



Open Research Online

Citation

Law, Effie L-C.; Thanydait, Santawat; Heintz, Matthias; Campbell, Abey and Wild, Fridolin (2021). Beyond Questionnaires: Innovative Approaches to Evaluating Mixed Reality. In: Proceedings of the 34th British HCI Workshop and Doctoral Consortium 2021, BCS Learning and Development Ltd.

URL

<https://oro.open.ac.uk/82554/>

License

(CC-BY 4.0) Creative Commons: Attribution 4.0

<https://creativecommons.org/licenses/by/4.0/>

Policy

This document has been downloaded from Open Research Online, The Open University's repository of research publications. This version is being made available in accordance with Open Research Online policies available from [Open Research Online \(ORO\) Policies](#)

Versions

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding

Beyond Questionnaires: Innovative Approaches to Evaluating Mixed Reality

Effie L-C. Law
University of Leicester, UK
lcl9@leicester.ac.uk

Santawat Thanydait
HKUST, China
santawat@ust.hk

Matthias Heintz
University of Leicester, UK
mmh21@leicester.ac.uk

Abey Campbell
University College Dublin, Ireland
abey.campbell@ucd.ie

Fridolin Wild
Open University, UK
f.wild@open.ac.uk

Mixed Reality (MR) technologies, including augmented and virtual reality, are increasingly used in a number of sectors, thanks to their capabilities to immerse users in multisensory interaction environments. However, as indicated by some systematic reviews, questionnaire remains the main method for evaluating the interaction quality of MR. There is a lack of innovative approaches addressing unique features of MR. It can dampen the advances of MR, as evaluation feedback can inform its future development. In this workshop we aim to explore this issue by inviting participants to share their practical experiences or conceptual ideas of evaluating MR in various contexts, using different methods and tools. The ultimate aim is to produce a research agenda on this topic for the community to examine it further in the future.

Mixed reality. Augmented reality. Virtual Reality. Evaluation methods. Head-mounted display.

1. INTRODUCTION

The recent rapid development of immersive and spatial computing technologies such as Microsoft HoloLens/Mesh, Google Glass, or Magic Leap suggests an imminent paradigm shift towards Mixed Reality (MR), which is even predicted to replace mobile phones in the coming decade (Leswing, 2021).

According to the widely recognized taxonomy by Milgram and colleagues (1994). MR refers to the full reality-virtuality continuum, including augmented reality (AR) and virtual reality (VR). AR is traditionally defined as technology with three core characteristics (Azuma, 1994): it combines real and virtual content; it is interactive in real-time; it is registered in 3D. The first characteristic differentiates AR from VR, which involves a total immersion of its user in simulated worlds, completely masking the real-world environment. Nevertheless, the boundaries between MR, AR and VR are getting blurred, which are attributable to the varied usage of these terms in the industrial and academic venues (Speicher et al., 2019).

With enriched sensory experiences enabled by MR-based applications, which are mostly multimodal (i.e. visual, audio, haptics, taste/flavour, smell), interacting with them can elicit positive emotional responses such as fun and pleasure in users, contributing to rich interaction experience. This accounts for the ever-increasing use of MR in a

range of sectors such as education (e.g. ARETE, Masneri et al., 2020) and medicine (e.g. VOSTARS)

To ensure the uptake of MR applications, it is critical to ensure their usability and user experience (UX). In accordance with the ISO 9241-210 definitions, an interactive system is usable when it can support its users to achieve their goals by completing related tasks with low or no error-rate, using optimal resources in terms of time and mental effort, and feeling satisfied with comfort whereas UX puts emphasis on user affect and sensation, and the meaningfulness of such interactions in everyday life (Law et al., 2009).

In the recent decade, there has been a number of research studies on designing and evaluating MR applications in various contexts. Several systematic literature reviews were conducted to analyse and synthesize which and how usability and UX methods were employed in these studies with some focusing on AR (e.g. Akçayır & Akçayır, 2017; Santos et al. 2013) and some on VR (e.g. Kim et al. 2020; Radianti et al., 2020). Interestingly, these systematic reviews consistently point to the fact that *questionnaire* is the main evaluation method alongside with the other established ones such as interview and observation.

