A METAVERSE FOR THE GOOD

April 9th-10th, 2024

Barcelona

Editors:

Mel Slater

Michael Barngrover
Doron Friedman
Aurelio Lopez-Tarruella
Omar Niamut
Sylvia Pan
Matthias Quent
Mavi Sanchez-Vives

Thanks to

Thanks to Professor Joan Guàrdia Olmos, Rector, Universitat de Barcelona
For opening the conference.

EMRN

The European Metaverse Research Network

A Metaverse for the Good

Conference 9-10th April

CCCB, Barcelona, Spain

Editorial Committee

Mel Slater (Chair)
Michael Barngrover
Doron Friedman
Aurelio Lopez-Tarruella
Omar Niamut
Sylvia Pan
Matthias Quent
Mavi Sanchez-Vives

Support

Euridice Alvaro (IDIBAPS)
Alicia Ferrer (IDIBAPS)
Dario Bellido (UB)
Dr Cristina Pulido (UBNeuro)

Special thanks to Mavi Sanchez-Vives, Laura Hirvi and Hélène Verbrugghe

Front cover design by Mavi Sanchez-Vives rendered by ChatGPT4.

Background

The metaverse concept is that of a large-scale social environment based on the convergence of virtual and physical reality. Participants can take part in the metaverse through virtual, augmented or mixed reality. They can carry out everyday activities together, ranging from entertainment through to education, business, politics, health or even legal proceedings. At this time the metaverse is more of an idea than a reality, but a number of nascent systems exist, and all the technological components, including solutions involving AI, are available to bring this idea into existence.

The metaverse may have a profound impact on the way we live, bringing new beneficial ways of carrying out educational, social and economic activities, but may also introduce problems in the ethical, legal, business, psychological, political and social sphere. There will be opportunities and problems that at the moment we cannot envisage, since each new technology brings with it fundamental changes that are hard to foresee. The mobile phone in conjunction with the world wide web and social media have led to revolutionary changes in society including on the positive side deep connectivity between people, but on the other, a threat to democracy. The immersive nature of the metaverse will have the potential to enhance both the positive and negative impact by an order of magnitude.

The EMRN is a research group that studies these various aspects of the metaverse concept. The 2024 EMRN conference features some of the EMRN members and their work, but we invited other stakeholders from academia, industry and policy to contribute to the discussion, and most contributions are not from EMRN members, showing the widespread interest in this topic.

We invited submission of papers and posters, in order to foster new collaborations, and generate new ideas, moving towards a greater chance of envisaging the future. We also afforded the possibility of exciting and inspiring keynotes and panels to enable a multi stakeholder dialogue.

By bringing research, industry and policy makers together to engage in an open discussion we hope to enhance the probability of a ‘metaverse for the good’.

CONTENTS

Metaverse for Education and Industry:..11

09.20-09:35: Ágnes Karolina Bakk, Paul Ballack and Fernanda Parente - *Current Challenges of Using Metaverse-like Environments for Educational Purposes*
09:35-09:50: Martin Dobricki, Michael Rihs and Sina Shahmoradi - *EcoVR – A virtual reality for learning eco-friendly food choice*
09:50-10:05: Helena Ortiz-Gil, Fernando Canet and Xavier Ruyra - *Advances and Challenges in the Use of Virtual Reality for Surgical - Planning in Cardiology*

Towards an Intelligent Metaverse: The use and impact of AI:.................31

Moderator: Prof. Doron Friedman

Sebastian Rings, Lucie Kruse, Marvin Grabowski and Frank Steinicke - *Anything is Possible with an Intelligent Virtual Agent at the Table*
Marco Gillies - *AI for Movement Interaction in VR*
Mikołaj Łysakowski, Jakub Gapsa and Piotr Skrzypczynski - *Origami Unfolded: Augmented Reality and Machine Learning Integration for Manipulative Training Scenarios*

Poster Abstracts:..41

Moderator: Michele Di Dalmazi

Edurne Bernal-Berdun, Maria Plaza, Daniel Martin, Sandra Malpica, Belen Masia, Ana Serrano and Diego Gutierrez - *A Taxonomy for D-SAV360, a Dataset of Gaze*
Aileen O'Brien, Julia Hutchinson, Nik bin Fauzi, Michael Abbott, James Railton, Darren Bell, Sarah White, Jared Smith and Simon Riches - *A Controlled Trial of Virtual Reality Hypnotherapy in Students for Exam Anxiety*
Valerio Placidi, Matteo P. Lisi, Alisha Vabba, Althea Frisanco and Salvatore M. Aglioti - *Anti-mafia hero embodiment as a tool to fight mafias: an immersive virtual reality study*
Matteo P. Lisi, Martina Fusaro and Salvatore M. Aglioti - *Artificial agents delivering pain and touch on the embodied avatar of human participants are socially evaluated depending on the stimulus valence*
Provenzano Luca, Marina Scattolin, Giuseppina Porciello, Marina Serena

Panasiti and Salvatore M. Aglioti - *Bodily Self-Consciousness and dishonest behaviour in Digital Interactions*

Alisha Vabba, Valerio Placidi, Matteo P. Lisi and Salvatore M. Aglioti - *Cracking the "law of silence": immersive virtual reality to study behavioral and neurophysiological response to mafia-type crime*

Roberto Cuervo Rosillo, Telmo Zarraonandia Ayo and María Paloma Díaz Pérez - *Creating Immersive Collaborative Experiences: an Architecture prototype*

Ziwen Lu - *Exploring Outdoor Augmented Reality Registration Error Evaluation*

Michael Wiesing, Eckart Zimmermann, Eike A . Schmidt and Axel Buchner - *Fourteen days of visuospatial training do not reduce motion sickness susceptibility*

Tony Donegan, Beñat Amestoy-Alonso, Isaac Calvis, Edgar Rodriguez de la Calle, Ramón Oliva, Mel Slater, Andrés Combalía and Maria V. Sanchez.Vives - V*irtual reality-assisted therapy for low back pain management*

Artur Pilacinski, Sophia Bertoni, Gabriella Andrietta, Marita Metzler and Christian Klaes - *Phantom touch illusion: how the brain provides apparent touch in the metaverse?*

Martina Fusaro, Arianna Vecchio, Luca Provenzano, Salvatore M. Aglioti - *Exploring the analgesic effects of virtual and real touches: neurophysiological and behavioural evidence*

Matteo Dal Magro, Pablo Pérez, Jesús Gutiérrez - *eXtended Reality and Biosignals in the Therapy of Individuals with Intellectual and Physical Disabilities*

Daniela La Foresta and Ilaria Bruner - *Metaverse applications in learning processes: good practices and main challenges*

Tobias Bieseke - *Ndinguwe – I am you*

Jose Ferrer Costa, Maria-José Ciudad, Maribel González, José Luis Rodríguez García and Julisa Cortes - *Real-World VR Integration for Mild Cognitive Impairment Rehabilitation*

Shachar Maidenbaum - *Spatial features in human social interaction: emerging spatial patterns in online gatherings with spatial affordances*

Jesús Gutiérrez and Pablo Pérez - *Towards a recommendation for interactive test methods for subjective assessment of extended reality communications*

Rachel Simnett, Aileen O'Brien, Ursula James, Owain Rich, Russell Mulchansingh, Katie Grayson and Zillah Watson - *Developing an innovative VR solution for anxiety/stress relief using a co-design process with students*

Mariachiara Tirinzoni - *A Metaverse for a menstruating body - proposal for research*

Maria Lipińska and Renata Włoch - *Embodiment - in Virtual Reality: A Classical Sociological Theory Perspective*
Géraldine Fauville, Anaïs Voşki, Marijn Mado, Jeremy N. Bailenson, Annika Lantz- Andersson - *Underwater virtual reality for marine education and ocean literacy: technological and psychological potentials*

European Projects:..93

Moderator Prof. Eric Brousseau

Simona Neri - *GuestXR: A Machine Learning Agent for Social Harmony in eXtended Reality*
Eva Fringi, Thibault Vicente, Stephen Brewster, Lorenzo Picinali and Alessandro Vinciarelli - *Sonicom: Transforming auditory-based social interaction and communication in AR/VR*
Justyna Świdrak, Maria V. Sanchez.Vives, Ferran Argelaguet and Maria Roussou - *META-TOO: Investigating gender-based inappropriate social interactions in the Metaverse*

Ethical, legal and political implications of the Metaverse:......................97

Moderator Prof. Matthias Quent

Mariachiara Tirinzoni - *Metaverse, implemented A philosophical perspective on agenda-setting ethical tools for the development of immersive technologies*
Verity McIntosh - *What do policymakers need to know about harassment in the metaverse?*
Emmie Hine - *Extended Reality is a Fundamentally Mediated Reality*
Lucie Kruse, Catharina Rudschies, Sebastian Rings and Frank Steinicke - *Considering Psychotherapy in the Metaverse*
Giovanni Sorrentino - *Biometrics and Inferred data in the metaverse*

Creating the Metaverse: Technical Challenges and Solutions:..............125

Moderator: Prof. Sylvia Xueni Pan

Anthony Steed - *Some Technical Challenges of Scaling from Social Virtual Reality to Metaverse(s)*
Sebastian Friston and Anthony Steed - *A Brief Review of Asymmetric Social VR*
Joanna Luberadzka, Enric Guso Muñoz, Umut Sayin and Adan Garriga Torres - *Acoustic coherence in extended reality: challenges and strategies*
Ju Zhang, Yuke Pi, Haoyang Du , Fang Ma and Sylvia Xueni Pan - *Digital Twin Embodied Interactions Design: Synchronized and Aligned Physical Sensation in Location-Based Social VR*
Hossein Javidnia, Diego Gonzalez, Arman Najari, Cai Felip and Yolanda Gómez - *Digitizing Fashion: 3D Garment Try-On and Avatar Creation Try-On and Avatar Creation*

Medicine, Neurotechnology and the Metaverse:...................................157

Moderator: Prof. Maria V. Sanchez-Vives

Eduardo Garcia-Giler, Bill Zhang, Mariana Misawa, Yulia Pyatova, Monica Daibert-Nido, Atri Sen, Samuel N. Markowitz, Kyle Cheung, Danielle Tchao, Pi Nasir, Inci Yamana Bajin, Eric Bouffet, Lora Appel and Michael Reber - *Immersive Virtual Reality: a paradigm shift and a disruptive tool for visual neurorehabilitation*
Toinon Vigier, Didier Acier, Samuel Bulteau, Solène Thevenet and Yannick Prié - *Towards immersive teleclinics in the Metaverse: the case of neuropsychological consultation*
Paweł Śniatała, Szymon Baliński, Mariusz Głąbowski, Sławomir Michalak, and Ewa Kozielewska–Zwierska - *Virtual, Personalized Rehabilitation Environment*

Identity and the body - metaverse applications:....................................177

Moderator: Mel Slater

Georgiana Cristina Dobre, Marta Wilczkowiak, Marco Gillies, Sylvia Xueni Pan and Sean Rintel - *Avatars in mixed reality meetings: A field study of avatar facial realism on felt and perceived emotion*
Beñat Amestoy, Tony Donegan, Isaac Calvis, Justyna Świdrak, Edgar Rodriguez de la Calle, Andrés Combalía, Ramón Oliva and Maria V. Sanchez-Vives - *Focus groups in the Metaverse: the experience of shared virtual spaces for patients, clinicians and researchers*
Yuke Pi, Sylvia Xueni Pan, Mel Slater and Justyna Świdrak - *Embodied Time Travel in VR: Witnessing Climate Change and Action to Overcome It*
Pablo Pérez, Marta Orduna, Jesús Gutiérrez and María Nava-Ruiz - *Towards the Incluverse: an inclusive metaverse (for people with intellectual disability)*
Adrià Gabarnet, Adrián Montesano and Guillem Feixas - *Construing the virtual self: The psychological functions of virtual identity in online videogames*
Jonathan Giron, Ori Chajuss, Emi Schenkelbach Kliener and Einat Geva - *Navigating the Metaverse: Enhancing Non-Verbal Communication Through Externalized Interoceptive Signals*

Metaverse for Industry and Education

Current Challenges of Using Metaverse-like Environments for Educational Purposes

Ágnes Karolina Bakk[1*], Paul Ballack[2], Fernanda Parente[3]

[1]Immersion and Interaction Hub, Innovation Center, Moholy-Nagy University of Art and Design Budapest, Hungary

[2]Paul Ballack,

[3]Fernanda Parente, Goethe Institute e.V., Germany

*** Correspondence:**
Ágnes Karolina Bakk
bakk@mome.hu

Keywords: Metaverse1, education2, professionals3, assessment4, innovation5.

Abstract

This paper presents the partial findings of a qualitative study aimed at mapping how stakeholders in higher educational institutions explore metaverse technologies, identifying the affordances these technologies seek to exploit, and examining how these technologies affect teaching methodologies. Through the analysis of 20 in-depth interviews, the paper uncovers the education and research professionals' views on metaverse and explores its potential to introduce new approaches within educational methods, particularly in the context of the experiential turn.

1 Introduction

Since it acquired its initial hype in October 2021 with Mark Zuckerberg's statement about creating Meta and its own social VR space, the concept of metaverse offered a ground of speculation as well as hope for the XR community for creating new means and use cases of togetherness in virtual spaces. While there is no established consensus concerning what the concept of metaverse(s) might mean, there appeared a multitude of opinions and many whitepapers addressing the issue, as it seemed to many that digital or virtual worlds will be affecting many aspects of our life, including economy, industry, and governance. For instance, many EU strategies have already addressed the anticipated transitions ("An EU Initiative on Virtual Worlds: A Head Start in the next Technological Transition" 2023), the industrial relatedness and also citizens' recommendations about the directions of developing the virtual worlds ("Staff Working Document: Citizens' Panel Report on Virtual Worlds. Citizens' Report from the Citizens' Panel with the 23 Recommendations." 2023), and various industry whitepapers looked more closely into how digital ethics can manifest in the metaverse (Chi et al. 2023); (Arunov and Scholz 2023) or into what shape these spaces will take in the future and how can they provide a space for democracy (Anderson and Rainie 2023). On the other hand, many of these whitepapers do not focus on how the various stakeholders consider the role of education in the metaverse-like environments in the frame of higher education.

For this study we aim to scrutinize how current educational and research professionals who work in the area of VR, immersive storytelling, psychology, and technology see the effect of Metaverse on education or educational research and the speculations surrounding the issue. In order to understand this, we analyzed 20 interviews from a wider qualitative study conducted by Goethe Institute.

2 Background literature

2.1 The Metaverse(s)

Besides some of the well-known occurrences of the concept of a metaverse, such as in Neal Stephenson's famous novel *Snowcrash* (Stephenson 2022), researchers from informatics and computer science were already invoking the concept before 2021. In 2006, the Metaverse Roadmap Summit hold by the Stanford Research Institute International created scenarios concerning metaversal futures centered around two key continua: augmented-simulated and external-intimate, mentioning Augmented Reality, Lifelogging, Mirror Worlds and Virtual Worlds as four major scenarios for the metaverse (see Figure 1), which can still be encountered in recent literature (Smart, Cascio, and Paffendorf 2006). Virtual worlds increasingly augment the economic and social life of physical world communities, while mirror worlds serve as informationally-enhanced virtual models or "reflections" of the physical world. Augmented reality utilizes Metaverse technologies to enhance the external physical world for individuals through location-aware systems and layered networked information. Lifelogging employs augmentation technologies to record and report the intimate states and life histories of objects and users, supporting object- and self-memory, observation, communication, and behavior modeling. Among these four scenarios, the concept of Virtual Worlds emerged as the dominant understanding of the metaverse during the COVID-19 pandemic (Key 2021, 6), although in 2009 Davis et al. (2009) (91) had already defined metaverses as "immersive three-dimensional virtual worlds (VWs) in which people interact as avatars with each other and with software agents, using the metaphor of the real world but without its physical limitations" and which enables teams to overcome geographical or other types of challenges for collaboration.

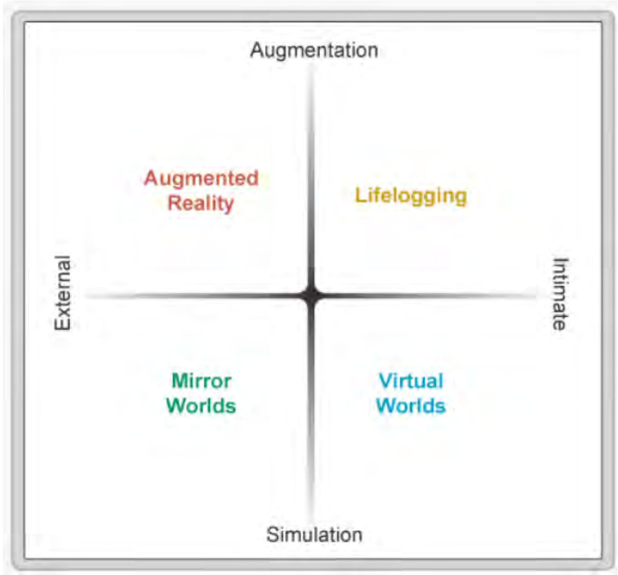


Figure 1. A diagram of the 4 types of Metaverse scenarios according to Metaverse Roadmap Summit 2006

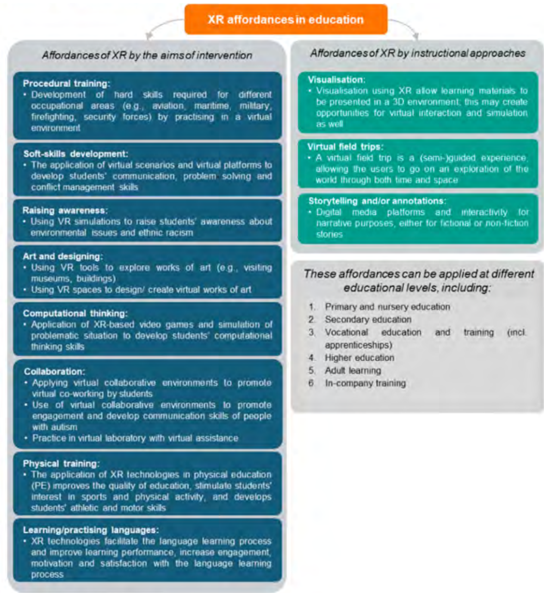
According to Go et al. (Go et al., 2021) the metaverse is “a 3D-based virtual reality in which daily activities and economic life are conducted through avatars representing the real themselves.” (as translated to English and cited in Kye et al. 2021), while Lee et al. define it as “a world in which virtual and reality interact and co-evolve, and social, economic, and cultural activities are carried out in it to create value.” (Lee et al. 2021) An often quoted definition of the metaverse is by Matthew Ball, according to whom the metaverse is: “A massively scaled and interoperable network of realtime rendered 3D virtual worlds that can be experienced synchronously and persistently by an effectively unlimited number of users with an individual sense of presence, and with continuity of data, such as identity, history, entitlements, objects, communications, and payments.” (Ball 2022)

As one can see, all of these definitions suggest that metaverse(s) are virtual worlds that can often connect users either by connecting their physical and digital life or help them overcome various challenges to collaboratively conduct various activities and create value as well. Mark Zuckerberg listed eight fundamental building blocks that Metaverse should be constituted by: 1. Feeling of presence; 2. Avatars; 3. Personal space (home space); 4. Teleportation; 5. Interoperability; 6. Privacy and safety; 7. “Virtual good”; 8. Natural interfaces (*The Metaverse and How We’ll Build It Together. Mark Zuckerberg’s Talk. Connect Conference 2021.*, n.d.). In the current study we will also map what professionals mean by the metaverse and how they perceive the educational challenges.

2.2 Education in the Metaverse(s)

Value-creation can be manifested through personal meetings and also educational activities that involve a sense of togetherness. This sense of togetherness, even though differing from the physical learning environments, enable social VR experiences suitable for educational activities. (Gunkel et al. 2018) There are several reviews that aim to map how the metaverse can contribute to various educational activities, but due to the undecisive usage of the term, these reviews usually offer a broader view reflecting the broad understanding of the researchers. In their systematic review, Tlili et al. (Tlili et al. 2022) outline articles that approach the question of education and metaverse in the framework of the Metaverse Roadmap Summit scenario, and they identify three waves that the publications thematize: social aspect, the potentials of technology-mediated presence and immersive technologies, and the third wave concerns "self-organized AI-powered virtual learning ecologies". The authors conclude that the Metaverse is not a new technology, but has been here with us for two decades in various forms, (Tlili et al. 2022, 11) and they also claim that the threats of the technology are not mapped in the current literature, only the possibilities are described.

Metaverse like environments manifested in VR can actually be useful for language learners (Thrasher 2023), where a pedagogical approach that incorporates game mechanics and gamification can create an immersive and participatory context for developing language competences (Wu, Zhang, and Lee 2024). According to Tlili et al. (within the broader sense that they frame the metaverse) the subject matters that are taught are mainly related to natural sciences, mathematics, and engineering (53%), and to a lesser extent concern general education (15%) and arts and humanities (11%). Discussing the relation of metaverse and education, Key et al. (Kye et al. 2021) draw our attention to the fact that we still need to analyze students' understanding of and demands from the metaverse, and also the effectiveness and attractiveness of the metaverse, while the instructors should also gain a deeper understanding of the platform's possibilities and pitfalls. They also point out the need to develop an educational platform that prevents the misuse of students' data (ibid.). Several studies show that VR-based use cases of the Metaverse-like environments are very suitable for specific trainings, such as aviation training (Fussell and Truong 2020), maritime educational training (Renganayagalu, Mallam, and Hernes 2022). It is also a suitable tool for teaching soft skills and training for arts and humanities. (Tlili et al. 2022, 22)The metaverse, understood as virtual worlds, can be fully accessed and experienced in its immersive capacities mainly by VR headsets. The tools that enable users to enter immersive virtual worlds can be used for instructional or interventional purposes (European Commission. Directorate General for Communications Networks, Content and Technology and Visionary Analytics. 2023), and therefore can allow the application of these technologies for education, from school education up until in-company education.



Source: Visionary Analytics, 2022

Figure 2. European Commission's report taxonomy on the XR affordances that can be used in education

Despite the multifaceted understanding of the Metaverse and its possibilities in the field of education, there is a research gap in comprehending the needs and the understanding of the stakeholders belonging to the cultural sector, mainly from the field of the education and research in arts and culture. What significance does the metaverse currently have for cultural and educational institutions? How relevant is it for them? What are the potentials for cultural and educational institutions in the metaverse? To what extent can they adapt to the recent trends? What are the various types of constraints or dangers that they encounter during possible use cases in education or research activities? As the metaverse-related technologies become extremely ubiquitous, our aim is to map the opportunities and the challenges of using or experimenting with metaverse-like environments or virtual worlds in education. Our study aims to serve as a springboard for the development of educational platforms or social virtual worlds and to offer specific key points for the designers of such platforms. To this end, the conducted interviews with professionals from the field of education and research will help us to explore the following research questions:

RQ1: How do professionals define (the) metaverse(s) and how do they situate it within the present and the future?

RQ2: What are the specific affordances that technologies related to the 'metaverse' provide?

RQ3: What are the possible effects caused by the metaverse-related technologies on the education?

3 Methodology

The interview guideline helps us inquire into various interest areas of the stakeholders from cultural institutions in order to map possible practices and needs from the art, culture and educational field. The aim was to examine the motivations for using metaverse-related technologies and solutions as well as the way these stakeholders understand metaverse platforms, that is, how they see the current status of the related technologies and how, in their view, these technologies can be used by cultural and educational institutions.

3.1 Participants

The qualitative study was conducted by the Goethe Institute between December 2022 and March 2023, via 41 in-depth semi-structured interviews with international professionals working in established cultural and educational institutions or VR-related companies. The participants were recruited either individually or via the snowball technique. Among the interviewees, 20 were involved in a higher educational institute or research institute; in the current study we analyzed the interviews made with these 20 participants.

3.2 Analysis

We conducted thematic analysis on the data using Braun and Clarke's protocol Braun and Clarke 2022). Thematic analysis allows identifying themes from the data, capturing participants' subjective experiences, and helps with analyzing emerging digital experiences and practices such as those related to using metaverse-related technologies in education. First, each interview was read by one member of the research team, coding them for important information related to the study research questions. Codes that represented similar concepts across different interviews were synthesized into themes, such as the variety of aspects around the concept of metaverse, the technological characteristics that allow them to be used in education and how these various characteristic, such as embodiment and simulation, can motivate us to rethink education in metaverse-like environments.

4 Findings

4.1 Future is not what it used to be: Understanding the Metaverse and its shortcomings in education

4.1.1 Technologies

Professionals working in higher educational institutions who use new technologies (or technologies related to the metaverse, e.g. VR or AR) often differ in their understanding of what metaverse means. P3 stated that metaverse "is a bit like the extended reality-mode, but real". Similarly, P1 thinks that the metaverse is a "fluid concept", while P11 and P18 mentioned that the metaverse is the (new) Internet. P9 suggested that the term 'media sphere' should be used for metaverses. According to many of the participants we do not fully grasp the potential of these technologies, but many of them note how the metaverse is connecting the physical reality with the virtual, and often quote Matthew Ball's definition as well (e.g. P19, P20). When discussing the related technologies, the participants are mentioning mainly VR (P5, P11, P16), AR (P8, P15) as well as mobile application developing (P7). Only P1 mentioned more in-depth technical algorithms that can be used in the metaverse, namely the blockchain, although P20 also emphasized several times the economical dimension of the metaverse. These statements also point out the enduring applicability of the four scenarios that were developed in 2006 in the Metaverse Roadmap. AR and VR technologies enable the users to create mirror world and virtual worlds as well, both serving different purposes. However, many participants mentioned that the Metaverse itself is not here yet (e.g. P1) or that the potentials of the Metaverse "are unknown yet" (P9). Some participants also expressed their concerns about data protection (P8) as many of these tools can record data about the user, which is also very much connected to the lifelogging aspect of the metaverse scenarios described in the 2006 roadmap (Smart, Cascio, and Paffendorf 2006).

4.1.2 Affordances

Several participants have mentioned particular characteristics of metaverses, namely simulation and embodiment. Simulation-wise, the opinions are manyfold: while P7 was not sure why to always use 'fantasy-worlds' to teach real-world scenario-based skills for the students, similarly to P4 who emphasizes that real-world situations and their re-enactment can offer authentic learning experiences, many other professionals emphasized how important is to create imagination-based worlds (P8, P15). The simulation aspect of the metaverse-like environments can also be approached from the notion of 'spatiality'. P19 mentioned that metaverse means moving from 2-dimensional space to 3 dimensions. Spatiality is also emphasized by P2 who described how learning in 3D and experimenting with new physicalities (e.g. folding shapes in VR) could be important assets, but also pointed at the importance of preparing the students to not to confuse virtuality with the reality (P11).

