs et al., 2022). Against this backdrop, multimodal interaction technologies hold promise for the digital monitoring and support of relevant stakeholders in conjunction with existing protocols and practices within the healthcare system.

Multimodal interaction technologies are artificial intelligence (AI) applications with automated emotion sensing and recognition capabilities. They rely on information sources such as speech, gestures, eye-movements and physiological reactions to situational stimuli. In recent years, progress has been made in the promotion of mental wellbeing via the rich, unstructured and often real-time data that these technologies generate [see for instance recent work on virtual reality (VR) exposure therapy environments (Hawajri et al., 2023) and mobile health (m-health) monitoring devices (Linardon et al., 2024)]. Notwithstanding these developments, the adoption of multimodal interaction technologies for mental wellbeing is hampered by several technical and societal issues.

The following challenges we consider most critical to the promotion of mental wellbeing via multimodal interaction technologies: first, and from a technical standpoint, ensuring the accuracy and reliability of emotion sensing and recognition algorithms across modalities presents a significant hurdle, especially for the effective integration of various information sources (Senaratne et al., 2022). Second, designing intuitive and user-friendly modes of interaction between technology and stakeholders—including effective communication under acknowledgment of diverse needs and preferences—is demanding (Miloff et al., 2020). In addition, facilitating mediated communications between patient and clinician is taxing, given that interactions must be empathetic, supportive, responsive,

and constructive. Finally, it is challenging to ensure user trust and acceptance in light of today's societal concerns regarding the ethical use of multimodal interaction technologies (Luxton and Hudlicka, 2022). Balancing innovation with regulatory frameworks that protect individuals' rights and mental wellbeing is currently an open challenge.

## 2 Contributions

Our call for papers aimed to collect use cases demonstrating the design and application of multimodal interaction technologies to promote mental wellbeing. We solicited interdisciplinary contributions focusing on specific user-technology interactions or the adoption of multimodal interaction technologies. Four papers were accepted for the Research Topic (Triantafyllopoulos et al., Lomas et al., El Kamali et al. and Zhong et al.).

The paper by Triantafyllopoulos et al. proposes a new method to detect speaker emotion, which is crucial for supporting personalized and timely mental health interventions. Using a deep neural network approach, the authors show that accounting for linguistic and paralinguistic (acoustic) information sources in integrated fashion leads to better prediction of emotion from speech compared to traditional—unimodal—approaches. This finding has important ramifications for the natural processing of written and spoken language, and for development of new mental health applications grounded in automated speech-based emotion recognition.

Lomas et al. address the issue of automated emotion recognition in multimodal interaction technologies from a different angle. In their paper, the authors zoom in on the problem how to align AI-generated and human emotions. They find that generative AI models are capable of generating emotional expressions that are well-aligned with a wide range of human emotions. However, this alignment depends on the design parameters of the AI model and on individual human perception. The authors argue that the emotions such technologies express must match how humans perceive them.

El Kamali et al. explore how older adults evaluate different modes of interaction of a virtual agent designed to support their wellbeing. A chatbot operating using text and images, and a tangible coach communicating via speech and light, were evaluated separately as well as together in several interaction models. In two empirical studies, the authors find that this user group appreciates virtual coaches that allow for multiple, parallel, and potentially redundant modes of interaction. In other words, redundancy-complement is a critical aspect of a conversational virtual agent from the perspective of the older adult. This finding points to the need to reckon with this often overlooked forms of multimodality in interaction design for mental wellbeing targeted at older adults.

Finally, Zhong et al. address the responsible use of anthropomorphism in the design of socially assistive robots

under the European Commission's current ethical guidelines for trustworthy AI. Through a perinatal depression screening scenario, the authors explore the role of AI transparency and degree of anthropomorphic conversation in patient-robot interaction. The authors argue that the degree of information transparency a robot shares with the patient matters under such circumstances. This paper serves as a reminder of the challenges arising from direct translation of ethical guidelines to real-life applications of multimodal interaction technologies for mental wellbeing.

Taken together, the contributions in this Research Topic emphasize that enhancing the performance of multimodal systems remains an ongoing challenge. The quest for efficient and convincing modes of communication and interaction between multimodal systems and their stakeholders will likely remain a topic of thorough investigation. We conjecture that in years to come, especially regulatory and ethical restrictions will gain prominence in the projects of researchers, practitioners, and developers. Regulatory constraints—and their likely tensions with intended system designs—should be seen as opportunities to eventually deliver better multimodal interaction technologies for mental wellbeing.

## Author contributions

IL: Writing – original draft, Writing – review & editing. LR: Writing – original draft, Writing – review & editing. TC: Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Hawajri, O., Lindberg, J., and Suominen, S. (2023). Virtual reality exposure therapy as a treatment method against anxiety disorders and depression—a structured literature review. *Issues Ment. Health Nurs.* 44, 245–269. doi: 10.1080/01612840.2023.2190051
- Holmes, E. A., Craske, M. G., and Graybiel, A. M. (2014). Psychological treatments: a call for mental-health science. *Nature* 511, 287–289. doi: 10.1038/511287a
- Linaron, J., Torous, J., Firth, J., Cuijpers, P., Messer, M., Fuller-Tyszkiewicz, M., et al. (2024). Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. *World Psychiatry* 23, 139–149. doi: 10.1002/wps.21183
- Luxton, D. D., and Hudlicka, E. (2022). “Intelligent virtual agents in behavioral and mental healthcare: ethics and application considerations,” in *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical and Policy Issues* (Cham: Springer), 41–55. doi: 10.1007/978-3-030-74188-4\_4
- Miloff, A., Carlbom, P., Hamilton, W., and Andersson, G., Reuterskiöld, L., and Lindner, P. (2020). Measuring alliance toward embodied virtual therapists in the era of automated treatments with the virtual therapist alliance scale (vtas): development and psychometric evaluation. *J. Med. Internet Res.* 22:e16660. doi: 10.2196/16660
- Roefs, A., Fried, E. I., Kindt, M., Martijn, C., Elzinga, B., Evers, A. W., et al. (2022). A new science of mental disorders: Using personalised, transdiagnostic, dynamical systems to understand, model, diagnose and treat psychopathology. *Behav. Res. Ther.* 153:104096. doi: 10.1016/j.brat.2022.104096
- Senaratne, H., Oviatt, S., Ellis, K., and Melvin, G. (2022). A critical review of multimodal-multisensor analytics for anxiety assessment. *ACM Trans. Comput. Healthc.* 3, 1–42. doi: 10.1145/3556980
- Wittchen, H.-U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., et al. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur. Neuropsychopharmacol.* 21, 655–679. doi: 10.1016/j.euroneuro.2011.07.018