This raises the concern whether the conventional methods and instruments, such as Jakob Nielsen's ten usability heuristics and System Usability Scale, are appropriate to evaluate MR technology or new approaches and bespoke scales addressing unique features of MR should be developed. The lack of

innovative methods and tools can dampen the advances of MR as evaluation feedback can inform its future development.

It also begs a more serious question **whether** and **what** MR-specific usability and UX evaluation methods are available, given that MR is so distinct from the traditional 2D-based technologies. The main goal of the workshop is to examine this key question (Section 3.1).

2. BACKGROUND

It is imperative to recognize the usefulness of questionnaires for evaluating the subjective perception of interaction quality. However, the lack of innovative methods for addressing the unique features of MR technology calls for more research efforts. Indeed, one distinct characteristic of MR is the use of head-mounted display (HMD), which is essential for VR. While many of AR applications are marker-based, their markerless counterparts, especially HMD-based, are on the rise.

HMD is a type of computer display device, worn on the head or is built in as part of a helmet, has a small display optic in front of one or each eye. The sophistication of HMDs can lead to their wider application and adoption, though high costs remain a challenge. Nonetheless, innovative methods are required to evaluate these emerging interaction devices. In the following, we present a review on specific measures related to the use of HMDs.

2.1 Location and movement

Since most HMDs include a spatial mapping feature and camera, participants' location and movements can be easily acquired and recorded through experiments for analysis later. For instance, Piumsomboon et al. (2019) calculated the total movement of participants in the environment to indicate physical load of each test condition (Figure 1). In collaborative scenarios, distance between collaborators can be used to investigate behavioral

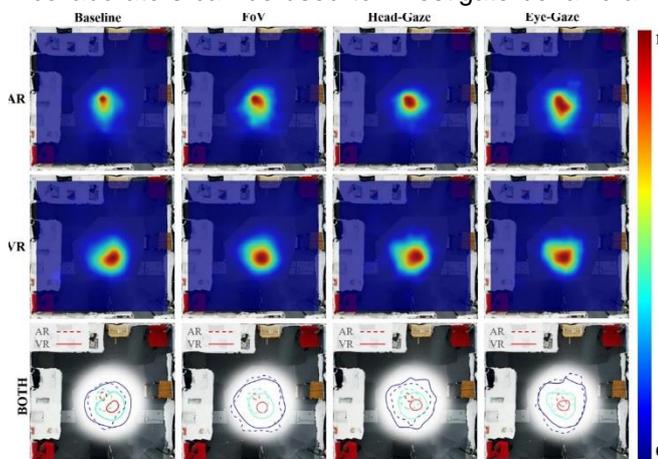


Figure 1. Piumsomboon et al. (2019) used heat-map to visualize physical movement between collaborators.

differences or proxemics interaction (Chow et al. 2019; Piumsomboon et al., 2019).

2.2 Head orientation, eyes gaze, and field of view

Head movements are commonly used as an alternative selection technique for HMD to provide hand-free selection for the users. With a portable eye tracker product such as Pupil Labs, eye gaze tracking is also possible to use in combination with the head movements to quantify user's focus and attention (Rahman et al., 2020). For example, Kytö et al. (2018) investigated head movements and eye tracking in their user study to find accurate selection techniques with HMD. Parr et al. (2020) used eye tracking equipment to investigate gaze behavior in children with developmental coordination disorder. In collaborative scenarios (Piumsomboon et al. 2019; Dey et al., 2017) measuring the time users spent sharing gaze or field of view can be used to examine communicative behavior or common ground establishing process.

2.3 Controllers and hands

Modern HMDs are often accompanied with controllers, which allow the users to interact with virtual objects in 3D spaces. The controllers' movement is a valuable measurement for user's physical load. Thus, multiple studies (Nguyen et al. 2017; Yan et al., 2018; Pontonnier et al. 2014) tracked the controller's movement to study ergonomic and task accuracy in the virtual environments. Hand gestures detection on the HMD is also possible by using additional hardware such as Leap Motion Sensor. Hand gestures, such as pointing gestures, are essential non-verbal cues during a collaborative scenario, and number of hand gestures is another indicator of communicative behavior or common ground establishing process (Piumsomboon et al. 2019; Yoon et al. 2020). Recent HMDs, such as Hololens2 and Oculus Quest2, support hand gesture detections, which make hand gestures detection accessible without additional hardware.