P9 mentioned that "technology is embodied knowledge", meaning that corporeality and embodiment is a very important asset of being in a virtual world in the form of an avatar; they allow not only practicing interactions or simulating situations in various types of trainings, but they can also make getting in touch with each other in social virtual spaces much easier, while P12 acknowledged that the appearances are not hindering us to reach out. The embodies nature of presence also means that the participants have to be offered a safe space (P11), where they can also assess how to present their digital identity and how to deal with anonymity in an ethical way (P14, P17).

4.2 Future is not yet here

There is a deviation between the expectation of the students and of the teacher, therefore the experiential turn in the education is being addressed more and more but "is not here yet" mainly due to technical limitations. On the one hand, for some educators–leaving aside the technical difficulties–metaverse promises joyful situations to interact with virtual characters (P20). On the other hand, assessing the students' knowledge is a challenge, as there can be a huge gap between what the students really learned, and what they feel that they learned (P20). While the togetherness is an important asset of the social virtual worlds, according to P10 developing social competence in the metaverse is still a question; this is also because young users have more difficulties to understand each other's emotional facial expressions (P9). It is important to recognize that there is not a generalized answer to various educational methodological challenges, and one always has to look at the given objectives (P7). The learning process can include knowledge acquisition (P7), and it offers a great opportunity to nurture interdisciplinary methods, but one also has to address the question of what happens to the students' motivation after the initial hype and how to maintain their attention and their willingness to learn (P7). Therefore, it is important to reframe the educational settings into a more 'experiential learning'(P12), where, besides the simulation-based trainings, other soft skills (e.g. social skills) would also be developed, for example with the help of role-playing (P16).

5 Conclusion

In this study, we mapped (only partially, due to the limited space) the current status of how the metaverse(s) and other related technologies are assessed and deployed by professionals working in higher educational institutions, and we grouped the information gathered from 20 in-depth interviews around two themes, "how the future is not what it used to be" and "how it is not here yet". With this we aimed to uncover possible avenues that can help designers of these spaces to re-think or further develop spaces for education, where the experience-based learning can be assessed, many skills can be taught in parallel, and the new sense of embodiment can become a tool for critical thinking about virtual identities.

6 Acknowledgments

The interviews were co-read by colleagues in the Moholy-Nagy University of Art and Design: Sebastian Gschanes, Samuel Chovanec, Edwina Portocarrero, Borbála Tölgyesi.

7 Attachment

Number	Position	Country	Gender
P1	professor	Norway	Male
P2	coordinator	Germany	Male
P3	coordinator	Germany	Male
P4	professor	Germany	Male
P5	professor	Germany	Male
P6	professor	Germany	Male
P7	researcher	Germany	Female
P8	coordinator	Germany	Female
P9	researcher	Germany	Female
P10	researcher	Germany	Female
P11	coordinator	Ecuador	Female
P12	professor	Israel	Female
P13	professor	Hong Kong	Female
P14	lecturer	The Netherlands	Female
P15	professor	Peru	Male
P16	coordinator	US	Male
P17	coordinator	non-disclosed	Female
P18	professor	Germany	Male
P19	professor	US	Male
P20	professor	US	Female

Figure 3. The list of the interview study participants

8 List of references

"An EU Initiative on Virtual Worlds: A Head Start in the next Technological Transition." 2023. Communication COM(2023) 442. Strasbourg. https://digital-strategy.ec.europa.eu/en/library/eu-initiative-virtual-worlds-head-start-next-technological-transition.

Anderson, Janna, and Lee Rainie. 2023. "The Future of Digital Spaces and Their Role in Democracy. Pew Research Center." https://www.pewresearch.org/internet/2021/11/22/the-future-of-digital-spaces-and-their-role-in-democracy/.

Arunov, Elmar, and Maike Scholz. 2023. "Digital Ethics in The Metaverse. Position on Ethical Guidelines. Whitepaper." Deutsches Telekom. https://www.telekom.com/resource/blob/1056852/634487d5eb9f615228c0dca33b8cfece/dl-whitepaper-metaverse-data.pdf.

Ball, Matthew. 2022. *The Metaverse: And How It Will Revolutionize Everything*. First edition. New York, NY: Liveright Publishing Corporation, a division of W.W. Norton & Company.

Braun, Virginia, and Victoria Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2): 77–101. https://doi.org/10.1191/1478088706qp063oa.

———. 2022. *Thematic Analysis: A Practical Guide*. London ; Thousand Oaks, California: SAGE.

Chi, Cheng, Daozhuang Lin, Ramesh Ramadoss, Yu Yuan, Zihang Yin, Can Luo, Wenqu Chen, Bin Yang, Ling Wei, and Ralf Ma. 2023. "The Industrial Metaverse Report." *The Industrial Metaverse Report*, 1–20.

European Commission. Directorate General for Communications Networks, Content and Technology. and Visionary Analytics. 2023. *Extended Reality: Opportunities, Success Stories and Challenges (Health, Education) : Final Report.* LU: Publications Office. https://data.europa.eu/doi/10.2759/121671.

Fussell, Stephanie G., and Dothang Truong. 2020. "Preliminary Results of a Study Investigating Aviation Student's Intentions to Use Virtual Reality for Flight Training." *International Journal of Aviation, Aeronautics, and Aerospace* 7 (3): 2.

Go SY, Jeong HG, Kim JI, Sin YT. 2021. "Concept and Developmental Direction of Metaverse." *Korea Inf Process Soc Rev [Internet]* 28: 7–16.

Gunkel, Simon, Hans Stokking, Martin Prins, Omar Niamut, Ernestasia Siahaan, and Pablo Cesar. 2018. "Experiencing Virtual Reality Together: Social VR Use Case Study." In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*, 233–38. SEOUL Republic of Korea: ACM. https://doi.org/10.1145/3210825.3213566.

Kye, Bokyung, Nara Han, Eunji Kim, Yeonjeong Park, and Soyoung Jo. 2021. "Educational Applications of Metaverse: Possibilities and Limitations." *Journal of Educational Evaluation for Health Professions* 18 (December): 32. https://doi.org/10.3352/jeehp.2021.18.32.

Lee, Lik-Hang, Tristan Braud, Pengyuan Zhou, Lin Wang, Dianlei Xu, Zijun Lin, Abhishek Kumar, Carlos Bermejo, and Pan Hui. 2021. "All One Needs to Know about Metaverse: A Complete Survey on Technological Singularity, Virtual Ecosystem, and Research Agenda." arXiv. http://arxiv.org/abs/2110.05352.

Renganayagalu, Sathiya Kumar, Steven Mallam, and M. Hernes. 2022. "Maritime Education and Training in the COVID-19 Era and Beyond." *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation* 16 (1): 59–69. https://doi.org/10.12716/1001.16.01.06.

Smart, John, Jamais Cascio, and Jerry Paffendorf. 2006. "Metaverse Roadmap. Pathways to the 3D Web." https://www.w3.org/2008/WebVideo/Annotations/wiki/images/1/19/MetaverseRoadmapOverview.pdf.

"Staff Working Document: Citizens' Panel Report on Virtual Worlds. Citizens' Report from the Citizens' Panel with the 23 Recommendations." 2023. SWD(2023) 250 final. https://digital-strategy.ec.europa.eu/en/library/staff-working-document-citizens-panel-report-virtual-worlds.

Stephenson, Neal. 2022. *Snow Crash*. Great Britain: Penguin UK.

The Metaverse and How We'll Build It Together. Mark Zuckerberg's Talk. Connect Conference 2021. n.d. Meta. Accessed March 3, 2023. https://www.youtube.com/watch?v=Uvufun6xer8&ab_channel=Meta.

Thrasher, Tricia. 2023. “Meeting in the Metaverse: Language Learners’ Insights into the Affordances of Virtual Reality.” In *Reimagining Education: Studies and Stories for Effective Learning in an Evolving Digital Environment*, edited by Deborah Cockerham, Regina Kaplan-Rakowski, Wellesley Foshay, and Michael J. Spector, 179–94. Educational Communications and Technology: Issues and Innovations. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-25102-3_17.

Tlili, Ahmed, Ronghuai Huang, Boulus Shehata, Dejian Liu, Jialu Zhao, Ahmed Hosny Saleh Metwally, Huanhuan Wang, et al. 2022. “Is Metaverse in Education a Blessing or a Curse: A Combined Content and Bibliometric Analysis.” *Smart Learning Environments* 9 (1): 24. https://doi.org/10.1186/s40561-022-00205-x.

University of Nebraska at Omaha, Alanah Davis, John Murphy, University of Nebraska at Omaha, Dawn Owens, University of Nebraska at Omaha, Deepak Khazanchi, University of Nebraska at Omaha, Ilze Zigurs, and University of Nebraska at Omaha. 2009. “Avatars, People, and Virtual Worlds: Foundations for Research in Metaverses.” *Journal of the Association for Information Systems* 10 (2): 90–117. https://doi.org/10.17705/1jais.00183.

Wu, Junjie Gavin, Danyang Zhang, and Sangmin Michelle Lee. 2024. “Into the Brave New Metaverse: Envisaging Future Language Teaching and Learning.” *IEEE Transactions on Learning Technologies* 17: 44–53. https://doi.org/10.1109/TLT.2023.3259470.

EcoVR – A virtual reality for learning eco-friendly food choice

Martin Dobricki[1*], Michael Rihs[1,2], Sina Shahmoradi[1]

[1] Education & Digital Technologies Lab, Institute for Research, Development and Evaluation, Bern University for Teacher Education, Bern, Switzerland

[2] Cognitive Psychology, Perception and Research Methods Lab, Institute for Psychology, University of Bern, Bern, Switzerland

*** Correspondence:**
Corresponding Author
martin.dobricki@phbern.ch

Keywords: learning, food choice, sustainability, human-environment relation, virtual reality.

Abstract

Eco-friendly food choice requires that human individuals relate themselves and their natural environment to each other. They may establish this human-environment relation by experiencing their food choices being accompanied by negative or positive changes of the eco-system. Hence, educating human individuals by enabling them to such human-environment experiences may foster their eco-friendly food choice. Yet, this would require inducing climate-related environmental changes on purpose that immediately follow one's choice of food. We have, therefore, developed a VR consisting in a life-sized virtual environment in which grabbing food is accompanied by climate-related environmental changes or their reversal depending on the food's carbon footprint. Moreover, we have started to investigate the educational utility of this *EcoVR* by asking experienced lower secondary school teachers to rate it in controlled experiments. Our findings suggest that the teachers regard EcoVR to be useful for teaching climate-friendly food consumption. Moreover, they regard this educational usefulness of EcoVR to be higher than that of its PC version. Hence, in lower secondary education EcoVR may have the potential to serve pupils for learning climate-friendly food choice. Moreover, it may serve as a digital bridge or «metaverse» between the learning of relevant skills at school and the application of these skills in real life.

1 Introduction

The ecosystems of earth are important for human well-being (Millennium-Ecosystem-Assessment, 2005). Climate-related changes of these ecosystems may therefore be regarded as a threat. Such climate-related environmental changes are driven by greenhouse gas emissions (Yue and Gao, 2018). Almost 30 % of these emissions are caused by the production of food (Poore and Nemecek, 2018). Hence, eating more food with a low carbon footprint and less food with a high carbon footprint has the potential to reduce greenhouse gas emissions significantly. Such eco-friendly food consumption requires that human individuals relate themselves and their natural environment to each other (Davis et al., 2009). They may establish this human-environment relation (Kaufmann-Hayoz, 2006) by experiencing their food choices being accompanied by negative or positive environmental changes. Hence, educating human individuals by enabling them to such human-environment experiences may foster their eco-friendly food choice. Yet, this would require inducing climate-related environmental changes on purpose that immediately follow one's choice of food. Being impossible in physical reality this can be accomplished in virtual reality (Markowitz and Bailenson, 2021). In fact, climate-related environmental changes can be programmed in virtual reality (VR) in a way that they occur or disappear as the immediate consequence of one's own behavior (Dobricki et al., 2021). Accordingly, we have developed a VR consisting in a life-sized virtual environment in which grabbing food is accompanied by climate-related environmental changes or their reversal depending on the food's carbon footprint. A promising opportunity for using this VR to teach eco-friendly food choice early to a large part of society may be in lower secondary education (Keller and Brucker-Kley, 2021). Here we have, therefore, started to investigate the utility of EcoVR for this educational level. For this purpose, we have asked experienced secondary school teachers to assess the educational utility of EcoVR in controlled experiments.

2 Material and methods

2.1 Participants

Ten teachers (3 males, mean age = 44.9 yrs, SD = 9.2 yrs) with normal or corrected-to-normal vision participated. This sample size was suggested by a priori power analysis (Brysbaert, 2019) specified as follows: $f = 0.3$, alpha error = 0.05, power = 0.8, corr. = 0.8. All participants had been teaching in lower secondary school for at least 7 years. All participants gave their written informed consent and could have withdrawn from the study at any time.

2.2 Stimuli and apparatus

The participants were presented with a 3D virtual environment from a first-person perspective (see here). This was accomplished by using the graphics engine Unity3D on a Lenovo Legion 7 computer with an AMD Ryzen 9 processor and an NVIDIA GeForce RTX 3080 graphics card. The virtual environment consisted of an island on a lake with a wooden house on one side of it and a market stand on its other side, and two other small islands with a market stand on each of them. One could move around on the different islands using a controller in the right hand and place food items into a basket or for removing them from this basket using a controller in the left hand. At each market stand one food item with a high carbon footprint and one with low carbon footprint were available. When placing a high carbon footprint food item into the basket, the sky becomes immediately cloudy, it starts to rain heavily, and the level of the lake rises such that everything is flooded. When replacing the high carbon footprint item with a low carbon footprint item, these climate-related environmental changes are immediately reversed. These were the two food items at the market stand on the three islands: Minced beef and potato (first island), cheese and tomato (second island), chocolate and apple (third island). The virtual environment was presented either with the desktop display of the computer or with a stereoscopic motion-tracked Reverb G2 head-mounted display (HMD) from Hewlett-Packard. Wearing this HMD participants were able to move around and place food items into the basket with the two controllers belonging to the HMD. Being presented with the virtual environment on the desktop screen they could do this using two Joy-Con controllers from Nintendo.

2.3 Problem-based learning task

On each of the three islands the following problem solving task (van Merriënboer, 2013) had to be accomplished. First, the high carbon footprint food item had to be placed into the basket. This triggered the rise of the sea level, which prevented that one could get to any of the other island. Second, this problem had to be solved by replacing the high carbon footprint food item with the low carbon footprint item, which caused the sea level to drop. The instruction of this task was as follows: Please imagine that your grandparents are in the wooden house and want to cook there. They want to cook a stew and prepare a dessert. However, they do not have everything they need for this. Therefore, they send you to get the following three products on the islands in front of you: minced beef and cheese for the stew and chocolate for dessert. Please get these products on the small islands in front of you and come back here. Under no circumstances should you return with an empty basket. The task is only completed when you have placed three food items on the table in front of the wooden house.

2.4 Experimental design

The study was pre-registered (https://osf.io/2hv5j). We used a within-subjects crossover design with two experimental conditions. Hence, all participants were exposed to both experimental conditions. The order of the conditions was determined by the crossover design. In one condition, the 3D virtual environment spatially included its observer, as it was viewed within the head-mounted display. This was the EcoVR condition. In the other condition, the 3D virtual environment spatially excluded its observer, as it was viewed on the desktop display. We named it the EcoPC condition.

2.5 Procedure

The procedure was the same in both experimental conditions: First, the use of the virtual environment was explained and briefly practiced. Subsequently, the teachers were informed about the learning task described above and asked to accomplish the task at the various market stalls using the controllers. This experience phase was followed by an evaluation phase. In this phase, the teachers were asked to assess EcoVR and EcoPC using the three psychometric questionnaire scales described in the next section.

2.6 Measures

The teachers were asked to assess EcoVR and EcoPC using three psychometric questionnaire scales (response format: 100 mm visual analog scale, minimum: not at all; maximum: very much). The first two of these scales corresponded to those of the technology acceptance model (TAM) from Davis (1989) and were each assessed using 6 items. One of these two scales was used to assess the perceived usefulness of EcoVR and EcoPC for teaching climate-friendly food consumption. The other scale was used to assess the ease of use of EcoVR and EcoPC. The third scale was used to evaluate spatial presence, i.e., the feeling of being present in the virtual environment. This scale was assessed using 8 items taken from the MEC Spatial Presence Questionnaire (Rössler, 2011).

2.7 Data analysis

The participants' individual scores on the TAM scales named perceived usefulness and ease of use, and on the spatial presence scale, were determined by calculating their mean rating of the questionnaire items used to assess each of these psychometric scales. Subsequently, these scale scores were compared across the two experimental conditions by one-way repeated-measures analyses of variance (ANOVAs) and by calculating the effect size η_p^2. As for descriptive statistics, we calculated the median (*Md*) and interquartile range [IQR] for all ratings. The statistical analyses were performed with the statistical software SPSS. The visualization of the statistical results was generated with the ggplot2 package within the statistical software R.

3 Results

The statistical analysis of the teachers' ratings in the two experimental conditions yielded the following results: As can be seen in Figure 1 the teachers rated the usefulness of EcoVR for teaching, *Md* = 69.5, IQR [55.3, 81.2], significantly higher, F (1, 9) = 6.78, p = 0.029, $\eta_p^2 = 0.430$, than the usefulness of the PC version, *Md* = 54.0, IQR [27.6, 71.9]. They also rated the ease of use of EcoVR, *Md* = 73.5, IQR [61.3, 91.3], significantly higher, F (1, 9) = 9.24, p = 0.014, $\eta_p^2 = 0.507$, than the ease of use of the PC version, *Md* = 47.3, IQR [17.8, 61.6]. In addition, the teachers' feeling of being present in the virtual environment, i.e., their sense of spatial presence (Figure 1) in EcoVR, *Md* = 83.1, IQR [70.0, 95.3], was significantly stronger, F (1, 9) = 61.68, $p < 0.001$, $\eta_p^2 = 0.873$, than in the PC version, *Md* = 29.1, IQR [12.0, 34.9].

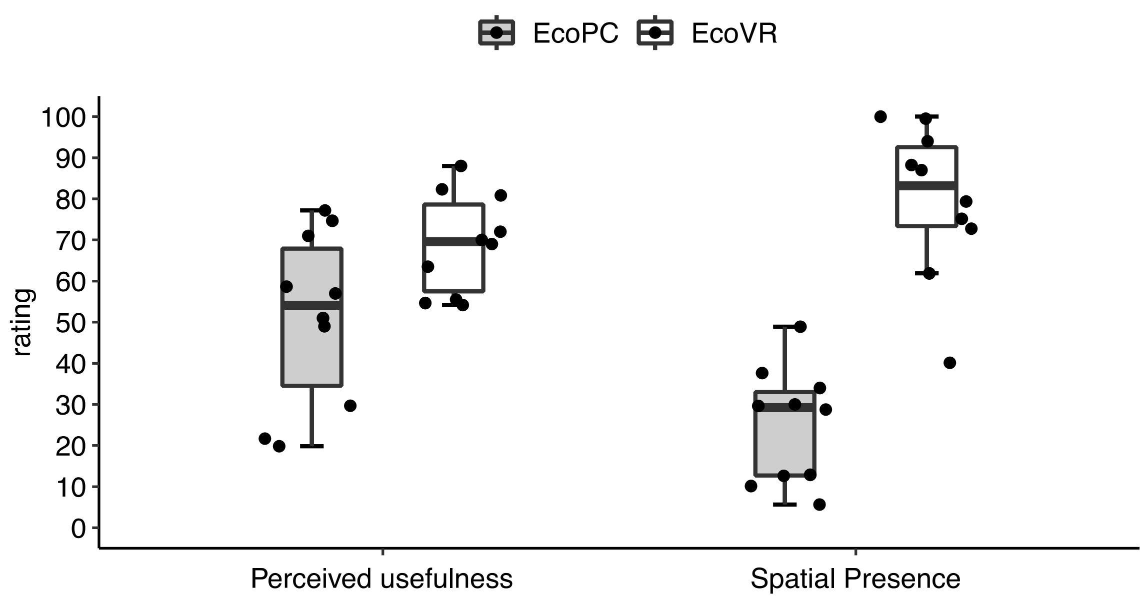


Figure 1. Comparison of EcoVR and EcoPC. Box-whisker plots of the ratings of EcoVR and the PC version in terms of their perceived usefulness for teaching and in terms of sensed spatial presence. Bold horizontal lines show the median of the ratings; boxes show the lower and upper quartiles; whiskers show the furthest data points within 1.5 times the distance to the lower and upper quartiles. Dots depict the individual ratings of the 10 teachers.

4 Discussion

The ratings of EcoVR by experienced teachers working in lower secondary education suggest that they regard EcoVR to be useful for teaching climate-friendly food consumption. Moreover, their ratings suggest that they regard the educational usefulness of EcoVR to be higher than that of its PC version.

EcoVR is programmed in a way that climate-related changes of the ecosystem such as heavy rainfall, immediately arise or disappear depending on the choice of food and its carbon footprint (Poore and Nemecek, 2018). Hence, the core feature of EcoVR is to enable learners to directly experience their food choices and such climate-related environmental changes relative to each other. Experiencing this human-environment relation (Davis et al., 2009) may serve to learn the information processing skills required for engaging into climate-friendly food consumption.

There are two important questions regarding the educational utility of EcoVR that soon need to be addressed in future studies. First, it must be investigated if EcoVR really can serve to foster the learning of eco-friendly food consumption. Second, it has to be investigated if pupils are able to transfer their learning in EcoVR to their everyday life. Most likely, this transfer may be facilitated by a learning task that has to be accomplished by pupils at home and then reported and reflected in school. Hence, when using EcoVR in the school context, the learning in EcoVR may be understood as part of a process that takes pupils from school to home and back to school again. In such a school-life cycle, EcoVR may serve as a digital bridge (Schwendimann et al., 2015) or «metaverse» between the learning of relevant skills at school and the application of these skills in real life.

5 Funding

This study was supported by BeLEARN (www.belearn.swiss).

6 Acknowledgments

The authors would like to thank to Jasin Sahraoui for helping with the design, development and programming of EcoVR and EcoPC.

7 References

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition* 2**,** 1-38.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13**,** 319-340.

Davis, J.L., Green, J.D., and Reed, A. (2009). Interdependence with the environment: Commitment, interconnectedness, and environmental behavior. *Journal of Environmental Psychology* 29**,** 173-180.

Dobricki, M., Kim, K.G., Coppi, A.E., Dillenbourg, P., and Cattaneo, A. (2021). Perceived educational usefulness of a virtual-reality work situation depends on the spatial human-environment relation. *Research in Learning Technology* 29**,** 2453.

Kaufmann-Hayoz, R. (2006). Human action in context: A model framework for interdisciplinary studies in view of sustainable development. *Umweltpsychologie* 10**,** 154–177.

Keller, T., and Brucker-Kley, E. (2021). "Experiences with virtual reality at secondary schools : is there an impact on learning success?," in *Balancing the tension between digital technologies and learning sciences,* eds. D. Ifenthaler, D.G. Sampson & P. Isaías. (Cham: Springer), 77-94.

Markowitz, D.M., and Bailenson, J.N. (2021). Virtual reality and the psychology of climate change. *Current Opinion in Psychology* 42**,** 60-65.

Millennium-Ecosystem-Assessment (2005). *Ecosystems and human well-being: synthesis.* Washington, DC: Island Press.

Poore, J., and Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science* 360**,** 987-992.

Rössler, P. (2011). *Skalenhandbuch Kommunikationswissenschaft.* Wiesbaden: Springer VS.

Schwendimann, B.A., Cattaneo, A., Zufferey Dehler, J., Gurtner, J.-L., Bétrancourt, M., and Dillenbourg, P. (2015). The 'Erfahrraum': a pedagogical model for designing educational technologies in dual vocational systems. *Journal of Vocational Education & Training* 67**,** 367-396.

Van Merriënboer, J.J.G. (2013). Perspectives on problem solving and instruction. *Computers & Education* 64**,** 153-160.

Yue, X.-L., and Gao, Q.-X. (2018). Contributions of natural systems and human activity to greenhouse gas emissions. *Advances in Climate Change Research* 9**,** 243-252.

8 Data Availability Statement

The datasets for this study can be found here: https://osf.io/6tsze/

Advances and Challenges in the Use of Virtual Reality for Surgical Planning in Cardiology

Helena Ortiz-Gil[1 2], Fernando Canet[2], Xavier Ruyra[3]

[1] Techer Team. Spain.

[2] Universitat Politècnica de València. Spain.

[3] Cardiac Surgery Department. Quironsalud Teknon Heart Institute. Barcelona. Spain

*** Correspondence:**
Helena Ortiz-Gil
hortiz@techerteam.com - Techer Team, SL. Avda. Marqués de Sotelo, 9 – 46002 Valencia (SPAIN)

Keywords: virtual reality (VR); immersive virtual reality (IVR); cardiac surgical planning; cardiology; presurgical planning; surgical simulation; metaverse;

Abstract
Immersive virtual reality (IVR) is emerging as a transformative set of technologies in the medical field, offering novel methods for interacting with and analyzing medical data. Notably, IVR holds significant promise as a tool for surgical planning, particularly in cardiac surgery. This study explores the specific advantages, applications, and existing challenges associated with IVR in cardiac surgical planning. We also discuss advancements that pave the way for future improvements in this technology. By examining relevant available tools and academic publications, we aim to provide an overview of the current usefulness and vast potential of IVR in this critical surgical domain.

Introduction
Immersive virtual reality (IVR) represents a collection of technologies that enables, in real-time, interactive exploration within realistic, three-dimensional (3D) virtual environments (Abbas, 2023). This innovative approach fundamentally transforms how clinicians engage with medical data and images. By facilitating the creation of detailed, 3D representations of patient anatomy, IVR fosters a deeper understanding of human structures and empowers surgeons in pre-operative planning (Mathab, 2022).
This novel perspective provides surgeons with comprehensively preparation, identify potential surgical complexities before entering the operating room, and ultimately optimize both procedural accuracy and patient safety.

Objetives
This study conducts a narrative review of the current academic literature and commercially available solutions relevant to the cardiology field. Our objective is to identify key findings, existing solutions, and ongoing challenges within this area. This analysis will inform future research directions and areas requiring further development.

Methododology
This study is based on a narrative review methodology and a evaluation of surgical planning tools using VR technology. We conducted a search of the academic literature, focusing on articles that explore VR applications in cardiac surgical planning. Additionally, we reviewed commercially available VR surgical planning tools in the current market landscape. This combined approach allowed us to identify key trends, potential benefits, and existing limitations within this field.

Analysis of current Virtual Reality Applications in Cardiology

This review has analyzed current VR tools specifically designed for surgical planning within the cardiology field, based on two systematic approaches published in academic literature (Lan, 2023; Deng, 2021) plus other recently emerged commercial offerings. Our analysis focused on key aspects of these VR tools, including usability, user interaction design, the degree of immersion provided, and the underlying hardware requirements. Information was gathered through a search of relevant comercial webpages conducted on March 29, 2024: https://echopixeltech.com/, https://www.medicalholodeck.com/en/, https://www.mevislab.de/, http://imhotep-medical.org/; https://www.medicalvr.eu/; https://holoeyes.jp/en/; https://surgicaltheater.com/; https://www.visuamed.com/

Our review process employed strict inclusion criteria. We excluded tools that are no longer commercially available or lack functionalities specifically designed for cardiac surgery planning. The analysis of the included solutions and their accompanying marketing materials revealed some areas of differentiation. These primarily encompass the level of interactivity between the tool and patient digital twins, the degree of immersion offered, and hardware requirements. Despite these variations, all reviewed applications highlight the advantages of VR technology in this domain, which will be discussed in the following section.

Analysis of Applications and Advantages of Virtual Reality

Following the analysis of commercially available solutions and relevant academic literature, this review highlights the significant advantages offered by IVR and VR in cardiac surgical planning. Firstly, VR facilitates the creation of detailed 3D visualizations of the patient's cardiac anatomy, encompassing both vascular structures and any present anatomical anomalies (Silva, 2018). This enhanced visualization empowers surgeons to accurately identify lesions and improve plan surgical strategies.

Additionally, VR enables the simulation of complex cardiac procedures, including coronary revascularization surgery and valve repair surgeries (Bouraghi et al., 2023). This simulation capability fosters improved procedural safety and efficacy, ultimately leading to a reduction in perioperative complications and enhanced patient outcomes (Romero et al., 2023). The benefits of IVR are pronounced in intricate interventions, such as pacemaker lead extraction (Carretero, 2024). Beyond enhancing surgical accuracy and safety, the integration of VR in surgical planning offers valuable opportunities for medical education and surgeon training. By enabling the simulation of complex procedures within a virtual environment, VR provides medical students with a risk-free, hands-on learning experience to develop and refine surgical skillset (Arjomandi, 2023).

Future Directions: Enhancing Graphical representation and Interoperability are key

The analysis revealed key areas requiring further development in VR-based cardiac surgical planning. One critical challenge lies in improving the graphical representation of cardiac dynamics, specifically blood flow and contractility. Potential solutions include advancements in computational fluid dynamics (CFD) simulations (Venn, 2023), including the potential integration of holographic display techniques (D'Aiello et al., 2023). These advancements would provide surgeons with a more realistic and dynamic virtual environment for surgical planning.

Another significant challenge is related to interoperability between various digital health tools. Seamless data exchange between virtual platforms holds immense potential for enhancing medical collaboration, education, and research (Paul et al., 2023). However, achieving this interoperability presents a complex hurdle that requires active collaboration between the scientific community, healthcare professionals, and technology developers. Standardization initiatives are crucial to foster the adoption of common protocols and facilitate the integration of currently disparate systems.

Conclusions: This promising technology needs to be advanced.

Immersive virtual reality (IVR) has emerged as a promising technology with demonstrably positive applications in cardiac surgical planning. The ability to generate detailed 3D visualizations of patient anatomy and simulate complex procedures offers substantial benefits for surgeons (e.g., improved preoperative planning, enhanced understanding of anatomical complexities). However, to fully realize the transformative potential of IVR in cardiac surgery, several key challenges require focused attention.

Furthermore, conducting well-designed clinical trials to evaluate the long-term impact of IVR on surgical outcomes and patient care is warranted. Continued research efforts exploring the diverse applications of IVR within cardiac surgery are highly desirable.

References:

Abbas J.R., O'Connor, A., Ganapathy E., Isba R., Payton, A., McGrath, B., Tolley, N., Bruce, I.A. (2023) *What is Virtual Reality? A healthcare-focused systematic review of definitions.* Health Policy and Technology. Volume 12, Issue 2. https://doi.org/10.1016/j.hlpt.2023.100741

Arjomandi,R.A., Hajzamani, D., Sardari N.P. (2023). *Simulation-based training in cardiac surgery: a systematic review. Interdisciplinary CardioVascular and Thoracic Surgery, Volume 37, Issue 2, August 2023, ivad079*; https://doi.org/10.1093/icvts/ivad079

Bakhuis, W., Max, S., Maat, A., Bogers, A., Mahtab, E., Sadeghi, Amir H. (2023). *Preparing for the future of cardiothoracic surgery with virtual reality simulation and surgical planning: a narrative review*. Shanghai Chest; Vol 7 (July, 2023) https://shc.amegroups.org/article/view/8115/html

Bouraghi, H., Mohammadpour, A., Khodaveisi, T., Ghazisaeedi, M., Saeedi, S., & Familgarosian, S. (2023). *Virtual Reality and Cardiac Diseases: A Systematic Review of Applications and Effects.* Journal of healthcare engineering. 2023, 8171057. https://doi.org/10.1155/2023/8171057

Carretero, E.G., de Alarcón González, A., de la Borbolla, M.G. et al. (2024). *Infections of cardiac implantable electronic devices and how to improve transvenous lead extraction by the virtual reality*. Indian J Thorac Cardiovasc Surg https://doi.org/10.1007/s12055-023-01663-9

D'Aiello, A.F., Bognoni, L., Bevilaqua, F. Negura, D.G., Ferrero, P. Micheletti, A., Giamberti, A., Militaru, C., Chessa, M. (2023). *Holographic Techniques as a Novel Method for Intervention Planning: A Tertiary Centre's Experience*. Current Health Science; vol 49 (4-2023) DOI:10.12865/CHSJ.49.04.15 https://www.chsjournal.org/article/49/4/15/

Deng, S., Wheeler, G., Toussaint, N., Munroe, L., Bhattacharya, S., Sajith, G., Lin, E., Singh, E., Chu, K.Y.K., Kabir, S., et al. (2021) *A Virtual Reality System for Improved Image-Based Planning of Complex Cardiac Procedures*. J. Imaging 2021, 7, 151. https://doi.org/10.3390/jimaging7080151

Lan L, Mao RQ, Qiu RY, Kay J, de Sa D. (2023) Immersive Virtual Reality for Patient-Specific Preoperative Planning: A Systematic Review. *Surgical Innovation*. 2023;30(1):109-122. https://doi.org/10.1177/15533506221143235

Mahtab, E. A. F., & Egorova, A. D. (2022). *Current and future applications of virtual reality technology for cardiac interventions.* Nature reviews. Cardiology, 19(12), 779–780. https://doi.org/10.1038/s41569-022-00789-4

Paul, P.K., Chatterjee, R., Chakraborty, S., Tiwary, K.S., Aithal, P.S., Sharma, S. (2023) *Advanced healthcare informatics practice. Vis-à-vis Virtual Reality Implications - A scientific overview.* 1 ed. s.l.:CRC Press. ISBN9781003340133

Romero, D. P., Latorre-Rojas, C. J., Latorre Quintana, M., Velasco Morales, M. L., Pardo Nino, L. S., & Arango, M. L. (2023). *Use of Virtual Reality and Three-Dimensional Printing in the Surgical Planning of Slide Tracheoplasty*. World Journal for Pediatric and Congenital Heart Surgery, 14(4), 503-508. https://doi.org/10.1177/21501351231163532

Silva, J, Southworth, M, Raptis, C. et al. (2018) *Emerging Applications of Virtual Reality in Cardiovascular Medicine.* J Am Coll Cardiol Basic Trans Science. 2018 Jun, 3 (3) 420–430.https://doi.org/10.1016/j.jacbts.2017.11.009

Venn, J., Larkee,C.E, Garcia, G. J. M., Rayz,V.L., LaDisa, J.F., (2023) *A workflow for viewing biomedical computational fluid dynamics results and corresponding data within virtual and augmented reality environments* .Frontiers in Medical Technology. February 2023. https://doi.org/10.3389/fmedt.2023.1096289

Towards an Intelligent Metaverse: The use and impact of AI

Meeting Intelligent Agents: How Virtually Anything is Possible with an Intelligent Virtual Agent at the Table

Sebastian Rings [1,*], **Lucie Kruse** [1], **Marvin Grabowski** [2], **and Frank Steinicke** [1]

[1] *Human-Computer Interaction, Universität Hamburg, Germany*
[2] *Industrial and Organizational Psychology, Universität Hamburg, Germany*

Correspondence*:
Sebastian Rings
sebastian.rings@uni-hamburg.de

ABSTRACT

Remote participation in online meetings has become an integral part of the modern work environment, enabling collaboration across distances and through intermediate restrictions. Virtual meetings, using virtual (VR) or augmented reality (AR) technology, are especially promising and part of major current technical developments such as Apple's Vision Pro[1] and Meta's Quest 3[2]. However, there is a growing realization that the potential of immersive meetings has yet to be fully unlocked. Augmenting meetings by introducing intelligent virtual agents (IVA) at the table as virtual participants could significantly enhance the efficacy of such meetings. This paper explores integrating IVAs into virtual meetings, possible roles, benefits, and the ethical considerations that accompany their usage. Moreover, we delve into the potential future implications of such integrations on workplace dynamics and meeting productivity.

Keywords: metaverse, virtual reality, meetings, intelligent virtual agents, ethics in IT

1 INTRODUCTION

The shift in interaction modalities persists beyond the pandemic, as numerous organizations have embraced the potential of virtual or hybrid collaboration. The growing demand for flexibility leads to a permanent transformation in organizational interaction dynamics, notably in meetings. They fulfill various essential functions, including problem-solving, idea exchange, consensus-building, and collective decision-making (Allen and Lehmann-Willenbrock (2023)).

The adoption of remote participation technologies has reshaped the landscape of professional meetings, emerging as a critical component of such. Moreover, there now appears to be a shift towards extended reality (XR) technologies, accelerated by advancements in VR, AR, and mixed reality (MR), requiring research to focus on this paradigm. XR technologies present a promising solution to address the limitations of virtual meetings and to reinvent collaboration practices. The emergence of the 'metaverse' introduces novel opportunities to leverage previously unimaginable technological capabilities within a new interaction space (Hennig-Thurau et al. (2023)). In XR meetings, participants converge within a shared virtual space

[1] https://www.apple.com/apple-vision-pro/
[2] https://www.meta.com/de/quest/quest-3/

facilitated by VR technology and head-mounted displays. This illusion of place and plausibility is realized through the utilization of 3D embodied avatars (Kim et al. (2023)). With high levels of immersion alone, these types of meetings can improve through better collaboration, yet there remains significant research potential when it comes to exploring the integration of IA and IVAs. This paper focuses on the potential, opportunities, and concrete roles of IVAs, as well as ethical considerations, which can inform and drive future work.

2 ROLES AND ADVANTAGES

IVAs are often presented as humanoid and somehow anthropomorphic virtual characters, with variations regarding realism or the depiction of different body parts (Weidner et al. (2023)), but often identical with human participants in VR meetings. Once added to a virtual meeting, an IVA can support the meeting dynamics in a myriad of ways. While IVAs can take both active and passive roles in a meeting, their mere presence can already affect the participants, e.g. by invoking social facilitation effects (Park and Catrambone (2007)).

Passive Role: In their passive role, they can serve as silent note-takers or summarizers, often an obligation in productive meetings, and create protocols, action items, and lists of attendance. Further, they can analyze participant engagement where applicable, and make suggestions that can support the current as well as future meetings, e.g., by identifying less productive points on the agenda in post. Overall, letting an IVA take over such roles can free other participants from these tasks, and let them focus on the discussion and engaging in the meeting.

Active Role: In their active role, IVAs can directly take part in the discussion and actively influence decisive factors to tune the efficacy of the current meeting. Agents can act as idea generators both by speaking out and creating engaging visual aids on the fly, which suit the current discussion (Fuchs et al. (2023)). Moreover, they can act as technical support, adjusting the virtual environment with regard to background noise, lighting, furniture, and retrieving relevant information from databases or the internet. They can actively identify issues and act to find suitable fixes by notifying participants, making adjustments, thus fixing microphone issues, reverberations, and positioning. Furthermore, by offering their entity as an interface for communication, they can improve the accessibility of a meeting by explaining complicated terms, acting as a translator for foreign languages, or sign language (Atasoy et al. (2023)).

3 ETHICAL CONSIDERATIONS

While the integration of AI into virtual meetings presents significant advantages, it also raises several ethical considerations. Especially in roles that involve documentation, real-time analysis, and personalized support through an IVA, significant privacy concerns are raised. The collection, analysis, and storage of potentially sensitive discussions necessitates robust data protection measures. Questions regarding who has access to this data, how long it is stored, and the measures in place to protect it from unauthorized access are paramount. Furthermore, ensuring that participants are fully aware of and consent to the data being collected and analyzed by IVAs is crucial for maintaining trust and ethical integrity (Rudschies and Schneider (2023)). Linked to Large Language Models (LLM), IVAs are also prone to biases (Bender et al. (2021)) which can manifest in IVAs by disproportionately valuing contributions from certain participants over others or misinterpreting information altogether. It is essential to ensure fairness throughout the meeting.

There are also notable ethical advantage worth highlighting: the transparency and naturalization of meeting recordings, e.g. by having the agent visually writing down notes. Traditionally, the act of recording digital meetings, e.g. for documentation purposes, has been a point of contention, raising concerns over privacy and misuse of recorded content. However, the presence and actions of IVAs can offer a more transparent and intuitive approach to this necessity.

4 IMPLICATIONS AND FUTURE WORK

Looking ahead, AI's role in virtual meetings is poised to grow, with agents as active participants or observers potentially becoming the norm. Future research should focus on integrating advanced AI developments, like LLMs, to further explore how IVAs can dynamically interact with human participants, thereby augmenting the meeting experience and environment. Considering these advancements, it is important to keep in mind the challenges and requirements, such as data safety, privacy and unbiased information processing. Initial findings indicate that XR meetings can enrich virtual 2D meetings in specific contexts (Steinicke et al. (2020)), however, we know little about how AI can provide support here. IVAs could have a positive influence on the dynamics of the meeting in this new interaction space, thus more research is required.

Funding: Funded by the BMBF within the scope of the HIVAM project, Grant Number: 16SV8878.

REFERENCES

Allen, J. A. and Lehmann-Willenbrock, N. (2023). The key features of workplace meetings: Conceptualizing the why, how, and what of meetings at work. *Organizational Psychology Review* 13, 355–378

Atasoy, M., Şilbir, L., Erümit, S. F. İ., Bahçekapili, E., Yildiz, A., and Karal, H. (2023). Visual appearance features of sign language avatars. *Kastamonu Eğitim Dergisi* 31, 386–403

Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. 610–623

Fuchs, A., Appel, S., and Grimm, P. (2023). Immersive spaces for creativity: Smart working environments. In *2023 International Electronics Symposium (IES)*. 610–617. doi:10.1109/IES59143.2023.10242458

Hennig-Thurau, T., Aliman, D. N., Herting, A. M., Cziehso, G. P., Linder, M., and Kübler, R. V. (2023). Social interactions in the metaverse: Framework, initial evidence, and research roadmap. *Journal of the Academy of Marketing Science* 51, 889–913

Kim, D. Y., Lee, H. K., and Chung, K. (2023). Avatar-mediated experience in the metaverse: The impact of avatar realism on user-avatar relationship. *Journal of Retailing and Consumer Services* 73, 103382

Park, S. and Catrambone, R. (2007). Social facilitation effects of virtual humans. *Human factors* 49, 1054–1060

Rudschies, C. and Schneider, I. (2023). Ethical, legal, and social implications (elsi) of virtual agents and virtual reality in healthcare. *Social Science & Medicine* , 116483

Steinicke, F., Lehmann-Willenbrock, N., and Meinecke, A. L. (2020). A first pilot study to compare virtual group meetings using video conferences and (immersive) virtual reality. In *Proceedings of the 2020 ACM Symposium on Spatial User Interaction*. 1–2

Weidner, F., Boettcher, G., Arboleda, S. A., Diao, C., Sinani, L., Kunert, C., et al. (2023). A systematic review on the visualization of avatars and agents in ar & vr displayed using head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*

AI for Movement Interaction in VR

Marco Gillies [1,*]

[1] *SeeVR Lab, Department of Computing, Goldsmiths, University of London, London, UK*

Correspondence*:
Corresponding Author
m.gillies@gold.ac.uk

1 KEYWORDS:

virutal reality, artificial intelligence, machine learning, movement interaction design, embodied interaction

Movement is key to virtual reality (VR), perhaps more so than displays such as graphics and audio. Slater's theory of Place Illusion (Slater, 2009) states that the feeling of place illusion is triggered by VR reproducing real-world sensorimotor contingencies (O'Regan and Noë, 2001). Place Illusion is *""the strong illusion of being in a place in spite of the sure knowledge that you are not there"* (Slater, 2009), is a key component of the sense of presence, which is itself core to the experience that VR creates. Sensorimotor contingencies are the relationships between our movements and our sensations. For example, as we turn our head our view of the world changes, giving us a different perspective. O'Regan and No e's Sensorimotor Theory (O'Regan and Noë, 2001) proposes that these sensorimotor contingencies lie at the core of our perception and experience. Slater, in turn, proposes that VR is so effective because it reproduces the same sensorimotor contingencies the we have learned from our experience of the real world. Turning our head in a head-tracked head-mouted display (HMD) will result in our view of the virtual world changing in exactly the same way that the real world would change if we turned our head without an HMD. This is in contrast to screen based displays, where we would have to perform button presses or mouse movements to turn our view. The fact that VR responds to our movements in the same way that the real world does makes us feel that we are really in the virtual world, not simply viewing it.

This theory places movement at the centre of the VR experience. It means that movement and displays are inseparable. It also suggests that movements in VR should, as far as possible, be similar to those we perform in the real world, so that our sensorimotor contingencies are, in turn, similar to those in the real world. This is supported by the theory of Plausibility Illusion, Slater's second component of presence. A key element of plausibility is that events should respond to you in ways that are credible (in many cases this means in ways similar to the real world. Plausibility is therefore a more cognitive analog of place illusion. Whereas place illusion means that low level sensory stimulus should respond realistically to our low-level movements, plausibility means that complex actions should respond realistically to our complex behaviours. Both concepts place our movements and actions at the centre of the experience and indicate that these actions and movements should be realistic so as to enable realistic interaction.

2 DESIGNING MOVEMENT INTERACTION

What does this mean for the design of VR experiences? We must have a much greater focus on movement in VR design than in other digital media. This is particularly true as movement has largely been ignored in mainstream digital products. It is reduced to button presses and finger touches. The greater focus on movement in turn requires a fundamental shift in design processes. Current processes often are limited to purely 2D design unsuitable to VR (e.g. paper prototyping)and even those in 3D focus on the design of objects and environment, rather than the movements of the people in those environments. However, there has been a recent interest in design approaches that focus on the body and its movements, for example Soma Design (Höök, 2018) or Embodied Sketching (Márquez Segura et al., 2013). These techniques allow designers to focus on their bodies in the design process and therefore focus on the experience of moving.

These body-focused design approaches has great potential for raising the importance of movement in the design of VR experiences. However, they are largely limited to the ideation phase of design. The prototyping and implementation of interactions is very reliant on coding. This can be extremely challenging for movement interaction, because our movement knowledge is embodied and implicit (Gillies, 2019). We know how to move and act in the world, but often cannot explain how we do so in words. A classic example is riding a bicycle: once we have learned to ride, we can effortlessly make the subtle movements required to balance a steer a bike, but we cannot teach this knowledge to some one else simply through a verbal explanation. Simple movement interactions can be implemented using standard game engine techniques such as colliders, however, these fail to scale to the kind of complex movements that characterise our richest interactions with the world, including dance, sports and body language.

In this paper I propose that Machine Learning, often now referred to as Artificial Intelligence (AI) offers an embodied way of prototyping and implementing movement interactions, because it makes it possible to prototype interactions by providing examples of movement rather than formulating mathematical rules in code. This prototyping by moving makes it possible to use embodied movement knowledge throughout the design process. This use of machine learning is potentially one of the most important uses of AI for VR, but it can also fundamentally change our view of AI technology, away from being an autonomous agents, to being a create tool and process. The rest of this paper will briefly outline two case studies, by my collaborators, of using machine learning for VR interaction design.

3 INTERACTML

InteractML (Hilton et al., 2021) is a machine learning platform for movement interaction design in VR, aimed at artists and dancers. Designers can prototype movement interactions by providing examples of movement. It provides a graph-based interface for setting up machine learning models (e.g. choosing a learning algorithm and data features) allowing a workflow without programming. The developers ran a series of workshops with creative practitioners, culminating with 3 longer residencies, which ran over 6 week and in which participants developed a project with the tool. The tool was co-designed with an ideation and design process. Participants were guided through an ideation process in which they were encouraged to design by moving and focus on their movements while designing. This was continued into the prototyping phase by having participants perform movements that were used as training data for a machine learning algorithm. Participants readily adopted the design by moving process, and commented that the use of machine learning enabled a more embodied design approach that allowed for more creativity in movement design. It was particularly attractive for participants with dance background or other movement practice, as it allowed them to use their embodied knowledge within the design of technology.

4 IMMERSIVE MACHINE LEARNING FOR GAME CHARACTERS

Social interaction is one of the most complex and important of human behaviours. Body movement, in the form of “body langauge” is a core part of how we interact with other people, and one of the most complex and subtle uses of the human body. Body language interaction with virtual humans has the potential to be one of the most compelling and powerful forms of interaction in VR (e.g. the work of Pan et al. (2016) on social skills training in medicine). However, the complexity of body language makes it particularly challenging no to design and a good example for design by movement. Dobre et al. (2022) used machine learning to recognise game player's level of engagement with a virtual character, enabling a form of gameplay that relies on social interaction rather than typical violence or puzzle solving. Though the data came from game players rather than designers, they developed a data labeling process aimed at the needs of creative designers.

5 CONCLUSION

Machine learning and AI hold the possibility of enabling compelling, movement-based forms of interaction that give a strong sense of presence. However, using it this ways fundamentally changes how we view AI. It is a creative design tool to be used by movement experts, rather than a technical skill that is the sole preserve of data scientists. This will change how we design machine learning tools and how they are integrated into VR systems, making them more creatively focus, but, I hope also more ethical and focused on the needs of people.

To end this paper, I would like to express the hope that the use of design by moving can help VR to reverse the trend of technology to disembody us. We now spend most of our time in front of screens, physically inactive and socially isolated. Let's make the metaverse an embodied medium that is health for our bodies, minds, relationships and socieity.

REFERENCES

Dobre, G. C., Gillies, M., and Pan, X. (2022). Immersive machine learning for social attitude detection in virtual reality narrative games. *Virtual Reality* 26. doi:10.1007/s10055-022-00644-4

Gillies, M. (2019). Understanding the Role of Interactive Machine Learning in Movement Interaction Design. *ACM Transactions on Computer-Human Interaction* 26, 1–34. doi:10.1145/3287307

Hilton, C., Plant, N., Fiebrink, R., Perry, P., Diaz, C. G., Gibson, R., et al. (2021). InteractML: Making machine learning accessible for creative practitioners working with movement interaction in immersive media. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*. doi:10.1145/3489849.3489879

Höök, K. (2018). *Designing with the Body - Somaesthetic Interaction Design* (MIT Press). doi:10.5753/ihc.2018.4168

Márquez Segura, E., Waern, A., Moen, J., and Johansson, C. (2013). The Design Space of Body Games: Technological, Physical, and Social Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA: ACM), CHI '13, 3365–3374. doi:10.1145/2470654.2466461

O'Regan, J. K. and Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *The Behavioral and brain sciences* 24, 939–973

Pan, X., Slater, M., Beacco, A., Navarro, X., Swapp, D., Hale, J., et al. (2016). The Responses of Medical General Practitioners to Unreasonable Patient Demand for Antibiotics - A study of medical ethics using immersive virtual reality. *PLoS ONE*

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos Trans R Soc Lond B Biol Sci* 364, 3549–3557. doi:10.1098/rstb.2009.0138

Origami Unfolded: Augmented Reality and Machine Learning Integration for Manipulative Training Scenarios

Mikołaj Łysakowski [1,*], Jakub Gapsa [2] and Piotr Skrzypczyński [1]

Poznań University of Technology: [1] Center for Artificial Intelligence and Cybersecurity, [2] Faculty of Mechanical Engineering, Poznań, Poland

Correspondence*:
Mikołaj Łysakowski
mikolaj.lysakowski@put.poznan.pl

ABSTRACT

This work explores the fusion of machine learning and augmented reality (AR) to guide origami folding. By deploying a YOLOv8 model, our AR system offers real-time detection and assessment of fold stages, providing users with instructions and feedback. We present an innovative dataset approach that enhances accuracy in origami model detection. Furthermore, our findings highlight the robust performance of the AR system across various origami models, advancing AR-based learning through automated task assessment with potential applications in training and education.

Keywords: extended reality, machine learning, edge computing, assembly task, education

1 INTRODUCTION

We investigate integrating Augmented Reality (AR) and Machine Learning (ML) to support individuals in independently learning manual skills, especially during prolonged periods of isolation (e.g., the COVID-19 pandemic). AR finds applications in both entertainment and in the training of manual assembly tasks (Zogopoulos et al., 2021). We explore the use case of origami folding, which is similar to assembly tasks in terms of skill requirements, problem-solving approach, and the sense of achievement it offers to participants. Existing AR-based origami tutorials lack the ability to recognise user actions or provide constructive feedback (Wiwatwattana et al., 2016), hindering skill acquisition. To address this limitation, we propose embedding action and outcome assessment directly into AR applications using edge-based machine learning methods. Through neural network technology, we aim to anticipate folding steps from camera feeds, providing users with feedback on correct execution and enhancing their learning experience for independent skill acquisition.

Our system is implemented within the Unity game engine. Upon initiation, users can designate the primary colour of the paper sheet (assuming the opposite side is white) and select one of the predetermined origami models to fold. Subsequently, they encounter an animation illustrating the entire folding process, succeeded by interactive step-by-step instructions. The users are tasked with folding actual paper in accordance with the presented instructions, and upon completing a step, they are prompted to activate a virtual button to validate their attempt. The algorithm forecasts the state of the folded paper based on the completed step of the process, and if the folded model is deemed correct by the employed neural network, the user can proceed. A demo of this application is available on YouTube[1].