2.4 Emotions and Expressions

With the advancement of Deep Learning, facial expressions and emotions recognition are becoming faster and more reliable. Recently, multiple telepresence studies used facial expressions and emotions detection to provide feedback for the users, since facial expressions and emotions can provide important information for social interaction. For example, Zeng et al. (2019) detect students' facial expressions and emotions during an online classroom to provide feedback for the instructor. Samrose et al. (2021) use facial expressions and emotion detection tools as a communication coach to improve user's behavior in an online meeting. In AR HMDs, facial expressions and emotion detection are possible since the HMDs are optical see-through. For instance, "Empathy Glass" (Masai et al.

2016; Lee et al. 2016) used facial expression and emotion detection to study empathy between users in a remote collaborative task (Figure 2).

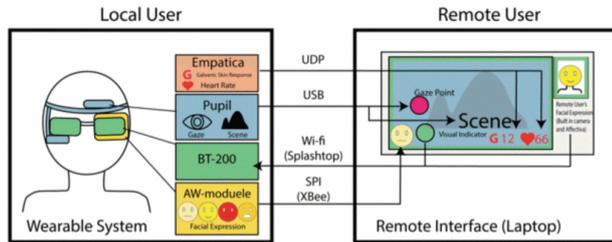


Figure 2. Empathy glass (Masai et al. 2016; Lee et al. 2016) shows local user's workspace, gaze, emotions (facial expression), heart rate, and galvanic skin response in real-time to the remote user.

2.5 Physiological data

With wearable devices, physiological data such as heart rate or oxygen level can be acquired to investigate the users' level of excitement during the tasks. As an example, Dey et al. (2017, 2018) measure heart rate in a collaborative VR game to understand empathic connection between users. Recently, neuroimaging methods such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), and functional near-infrared spectroscopy (fNIRS), are used to monitor social interaction in collaborative scenarios. A review by Barde et al (2020) describe a possible usage of neuroimaging methods in collaborative virtual environment.

3 WORKSHOP PLAN

The above concise literature review lays the groundwork for the workshop of which we present the details in the following sub-sections.

3.1 Aim and objectives

The main aim of the workshop is to explore the key question: **whether** and **what** MR-specific usability and UX evaluation methods are available, **how** these methods are applied, **which** factors facilitate or hinder their wider adoption.

The aim informs three objectives:

- To invite participants to share their practical experiences and insights on deploying different evaluation methods and tools for MR technologies in different contexts, especially in relation to the use of HMDs.
- To identify strengths and weaknesses of such methods and tools and challenges of applying them.
- To develop a research agenda for developing innovative evaluation methods for MR technologies.

Note that while we strongly invite submissions on HMD-based applications, others MR technologies

(e.g. AR markers) being evaluated with new or traditional approaches are welcome as well. We aim to analyse and discuss a range of research and practice.

3.2 Benefits and significance

The topic of the workshop is timely and relevant, given the ever-increasing interests and efforts in harnessing the power of immersive technologies in many sectors. MR, as emerging technology, entails innovative methods and tools to evaluate their interaction quality. While the prevailing usability and UX approaches like questionnaire are generally applicable, more insightful evaluation outcomes that can inform the design of MR will be obtained. For instance, to what extent MR is effective in detecting emotions through multisensory data and how to improve the accuracy.

The workshop will bring together researchers and practitioners who are interested in MR technology to engage in exchanges of knowledge. Sharing practical experiences of using different evaluation methods for different MR technologies in different contexts can lead to insights that will stimulate further work. This will be realised through a research agenda on innovative methods for evaluating MR technology, which will be the main output of the workshop.

3.3 Overview of planned activities

3.3.1 Prepare and distribute call for papers.

Papers on applying methods and tools for evaluating MR technologies will be solicited and should address the following questions:

- What MR technology is used in which application domain?
- What is the evaluation method used (or conceptualized)?
- How is the evaluation method applied? What challenges are encountered and how they are overcome?
- What are the strengths and weaknesses of the evaluation method?
- How will the evaluation method be enhanced?