[1] https://www.youtube.com/watch?v=RrSPt5YUIg8&ab_channel=CAICS

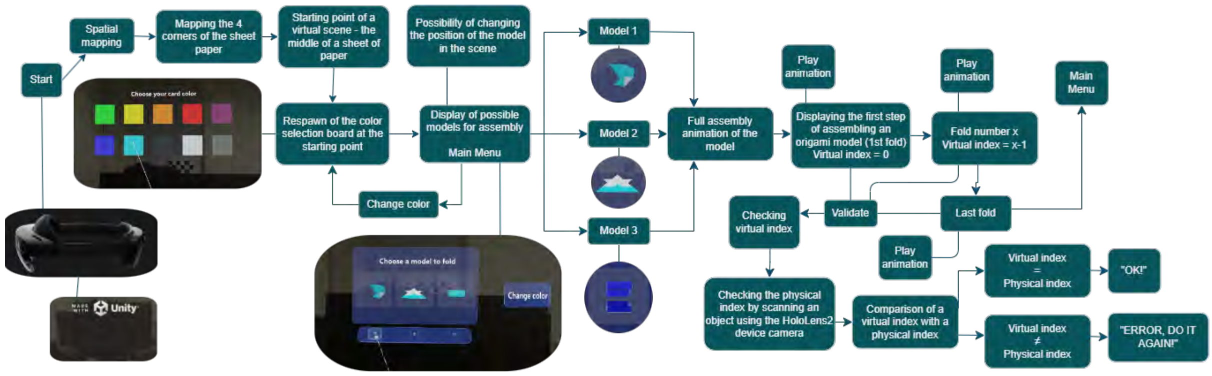


Figure 1. Data and control flow in the origami folding application with ML-based outcome assessment.

2 IMPLEMENTATION AND RESULTS

The origami models were created in the Blender 3D software[2] and then exported to the Unity environment. Each folding step was divided into a separate model, with appropriate events triggering the next stage after the previous one was completed. Drawing upon our prior work (Łysakowski et al., 2023b), we adopted a YOLOv8 model to detect and categorise the stages of the paper folding process. To execute inference on-board the HL2, we utilised Unity Sentis, the successor of the Barracuda neural network inference library previously employed (Łysakowski et al., 2023b). As real-time capabilities are unnecessary, given that prediction is triggered on demand, our solution employs a compact neural network model with images sized at 320×320 pixels, facilitating an inference time of approximately 500 ms, thereby striking an acceptable balance between accuracy and performance for our application.

For the experiments outlined herein, we compiled a dataset featuring a "samurai hat", "bird" and "box" origami models. Folding these intelligible models entails respectively 11, 10 and 20 stages, with each stage culminating in a predefined correct paper shape, representing the outcome of the process. The images of every stage were gathered using an HL2 device with different view angles and a distance to a paper sheet, imitating the validation process during folding origami while sitting at a desk. Following dataset acquisition, we applied the *Track Anything* algorithm (Yang et al., 2023), predicated on the Segment Anything Model, to generate ground truth data, manually labelling solely the first frame in each stage with a bounding box. Subsequently, we augmented the dataset by segmenting the primary paper colour in the HSV (Hue Saturation Value) space and substituting it with alternative hues. Consequently, we obtained around 150 training images and 70 validation images per stage for every origami model. To guarantee a dependable benchmark, we gathered an additional 30 images per folding stage for the test set, featuring previously unseen paper colours. Compared to the training set, these images were taken on various surfaces and under diverse lighting conditions, thus ensuring the assessment of generalisation capabilities. The created dataset encompasses all the folding stages of the model. During the training process, we also applied standard augmentation operations: image rotation, translation, scale, HSV change and mosaic. The test set was not subjected to colour replacement augmentation. Illustrations from the training process and detection outcomes for all three origami models are depicted in Fig. 2, whilst numerical results of AP (Average Precision) and AP_{50} for the Intersection over Union threshold set at 50%, on both validation and test sets, are presented in Tab. 1.

[2] We used resources available on the website https://origami.guide/

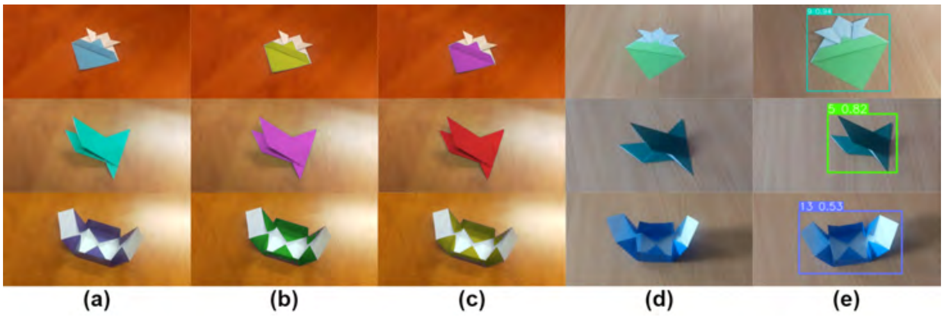


Figure 2. Example images for the samurai hat, bird and box origami models. Images from the train/val set with colour augmentation (a,b,c). Examples from the test set (d) and final detection (e).

Model/Dataset	AP_{50}^{val}	AP^{val}	AP_{50}^{test}	AP^{test}
Samurai Hat	0.995	0.966	0.973	0.934
Bird	0.988	0.968	0.969	0.896
Box	0.995	0.982	0.927	0.86

Table 1. Numerical results of YOLOv8s on the three origami models averaged over all stages of folding.

3 CONCLUSION

Our work illustrates that AR support, which provides immersive and interactive experiences for training in various domains, can be augmented by a machine learning model to automate task evaluation upon its completion. This approach may extend beyond origami folding to manual assembly or maintenance processes and analogous hands-on tasks. While promising, the model may still lack accuracy and robustness, especially when dealing with variations in incorrectly folded patterns. Future efforts will focus on improving the model's resilience to variations in incorrectly assembled patterns, lighting conditions, and backgrounds, thereby enhancing its generalisation capabilities. We also plan to validate our approach in user tests, as we did with the HL2 object detector (Łysakowski et al., 2023a).

ACKNOWLEDGMENT

This work was done at the Poznań University of Technology Centre for Artificial Intelligence and Cybersecurity. The study was supported by funding provided through an unrestricted gift by Meta.

REFERENCES

Wiwatwattana, N., Laphom, C., Aggaitchaya, S., and Chattanon, S. (2016). Origami guru: An augmented reality application to assist paper folding. In *Information Technology: New Generations* (Springer), 1101–1111

Yang, J., Gao, M., Li, Z., Gao, S., Wang, F., and Zheng, F. (2023). Track anything: Segment anything meets videos, arXiv, cs.cv 2304.11968

Zogopoulos, V., Birem, M., De Geest, R., Hofman, R., Jorissen, L., Vanherle, B., et al. (2021). Image-based state tracking in augmented reality supported assembly operations. *Procedia CIRP* 104, 1113–1118

Łysakowski, M., Żywanowski, K., Banaszczyk, A., Nowicki, M. R., Skrzypczyński, P., and Tadeja, S. (2023a). Using AR and YOLOv8-based object detection to support real-world visual search in industrial workshop: Lessons learned from a pilot study. In *IEEE Int. Symp. on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 154–158

Łysakowski, M., Żywanowski, K., Banaszczyk, A., Nowicki, M. R., Skrzypczyński, P., and Tadeja, S. K. (2023b). Real-time onboard object detection for augmented reality: Enhancing head-mounted display with YOLOv8. In *IEEE Int. Conf. on Edge Computing and Communications*. 364–371

Poster Abstracts

A Taxonomy for D-SAV360, a Dataset of Gaze Scanpaths on 360º Ambisonic Videos

Edurne Bernal-Berdun [1,†,*], **Maria Plaza** [1,†], **Daniel Martin** [1], **Sandra Malpica** [1], **Belen Masia** [1], **Ana Serrano** [1], **and Diego Gutierrez** [1]

[1] *Graphics and Imaging Lab, Universidad de Zaragoza - I3A, Zaragoza, Spain*
† Joint first authors.

Correspondence*:
Edurne Bernal-Berdun
edurnebernal@unizar.es

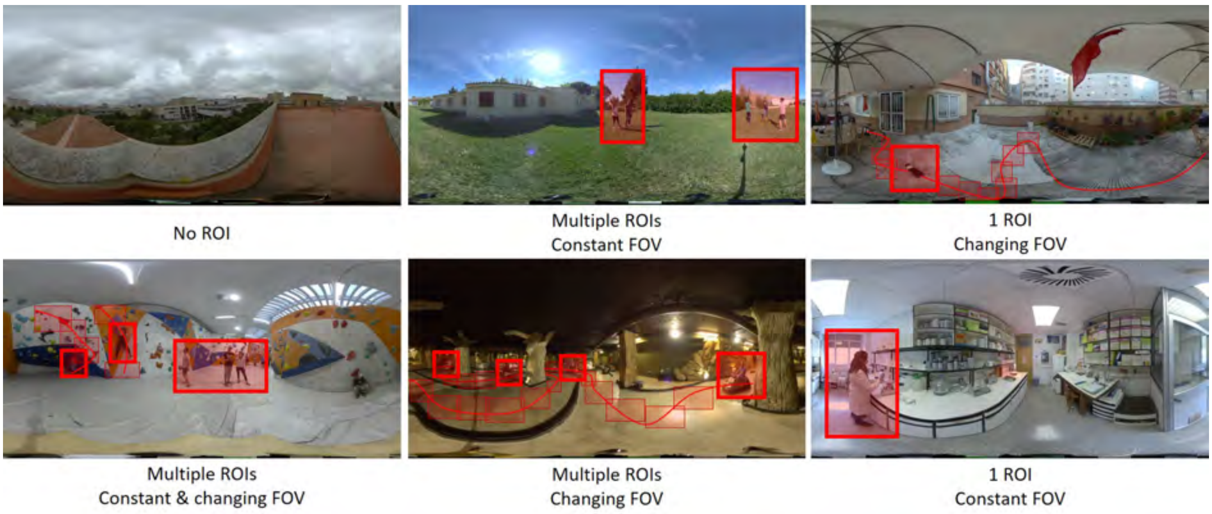


Figure 1. We present a taxonomy for D-SAV360 (1), an extensive dataset of viewing behavior on 360º ambisonic videos. Our taxonomy is a structured framework that encourages a more comprehensive and exhaustive analysis for future research on visual behavior attention in virtual reality. We believe it could be used as a benchmark and valuable resource for training novel audiovisual attention prediction models. Here we show sample frames from six of our captured videos that present six different categories of the proposed taxonomy. We show a representative frame for each video, highlighting in red the regions of interest (ROIs).

ABSTRACT

Understanding how humans visually behave within virtual reality environments is essential to fully leverage its potential. Addressing this, recently, the dataset D-SAV360 (1) was published in IEEE Transactions on Visualization and Computer Graphics (ISMAR 2023). This dataset comprises the most extensive dataset of viewing behavior on 360º ambisonic videos to date, with 4609 head and eye scanpaths for 360º videos with directional ambisonic sound, along with the analysis of the collected scanpaths from a total of 87 participants viewing 85 different videos. In this paper, we present a novel taxonomy for this dataset. Our taxonomy provides a structured framework that encourages a more comprehensive analysis for future research on visual behavior attention in virtual reality. By categorizing the data in a specific manner, our taxonomy allows to

discern conditions that may affect eye movement statistics. Furthermore, by making our taxonomy, methodology, and data collection system publicly available, we hope to facilitate the development, evaluation, and support of future research in this exciting field.

Keywords: Audio-visual Saliency, Virtual Reality, 360º Videos, Taxonomy

1 INTRODUCTION

Creating engaging experiences that exploit the potential of virtual reality (VR) becomes increasingly important as VR techniques and applications continue to blossom. Consequently, understanding and being able to systematically predict human visual behavior plays a crucial role in achieving this. For instance, a detailed understanding of visual behavior in VR can enhance storytelling by enabling creators to design more compelling experiences (2), or inform the development of efficient content-aware compression (3) and rendering techniques that take into account visually attractive regions to reduce computational costs. To accomplish this, it is necessary to have the availability of a substantial dataset containing a wide range of scenarios, along with corresponding gaze data collected from a diverse and extensive group of observers.

One of the earliest attempts to create a comprehensive dataset of gaze data in VR is the dataset by Sitzmann et al. (3). This dataset consists of twenty-two static scenes of visual stimuli, thus presenting two main limitations: firstly, static scenes imply that no movement or plot dynamics could influence the observers' attention; and secondly, it overlooks that our perception of the surrounding world is not based on visual stimuli but is multimodal, involving inputs form multiple senses (4).

The existing follow-up datasets have tackled these limitations by collecting gaze data for 360º videos. Despite that, they either do not include audio sources (5, 6) or overlook sound directionality (7), which is a key aspect. Humans inherently perceive sound through a combination of frequency, amplitude, and direction. This ability to locate the sound source significantly impacts our visual behavior, often directing our attention toward the source's direction (8). Therefore, gaze data collected without considering sound directionality fails to capture the full extent of the multimodal interactions that drive human visual behavior (9).

In this work, we present a taxonomy for the dataset D-SAV360. We have analyzed multiple aspects of the D-SAV360 dataset, reflecting in a categorized manner the scene diversity of 360º videos. Our structured taxonomy features dimensions such as content, acoustic information, or spatial regions of interest (ROIs). We hope this taxonomy helps researchers to better understand visual attention, and thus derive more accurate models of human visual behavior. This dataset is available at https://graphics.unizar.es/projects/D-SAV360.

2 RELATED WORK

The interest in 360º videos has grown significantly over the past few years, leading to the generation of multiple datasets. For instance, Morgado et al. (10) collected a dataset of 360º videos to generate and align spatialized audio taking into account the visual content (10, 11). Although this dataset contains a large number of videos, they were batch downloaded from YouTube, leading to inconsistencies in length, resolution, and frame rate, as well as a limited variety of scenes due to a lack of curation. Similarly, Rana et al. (12) created 360AVD, a dataset for learning to generate ambisonics from visual cues featuring short 10-second clips. De Coensel et al. (13) gathered a dataset of immersive audiovisual recordings of cityscapes to evaluate the perceptual influence of noise control and soundscaping measures through auralization. However, despite the usefulness of these datasets, none of them include head and gaze data, which is fundamental for gaining insights into how users perceive and process auditory and visual stimuli in immersive environments.

Along these lines, some works have gathered datasets of 360º videos with associated eye and head movement data to analyze the exploration behavior of users (5, 14, 6). However, their videos were played without directional audio or even without sound, which is an important element for immersion and has been shown to affect participants' visual behavior (15). More recently, Zhang et al. (7) introduced ASOD60K, an audiovisual 360º dataset that also captured gaze and head movements. Nevertheless, in their studies, their videos were presented only with mono sound, which limits the immersive experience (16, 17). Chao et al. (15) introduced a dataset that included ambisonics, which they used to compare viewing behavior between muted, mono, and ambisonic audio. While their dataset does provide head data, it does not include gaze data and is limited to fifteen videos and fifteen participants. Although some of these datasets present a high-level categorization of their videos, none of them include a comprehensive taxonomy that facilitates the investigation of how different environments and stimuli influence visual behavior within VR environments.

3 D-SAV360: DATASET OF GAZE SCANPATHS ON 360º AMBISONIC VIDEOS

In this section, we offer a summarization of the content of D-SAV360 (1). This is the most extensive dataset of viewing behavior on 360º ambisonic videos to date. It comprises a total of 85 monoscopic 360º videos with first-order ambisonic sound. This dataset also incorporates head and eye tracking data. For each of the 85 videos, it captures participants' visual behavior while watching the videos for 30 seconds. In total, 4,609 head and gaze trajectories for 360º videos with first-order ambisonics were recorded, providing a rich source of information for future research. The collected tracking data includes head position and orientation, pupil diameter, eye openness, eye gaze vector, and gazed image coordinates, all sampled at 120Hz. Moreover, it provides an analysis of the collected scanpaths and shows that various factors such as viewing mode, content type, and gender significantly impact eye movement statistics.

Video Collection: The dataset contains 50 stereoscopic videos recorded using a Kandao Obsidian S, equipped with six fish-eye lenses that present varied content and scene configurations. Additionally, it includes 35 monoscopic videos curated from the dataset of Morgado et al. (11) carefully selected ensuring uniform resolution, adequate duration, and semantic diversity, along with the aforementioned 50 videos also in monoscopic format. Furthermore, it provides an estimation of the optical flow for each video using the deep learning model called RAFT (18) and provides the computed audio energy maps (AEM) obtained with the decoder employed by Morgado et al. (11). The videos were standardized before recording gaze data by downsampling their resolution, unifying the frame rates, and limiting them to the most relevant 30 seconds.

Gaze Data Collection: The stimuli were presented on an HTC Vive Pro Eye head-mounted display (HMD) with a horizontal field of view (FoV) of 110 visual degrees and a vertical FoV of 110 visual degrees, a resolution of 1440x1600 pixels per eye, and a frame rate of 90 fps. Three HTC sensors were used to track participants' position, which was logged at 120Hz. To collect eye-tracking data, the SRanipal Unity SDK 4 developed for the Tobii eye-tracker integrated into the HTC Vive Pro Eye was used.

Procedure: Participants explore freely each of the videos by rotating in place while in a standing position. The experiment was split into two main blocks: one with 25 stereoscopic videos and another with 30 monoscopic videos, each divided into three phases of ten to five videos. Between phases and blocks, there were breaks. At the beginning of each block and phase, the eye calibration procedure was performed to ensure correct eye-tracking measurements throughout the whole experiment. All participants started at the same position for a given video and the participants' engagement was verified throughout the experiment with non-intrusive four-alternative forced-choice sentinel questions.

4 TAXONOMY

The taxonomy presented here aims to provide a structured classification system for organizing and categorizing the D-SAV360 dataset. Our taxonomy encompasses the dimensions of Content, Acoustic, and Spatial Regions of Interest (ROIs), which capture different aspects that impact visual behavior. Both the Content and Acoustic dimensions exhibit a transversal relationship, influencing and interacting with the ROIs dimension. However, the Content dimension interacts on a broader level with the other dimensions, allowing all possibilities to occur within each category of Content dimension (see Figure 2). In the following, we explain each dimension and its categories.

4.1 Dimension 1: Content

This first dimension of the taxonomy focuses on scene content. Scenes are categorized based on their thematic characteristics, providing insights into how different types of scenes influence gaze behavior and attentional processes.

Outdoor: Environments that primarily occur in open spaces, such as landscapes, streets, or parks.

- *Urban:* Settings characterized by man-made structures, buildings, roads, and urban infrastructure.
- *Natural:* Settings characterized by natural elements such as forests, mountains, and parks.

Indoor: Environments that primarily occur within enclosed spaces, such as rooms or buildings.

- *Complex Room:* Scenes with high visual complexity, containing numerous objects, textures, and details.
- *Simple Room*: Scenes with minimal visual complexity, containing few distinct objects.

4.2 Dimension 2: Acoustic

This dimension addresses auditory cues, allowing us to examine how auditory stimuli impact attention processes.

- *Ambient only:* These scenes feature ambient audio without any distinct ROI-related sound sources.
- *Without visual ROI source:* Scenes where there is no visual ROI but contains other auditory cues related to ROI.
- *With visual ROI:* Includes a specific auditory cue or sound source that corresponds to a ROI seen within the scene.
- *With and without visual ROI:* Setting containing both visual ROI sources and non-visual ROI sources.

4.3 Dimension 3: Spatial Regions of Interest (ROIs)

ROIs are defined as specific areas within the field of view (FOV) that are likely to attract viewers' attention. This dimension allows for the identification and analysis of gaze patterns within different regions and view fields.

One ROI or multiple concentrated in the same FOV: This category includes scenarios where viewers' attention is focused on either a single region of interest or multiple regions concentrated within the same field of view. To provide insights into the stability and dynamics of visual attention, we also distinguish: *Constant FOV* from *Changing FOV*.

Multiple ROIs: Scenes with multiple ROIs distributed across the FOV. In this category we distinguish: *Constant FOV*, *Changing FOV* and *Constant and changing FOV.*

No ROI present or infinite ROIs: Settings characterized by no specific region of interest defined, or where there are infinite potential ROIs within the scene. Analyzing gaze behavior in such contexts provides insights into spontaneous visual exploration and attentional processes in the absence of predefined stimuli.

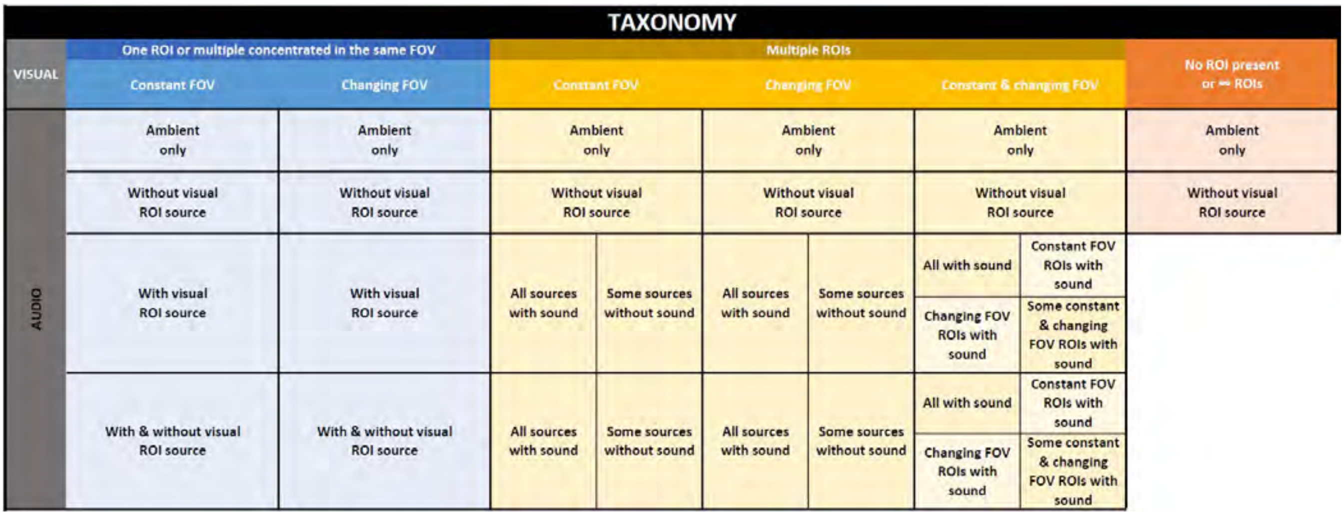


Figure 2. Our taxonomy framework represents the transversal relationship and interaction between the ROIs and Acoustic dimension. The Content dimension, which encompasses Outdoor (natural and urban) and Indoor (complex and simple room) environments, interacts on a broader level with the table presented, allowing all these possibilities to occur within each category of Content dimension.

5 CONCLUSIONS

We have presented a taxonomy for D-SAV360, a comprehensive dataset that includes head and gaze tracking data from 87 participants observing 85 different 360º ambisonic videos in VR, including both stereoscopic and monoscopic videos. The dataset jointly with the proposed taxonomy makes it a valuable resource for studying and modeling human visual behavior in immersive environments.

Looking ahead, we see the potential for the data collection jointly with our taxonomy to be leveraged in exploring new scenarios such as interactive scenes or social VR, investigating the influence of sound and movement, or the cognitive load derived by cluttered scenes. Additionally, we believe it could be a useful complement to benchmark audiovisual saliency models, allowing the identification of failure configurations.

ACKNOWLEDGMENTS

We extend our gratitude to the members of the Graphics and Imaging Lab for their support and collaboration in the video recordings. This work has received funding from Grant PID2022-141539NB-I00 funded by

MICIU/AEI/10.13039/501100011033 and by ERDF/EU; by the European Union's Horizon 2020 research and innovation program through the PRIME project (MSCA-ITN, grant agreement No. 956585); and by the Government of Aragon's Departamento de Ciencia, Universidad y Sociedad del Conocimiento through the Reference Research Group "Graphics and Imaging Lab (ref T34_23R). Additionally, Sandra Malpica, Daniel Martin, and Edurne Bernal-Berdun were supported by a Gobierno de Aragon predoctoral grant (2018-2022, 2020–2024, and 2021–2025, respectively).

REFERENCES

1 .Bernal-Berdun E, Martin D, Malpica S, Perez PJ, Gutierrez D, Masia B, et al. D-sav360: A dataset of gaze scanpaths on 360° ambisonic videos. *TVCG* (2023) 1–11.

2 .Serrano A, Sitzmann V, Ruiz-Borau J, Wetzstein G, Gutierrez D, Masia B. Movie editing and cognitive event segmentation in virtual reality video. *ACM Trans. on Graphics* **36** (2017).

3 .Sitzmann V, Serrano A, Pavel A, Agrawala M, Gutierrez D, Masia B, et al. Saliency in VR: How do people explore virtual environments? *TVCG* **24** (2018) 1633–1642.

4 .Martin D, Malpica S, Gutierrez D, Masia B, Serrano A. Multimodality in VR: A survey. *ACM Computing Surveys (CSUR)* **54** (2022) 1–36.

5 .David EJ, Gutiérrez J, Coutrot A, Da Silva MP, Callet PL. A dataset of head and eye movements for 360° videos. *Proc. of ACM Multimedia Systems Conference* (2018), 432–437.

6 .Zhang Z, Xu Y, Yu J, Gao S. Saliency detection in 360° videos. *Proc. European Conference on Computer Vision (ECCV)* (2018).

7 .Zhang Y, Chao FY, Zhang L. ASOD60K: An audio-induced salient object detection dataset for panoramic videos. *ArXiV (Preprint)* (2021).

8 .Noesselt T, Bergmann D, Hake M, Heinze HJ, Fendrich R. Sound increases the saliency of visual events. *Brain Research* **1220** (2008) 157–163.

9 .Warp R, Zhu M, Kiprijanovska I, Wiesler J, Stafford S, Mavridou I. Moved by sound: How head-tracked spatial audio affects autonomic emotional state and immersion-driven auditory orienting response in VR environments. *Journal of the Audio Engineering Society* (2022).

10 .Morgado P, Li Y, Vasconcelos N. Learning representations from audio-visual spatial alignment. *Advances in Neural Information Processing Systems* (2020).

11 .Morgado P, Nvasconcelos N, Langlois T, Wang O. Self-supervised generation of spatial audio for 360°video. *Advances in Neural Information Processing Systems* (2018), vol. 31.

12 .Rana A, Ozcinar C, Smolic A. Towards generating ambisonics using audio-visual cue for virtual reality. *International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2019), 2012–2016.

13 .De Coensel B, Sun K, Botteldooren D. Urban soundscapes of the world: Selection and reproduction of urban acoustic environments with soundscape in mind. *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (2017), vol. 255, 5407–5413.

14 .Xu Y, Dong Y, Wu J, Sun Z, Shi Z, Yu J, et al. Gaze prediction in dynamic 360° immersive videos. *CVPR* (2018), 5333–5342.

15 .Chao FY, Ozcinar C, Wang C, Zerman E, Zhang L, Hamidouche W, et al. Audio-visual perception of omnidirectional video for virtual reality applications. *Intl. Conf. on Multim. & Expo Workshops* (2020).

16 .Masia B, Camon J, Gutierrez D, Serrano A. Influence of directional sound cues on users exploration across 360° movie cuts. *IEEE Computer Graphics and Applications* (2021).

17 .Min X, Zhai G, Zhou J, Zhang XP, Yang X, Guan X. A multimodal saliency model for videos with high audio-visual correspondence. *IEEE Trans. on Image Processing* **29** (2020) 3805–3819.

18 .Teed Z, Deng J. RAFT: recurrent all-pairs field transforms for optical flow. *ECCV* (2020).

A Controlled Trial of Virtual Reality Hypnotherapy in Students for Exam Anxiety

Aileen O'Brien [1] Nik bin Fauzi, [2] Michael Abbott [3,] James Railton Darren Bell [3] Sarah White [1] Jared Smith [1], Nadia Mantovani [1] and Simon Riches [4]

1. St George's University of London, Barnet, Enfield and Haringey Mental Health NHS Trust 2, Southwest London and St George's Mental Health NHS Trust 3. King's College London, Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, London, 4.

Corresponding Author

aobrien@sgul.ac.uk

Keywords: virtual reality, student mental health, anxiety, hypnotherapy

Introduction

Virtual reality Hypnotherapy (VRH) describes the combination of a Virtual Reality (VR) experience with audio hypnotherapy. There is some early evidence that hypnotherapy can be highly effective when combined with VR for anxiety.

Materials and Methods

St. George's University of London (SGUL) is the only specialist healthcare university in the UK. An unblinded randomized controlled trial of 97 students was performed in November 2024 comparing VR hypnotherapy over five days with the same experience watched on a smartphone. All participants completed a number of wellbeing scales repeated two weeks post-trial. Before and after each session participants were also asked to complete simple Visual Analogue Scales asking how stressed, happy, sad, calm, and anxious they were (on a scale of 0 to 10). They were asked at the end of the five days which group they would have preferred to have been in and were asked about their views of the experience. Qualitative responses were analyzed by content analysis to identify themes.

Results

There were no adverse incidents or side effects. Quantitative data is currently being analyzed. Over 90% of those in the video group said they would have preferred to have been in the VR group. In the VR group, over half the sample said the experience made them feel 'calmer' or 'more relaxed'. Compared to the VR group, a smaller number of students in the video group reported benefit. Themes in the VR group included feeling more calm, confident, appreciating the opportunity to regulate emotions, and feeling "recharged".

Discussion

Results of the quantitative scales across the two groups are being analyzed. The students were receptive to the VR experience and most would have preferred that as an intervention. Initial qualitative feedback was very positive but highlights the need to run a more formal and structured over-time study to measure longer-term effectiveness and impact of the VR experience.

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. Dr O'Brien is an unpaid adviser for the development company, Phasespace.

Author Contributions

All authors were involved in the acquisition, analysis or interpretation of the work and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding

Funding was provided by a UK Innovate Grant (1005501)

Data Availability Statement

Datasets are available upon request

Anti-mafia hero embodiment as a tool to fight mafias: an immersive virtual reality study

Valerio Placidi[1,2*], Matteo P. Lisi[1], Alisha Vabba[1], Althea Frisanco[1], Salvatore M. Aglioti[1]

[1]CLN2S@Sapienza, Fondazione Istituto Italiano di Tecnologia (IIT) and Department of Psychology, Sapienza University of Rome, Viale Regina Elena 291, 00161 Rome, Italy

[2]International School of Advanced Studies, Department of Experimental Medicine and Public Health, University of Camerino, Camerino, Italy

*** Correspondence:**
Corresponding Author
valerio.placidi@unicam.it

Keywords: anti-mafia, attitudes, embodiment, proteus effect, psychophysiology, immersive virtual reality

Abstract

Mafias are criminal organizations, that chase wealth and power not only through the commission of crimes, but also by taking advantage of the passive acceptance of the local populations. Although studies have investigated the psychological predictors of attitudes and behaviors toward mafias, little attention has been given to practices potentially able to promote anti-mafia action. The aim of the present project is to evaluate whether specific immersive virtual reality (IVR) experiences may foster anti-mafia attitudes and behavioral intentions. In a pilot study, participants embodied - through visuomotor synchrony - an avatar resembling an anti-mafia hero (i.e., Giovanni Falcone, experimental group, n=6) or an unfamiliar person (control group, n=6). To familiarize with the virtual body, both groups completed motor tasks and then gave the same neutral speech expressing gratitude to a group of neutral-looking collaborators. Five days prior to the experiment, and immediately after the IVR exposure, participants completed an affective misattribution procedure and a mouse-tracking task measuring anti-mafia attitudes and action intentions, respectively. Preliminary results seem to indicate that embodying the anti-mafia hero increases negative misattribution elicited by mafia and the intention to act against mafias, compared to the baseline measurement. In opposition, no changes in misattribution and intentions are present after the control avatar embodiment. Despite being non-definitive, our results underscore the potential benefits of embodying an anti-mafia hero in fostering negative attitudes against organized crime. Importantly, this project represents a first attempt to apply IVR to fight mafia-type organized crime by fostering anti-mafia action. Since data collection is still ongoing, further results will be presented during the conference.

Artificial agents delivering pain and touch on the embodied avatar of human participants are socially evaluated depending on the stimulus valence

Matteo P. Lisi[1,2*], Martina Fusaro[1,2], Salvatore M. Aglioti[1,2]

[1]CLN[2]S@Sapienza, Fondazione Istituto Italiano di Tecnologia (IIT) and Department of Psychology, Sapienza University of Rome, Viale Regina Elena 291, 00161 Rome, Italy

[2]IRCCS, Santa Lucia Foundation, Via Ardeatina 306, 00179 Rome, Italy

*** Correspondence:**
matteo.lisi@iit.it

Keywords: Virtual agents, Social Evaluation, Virtual pain and touch, Interpersonal distance, Psychophysiology, Eye-tracking, Mouse-tracking.

Abstract

Digital immersive environments represent a unique opportunity to establish and maintain social connections when others are not physically present. Previous research has demonstrated the ability of Immersive Virtual Reality to induce embodiment facets like for example a feeling of ownership over a humanoid avatar observed from a first-person perspective. Furthermore, the mere observation of a syringe or a gentle caress delivered to the embodied virtual body can trigger derivative feelings of pain and pleasure, respectively. Nevertheless, it is unclear how the type and strength of vicarious sensations induced by stimuli delivered by a virtual agent on one's own virtual body can shape the human participants' attitudes towards the virtual agent and affect subsequent social interactions with it. In a within-participants design, 40 neurotypical individuals observed virtual agents delivering either a painful (virtual stab) or a pleasant virtual stimulus (virtual caress) on the right hand of a virtual body perceived as their own. We collected pre-post measurements to assess the effects of the virtual somatosensory stimuli on 1) explicit judgments of trustworthiness and attractiveness towards the virtual agent, on 0-100 visual analogue scales (VAS); 2) implicit evaluation of the virtual agents' facial attractiveness using a mouse-tracking procedure and 3) physiological arousal and gaze behavior during the virtual agents' approach in an interpersonal distance task. We found a non-linear pattern in explicit judgments' change, as the virtual pleasant touch improved the virtual agents' evaluation only when their baseline (pre-touch) levels were low (below 50, the midpoint of the VAS), suggesting a ceiling effect. In contrast, the virtual stab decreased the virtual agents' attractiveness and trustworthiness only when their baseline levels (pre-stab) were high (above 50, the midpoint of the VAS), indicating a possible floor effect. Caressing vs stabbing agents showed an increase in implicit attractiveness as indexed by the mouse tracking test, and induced a reduction of physiological arousal (SCR, HR) during their approach in the interpersonal distance task. Moreover, participants' gaze towards the stabbing agents' eyes increased during the interpersonal distance task already at far distances, hinting at increased alertness towards them. This study suggests that the quality of the somatosensory interactions experienced through one's embodied avatar can affect subsequent social interactions with virtual agents, both at the behavioral and autonomic levels. These findings might be relevant for understanding the psychological effects of embodied interactions in immersive digital spaces. Understanding the impact of both pleasant and unpleasant somatosensory

experiences on social interactions paves the way for developing more immersive and engaging virtual environments that foster positive interpersonal exchanges.

Bodily Self-Consciousness and dishonest behaviour in Digital Interactions: insights from the "Roll & Tell" dice social game applications for smartphone and immersive virtual reality

Provenzano Luca[1], Marina Scattolin[2,3], Giuseppina Porciello[2,3], Marina Serena Panasiti[2,3], Salvatore Maria [1]

[1]Sapienza University of Rome and CLN2S@SAPIENZA, Istituto Italiano di Tecnologia, Rome, Italy

[2]Dipartimento di Psicologia, Sapienza Università degli studi di Roma, Rome, Italy.

[3]IRCCS Fondazione Santa Lucia, Rome, Italy

*** Correspondence:**
Luca Provenzano
luca.provenzano@iit.it

Keywords: Immersive Virtual Reality, Morality, Bodily Self-Consciousness, Social interactions, Online gaming, Moral decision making

Abstract

Studies suggest that bodily self-consciousness (BSC) significantly influences dishonest behaviour. Given the ongoing digitalization of our social interactions, it is crucial to explore how these transformative technologies influence BSC and, consequently, dishonesty.

To address this issue, we developed the "Roll & Tell" smartphone app, through which dyads of participants (N= 43) played 60 matches of an online dice game. Each match involved participants rolling a virtual dice for 10 turns and reporting the outcomes to their opponent. The player reporting the highest outcome won the round, regardless of the actual outcome. Reporting the same outcome resulted in a loss for both players. Winning most of the rounds determined the overall match winner. Each round and match win were considered for compiling a global ranking, where the top 3 players received an additional (monetary) reward. Therefore, participants could increase their chances of winning by dishonestly reporting higher outcomes than those shown by the die roll. The difference between the reported vs. the real outcome was considered an index of dishonesty. We assessed individual Sense of Ownership (SoO) and Sense of Agency (SoA) through questionnaires.

Results show that participants' dishonest behaviour correlates positively with their ranking position, and negatively with SoA. This suggests that feeling responsible for one's own actions can facilitate moral behaviour even during digital social interactions, where the body is not prominently involved.

To further investigate the role of BSC during online interactions, we are developing an immersive virtual reality (IVR) version of the "Roll & Tell" app. Participants will access an online virtual environment to play the same dice game. To toss the dice, participants will use a virtual body whose movements they will fully be in control of. To investigate whether alterations of SoO and SoA affect participants' decisions during digital interactions, one condition will involve the fading of the virtual body until no longer visible, making it more difficult for participants to interact with the immersive environment. We expect that this embodiment manipulation will alter participants' BSC and, consequently, affect their propensity to behave dishonestly.

Cracking the "law of silence": immersive virtual reality to study behavioral and neurophysiological response to mafia-type crime.

Alisha Vabba[1], Valerio Placidi[1,2], Matteo P. Lisi[1], Salvatore M. Aglioti[1]

[1]CLN2S Sapienza, Fondazione Istituto Italiano di Tecnologia (IIT) and Department of Psychology, Sapienza University of Rome, Viale Regina Elena 291, 00161 Rome, Italy

[2]International School of Advanced Studies, Department of Experimental Medicine and Public Health, University of Camerino, Camerino, Italy

***Correspondence:**
Alisha Vabba, alisha.vabba@iit.com or alishavabba@gmail.com

Keywords: anti-mafia, bystander intervention, neurophysiology, virtual reality, prosocial behavior, mafia-type organized crime

Abstract

Mafia-type criminal infiltrations are on the rise, corrupting enterprises and damaging legal economies at the national and transnational level. As these operations thrive on complicity within communities and businesses, understanding the mechanisms at play when individuals witness and turn a blind eye to mafia-type crime is crucial for informing policies aimed at breaking the pervasive "law of silence" and combating mafias effectively. In two preliminary surveys, we observed that personality dimensions, demographics, and social context predicted anti-mafia inclinations, as measured via tasks developed in our lab - the Attitudes towards Italian Mafias Scale (AIMS; $N = 211$) and the Anti-Mafia Collective Action intention task (AMICA; $N = 360$). Building on these findings, we are developing an innovative serious game, using immersive virtual reality (IVR) to replicate criminal scenarios in a controlled, ethical, and safe environment, in combination with neurophysiological and behavioral investigations of what happens in the body and brain when individuals witness and react to mafia-type crime. Participants embody an avatar resembling themselves and witness multiple crime scenes (e.g., extorsion, bribery, money laundering, corruption, fraud, etc.), followed by decision nodes, in which they face moral choices (e.g., report the crime, remain silent, accept/refuse bribery), the consequences of which are played out. The game allows to investigate cognition (e.g., empathy, sense of agency, and cognitive conflict) via neurophysiological and behavioral tools like motion tracking, facial mimicry analysis, physiological activity, and electroencephalography. It enables manipulation of decision-influencing factors, such as personal risk and benefit or harm to others, and can help assess how variables like demographics, moral beliefs, and trust in institutions modulate behavior and cognition. Targeting non-WEIRD populations and sectors heavily impacted by Mafia infiltration (e.g., local politics, construction, transport, waste disposal and humanitarian aid), we ultimately aim to inspire effective interventions promoting prosocial behavior (i.e., bystander intervention, collective action, incorruptibility) against mafias.

On the Way of an Architecture for Supporting XR Collaborative Experiences

Roberto Cuervo-Rosillo [1,*], Telmo Zarraonandia [1] and Paloma Díaz [1]

[1] *DEI+ Group, University Carlos III of Madrid, Leganés , Madrid, Spain*

Correspondence*:
rcuervo@pa.uc3m.es

ABSTRACT

Extended reality (XR) encompasses a convergence of immersive technologies, seamlessly merging the boundaries between the physical and virtual realms. Social XR utilizes XR technology to facilitate social interactions transcending physical distances within an immersive environment where multiple persons can share, communicate and collaborate.

This innovative approach shows potential for overcoming the constraints of collaborating solely in physical or virtual environments. This facilitates scenarios where remote collaborators begin interacting in a shared virtual space and subsequently transition to continuing their collaboration in the physical world, for example. While recent technological advancements are enabling the delivery of such experiences, the design of these environments remains largely unexplored.

The objective of this work is to establish an architecture that facilitates collaborative experiences within XR environments.

Specifically, the system architecture should be capable of enabling:

- Collaboration in immersive virtual environments, allowing distant collaborators to interact within a completely synthetic environment using avatar representations.
- Collaboration in physical settings augmented with virtual elements, encompassing situations where collaborators in the same location interact directly while virtual elements enhance the environment to aid in the task at hand.
- Collaboration in hybrid settings, which involve scenarios where some collaborators interact physically while others prefer to engage within a virtual environment, represented by avatars.
- Smooth transitions between these three configurations. To initiate progress toward this objective, our focus is on identifying the fundamental components essential for crafting a collaborative experience and delineating the characteristics and critical aspects of each of them.

Virtual World Manager: The central component of the architecture is the software responsible for generating a replica of the virtual environment for each user and ensuring their synchronization and overall coherence throughout the collaborative experience. Network libraries like *Photon*[1] offer a set of components and infrastructure for generating instances of the same virtual environments

[1] https://doc.photonengine.com/pun/current/getting-started/pun-intro

and interconnecting them through the network. Other platforms like *Ubiq*[2] offer a broader array of services supporting avatar representation, communication, logging, among other features.

Representation Manager: To meet the needs of users interacting from various types of XR environments, it may be necessary to offer each of them different representations of the virtual elements within the shared experience. For instance, users interacting from VR environments might prefer background elements that enhance the sense of presence, while those in AR environments may prefer to have such components removed for a more seamless interaction. Similarly, full avatars may be more suitable to represent users in VR, whereas in AR, representing only parts of the avatars such as their hands might feel more appropriate for co-located users. The same applies to the rest of elements of the collaborative experience. This underscores the necessity of a representation manager, which determines how each element participating in the experience should be represented based on the requirements of the user interacting with it and the environment she is using.

Awareness Services Manager: To effectively collaborate in virtual environments, it's essential to implement mechanisms that enable awareness among collaborators. This entails knowing who is present in the workspace and what activities they are engaged in. Similar to the representation of virtual elements, awareness requirements may vary depending on the type of realm from which the user is interacting. The awareness services manager will be responsible for generating the appropriate cues for each collaborator based on their tasks, realm type, and preferences.

Virtual Viewpoints Alignment: It is essential for co-located AR collaborators to share common references in their physical space. This ensures that their virtual spaces are aligned with each other and with the physical room, allowing all users to visualize the same virtual objects in the same physical locations. This highlights the necessity for an architectural component tasked with enabling the setting, retrieval, and sharing of a series of anchors within the physical space where collaboration takes place. This guarantees that the perspectives of the collaborators remain synchronized.

In this poster, we present an architecture that enables the creation of collaborative experiences in mixed (XR) immersive environments by abstracting the device used. This architecture presents four basic components, but the need to include additional components as well as a set of services to develop its functionalities is taken for granted.

[2] https://ucl-vr.github.io/ubiq/

Exploring Outdoor Augmented Reality Registration Error Evaluation

Ziwen Lu [1,*], **Simon Julier** [1]

[1] *Department of Computer Science, University College London, United Kingdom*

Correspondence*:
ziwen.lu@ucl.ac.uk

ABSTRACT

We are motivated by the problem of how to make Augmented Reality (AR) more useful in outdoor urban environments. We propose to tackle two themes:

- Explore and attempt to clarify what ”good registration” means, especially in outdoor AR
- Explore methods to evaluate registration errors for outdoor AR

Registration errors are probably inevitable in AR. They are caused by many factors, including tracking, calibration, and modelling (authoring). Outdoor environments aggravate the impact of these factors through harsher lighting conditions, dynamic objects, larger space, and greater depth. However, even in the presence of such registration errors, AR systems can still be successfully used, as evidenced by the numerous number of products and prototype systems within research papers. On the other hand, AR systems experiencing more severe registration errors, such as tracking failure in challenging environments, will lead to users giving up on the AR system in this situation.

We are exploring the impact of these errors with two outdoor social mixed reality experiences in real world contexts. Both experiences are built to allow asymmetric collaboration between outdoor AR and virtual reality (VR) users. We identified the existence of registration errors from the aforementioned sources in those experiences. We have chosen these application domains because they allow us to investigate the effects of registration errors over multiple scales. These range from being able to correctly highlight a small-scale structure, such as a window, to providing cues which only provide a rough indication of a general area.

Additionally, we conducted a background survey to examine existing methods for registration error evaluation. We concluded from the survey that the impact of registration error is still largely unexplored in the context of outdoor AR.

Keywords: Augmented Reality, Registration Error, Tracking, Perception Study, Outdoor, User Study

ACKNOWLEDGMENTS

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 739578

Fourteen days of visuospatial training do not reduce motion sickness susceptibility

Michael Wiesing[1,2*], Eckart Zimmermann[1], Eike A. Schmidt[1], Axel Buchner[1]

[1] Institute for Experimental Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany

[2] Faculty of Psychology, Universitat de Barcelona, Barcelona, Spain**

*Corresponding Author
Email: m.wiesing@ub.edu

Keywords: Virtual Reality, Motion Sickness, Visuospatial Skills, Motion Sickness Susceptibility, Visuospatial Training

Abstract

Recent research by Smyth et al. (2021) was interpreted to suggest that visuospatial training is an effective method for reducing motion sickness susceptibility. The present study is a conceptual replication of Smyth et al.'s study with a larger sample size to test the replicability and robustness of the original findings. Participants were exposed to a 30-minute simulated drive in a virtual environment designed to induce motion sickness. Motion sickness was measured before, during, and after the drive. Subsequently, participants were divided into two groups: one received daily visuospatial training for 14 days, mirroring the intervention described by Smyth et al., while the control group did not engage in any form of training. Following the training period, participants were exposed to a second simulated drive to evaluate changes in their susceptibility to motion sickness.

Contrary to the promising findings of Smyth et al., there was no statistically significant difference in motion sickness reduction between the visuospatial training group and the no-training control group. Both groups exhibited a general habituation to motion sickness over time, suggesting that repeated exposure to the motion sickness-inducing stimulus might itself contribute to a natural reduction in motion sickness susceptibility.

References

Smyth, J., Jennings, P., Bennett, P., & Birrell, S. (2021). A novel method for reducing motion sickness susceptibility through training visuospatial ability–A two-part study. *Applied Ergonomics*, *90*, 103264.

**Note: Michael Wiesing is now affiliated with the EventLab, Faculty of Psychology, Universitat de Barcelona, Barcelona, Spain. The work reported in this abstract was conducted while he was at Heinrich Heine University, Düsseldorf, Germany.

Virtual reality-assisted therapy for low back pain management

Donegan, T.[1*] Amestoy, Beñat[1*], Calvis, I[1]., Rodriguez, E[1]., Oliva, R.[2], Slater, M[2]., Combalia, A. [1,3,4,5], Sanchez-Vives, M.V.[1,6]
[1] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), c. Villarroel, 170, 08036 Barcelona, Spain.
[2]Event Lab, Faculty of Psychology, University of Barcelona, Barcelona, Spain
[3] Servei de Cirurgia Ortopèdica i Traumatologia, Hospital Clínic de Barcelona, Universitat de Barcelona (UB), c. Villarroel, 170, 08036 Barcelona, Spain
[4] Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona (UB), c. Casanova, 143, 08036 Barcelona, Spain
[5] Departament de Cirurgia i Especialitats Medicoquirúrgiques, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona (UB), c. Casanova, 143, 08036 Barcelona, Spain.
[6] Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig de Lluís Companys, 23, 08010 Barcelona, Spain

*** Correspondence:**
Corresponding Author
amestoy@recerca.clinic.cat; donegan@recerca.clinic.cat

Keywords: virtual reality, AI, chronic pain, medical applications, rehabilitation, clinical VR

Abstract

Chronic low back pain is the leading cause of disability worldwide [1]. Current medical treatment and management strategies are expensive and often ineffective, and there have been recent calls to reduce the overmedicalization of low back pain, in particular in response to the opioid crisis in the US. Alternative strategies gain importance such as "digiceuticals", including reassurance and education to address unhelpful cognitions and behaviors, and strategies to support self-management and resumption of normal activities. Exercise forms a core part of these recommendations, but often there is little guidance as to the optimal type and dosage, and there is often kinesiophobia, or fear of movement, as a consequence of the low back pain. We have tried to overcome some of these limitations by using a virtual environment for the home training of patients with chronic low back pain. The first principle has been to develop embodiment of the virtual body [2,3], a virtual body that can be transformed to reduce chronic pain [4], and carry out movements and exercises that may correlate or not with those performed by the real body [5,6]. In this poster we will revise the challenges and opportunities of the domestic use of immersive rehabilitation programs, the potential AI-controlled adaptation to the patient's evolution, and the effectiveness of the pain ratings reduction and persistence of it after the experience in virtual reality.

Funding

Supported by XR-PAIN Grant Agreement No 101070533, funded by the European Media and Immersion Lab project financial support to third parties, which is funded by the European Union, and by Departament de Recerca i Universitats de la Generalitat de Catalunya (AGAUR 2021-SGR-01165- NEUROVIRTUAL), supported by FEDER.

Acknowledgments

We would like to thank the patients that participated in the experience.

References

1. Global Burden of Disease Study, C., Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*, 2015. **386**(9995): 743-800.
2. Slater, M., et al., Inducing illusory ownership of a virtual body. *Front Neurosci*, 2009. **3**: 214-220.
3. Sanchez-Vives, M.V., et al., Virtual hand illusion induced by visuomotor correlations. *PloS one*, 2010. **5**(4): e10381.
4. Matamala-Gomez, M., et al., Decreasing pain ratings in chronic arm pain through changing a virtual body: different strategies for different pain types. *The Journal of Pain*, 2019. **20**(6): 685-697.
5. Álvarez de la Campa Crespo, M., et al., Virtual embodiment for improving range of motion in patients with movement-related shoulder pain: an experimental study. *Journal of Orthopaedic Surgery and Research*, 2023. **18**(1): 729.
6. Bourdin, P., M. Martini, and M.V. Sanchez-Vives, Bourdin, P., Martini, M., & Sanchez-Vives, M. V. (2019). Altered visual feedback from an embodied avatar unconsciously influences movement amplitude and muscle activity. *Scientific reports*, 2019, **9**(1): 19747.

Phantom touch illusion: how the brain provides apparent touch in the metaverse?

Artur Pilacinski[1,*], Sophia Bertoni[2], Gabriella Andrietta[2], Marita Metzler[1], Christian Klaes[1]
[1] Chair for Neurotechnology, Ruhr-University Bochum, Germany
[2] CINEICC, University of Coimbra, Coimbra, Portugal
***Corresponding author:** AP: art.pilacinski@gmail.com

Keywords: phantom touch, virtual reality, tactile prediction, shared spaces, social touch, virtual touch

Introduction: Users in shared virtual spaces (such as VRChat or Metaverse, etc.) anecdotally report a weak sensation of touch when touching objects or other users are touching them, even though there is no real touch present. On the other hand, the nervous system attenuates the predicted sensory input caused by own movements (Blakemore et al., 2000). This mechanism is called tactile suppression. But what happens with tactile suppression if there is no afferent tactile signal but visual touch cues are present? Recently, (Pilacinski et al., 2023) we answered this by showing that touching objects in virtual reality results in users consistently reporting a weak tingling sensation we called the phantom touch illusion (PTI).

Methods: We used an immersive virtual reality scenario in which subjects touched their body using a virtual object. This touch resulted in a tingling sensation corresponding to the location touched on the virtual body. The subjectively reported intensity of the illusion has different strength across different parts of the hand. Interestingly, the illusion was also present when subjects touched invisible (inferred) parts of their limb. We reason that PTI results from the tactile suppression process. We additionally tested this by comparing PTI to self-touch using different laser pointers and pantomimed (no visible effector) touch. These conditions were reported as tingling touch by a dramatically lower proportion of subjects.

Discussion: Touching objects without physical touch available results in a weak tingling sensation. The presence of PTI when touching invisible body parts suggests that tactile suppression is not exclusively based on vision, but rather on multi-sensory input involving body schema. This finding indicates that, in virtual reality, the human brain uses available visual information to construct a multi-sensory representation of the apparent physical world, including objects (c.f. Brogni et al., 2011; Ehrsson, 2005) or other users (Alexdottir et al., 2022). How this multi-sensory representation is influenced by top-down cognition or multisensory integration remains a topic for discussion (see: Lush et al., 2020; Slater and Ehrsson, 2022), but our data suggests that phantom touch might be an important gateway to approach embodied experiences in virtual spaces.