3.3.3 Workflow

Time	Activity
9:00 – 9:15	Welcome & Introduction
9:15 – 10:00	Keynote
10:00 – 11:00	Paper presentations Session 1
11:00 – 11:15	Break
11:15 – 12:15	Paper presentations Session 2
12:15 – 13:00	Lunch
13:00 – 14:15	Group work Session 1
14:15 – 14:30	Break
14:30 – 15:45	Group work Session 2
15:45 – 16:00	Wrap up and Closing

Welcome & Introduction: The organisers will introduce the aim and objectives of the workshop and how it will be run. Each participant will be asked to briefly introduce themselves.

Keynote: An experienced MR researcher will be invited to give a keynote to highlight the state-of-the-art of MR technology.

Paper presentations Session 1 & 2: Each presenter will be given 10 to 15 minutes to present their paper, followed by Q&A.

Group work Session 1: Participants will be divided into groups of 4 to discuss the following questions:

- Which features of MR technology need to be further enhanced to make it become the future mainstream digital communication device (cf. mobile phones)?
- What are other use scenarios of the to-be-enhanced MR technology?

After discussing the questions for about 45 minutes, the plenum will be reconvened. Each group will present their discussion outputs to get feedback.

Group Work Session 2: The same groups of 4 will continue the discussion based on the outputs of Group Work Session 1. The new questions to be discussed are:

- What innovative methods and tools need to be developed to evaluate the to-be-enhanced MR features that can support the proposed use scenarios?
- How to consolidate the above outputs into a research agenda?

Similar to Session 1, the group discussion will last about 45 minutes. The discussion outputs will be presented in the plenum to invite feedback.

Wrap up and Closing: The organisers will discuss with the participants about the future post-workshop research activities. They can include a follow-up workshop in another venue, a joint publication, and a journal special issue.

4 ORGANISERS

Effie Lai-Chong Law is a full professor in Human-Computer Interaction (HCI), specialising in usability and user experience (UX) methodologies. Effie's recent research foci are Mixed Reality, Affective Computing and Trustworthy Autonomous System. She has rich experience in organising research events. She is an associate editor of *IJHCS* and *lwC*.

Santawat Thanyadit is a post-doc researcher in computer science and HCI, specialising in collaborative virtual environment for education. He received his PhD from the Hong Kong University of Science and Technology (HKUST). He investigates teaching and learning methods inside a virtual lab that allows instructors to customize lessons.

Matthias Heintz is a Teaching Fellow in Informatics with a focus on Human-Computer Interaction and Mobile & Ubiquitous Computing. His recent research work focuses on Participatory Design, usability, and user experience, especially in the context of technology-enhanced learning (TEL), without and with Augmented Reality (AR).

Abraham Campbell is an Assistant Professor for University College Dublin (UCD), Ireland, where coordinates the VR lab to explore tele-presence apps. He is an investigator for the CONSUS SFI Centre, exploring AR applications in farming. He is also a Chief Research officer with MeetingRoom, an online VR collaborative meeting software company.

Fridolin Wild is professor at the Institute of Educational Technology of The Open University, leading the Performance Augmentation Lab (PAL). Fridolin seeks to close the dissociative gap between abstract knowledge and its practical application, speeding knowledge refinement and integration into polished performance.

5 REFERENCES

- Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review, 20*, 1-11.
- ARETE: Augmented Reality Interactive Educational Systems. <https://www.aretoproject.eu/>
- Azuma, R. (1997). A Survey of Augmented Reality, In: *Presence: Teleoperators and Virtual Environments 6*, 4 (August 1997), 355-385.
- Barde, A., Gumilar, I., Hayati, A. F., Dey, A., Lee, G., & Billinghamurst, M. (2020). A Review of Hyperscanning and Its Use in Virtual Environments. In *Informatics*, 55.
- Chow, K., Coyiuto, C., Nguyen, C., & Yoon, D. (2019). Challenges and Design Considerations for Multimodal Asynchronous Collaboration in VR. *Proceedings of the ACM on Human-Computer Interaction, 3*, 1–24.
- Dey, A., Chen, H., Zhuang, C., Billinghamurst, M., & Lindeman, R. W. (2018). Effects of sharing real-time multi-sensory heart rate feedback in different immersive collaborative virtual environments. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 165–173.
- Dey, A, Piumsomboon, T., Lee, Y., and Billinghamurst, M. (2017). Effects of sharing physiological states of players in a collaborative virtual reality gameplay. In *Proceedings of Conference on Human Factors in Computing Systems (CHI'17)*, 4045–4056.
- Kim, Y. M., Rhiu, I., & Yun, M. H. (2020). A systematic review of a virtual reality system from the perspective of user experience. *International*