References

Alexdottir, S., & Yang, X. Phantom touch phenomenon as a manifestation of the visual-auditory-tactile synaesthesia and its impact on the users in virtual reality. In 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). 727–732. (IEEE, 2022).
Blakemore, S. J., Wolpert, D. & Frith, C. Why can't you tickle yourself?. NeuroReport 11(11), R11–R16. https://doi.org/10.1097/00001756-200008030-00002 (2000).
Brogni, A., Caldwell, D.G., Slater, M. Touching sharp virtual objects produces a haptic illusion. In Virtual and Mixed Reality—New Trends. VMR 2011. Lecture Notes in Computer Science. Vol. 6773 (Shumaker, R. eds.). https://doi.org/10.1007/978-3-642-22021-0_26 (Springer, 2011).
Ehrsson, H. H., Holmes, N. P. & Passingham, R. E. Touching a rubber hand: Feeling of body ownership is associated with activity in multisensory brain areas. J. Neurosci. 25(45), 10564–10573. https://doi.org/10.1523/JNEUROSCI.0800-05.2005 (2005).
Lush, P. et al. Trait phenomenological control predicts experience of mirror synaesthesia and the rubber hand illusion. Nat. Commun. 11(1), 1–10 (2020).
Slater, M. & Ehrsson, H. H. Multisensory integration dominates hypnotisability and expectations in the rubber hand illusion. Front. Hum. Neurosci. 16, 834492 (2022).

Exploring the analgesic effects of virtual and real touches: neurophysiological and behavioural evidence

Fusaro Martina[1,2,*], Vecchio Arianna[2], Provenzano Luca[1,2], Aglioti Salvatore Maria[1,2],

[1]CLN2S@Sapienza, Fondazione Istituto Italiano di Tecnologia (IIT) and Department of Psychology, Sapienza University of Rome, Viale Regina Elena 291, 00161 Rome, Italy

[2]IRCCS, Santa Lucia Foundation, Via Ardeatina 306, 00179 Rome, Italy

*** Correspondence:**

Corresponding Author: martinafusaro@gmail.com

Keywords: immersive virtual reality, pleasant touch, pain modulation, EEG

Research findings suggest that both subjective and electroencephalography (EEG) responses to pain can be influenced by the experience of pleasurable touch. Using immersive virtual reality (IVR), we showed that slow (3 cm/sec) virtual touches (in absence of tactile feedback) administered to an avatar observed from a first-person perspective can evoke a sense of pleasantness resembling the one evoked by real-life stimulation. To further explore this phenomenon, we conducted an experiment in which 30 participants received blocks of slow and fast (30 cm/sec) touches in both real-world and IVR scenarios, followed by laser induced painful stimulations. Our objective was to examine the processing of event related potentials evoked by real and virtual touches and to assess their analgesic effect on the induced pain. Additionally, we gathered subjective ratings regarding the perceived pleasantness of touches and the unpleasantness of the laser. Subjective reports revealed that slow touches were consistently rated as more pleasant than fast touches, both in real ($p<.001$) and IVR scenarios ($p<.001$). EEG results demonstrated that slow touches, whether real or virtual, elicited a higher negative late potential on frontal-central sites (500-1200ms) compared to fast touches ($p=.006$). Furthermore, the P2 components of laser-evoked potentials over central electrodes (230-350ms), indicative of pain processing, were reduced when preceded by virtual touches (no matter whether slow or fast) compared to real touches ($p=.01$). Importantly, participants rated the unpleasantness of the laser to be lower when preceded by slow touches (slow vs fast touches, $p=0.03$), regardless of the scenario. In conclusion, our study highlights the remarkable potential of both real and virtual slow touches in influencing subjective pleasantness and neural responses, presenting promising opportunities for leveraging immersive virtual reality as a tool for pain management and rehabilitation.

eXtended Reality and Biosignals in the Therapy of Individuals with Intellectual and Physical Disabilities

Matteo Dal Magro [1,2*], **Pablo Pérez** [2], **Jesús Gutiérrez** [1]

[1] *Grupo de Tratamiento de Imágenes, Universidad Politécnica de Madrid, Spain*
[2] *Nokia XR Lab, Madrid, Spain*

Correspondence*:
Corresponding Author
m.dalmagro@upm.es

ABSTRACT

The majority of existing eXtended Reality (XR) applications and research studies target neurotypical non-disabled individuals, there is thus a lack of studies in the realm of inclusive XR (Halbig and Latoschik (2021); Long et al. (2022)). This PhD will leverage the benefits and opportunities offered by immersive scenarios, integrating them with objective and reliable measures obtained through biosensors, seeking to develop effective therapy paradigms targeted to individuals facing vulnerable situations. Simultaneously, the use of these technologies will enable the study of human interactions within XR, providing valuable insights on the behaviors, emotions and experiences of underrepresented categories.

The main project, Incluverso5G[1], carried out in collaboration with Nokia XR Lab (Spain) and Fundación Juan XXIII (Spain), will involve the development of therapy programs aimed at individuals with intellectual disabilities.

The first use case aims to develop immersive experiences for the treatment of bathmophobia using exposure therapy techniques (Cortés et al. (2024)). Various biosignals, such as Electrodermal activity (EDA), Heart Rate (HR) and Electromyography (EMG) will be collected. In subsequent analyses the biomarkers will be analyzed to detect cognitive states which will be correlated with the therapist's decisions during the sessions.

The second use case aims to train attention and inhibition skills using immersive serious games. This phase will specifically target subjects who are actively seeking employment. The application will feature various scenarios inspired by real life job situations in order for patients to develop various skills directly applicable to the workplace. Biosignals, including EDA, HR, EMG, eye tracking and head movements will be collected and analyzed to study behaviors, interaction and experiences of individuals with intellectual disabilities.

Through a multidisciplinary approach, this PhD will strive to bridge the gap between XR technology and inclusion, with a particular emphasis on enhancing the quality of life for those facing vulnerable situations.

[1] This work is partially supported by project TSI-064200-2022-009 (INCLUVERSO 5G) funded by program UNICO I+D 6G 2022 of the Spanish Government within the framework of the Recovery, Transformation and Resilience Plan and by project PID2020-115132RB (SARAOS) funded by MCIN/AEI/10.13039/501100011033 of the Spanish Government.

Keywords: inclusive metaverse, extended reality, intellectual disabiliy, biomarkers

REFERENCES

Halbig A, Latoschik ME. A systematic review of physiological measurements, factors, methods, and applications in virtual reality. *Frontiers in Virtual Reality* **2** (2021). doi:10.3389/frvir.2021.694567.

Long N, Lei Y, Peng L, Xu P, Mao P. A scoping review on monitoring mental health using smart wearable devices. *Mathematical Biosciences and Engineering* **19** (2022) 7899–7919. doi:10.3934/mbe.2022369.

Cortés C, Goyena M, Dal Magro M, Fernández-Alcaide A, Nava-Ruiz M, Gutiérrez J, et al. Immersive behavioral therapy for phobia treatment in individuals with intellectual disabilities. *2024 IEEE Conference on Virtual Reality (IEEEVR)* (2024), Accepted.

Metaverse applications in learning processes: good practices and main challenges

Daniela La Foresta[1], Ilaria Bruner[2*]

[1]University of Naples Federico II, Department of Political Science, Naples, Italy

[2]University of Naples Federico II, Department of Political Science, Naples, Italy

*** Correspondence:**
Ilaria Bruner
ilaria.bruner@gmail.com

Keywords: Education, Immersive Didactic, Metaverse application

Abstract

The Metaverse, hallmark technology par excellence of the fourth industrial revolution, constitutes a key transition to a new dimension of interconnected experiences. Often defined as the next frontier of Web 3.0 due to its various dimensions and implications in today's world, it represents a fertile ground for research and innovation.

The fields of application of the new technological paradigm are manifold, ranging from entertainment to commerce, from work to tourism. In particular, education appears to be one of the fields most affected by the new technology, which could be a real keystone for its innovation: traditional approaches, often criticized for being overly theoretical and rigid, stand to be significantly impacted by this innovative technology.

The research aims to analyse the main uses of the Metaverse in the didactic and educational sphere, through a systematic review of the literature on the subject and of the main initiatives implemented in the European sphere. More specifically, it intends to propose a reflection on the successful elements that could determine a replicability of the actions, while at the same time highlighting the potential negative effects deriving from the technological deficit, student involvement, teacher adaptability and ethical considerations.

Drawing upon the extensive body of academic research on teaching processes and the emerging literature on AR, VR and Metaverse applications, the aim is to provide a comprehensive overview of the issue. Through in-depth analysis, the contribution intends to demonstrate that while the use of the Metaverse in education is still nascent, it holds the potential to revolutionize didactic processes and generate positive learning outcomes.

Ndinguwe – „I am You“ How Do Recipients Integrate Experiences with Unfamiliar Forms of Perception in Their Imagination of Reality?

Tobias Bieseke [1,2*], Jan Schulten[1], Azziza El-Yabadri[1], Georg Trogemann[2], Harald Opel[1,]

[1] storyLAB kiU, Department of Design, Fachhochschule Dortmund, Dortmund, Germany

[2] Laboratory for Experimental Computer Science, Academy of Media Arts Cologne (KHM), Cologne, Germany

*** Correspondence:**
Corresponding Author
tobias.bieseke@fh-dortmund.de

Keywords: expanded realities (XR), virtual realities (VR), artistic research, metaverse, self-representation, mixed reality, haptically feedback, Ndinguwe,

Abstract

The research project "Ndinguwe"[1] investigates the interaction between cognition and narration in virtual space within the technical setting of a head-mounted display. It examines the self-representation of the recipients as avatars in virtual space, the analysis of the relationship between virtual objects and their influence on perception through haptic feedback and the exploration of the interface between the virtual environment and the real world in so-called mixed realities. The central research question is: How do the participants integrate unknown forms of perception into their personal space of experience?

1 Introduction

"Our cognitive domain is limited and infinite in the same way that our reality domain is limited and infinite" (Translated by the author/Maturana 2000, p.140)

The head-mounted display-based (HMD) artistic research project Ndinguwe was developed in the StoryLab kiU of the Fachhoschule Dortmund in the Dortmunder U in collaboration with the PhD program of experimental computer science of the KHM Cologne. The focus is on three main aspects:

- ➔ the self-representation of the recipient as an avatar in virtual space
- ➔ the relationship of virtual objects to haptic feedback
- ➔ and the connection of the virtual environment with the real environment in so-called mixed realities.

The word Ndinguwe is borrowed from the Xhosa language and means „I am you“. This is an interpretation of the Ubuntu philosophy often represented by Nelson Mandela, which simply means: "I am because you are". This experiment is about the transmission of self-perceptions within

[1] Here is a short trailer available: https://vimeo.com/915943769?share=copy filmed and edited by Stephan Hauptmann

discriminated people[2], which can be simplified to the factors of gender, age, ethnicity, and physical condition. The inner perspectives of such attributions are artificially generated and enable the experience of a new self. For example, there is the question if an avatar can function as a bridge for identification, empathy or rejection? These and other states of self-awareness are simulated in Ndinguwe, thus enabling experiences that cannot yet be generated in any other medium. In addition, the reference to haptic objects, such as a ball and a chair, is established, which are playfully used to become objects of action. In the application, no interfaces are used remotely, but the interactions culminate with the bodily behavior that the recipients know from their everyday life. This enables an intuitive hybridization of virtuality and reality into a mixed reality and establishes a cognitive connection. The potential of this way of working lies in this pattern. The experiment is examined for its results with the help of micro-phenomenology for artistic research.

2 Methods

The subject is immersed and enters a computer-generated environment into which their self-representation is depicted as a human imitation. For the evaluation of the experiment, it is exactly these microperceptual classifications that are of interest, which is why the method of microphenomenology was used for the evaluation. The evaluation was conducted with the support of the method of microphenemonology, which has its roots in the psychology of the Würzburg School. At the center of this method is the subjective "experience" of the participants, which is simultaneously expressed in language. Ideally, a kind of flow is created in which the perceptive mechanisms, associations, moods, reflections, or observations of the participants can be expressed in words. Thus Gerhard Benetka and Thomas Slunecko write in their text: “»Erleben«, das zur Sprache kommt“ ("»Experiencing« that comes to speech"/ translated by the author):

"Microphenomenologists want to induce with their questions a kind of experiential trance, in which one does not look back on what has been experienced, but in which the experience is held in presence and spoken from the epistemic authority of this current experience. They want to go, to put it in a phenomenological turn, to the experience itself - and not to be content with the memory of what has been experienced." (translated by the author/Benetka et al. 2021)

Of course, this method is misappropriated for the artistic research applied here. As described, it has a different structure and pursues its own production goals. The results varied accordingly. Some participants felt overwhelmed with the task of translating the experience into language. With others in turn, one could document numerous descriptions and associations in the event moments. There are approximately 70 evaluations. There are three parts of the experiment: The 1.) is about mixed reality (MR), 2.) is about haptical feedback and the 3.) is about self-representation in virtual reality (VR).

2.1 Mixed Realities

Question: How do participants experience their environment within an expanded reality? How does it change their relationship?
Process: This part of the experiment is examined in three different constellations.

1. the participants drive into a virtual tunnel in which holes show the real environment.
2. the participants experience their body overlaid with a virtual body in front of a mirror
3. the participants are attacked by a virtual bird after experiencing haptic feedback.

---

[2] It is inspired by the experiment of Mel Slater about the reducing of racial bias in black avatars (Slater et al. 2013)

In these different representations of mixed reality, the experience of the interaction of virtual reality and the environment on the participants is examined.

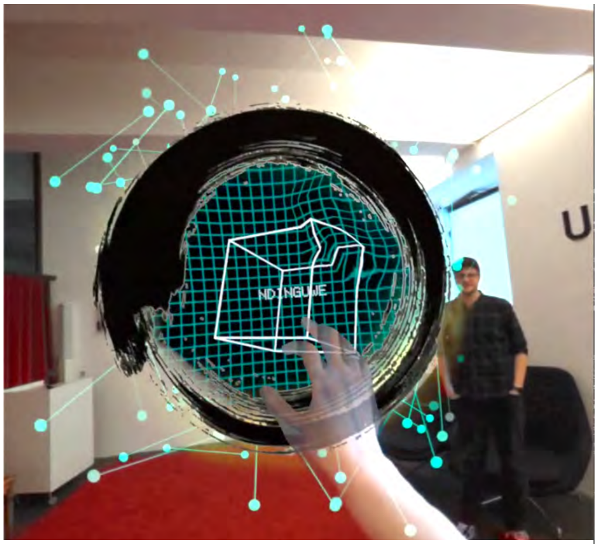


Figure 1 The Entrance of Ndinguwe.

Figure 2 MR Mirroring of real and virtual body.

Quickresults:

- ➔ As soon as the virtual space moves, the participants stand still, as soon as the space stands still, the participants move. (No Sickness)
- ➔ The participants experienced the overlaying of their own body with a virtual body as an intensification of the relationship with the avatar.
- ➔ The participants react to the bird, but do not consider it to be physically present.

2.2 Self-Representation

Question: What relationship do participants establish with self-representations that bear the hallmarks of discrimination?
Process: The participants could then try out their interaction in front of the mirror while listening to the narration of the character. Their visual representation was limited to gestures, their upper body, their position, and the movement of the fingers. Facial expressions were not displayed. When the participants crossed the starting line, their respective avatar transformed in the following sequence: a.) homosexual man, b.) Ukrainian woman without a hand, c.) old man with dementia or d.) African refugee. The choice of characters was based on characteristics that favored discrimination (age, gender, physical condition, and ethnicity).

Figure 3 Left: Participant claps her hands together. Right: What she see in VR.

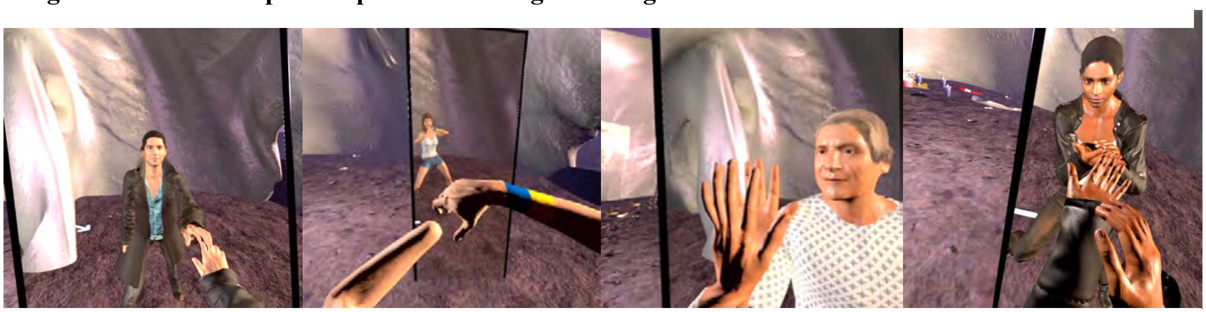


Figure 4 The four different avatars from left to right: a homosexual man, a woman without a hand, an old man and a colored refugee.

Quickresults:

- ➔ Identifikation 23 %
- ➔ Empathie 56 %
- ➔ Rejection 14 %
- ➔ Nothing happens 7 %
- ➔ Many different subjective descriptions of the experience. (e.g. the absence of the hand was described as phantom pain)

2.3 Haptical Feedback

Question: How does the participants' relationship to the virtual environment change with the integration of haptic feedback?
Process: This part of the experiment consists of three sub-stages.

1. interaction with virtual balls,
2. interaction with a virtual doppelganger and
3. interaction with physical objects with virtual representation.

The three stages are intended to make it possible to compare how the participants' subjective relationship to the experience changes. In the last part, when the participants place a chair at a certain point in the room, a flight through the virtual Dortmund city is started.

Figure 8 Participant bringst he physical chair to virtual place.

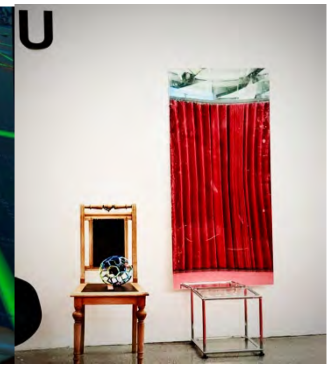


Figure 9 Physical objects.

Figure 10 Participant interacts with virtual balls.

Quickresults:

- ➔ A virtual doppelganger that does not behave like a mirror image cannot be controlled intuitively.
- ➔ Haptic contact increases the immersive experience.
- ➔ Artificially generated experiences (flying with a chair) become more subjectively believable through physical feedback.

3 Results - Three Types of Coherence between Perception and Virtual Reality

1. no coherence: If there is no coherence, the experiences are mostly perceived as irritation or error because they cannot be integrated into structures guided by the imagination (e.g.: The virtual ball flies through the participants).
2. a partial coherence: In the case of partial coherence, ambiguous interactions occur, which partially produce effects in which the brain tries to compensate for the experience difference and thus causes hallucinations (disembodied contacts are felt; illusion of seeing blood).
3. an approximate coherence with known experiences: In the case of approximate coherence, irritations are accepted, and differentiation is subjectively difficult. A form of multimodal perception emerges (net structure of the ball is accepted despite the optic difference; haptics increases confidence in the virtual environment). Unknown experiences are therefore integrated into known ones by reference to them.

4 Discussion and Conclusion

The research question posed at the beginning, how the participants aligned their personal space of experience with the experiences in the virtual world, can be answered to the effect that the integration takes place in different stages depending on the user: There is 1. no coherence, 2. a partial coherence, or 3. an approximate coherence with known experiences. If there is no coherence, the experiences are mostly perceived as irritation or error because they cannot be integrated into structures guided by the imagination (e.g.: The virtual ball flies through the participants). In the case of partial coherence, ambiguous interactions occur, which partially produce effects in which the brain tries to compensate for the experience difference and thus causes hallucinations (disembodied contacts are felt; illusion of seeing blood). In the case of approximate coherence, irritations are accepted, and differentiation is subjectively difficult. A form of multimodal perception emerges (net structure of the ball is accepted despite the optic difference; haptics increases confidence in the virtual environment). Unknown experiences are therefore integrated into known ones by reference to them. The interaction with the avatars shows that the half-identity allows for a confrontation with the alien self, whereby the participants can react either insituatively or extituatively, i.e. either they playfully let themselves in or they reject it. Both reactions create an expanded sense of self-awareness. In this regard, the philosopher Tom Poljanšek (Poljanšek 2022 p. 183), following the phenomenologists Johann Friedrich Herbart and Edmund Husserl, speaks of apperception. By this, he means the process of Hinzu-wahrnehmens (adding-perceiving/translated by the author), i.e. the addition of perceptions. This is what happened to the participants when they suddenly wore an avatar. The inner monologues of the avatars, on the other hand, are less relevant and are only perceived with attention when they are of interest. As soon as the participants were no longer spectators but participants, they wanted to act and take on the attitude of the listener only to a limited extent. Identification can therefore only be generated in the sense of "locating" the self-representation and not in the sense of transferring a persona. The degree of empathy depends on the ability to get involved.

The ball game showed that provoking a stressful situation results in discomfort but not cancellation. Nevertheless, physical reactions such as cries of pain or evasive reactions resulted, although it would have been enough to close the eyes to escape the optical bombardment. The haptic feedback, on the other hand, meant a clear increase in immersion. It turns what is seen into something material, even if it does not meet the expectations of haptics. The flight on the chair becomes more real because it has a physical basis. Here, the participants reverted from the state of the participants to the attitude of the spectator. The flight on the chair became more real because it had a physical basis. Accordingly, more of the narrative was perceived in this last phase. Overall, it remains to be noted that narration takes a backseat to interaction. For the narration to be received, according to the current state of knowledge, either the willingness of the participants must exist, or it must be stimulated (e.g. by reducing the possibilities for interaction). It is important to activate the inner attitude as a spectator or listener.

5 Author Contributions

Tobias Bieseke is the mainauthor and responsible for conceptualization, theoretical framework, text preparation and implementation. Jan Schulten was the programmer of the application of the experiment and Azziza El-Yabadri was the art director. Prof. Dr. Georg Trogemann and Harald Opel are the supervisors of thie experiment.

6 References

Maturana H (2000) Biologie der Realität.

Benetka G and Slunecko T (2021) „»Erleben«, das zur Sprache kommt", Journal für Psychologie 29, Nr. 2 (o. J.): 30, https://doi.org/DOI: 10.30820/0942-2285-2021-2-17

Poljanšek T (2022) „Nie ganz bei den Sachen", in Technik-Ästhetik: zur Theorie techno-ästhetischer Realität, ed. O. Ruf and L. C. (Grabbe, Medien- und Gestaltungsästhetik 12, Bielefeld: transcript), 183-202.

Slater M et al. (2013) Putting yourself in the skin of a black avatar reduces implicit racial bias. Consciousness and Cognition 22 (2013) 779–787. https://doi.org/10.1016/j.concog.2013.04.016

Real-World VR Integration for Mild Cognitive Impairment Rehabilitation

Jose Ferrer Costa[1], Maria-José Ciudad[1], Maribel González Abad[1], José Luis Rodríguez García[1], Julisa Cortes[2]

[1]Badalona Serveis Assistencials, Badalona, Spain.

[2]IES Health intern for the Liberal Arts & Business program

*** Correspondence:**
Jose Ferrer Costa.
jfcosta@bsa.cat

Keywords: virtual reality, augmented reality, mild cognitive impairment, cognitive rehabilitation, social interaction.

Abstract

This pilot explores the integration of Virtual Reality (VR) and Augmented Reality (AR) for enhancing cognitive rehabilitation in seniors with mild cognitive impairment (MCI). A customized VR program comprising interactive exercises and 360-degree videos was developed to enhance cognitive functions and encourage social interaction. The program's usability and the qualitative feedback from both participants and healthcare professionals were crucial, indicating strong acceptance and the potential to boost engagement. This feedback was essential for refining the VR content, demonstrating its adaptability and potential in geriatric cognitive rehabilitation.

1 Introduction

The aging population is increasingly confronted with the challenge of mild cognitive impairment, a condition that often precedes more severe dementia types (1, 2). Despite the effectiveness of traditional cognitive therapies, they frequently fall short in terms of engagement and accessibility (3, 4). This pilot study, conducted at the Intermediate Care Center El Carme in Badalona, Spain, explores the usability and acceptability of Virtual Reality and Augmented Reality as innovative interventions for seniors with MCI. In collaboration with Reality Telling, we have developed a customized VR program comprising interactive cognitive exercises (Figures 1-4 and immersive 360-degree videos (Figures 5 and 6). These exercises aim to enhance memory, executive function, and attention through engaging in daily life tasks, while the videos provide virtual tours to promote social interaction and emotional engagement. The assessment of this intervention's impact on cognitive functions, conducted using standardized scales, will be detailed in a forthcoming publication, focusing this presentation on the practical application and user experience of VR and AR in cognitive rehabilitation.

2 Methods

The pilot, registered under ClinicalTrials.gov ID NCT06155721, enrolled 45 senior participants with MCI to assess the impact of VR on various cognitive functions. The intervention involved 8 VR sessions using Meta Quest 2 headsets, integrating hand-tracking technology for immersive interaction. The study was designed to capture detailed daily feedback from both patients and therapists, allowing for the continuous adjustment and improvement of the virtual content created.

This process allowed the continuous adaptation of VR and AR content to meet user needs effectively, emphasizing safety and engagement (Figures 6 and 7). Final feedback was analyzed using the SUS and qualitative comments, providing a comprehensive evaluation of the VR program's complexity, entertainment value, and overall comfort. The findings from this feedback were instrumental in refining the VR intervention, demonstrating the value of a patient-centered and iterative design process in clinical VR applications.

3 Results

3.1 Patients´ Engagement and Satisfaction

An overwhelming 90.48% of participants felt comfortable using the VR system, and 76.19% expressed a strong desire to use the program frequently, showcasing high user satisfaction and acceptance. The majority found VR exercises more entertaining than traditional methods, with 50% agreeing or strongly agreeing, highlighting VR's potential as a more engaging cognitive stimulation tool. The complexity of the program was not a significant barrier, with a majority of participants reporting they found the program either agreeable or neutral in terms of complexity (figure 8).

3.2 System Usability and Acceptability by the patients

While some participants perceived the VR headset as heavy, the study's tailored content, developed following guidelines to reduce side effects, was effective—indicated by the high comfort levels reported. Notably, despite reports of discomfort, the majority did not have to discontinue the VR experience, suggesting successful mitigation strategies were in place (Figure 9).

3.3 Professional Evaluation: Usability and Impact on Patient Care Quality

Usability: Professionals unanimously praised the VR program's user-friendly interface, highlighting its ease of integration into daily clinical practice. While acknowledging a learning curve for some patients, it was considered manageable, pointing to the program's straightforward design and relevance to clinical settings (Figure 10).

Impact on Patient Engagement: Healthcare professionals observed an enhanced patient engagement with VR sessions compared to conventional methods. The immersive VR environment minimized distractions and discouraged skipping challenging tasks, leading to higher levels of active involvement and immersion. This contrasts with the lesser engagement seen with computer or paper-based tasks, where patients showed more distractibility.

Effectiveness in Cognitive Stimulation and Social Interaction: Professionals reported improvements in participants' mood, cognitive engagement, and willingness to engage in therapeutic activities. The use of 360-degree videos, in particular, was seen as beneficial for fostering socialization and making group sessions more impactful and relatable.

4 Discussion

4.1 Pilot Study Insights

Our pilot shows a significant shift towards using VR and AR for MCI treatment, indicating a more effective engagement method for seniors compared to traditional techniques. Customized VR content, tailored to seniors' cognitive levels and designed to minimize side effects, has seen broad acceptance among users and therapists, highlighting VR's potential for personalized cognitive rehabilitation.

4.2 User-Centered Design and Safety

Iterative content refinement, driven by users and professionals´ feedback, underscores our dedication to safety and experience, adapting the VR environment accordingly. These adjustments ensure the safety and immersion of seniors, showing the dynamic adaptability of VR/AR in meeting user needs .

4.3 Preference and Practicality

The transition from traditional methods to VR indicates a clear preference for immersive experiences that minimize distractions. The professionals´ choice between VR and AR is dictated by the nature of the task, with AR recommended for movement-involved activities for safety, and VR for focus-intensive tasks.

4.4 Future Directions

Ongoing research aims to further evaluate VR's therapeutic effectiveness, with an emphasis on accessibility and customization for diverse cognitive needs. This initiative marks a step towards a more engaging, personalized cognitive care model, promoting the continuous innovation and integration of immersive technology in healthcare.

5 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.

6 Author Contributions

JFC conceptualized the main idea, oversaw the development of VR environments, and, together with MJC, authored the theoretical framework and text preparation. MJC also contributed to the conceptualization and theoretical foundation. MGA, JLRC, and JC were pivotal in data collection and the practical implementation of the pilot, ensuring the study's operational success. All authors are committed to the integrity and accuracy of the work, adhering to the standards for authorship criteria and accountability.

7 Funding

The pilot study received principal support from Badalona Serveis Assistencials, with additional funding from a digital content grant to Reality Telling by Àrea Metropolitana de Barcelona (AMB).

8 Acknowledgments

We extend our heartfelt thanks to Reality Telling for their invaluable collaboration in developing the VR program that stands at the heart of this study. Special gratitude is also due to the AMB for their financial support through a grant for innovative social and technological entrepreneurship initiatives.

9 References

1- United Nations, Department of Economic and Social Affairs, Population Division. World population ageing 2015. [Internet]. 2015. Available from: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf

2- World Health Organization. Dementia: fact sheet. [Internet]. 2023 Mar 15. Available from: https://www.who.int/en/news-room/fact-sheets/detail/dementia

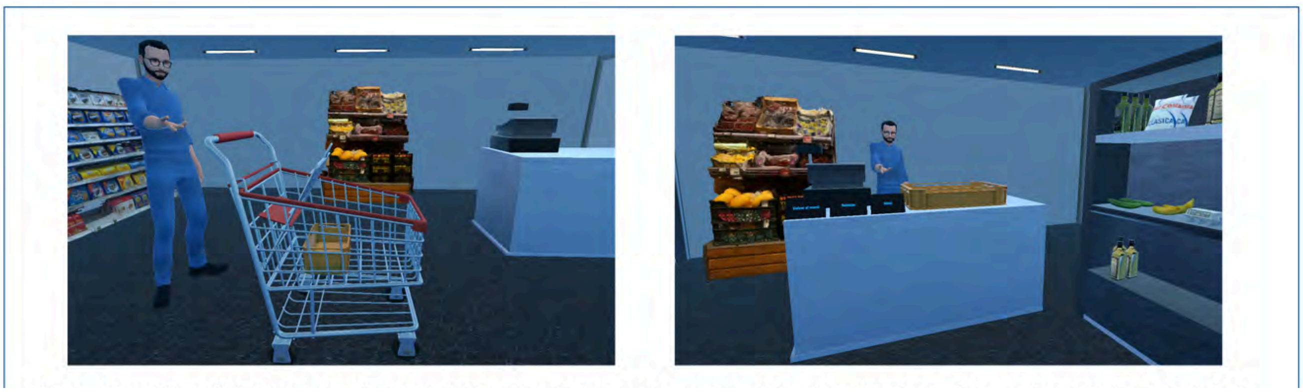
3- Brai E, Tonacci A, Brugada-Ramentol V, D'Andrea F, Alberi L. Intercepting dementia: Awareness and innovation as key tools. Frontiers in Aging Neuroscience. 2021 Oct 13;13:730727. https://www.frontiersin.org/articles/10.3389/fnagi.2021.730727/full

4- Gomez-Caceres B, Cano-Lopez I, Alino M, Puig-Perez S. Effectiveness of virtual reality-based neuropsychological interventions in improving cognitive functioning in patients with mild cognitive impairment: A systematic review and meta-analysis. The Clinical Neuropsychologist. 2023 Oct 3;37(7):1337-70. https://doi.org/10.1080/13854046.2022.2148283

10 Figures

Figures 1-4: VR cognitive exercises: Patients seated and standing, engaging in interactive content.

Figures 5 and 6: 360 video VR session: Exploring Badalona´s seaside virtually.



Figures 6 and 7: Initial VR supermarket task design posed safety concerns for seniors bending down to place objects in the cart. Responsive redesign introduced a basket on a table, enhancing safety and immersion, showcasing VR´s adaptability for senior-friendly therapeutic environments.

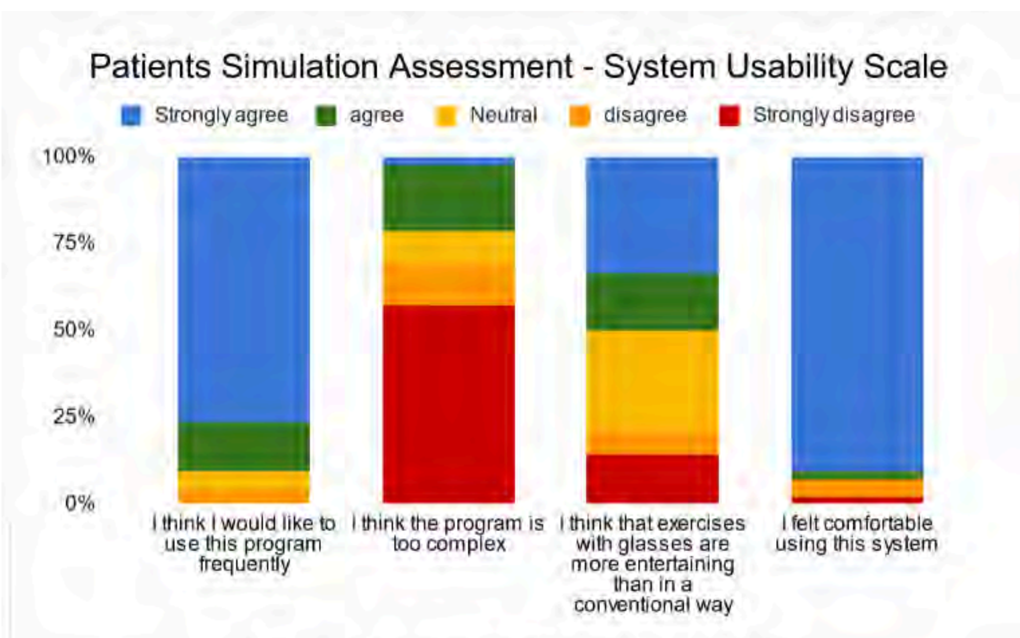


Figure 8: Participants' Usability Evaluation

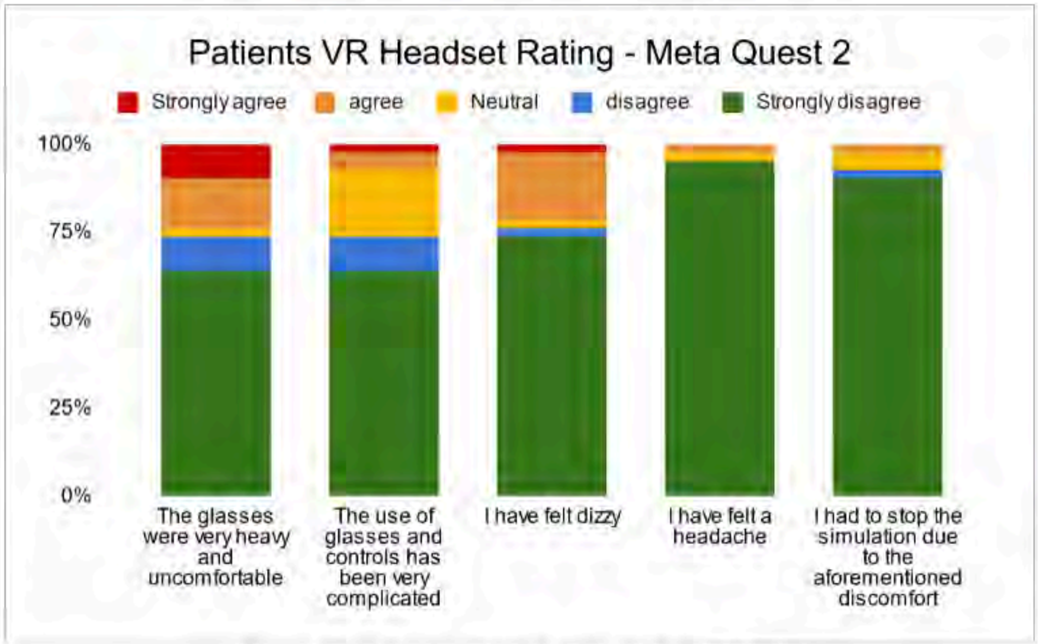


Figure 9: VR Headset Comfort Feedback

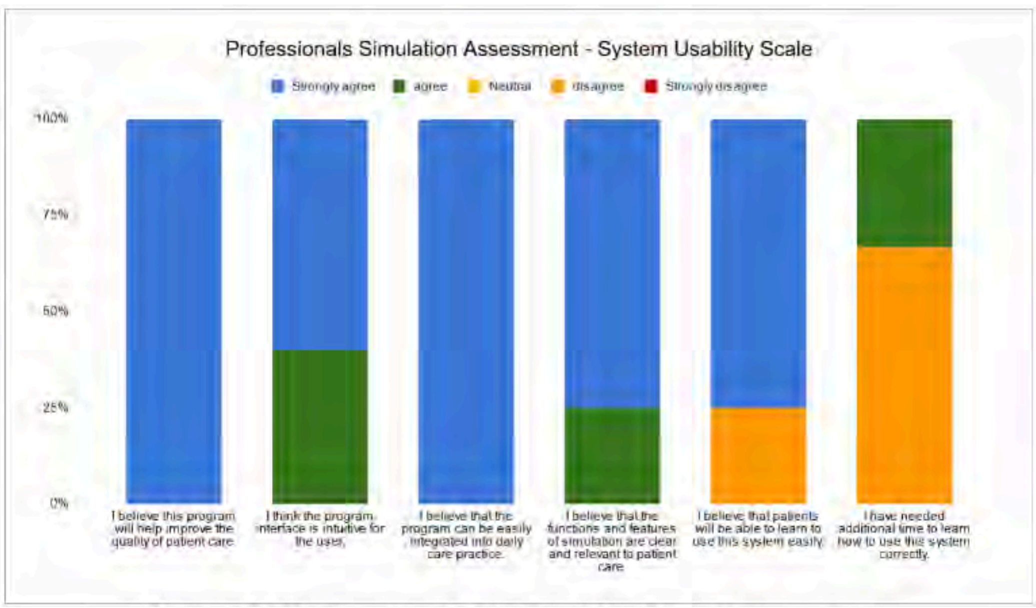


Figure 10: Professionals' System Usability Assessment

Spatial features in human social interaction: emerging spatial patterns in online gatherings with spatial affordances

Shachar Maidenbaum[1,2]

[1]Spatial Reality lab, Department of Biomedical Engineering, Ben-Gurion University, Be'er Sheva, Israel

[2] Spatial Reality lab, Zelman School of Brain Sciences, Ben-Gurion University, Be'er Sheva, Israel

*** Correspondence:**
Shachar maidenbaum
mshachar@bgu.ac.il

Keywords: remote communication, spatial cognition, virtual reality, academic conference, online interaction, virtual interaction

Abstract

Human social interaction naturally includes spatial elements. For example, we can approach people of interest or move away from others, we can turn to talk to specific people, talking volume is modulated by direction and distance etc. . These aspects are unfortunately limited during most remote interaction as video-based tools (e.g. Zoom) do not offer spatial affordances. We suggest that the lack of these spatial affordances may underlie many of the challenges and frustrations users have with these platforms. As different potential metaverse platforms do offer such affordances we asked whether participants would naturally utilize them, and if we would see the emergence of spatial behavior patterns.

To do so, we decided to focus on a minimalistic meeting tool - proximity chats which enable users to control avatars in a virtual environment. We chose academic conferences as our use case, and specifically focused on virtual academic poster sessions where spatial affordances may be especially critical. Thus, we ask, as you look around you in the current poster session, wil people behave similarly from a spatial aspect in a virtual setting?

We observed several poster sessions of medium sized crowds (N=200-300 participants per session, M=6 sessions), and test: (1) users' subjective preference and comparison to zoom based poster sessions (2) spatial behavior emerging during the use of these platforms.

We find that users were highly enthusiastic about the platforms which included spatial features, and we show the emergence of a series of spatial behavioral patterns which exist in real world poster sessions but are missing in video-conferencing based platforms. This natural utilization and emergence of these patterns, coupled with the enthusiastic user response highlights the importance of spatial affordances on the one hand, and the potential of spatial virtual environments on the other. Tracking these spatial patterns also offers in turn insights into the design of virtual spaces and types of spatial social interactions.

Towards a recommendation for interactive test methods for subjective assessment of extended reality communications

Jesús Gutiérrez [1,*], **Pablo Pérez** [2]

[1] *Grupo de Tratamiento de Imágenes, Universidad Politécnica de Madrid, Spain**
[2] *Nokia XR Lab, Madrid, Spain*

Correspondence*:
Corresponding Author
jesus.gutierrez@upm.es

ABSTRACT

Traditionally, the evaluation of communication systems has been mainly performed following standard methodologies described in recommendations by the International Telecommunication Union (ITU). These recommendations provide useful guidelines on assessment methodologies, how to gather the user opinions and measures, number of required participants to obtain robust conclusions, etc. Examples of these recommendations are the ITU-T P.920 (ITU-T, 2000), which covers the subjective evaluation of video-conferencing systems, or the ITU-T P.1301 (ITU-T , 2017), which addresses the subjective quality evaluation of audio and audiovisual multiparty tele-meetings. However, there is still no recommendation providing guidelines on the evaluation of eXtended Reality (XR) communication systems, which it's becoming essential given the recent emergence of numerous immersive approaches (Pérez et al., 2022).

In this sense, the Immersive Media Group (IMG) of the Video Quality Experts Group (VQEG)[1], in collaboration with the ITU-T Study Group 12, is currently working on a joint test plan among 14 international laboratories worldwide, both from academia and industry, which aims to provide results that can support a new recommendation for interactive test methods for subjective assessment of XR communications (ITU-T, 2023). In particular, this activity will provide a methodology to describe the test design (e.g., which system influencing factor to test, how to control context and human influencing factors, which user experience constituents to address, etc.), which should cover two types of designs: 1) systematically control a technical factor of the system and observe the effects on the user experience, and 2) test complete "blackbox" systems without exploring individual variables. In addition, a reduced set of communication-based interactive tasks that are suitable for testing XR in communication systems will be validated, focusing on audio communication, visual communication, object manipulation, and environment exploration. Finally, a subset of relevant measures (e.g., questionnaires regarding perceptual quality and user experience constituents, physiological measures, behaviour analysis, etc.) will be proposed that can be applicable to a wide range of use cases and systems. This poster will provide an overview of the status of the test plan and the details of the experiments being performed. Tentatively, the recommendation is expected to be published in the beginning of 2025.

Keywords: extended reality, subjective assessment, immersive, communications, recommendation

*This work was supported by the European Union's HORIZON Research and Innovation Programme through Project "XRECO: XR mEdia eCOsystem" under Grant 101070250 and through Project PID2020-115132RB (SARAOS) funded by MCIN/AEI/10.13039/501100011033 of the Spanish Government.

[1] https://vqeg.org/projects/immersive-media-group/ - The authors are co-chairs of VQEG-IMG.

REFERENCES

ITU-T (2017). Rec. P.1301: Subjective quality evaluation of audio and audiovisual multiparty telemeetings

ITU-T (2000). Rec P.920: Interactive test methods for audiovisual communications

ITU-T (2023). Work item p.ixc "interactive test methods for subjective assessment of extended reality communications

Pérez, P., Gonzalez-Sosa, E., Gutiérrez, J., and García, N. (2022). Emerging immersive communication systems: Overview, taxonomy, and good practices for qoe assessment. *Frontiers in Signal Processing* 2. doi:10.3389/frsip.2022.917684

Developing an innovative VR solution for anxiety/stress relief using a co-design process with students

Rachel Simnett[1], Zillah Watson[*1], Dr Aileen O'Brien[2], Katie Grayson[1], Ursula James[1], Russell Mulchansingh[1]

[1] Phase Space Ltd, London, UK
[2] St George's, University of London

*** Correspondence:**
Corresponding Author
zillah@phasespace.co.uk

Keywords: virtual reality[1], hypnotherapy[2], UX[3], therapy[4]

Abstract

Introduction

Phase Space VR is a unique solution for the relief of stress and anxiety which has been developed using a co-design process with medical students. The solution was initially created by an interdisciplinary team of experts from the fields of hypnotherapy, VR production and UX. It aims to combine the affordances of immersive technology with proven therapeutic techniques to create a highly effective solution. It was also vital that the VR experience was easy to use, enjoyable and accessible. This poster will examine the design and co-design processes adopted.

CO-DESIGN PROCESS

Phase Space partnered with St George's University of London (SGUL) in this co-design phase. Working with medical students, the VR was then developed and refined through a series of research activities including 3 design 'sprints'. A qualitative research methodology was used, with facilitated, one-on-one sessions that combined observation with a semi-structured interview. The sessions included exploratory questions around context and appetite for the VR as well as focused usability-based questions around the UX of the solution. This flexible approach enabled a robust data set to be collected whilst also allowing the sessions to be participant-led. 12 students took part in each round of testing. After each round, participants' feedback was analyzed and then discussed with the whole team. Design refinements were then agreed and implemented for the next round.

RESULT

Following this co-design process, the effect of the VR experience was assessed in in a randomised controlled trial of 100 SGUL students prior to their exams (O'Brien et al 2024). The VR experience has also been piloted for real world evaluation in schools, healthcare and corporate workplaces around the UK. This has enabled more data to be gathered to feed into the design of the implementation phase.

DISCUSSION

Is the failure to include effective design processes and user input in the development of therapeutic VR holding back adoption? Is the failure to include effective design processes and user input in the development of therapeutic VR holding back adoption? What challenges arise in university/industry partnerships and how can they be overcome?

A Metaverse for a menstruating body - proposal for research

Mariachiara Tirinzoni

Independent researcher, Milano, Italy

*** Correspondence:**
Corresponding Author
mariachiara.tirinzoni@gmail.com

VR, XR, Metaverse, Women Health, Period awareness

Abstract

XR experiences are used with remarkable results to help patients on different levels, but it seems to be lacking a focus on period health and menstrual well-being, both for awareness purposes and for treatment. I believe that creating metaverse experiences to inform and to offer an experience that allows user to life these conditions can help developing awareness, thus attention, thus more focus on the research on these topics.

Premises

We already have seen many applications of XR for human health, both mental and physical. However, the specific application field of women's health, the examples are quite limited, especially if we look into women's pain management examples. According to a review by Dr. Payal Ghatnekar - Immersive Technology Research Development Lead, Digital Futures Lab, Torbay and South Devon NHS Foundation Trust (2023), the focus of VR on women health has been on pregnancy, fitness and ovarian/breast cancer, with rare applications for other conditions and experiences. This is often linked to a severe lack of awareness about those experiences, together with a focus on gender related medicine that is still far from being fully developed: and that's what makes them basically invisible, with grave delays in diagnosis and overall lack of treatment of discomforts that could have been addressed otherwise.

Proposal

Since XR experiences are characterised by a higher level of emotional involvement, and they are demonstrated to create an environment for the mind where the user can feel "in other people shoes" much easier that through other kinds of narrative experiences, I propose to to create a set of metaverse experiences that help raise awareness on:

- how the management of menstruations impact the day to day practical life of a woman, both at home and at work
- Awareness on periods that are abnormally painful, individuating pain points both as a MD and as a patient, to make them fully aware that their condition is not "normal", thus not able to get a diagnosis i.e. of endometriosis in time.
- long -term/ chronic pain/ discomfort management for conditions like endometriosis, menopause, PMDD.

Embodiment in Virtual Reality: A Classical Sociological Theory Perspective

Maria Lipińska[1*], Renata Włoch[2]

[1]Department of Sociology, University of Warsaw, Warsaw, Poland

[2]Department of Sociology, University of Warsaw, Warsaw, Poland

*** Correspondence:**
ml.lipinska@student.uw.edu,pl

Keywords: virtual reality1, embodiment2, body3, metaverse4, sociology5.

Abstract

Virtual Reality (VR) has emerged as a powerful tool for immersive experiences, enabling users to inhabit digital environments and interact with virtual entities. In this paper the concept of embodiment in VR is presented from the perspective of five classic sociological theories (Marx, Durkheim, Weber, Simmel and Merleau-Ponty). We advocate for the applicability of these theories to enhance the sociological framework, aiming to bridge research gaps in the discipline. This is demonstrated by employing these theories to briefly analyze a contentious incident of virtual reality (VR) rape, thereby illustrating their practical utility in understanding complex social phenomena.

1 Introduction

Virtual reality (VR) has become a new technological dream of those creating computer-generated worlds (Sherman and Craig 2003). Virtual reality allows users to embody various avatars and interact with others in computer-generated worlds, creating slightly new social relations and structures. Thus, in the metaverse, it is possible to empirically explore virtual spaces and re-experience corporeality.

From the sociological perspective, the most important dilemma connected to embodiment in the metaverse is whether users create new social realities and norms or replicate the existing ones (Barricelli et al. 2016). The rules of virtual reality are significantly different from those of the physical world. If we want to speak in the metaverse, we must first turn on our microphone to be heard by others. We can also mute other users or even block their presence so that they are not visible to us as we travel through the virtual world. According to some researchers, this creates problems regarding the differences between the needs of the avatar and the biological body (Blackwell et al. 2019; Uspenski and Guga 2022). Such technological possibilities are entangled with both opportunities and risks. One of the latter is the case of sexual exploitation and harm in metaverse: in January 2024 (Sales 2024), the UK's police started the investigation in the case of a virtual gang-rape on underage girl. Although virtual rapes do not leave physical scars, the media reports suggest that the victim may experience psychological problems (Horne 2023).

One of the primary challenges in the research on VR arises from the lack of ontological precision regarding the nature of virtual beings and body boundaries. Thus far, researchers have struggled to establish a universally accepted definition for key terms within this research field, highlighting a gap in both empirical and theoretical methods, understanding regarding metaverse-related terminology and the problematic state of non-tangible body experience (Kilteni et al. 2012; Sampson 2019; Uspenski and Guga 2022). The aim of this study is to explore the concept of embodied self in VR, and create a sociological framework for researching embodiment, which will enhance the understanding of how individuals interact with their virtual bodies and other avatars. Moreover, such methods will develop the theoretical inputs as well as the social implications of embodiment in VR.

2 Methods

Digital sociology, being a constantly evolving discipline, routinely explores traditional theories in order to reinterpret social changes caused by technological progress (Zuboff 2019). Thus, following the pragmatic use of sociological theories (Selwyn 2019), we decided to study five classical theories - Marx, Weber, Durkheim, Simmel, Merleau-Ponty - in order to explain the embodiment in VR from the perspective of sociology.

Our starting point was to critically review the theories related to the body and embodiment as presented in the now-classic book by Schilling (2005). We agree with Shilling (2001), that traditional sociological approaches are limited in explaining the multidimensional aspects of embodiment. Thus, in this study we applied Critical Literature Review (CLR), including research outcomes from the analysis of five main sociological perspectives on the body (see: Table 1). The primary research problem is articulated through the following questions: In what ways can classical sociological theories be applied to the context of embodiment in virtual reality? What theoretical and methodological opportunities do they unveil for analyzing the domain of technology-mediated experiences of the body and embodiment?

3 Embodiment in VR from five sociological perspectives

Table 1 here

Embodiment in sociology is entangled with a broad spectrum of interpretations. According to Shilling, it is a relationship between body and social structure as well as a *medium for the constitution of society which has at its centre a concern with human experience* (Shilling 2001, p.327). Shilling's understanding of the term is based on the convergence theory based on conclusions from works of Marx, Durkheim and Simmel in which *body possesses properties that are a source for the creation of social life* and plays a role in locating the structural properties of society (Shilling 2005, p.10). Furthermore, this theory is enriched by other sociological theorists, such as: Haraway, Weber, Turner and Merleu-Ponty. The main aspects of Shilling's analysis are: the role of embodiment in social structures and power structures controlling bodies.

In the first criterion, the body is a location for various social features: economy (Marx), cultures with its rituals and symbols (Durkheim) and social forms (Simmel). The main similarity between Simmel's and Durkheim's approaches is treating the body as a crucial medium for constructing the society and as the main source of social symbols (Shilling 2001). According to Simmel, the body is a source of dispositions, forming the social structures, which could evoke socially binding emotions in individuals. Merlau-Ponty claimed that the body is one of the means of communication, which enables us to experience the world. This experience is possible due to sensory perception, creating visions of the surrounding environment (Shilling 2005). On the contrary, Weber sees embodiment through the perspective of culture. Moreover, in his view mechanized capitalism - involving technological tools, is creating an 'iron cage' for the body, which limits our biological mechanisms. The body in classical theories is not perceived through the perspective of interaction with technology. Nevertheless, those insights can be reinterpreted in the perspective of embodiment in VR.

Let's apply the theories outlined in Table 1 to the case of rape in virtual reality (VR). The concept that virtual embodiment influences psychological responses to traumatic events aligns well with the theories of Simmel and Merleau-Ponty. By the embodiment in VR through avatar, we are placed in a completely new environment, therefore self is situated in a new, unknown reality, where social rules are not clear. Virtual bodies determine our limitations, location and possibilities in the metaverse, and it is essential for entering it. The need of creating virtual spaces might be caused by the need of controlling matter. An avatar acts as an extension of the physical body, conveying our senses into the metaverse, potentially intensifying experiences of abuse, in line with the principles of embodied cognition. Moreover, although biological reactions to violence or brutal attacks are somewhat automatic, the technology-mediated environment disrupts defense and escape reflexes, a phenomenon highlighted by Weber's theory. In this context, the emotional impact on the victim of a virtual rape is no less intense or valid than that of a physical assault.

This situation also sheds light on the differences between social norms in the metaverse and the physical world. Virtual rape can be seen as a manifestation of the absence of social norms and a failure of the virtual social structure to prevent abuse, echoing Durkheim's perspective on social order and its effects on individuals. Additionally, applying Marxist theory, we recognize that major tech companies bear responsibility for creating safe spaces for users and addressing virtual abuses. This preliminary analysis suggests that classical sociological theories offer valuable insights and explanatory power for researching embodiment in virtual reality, highlighting their relevance and applicability.