- Journal of Human-Computer Interaction*, 36(10), 893-910.
- Kytö, M., Ens, B., Piumsomboon, T., Lee, G.A. & Billinghamurst, M. (2018). Pinpointing: Precise head- and eye-based target selection for augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14.
- Law, E. L. C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 719-728).
- Lee, Y., Masai, K., Kunze, K., Sugimoto, M., & Billinghamurst M. (2016). A Remote Collaboration System with Empathy Glasses. *IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, Merida, Mexico, 2016, pp. 342-343.
- Leswing, K. (2021). The tech industry is looking to replace the smartphone
<https://www.cnbc.com/2021/02/20/apple-facebook-microsoft-battle-to-replace-smartphone-with-ar.html>. Accessed on 29/03/21
- Masai, K., Kunze, K., Sugimoto, M., & Billinghamurst, M. (2016). Empathy Glasses. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA 2016), ACM, 1257–1263.
- Masneri, S., Dominguez, A., Wild, F., Pronk, J., Heintz, M., Tiede, J., Nistor, A., Chiazese, G., & Mangina, E. (2020). ARETE - An Interactive Educational System using Augmented Reality, In: *Proceedings of the 6th International Conference of the Immersive Learning Research Network (iLRN)*, pp. 283-286
- Milgram, P.; Takemura, H.; Utsumi, A.; Kishino, F. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. *Telematic and Telepresence Technologies*. 2351. 10.1117/12.197321.
- Nguyen, H., Pontonnier, C., Hilt, S., Duval, T., & Dumont, G. (2017). VR-based operating modes and metaphors for collaborative ergonomic design of industrial workstations. *Journal on Multimodal User Interfaces*, 11, 97–111.
- Parr, J. V. V., Foster, R. J., Wood, G., Thomas, N. M., & Hollands, M. A. (2020). Children with developmental coordination disorder show altered visuomotor control during stair negotiation associated with heightened state anxiety. *Frontiers in Human Neuroscience*, 14, 512.
- Piumsomboon, T., Dey, A., Ens, B., Lee, G., and Billinghamurst, M (2019). The effects of sharing awareness cues in collaborative mixed reality. *Frontiers in Robotics and AI*, 6, 5.
- Pontonnier, C., Duval, T., & Dumont, G. (2014). Collaborative virtual environments for ergonomics: Embedding the design engineer role in the loop. In *International Workshop on Collaborative Virtual Environments (3DCVE)*, 1–5.
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147, 103778.
- Rahman, Y., Asish, S. M., Fisher, N. P., Bruce, E. C., Kulshreshth, A. K., & Borst, C. W (2020). Exploring eye gaze visualization techniques for identifying distracted students in educational VR. In *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 868–877.
- Samrose, S., McDuff, D., Sim, R. et al. (2021). MeetingCoach: An Intelligent Dashboard for Supporting Effective & Inclusive Meetings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'21)*.
- Santos, M. E. C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J., & Kato, H. (2013). Augmented reality learning experiences: Survey of prototype design and evaluation. *IEEE Transactions on Learning Technologies*, 7, 38-56.
- Speicher, M.; Hall, B.; Nebeling, M. (2019). What is Mixed Reality? In: CHI 2019, ACM, New York, NY, USA, doi.org/10.1145/3290605.3300767
- VOSTARS. Video and Optical See Through Augmented Reality Surgical Systems. <https://www.vostars.eu/>
- Yan, Y., Yu, C., Ma, X., Huang, S., Iqbal, H., & Shi, Y. (2018). Eyes-free target acquisition in interaction space around the body for virtual reality. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13.
- Yoon, B., Kim, H., Oh, S. Y., & Woo, W. (2020). Evaluating remote virtual hands models on social presence in hand-based 3D remote collaboration. *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 520–532.
- Zeng, H., Shu, X., Wang, Y., Wang, Y., Zhang, L., Pong, T. C., & Qu, H. (2020). EmotionCues: Emotion-oriented visual summarization of classroom videos. *IEEE Transactions on Visualization and Computer Graphics*.