4 Discussion & Conclusions

The embodiment in VR can be fruitfully addressed from different theoretical perspectives. Most of the mentioned approaches focus on the relation between virtual body and the society, as well as between virtual and physical body. In the context of embodiment in VR, the most accurate understanding seems to be Simmel’s and Merleau-Ponty’s approaches, because they emphasize the role of the body in creating social experiences. Moreover, the example of embodiment in VR provides an opportunity to explore the complex relationship between body and technology, as well as the influence of the virtual body on self.

In our view, the classical theories are providing valuable perspectives on the embodiment process, allowing the sociological theory to smoothly integrate more recent approaches, which suggest that the embodiment in VR can be interpreted with the frame of a cyborg without divisions between physical and virtual (Haraway 2016), or as a socially constructed body (Shilling 2005).

Tables

Table 1. Classical sociological theories implemented to embodiment in VR.

Author	How is embodiment presented in this theory?	How this theory can be implemented to embodiment in VR?
Karl Marx	Embodiment is immersed in the capitalist dynamic of exploitation and oppression. Through structural domination, capitalism alienates embodied individuals from their practical activities	Embodiment is determined by capitalist practices unfolding within the virtual space, where much depends on the management of digital infrastructure by big tech companies.

	and denies them all but their most basic needs. Body is entangled with economic structures through consumption practices.	In parallel to buying goods in physical reality, users signal their status in VR social communities through purchase of non-material clothes and avatars.
Max Weber	Embodiment is a cultural practice where the body is embedded within a specific social context and culture. To analyze embodiment effectively, it is crucial to consider the diverse and contextual interpretations provided by various actors. Additionally, the influence of institutions, like capitalism, must be acknowledged for their role in establishing rationalized norms (referred to as the 'iron cage') that shape the social activities associated with the body.	Applying Weber's concept of interpretive understanding, or verstehen, to virtual reality (VR) enhances the analysis of users' perceptions and experiences regarding embodiment, identity expression, emotional expression, and social connection through avatars. To investigate user behavior in the metaverse effectively, it is essential first to comprehend our embodiment in the physical world and then merge this understanding with virtual embodiment concepts.
Émile Durkheim	Societal structures and collective consciousness - social facts - shape the individual's experiences and understandings, including those related to the body. Body is entangled with culture, including rituals and symbols. A man/woman is always a homo duplex, integrating biological drives and social needs within one body. The behavior and representation of the self through body is guided by social norms, as exemplified by the suicide study.	Social representations of embodiment may impact the way the body is experienced in VR. In social VR apps users create their rituals and social norms, symbols and ways of communication, which might replicate or diverge from these from physical reality. The social norms dictate the ways virtual bodies are being treated in VR. Without an avatar we can't enter the metaverse.
Georg Simmel	The body is foundational in constructing social structures, as it elicits emotions that foster social bonds. By transcending bodily limitations and exerting control over the material world, individuals enhance their sense of agency.	In the Simmel's lens embodiment can be seen as a part of constantly produced virtual sociality, since interactions are fundamental for creating communities in the metaverse. The process of embodiment in VR is another act of controlling matter. In the metaverse, users are able to create new worlds with their own preferences and rules.

Maurice Merleau-Ponty	The body serves as our primary means to interact with and explore the world; hence, all cognition is inherently embodied. Our capacities for thought, reasoning, and understanding are deeply connected to our physical experiences. Individuals experience embodiment through the body schema, which informs their spatial orientation and interaction with the environment. Technology, or tools, act as extensions of our bodies, further enhancing our interaction with the world. Thus, the body is not merely a passive object but a conscious subject actively engaging with its surroundings.	Our engagement with technology, tools, and other extensions of the body is a natural expansion of our bodily capabilities, further integrating our physical selves with the world around us. Thus, avatars and embodiment in VR are the extensions of our physical body and enable us to explore the metaverse. Experience in VR must be embodied to be fully immersive. Avatar places our virtual existence in space and virtual world.

5 Author Contributions

Maria Lipińska: main idea, conceptualization, theoretical framework, text preparation; the first author is responsible for 75% of work twith the article.

Renata Włoch: conceptualization, theoretical framework, editing.

6 References

Barricelli BR, Gadia D, Rizzi A, Marini DLR (2016) Semiotics of virtual reality as a communication process. Behaviour & Information Technology 35(11):879–896. https://doi.org/10.1080/0144929X.2016.1212092

Blackwell L, Ellison N, Elliott-Deflo N, Schwartz R (2019) Harassment in Social Virtual Reality: Challenges for Platform Governance. Proc ACM Hum-Comput Interact 3(CSCW):1–25. https://doi.org/10.1145/3359202

Haraway A (2016) Manifest cyborgów

Horne C (2023) REGULATING RAPE WITHIN THE VIRTUAL WORLD

Kilteni K, Groten R, Slater M (2012) The Sense of Embodiment in Virtual Reality. Presence: Teleoperators and Virtual Environments 21(4):373–387. https://doi.org/10.1162/PRES_a_00124

Sales NJ (2024) A girl was allegedly raped in the metaverse. Is this the beginning of a dark new future?

Sampson TD (2019) Transitions in human–computer interaction: from data embodiment to experience capitalism. AI & Soc 34(4):835–845. https://doi.org/10.1007/s00146-018-0822-z

Selwyn N (2019) What is Digital Sociology? Wiley

Sherman WR, Craig AB (2003) Understanding Virtual Reality

Shilling C (2001) Embodiment, Experience and Theory: In Defence of the Sociological Tradition. The Sociological Review 49(3):327–344. https://doi.org/10.1111/1467-954X.00335

Shilling C (2005) The Body in Culture, Technology and Society

Uspenski I, Guga J (2022) Embodying Metaverse as artificial life: At the intersection of media and 4E cognition theories. Filoz drus 33(2):326–345. https://doi.org/10.2298/FID2202326U

Underwater virtual reality for marine education and ocean literacy: technological and psychological potentials

Géraldine Fauville[1], Anaïs Voşki[2], Marijn Mado[3], Jeremy N. Bailenson[3], Annika Lantz-Andersson[1]

[1]Department of Education, Communication and Learning, University of Gothenburg, Sweden
[2]Doerr School of Sustainability, Stanford University, USA
[3]Department of Communication, Stanford University, USA

Given the crucial role the ocean plays in human and planetary health, ocean literacy—understanding the ocean's influence on us and our influence on the ocean—has been gaining momentum as a significant component and determinant of sustainable human-ocean interactions. First-hand experience in nature is essential in environmental education to promote learning and to develop care, understanding and interest in conservation. However, experiencing the ocean is challenging because of its remote nature. Immersive technologies can provide these encounters with nature and contribute to enhancing environmental awareness and pro-environmental behavior. Underwater Virtual Reality (UVR) constitutes the next frontier in immersive technology for exploring the ocean as it enables *double immersion*–being immersed visually in a 360-degree video of a dive in a VR headset, while being physically immersed in water. Through this enhanced level of immersion, UVR is offering new avenues of psychological access and pedagogical opportunities toward the ocean. This study is thus the first to explore UVR's potential for ocean literacy and marine environmental education.

We conducted semi-structured interviews and survey questionnaires with 19 marine scientists and education experts. Each participant experienced the same two 5-minute long, 360-degree computer graphic imagery videos displaying an underwater dive. In these videos, the users encounter, among other things, sharks, whales, and coral reef. Reflexive thematic analysis was employed to analysis the content of the interviews.

Our findings indicate UVR has important technological and psychological potentials. First, participants often described experiencing high level of presence as "*not being in the pool with a chain around the waist, but actually experiencing the ocean*". Second, motion sickness seemed infrequent with, for example, a participant who "*Had nothing at all*" in UVR despite a tendency to experience it in VR. Third, UVR also triggered positive affect, especially awe, empathy, and flow, which can provide the emotional basis for improving ocean literacy and reduce psychological distance to marine environmental issues. We also discuss UVR's current limitations, especially its accessibility and technological challenges, which must be addressed for the purposes of scaling these preliminary but promising findings. Based on these findings revealing unique potentials of UVR, we propose a new seven-dimensional ocean literacy framework including the dimensions of emotions, attitude, awareness, ocean connectedness, behavior, communication, knowledge, and find UVR to be uniquely advantageous and most promising for the first four. This study has been published in *Environmental Education Research*.

European Projects

GuestXR: A Machine Learning Agent for Social Harmony in eXtended Reality

Simona Neri[1*]

[1]Eurecat Centre Tecnòlogic de Catalunya

* simona.neri@eurecat.org

Keywords: virtual reality, machine learning, artificial intelligence, neuroscience, social psychology. (Min.5-Max. 8)

Abstract

Immersive social spaces are increasingly being adopted on a wide scale, but it is important to exercise caution when applying them within the framework of social media. While user-generated content fuels these spaces, it often ignites negative interactions, bullying or abuse, posing risks to vulnerable individuals. This pattern extends to virtual reality (VR) environments, where despite their promise, instances of misconduct are prevalent. GuestXR is a socially interactive multisensory platform that uses eXtended Reality as the medium to bring people together for immersive, synchronous interaction with positive social interactions. GuestXR employs artificial agents, powered by machine learning, to guide users towards constructive engagements. Drawing from neuroscience and social psychology, these agents are grounded in Agent-Based Modeling (ABM), fostering healthy group dynamics. GuestXR facilitates group meetings in extended reality, recognizing that each meeting has specific goals, from entertainment to conflict resolution. It addresses the prevalence of undesirable behaviors like discrimination and abuse, often fueled by anonymity. GuestXR employs an underlying agent to monitor and intervene in meeting dynamics, aiming to steer them towards positive outcomes, while minimizing negative behavior. The agent learns from its actions, adjusting its approach to maximize long-term rewards and increase the likelihood of successful outcomes.

At its core, GuestXR utilizes Reinforcement Learning (RL) to refine its algorithms. The project sets its foundation for optimizing interventions that align with meeting objectives, guided by real-time monitoring, conducting simulations rooted in social science principles, thereby providing a robust training environment for GuestXR agents. The project implements extended reality within GuestXR and assesses user responses during system operations while introducing multimodal haptics to improve the overall User Experience. The project undergoes testing and validation across four social-oriented Use Cases include a persistent virtual space in which multiple people can visit and discuss topical matters, the use of GuestXR by people with hearing disabilities, applications addressing climate change and conflict resolution, and an in-depth exploration of protests as both social assets and challenges.

The GuestXR project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101017884.

Sonicom: Transforming auditory-based social interaction and communication in AR/VR

Eva Fringi[1*], Thibault Vicente[2], Stephen Brewster[1], Lorenzo Picinali[2], Alessandro Vinciarelli[1]

[1]School of Computing Science, University of Glasgow, Glasgow, UK

[2]Dyson School of Design Engineering, Imperial College London, London, UK

*** Correspondence:**
Eva Fringi
Evangelia.Fringi@glasgow.ac.uk

Keywords: 3D audio, Personality Computing, Auditory Distance Perception, Social Signal Processing, Computational Paralinguistics.

Abstract

Immersive audio is our everyday experience of being able to hear and interact with sounds around us. The impact of immersive audio beyond perceptual metrics such as realism and localisation is still an unexplored area of research, specifically when related with social interaction, entering the behavioural and cognitive realms. SONICOM aims to revolutionise the way we interact socially within AR/VR environments and applications by leveraging Artificial Intelligence to design a new generation of immersive audio technologies and techniques, specifically looking at personalisation and customisation of the audio rendering. Amongst the themes that the project explores, is the interplay between social and interpersonal distance as perceived within a virtual immersive audio environment. More specifically, it is investigated how varying virtual proximity between speakers affects their judgement on personality trait attribution. A novel corpus comprising recorded speech from 120 subjects was collected and rendered at different virtual distances, which were annotated with respect to personality and distance perception. State-of-the-art deep learning methods were applied to perform automatic personality perception (APP) in the context of virtual audio, indicating that personality perception (at least for some traits) changes with the perceived physical distance from the speaker. To the best of our knowledge this is the first attempt at speech analysis focusing on virtual distance perception in conjunction with personality perception, and contributes to the development of a socially intelligent AR/VR environment which will cater for users' different social dynamics.

Acknowledgments

This project is supported by the Horizon2020 FET Proactive – Boosting emerging technologies (H2020-FETPROACT-2018-2020).

META-TOO: Investigating gender-based inappropriate social interactions in the Metaverse

Justyna Świdrak[1*], Maria V. SanchezVives[1], Ferran Argelaguet[2], Maria Roussou[3]

[1] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain

[2] Inria de l'Université de Rennes, IRISA, CNRS, Rennes, France

[3] Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Athens, Greece

*** Correspondence:**
Justyna Świdrak
swidrak@recerca.clinic.cat

Keywords: safe metaverse, cyber-harassment prevention, online misconduct reduction, responsible virtual interactions, virtual embodiment, knowledge transfer

Abstract

The goal of Metaverse is for people to socialise, work, and have fun, and platforms like ROBLOX and VR Chat prove its growing prominence and potential. In this digital sphere, avatars and their deep sense of embodiment raise concerns, especially with emerging reports of cyber-harassment. The META-TOO project, bridging VR/AR and behavioural research, collaborates with leading European institutions to tackle online misconduct in these virtual environments. It underscores the knowledge and skills transfer from Inria and IDIBAPS to the Project's Coordinator, the National and Kapodistrian University of Athens in Greece. With that aim, META-TOO's activities will include not only scientific research, but also a wide range of activities such as workshops, seminars, meetings, and networking. Scientifically, META-TOO will investigate inappropriate behaviours in Social VR, setting the stage for future research. It will identify physiological markers that hint at users sensing harassment, enriching our comprehension of human-computer interplay, and produce a series of tools to detect and cope with sexual (and other) harassment. Additionally, the Project will sensitise users to inappropriate behaviours by empathy and perspective-taking simulations. Its research activities will also include studies on bystander behaviours in the Metaverse as a source of user empowerment. Overall, META-TOO envisions a more inclusive digital domain, reducing the negative impacts of harassment and bolstering safety in the Metaverse. The three-year project will commence in June 2024 and is funded under the HORIZON-WIDERA-2023 (Twinning) call.

Ethical, legal and political implications of the Metaverse

Metaverse, implemented.
A philosophical perspective on agenda-setting ethical tools for the development of immersive technologies

Mariachiara Tirinzoni
Independent researcher, Milano, Italy

***Correspondence**:
Mariachiara Tirinzoni
mariachiara.tirinzoni@gmail.com

Keywords: VR, XR, Metaverse, Social networks, Ethics, Philosophy of science, Epistemology

Abstract

The specific application of immersive technologies known as 'metaverse' emerges as a potential substitute for current social networks. Given the social and anthropological effects we have been able to observe as a consequence of the mass adoption of social networks over the last 15 years, the following work aims to provide the initial sketch of a framework for structuring an approach to the design and deployment of the metaverse that safeguards the values of transparency, justice and fairness, avoiding simplistic but potentially distorting drifts already suggested by major actors. In particular, the reflection focuses on a shift of perspective that, drawing on the theory of affordances, sees an important role of ethical thinking not just in the design of the metaverse, but in its deployment and implementation too.

Article type: perspective

1. Introduction

The metaverse is not simply a connection and communication technology, but a concept, a theory, perhaps even a utopia that we have not yet really put into practice: a tool for interaction that enables a communication broader and deeper than the web 2.0 one, in the sense that it can carry more elements than just our online profile and our words, thus preserving at least in part the corporeality of the expression of ourselves without being restricted to a space defined by the laws of physics and biology, and ultimately to a precise and finite time, a world where we can "make memories that are impossible to create in real life"[1].

The objection that the metaverse is an 'artificial' environment and therefore 'it's not real life' is not tenable. As a matter of fact, judging from what we have seen to be the effects of the use of social media in the recent history, digital life, although it has its own characteristics and dynamics, although it is not 'physical' in the most concrete sense of our entire biological presence in the virtual environment, is *real life* to all intents and purposes, from the creation of relationships and networks, to the fruition of information, to its consequences on the political life of countries. I will also take it for granted that even in a virtual environment we may or may not be aligned with personal and social

[1] Rosa, 2023

ethical values. However, like any specific environment, the metaverse will offer itself for use with rules of its own, determined by its technological structure and the communities that inhabit it.

2. Would you harm someone in the metaverse?

Would you feel different about harm *because* you are in the metaverse? I am not questioning here the equivalence of value between a real human being and a virtual one, as long as the latter is endowed with consciousness and linked, as the term 'avatar' implies, to a real/physical/offline human being, we need, however, to question the neutrality of the digital means of communication.
If virtual life[2] were ethically neutral, the phenomenon we have come to know on social networks[3], in which people feel more 'free' to engage in offensive behaviour and modalities of communication, than they would in the real world (the phenomenon of 'keyboard warriors') should not occur: using the tool of online communication should not alter our behaviour. On the other hand, we know that such incidents do exist, not only on social media but in the metaverse too, and this suggests that digital media as they are designed have an impact not only, in this case, on our individual identities in a descriptive sense[4] but also on our 'ethical habits', just as the technologies that have become commonly used throughout history do. Why?

3. The metaverse: a non-neutral space-tool

I hypothesise that metaverse, as a technological tool, constitutes an artifact endowed with what, according to the phenomenological school[5] and then to psychological theories such as Gibson's[6], is defined as affordance, i.e. that it has the capacity, as an object designed to make its own use intuitive, to 'invite' so to speak to a specific action, to a specific use.
The metaverse, in my opinion, falls perfectly into this category, as an object designed from the ground up for general functions such as hosting of movement, fruition of contents and interactions with other users, all of this happening according to precise and defined UX, user experiences, together with specific functions, depending on the use that has been designed for it (e.g. a videogame, a marketing campaign, a place for interactive training, relaxation, mindfulness and so on). Exactly because of this it does not benefit from pure neutrality: I do not enter the metaverse to do 'whatever action I want', but with a set of actions that I will be more or less directed to carry out. Let's imagine for example a path designed to expose the user to the vision of a series of visual materials, where the user's range of movement is studied (the environment is designed to guide the user) to understand which ones attract the most attention. In short, the metaverse is an artifact that "projects around itself a script that can take possession of each onlooker and force them to play a role in the story by making only changes"[7].

But I also argue that the metaverse is a special kind of artifact, precisely an environment endowed by its very nature with affordances: not a place that simply hosts actions but also a place that serves to *make things happen* and that is designed and gets modified according to these instrumental functions, a space-tool.

---

[2] I am referring here to any form of online relationship, including social media.
[3] https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/
[4] Will the use of an avatar make me a different person? But this is not the subject of this paper.
[5] Pfänder, 1900
[6] Gibson, 1979
[7] Gibson, 1986

If the metaverse is a space-tool, then it has certain features that make it ethically non-neutral as a tool endowed with affordances, but also features that are specific to places that are inhabited or ready to be inhabited, and these two classes of features flow into each other: on the one hand, we see a place that influences the behaviour of those who inhabit it by presenting them with certain functions, but at the same time, as a manipulable space, it hosts and accommodates the modifications brought about by those who enter it, getting modified by them and thus recreating new conditions. To the point that it could be configured as a new habitat[8].
Think of the complete customisability of avatars, environments, and modes of movement, and the possibility of evading the laws of biology, the integrations with AI that determine almost endless types of interaction and the collection of behavioural data that can be gathered inside the virtual environments for the most diverse purposes.

The metaverse takes planning of the environment as a manipulation of the individual's experience to the extreme. And since manipulation is in turn the work of individuals, we end up being both human and to borrow Leibniz's words, the great clockmasters too.
This characteristic is also the crucial aspect whereby the metaverse departs significantly from social media: in the metaverse we recover, for better or worse, a 'spatialisation' of experience. Are we therefore more 'there', present, in the metaverse? How do we design this space?

4. A urbanology of the metaverse

We are thus faced with the challenge of building an ethical framework that resembles an urbanology of the metaverse than a more linear a priori regulation: the new social world we are going to inhabit virtually requires a structure that must *embody* the values we want to guarantee and protect, not just *be subject* to them, and it must do so at the very moment it is built and meets the user, that is, at the design stage, otherwise it will not be habitable.
It goes without saying that the first commandment of the ethical innovator's agenda concerns the manner in which the metaverse is designed. Since the metaverse is already inhabited, or rather is in the process of being populated, and populated in the way an ecosystem is populated: dynamically and in continuous change, due to the influence of those who inhabit/use it. Compared to what we have learnt (often painfully) from social media, there is a widespread feeling that we are still in time to draw some of the rules of the game before the metaverse enters fully the world and society, heavily influencing the life of individuals and democracies.
This is why we need, first of all, to speak about ethics by design, and not only ethics as a part of regulations and policies external to the tool itself.

As Schiaffonati argued recently[9], ethics is all too often only taken into account at the end of an innovation process, when the technology has proven to work as a product to be sold, when it has proven to have no technical flaws, when it has ensured that it can be adopted, used, exploited and thus can bring the expected economic benefits. On the contrary: we have no need for an ethics of 'running for cover', but for an approach intrinsic to the development of the technology itself. The framework I'm referring to is that of *ethics by design*: the design of the space-tool must contain within itself, just like the design of a city made for its inhabitants, streets, services and facilities in which values are 'affordable'; where it is not necessary to make additional efforts to 'live well'.

[8] Notte, 2002
[9] Bellini, Della Mura, 2023

For instance: where we identify a connection between the lack of corporeality of social media presence and the rise of ‘keyboard warrior’ attitudes, that is precisely the place where a design that enhances an individual presence as richly as possible could go in the direction of stimulating behaviour more akin to what we would have in person, and not behind the veil of pseudo-anonymity and detachment caused by a 2d approach, based on the isolated nature of words and almost entirely devoid of non-verbal languages.
Another case: an application engineered to autonomously gather and distribute personal data of its users, devoid of informed consent, breaches privacy in itself. If we incorporate privacy as a design feature, for example through explicit consent collection, we are ensuring that, at least as far as the tool is design, it does not “invite” (in the sense of affordance) the breach of privacy during its use.

A well-done metaverse “urban design” will not leave too many dark alleys, it will make the spaces for the use of common goods accessible, it will create the context and the right space (literally) for the flourishing of the individuals who inhabit it, studying their flows and habits: as in the adolescent question "what would you take with you on a desert island", it is important to understand what values we take with us into the metaverse[10], compared to those carried in 2d social networks and to those we make experience in the social ecosystems we enter every day.

5. Sketching the agenda for a responsible implementation

While an ethical design is very important to ensure that what we are putting into the world a technology that works according to our values, it is also clear that it’s not enough, since many challenges can come from the way which the tool is used, or better from the way we enter a given tool, with its affordances and its consequences in terms of influencing thinking, into the everyday world. In short, there are uses and use contexts where the value we are pursuing are not manifested or attained, not because of the tool design itself but because of the circumstances in which the tool is introduced into the environment, workplace, training, healthcare and virtually any other application field.

In the specific case of metaverse, we find ourselves in an advantaged position, since we are working exactly when first use cases and applications are appeariing, but larger adoption is still in the future.

This is why I would like to introduce an additional concept that can guide an ethical approach to the metaverse, collateral to that of ethics by design: I’m talking about responsibility in *implementation*.
Right of access, transparency of information and fairness of use depend in fact not only on the nature of the technological tool, its design and its goals, but also on the ways and circumstances in which the technology is translated into the world of human life and goals, and, on the other hand, on how humans, the users, are enabled to enter the metaverse, to make use of this space-tool.

I’m arguing that, if on one hand we have an artifact that ‘invites’ a certain use and behaviour, and on the other the subject who will make use of it, the very relationship between the two is determined by the implementation (and deployment) of the technology itself. That is, it is the implementation,

[10] It’s out of the scope of this work, but still important to underline that the current plague of cyberbulling in video-games should not be seen as a blocking objection to the pursuit of greater civility in the metaverse: instead, the mass of current research into the subject should be incorporated in the eethical design of the metaverse, in order to hampen the emergence of this phenomenon since the very beginning.

occurring when we (quite literally) put it in the hands of the users, that *activates* the latent relationship between subject and object. In turn, however, implementation is a technological act itself, which responds to the procedures and characteristics of the technology: we can see that there is a specific affordance that invites the implementer to act in a certain way. Also, in implementing a new tool inside a pre-existent structure (a business model, a way to live an experience), it is possible that we need the structure to change, in order to accommodate the implementation itself, and while this modification could be requested by the technology itself, its design, how this modification occurs depends on the implementation strategy.
For example, adopting a metaverse meeting solution in the project management routine of a company can affect the routine itself and the performance of the users: are we designing the implementation process to preserve the value of equal access, that is to accomodate diverse predispositions of our users, such as various impairments or simply lack of experience in moving through a virtual environment? Are we ensuring that everyone can contribute equally to the work done during the meeting or are we accepting epistemic injustice because not everyone has been enabled to use the medium of the virtual meeting at its full? An ethical approach to implementation will incorporate this issue in the process of adoption, not leaving the accessibility of the tool and its advantages to the design of the tool itself[11].

Implementation is defined as "the process of taking on a new software application and incorporating it into existing business workflows"[12]. Going back to analyse the affordance of it, we could witness to multiple 'invitations': the one that guides us to use the tool, the one to guides us in exploring the possibilities it offers, the one that dictates the way it interacts, for example, with the physical circumstances in which the user will need to be to enjoy the experience provided. In the case of a space-tool like the metaverse, it also means incorporating the possibilities of the virtual immersive reality into the existent reality, while entering a space we've never been to, using a tool we never used before, that nevertheless preserves some resemblances (thus affordances) with other tools and spaces: physical world and social media. Not just the possibilities: the consequences. What is going to happen when we add a new layer of reality through the implementation of the immersive experience of the metaverse? As we saw in the forementioned examples, should we attempt to make this process an ethical one, it is necessary to design it so that it responds not just to the characteristics of the technology itself (where this by design should be used), but also to the characteristics of the user, the circumstances of use and the values we wish to preserve.

That means that we are called to compare the affordances of this space-tool with the invites to action coming from the other tools and environments: is the metaverse inviting the same use or action? How can we adapt a completely manipulable tool-space, so that it invites us in a direction that protects

[11]Another example could be implementation of VR medical training: a great tool that allows anyone in any part of the world to access the training, in theory. In practice, however, it is not enough that the tool allows it: the conditions in which the metaverse can be correctly used for this purpose must be created and protected. For example, the implementation will need to secure a broadband network of sufficient power for the real-time experience and rendering.
A further question to reflect on is how and in which cases to define the "correct use" of a tool, a question that straddles the design of the object itself and the resulting affordance, and the construction of the implementation process.
[12] https://www.walkme.com/glossary/software-implementation/
Let me note that 'innovation' itself comes from the latin in-nova, a world that contains a strog reference to the fact that we are literally putting something new into the world: there's no way we can act like innovation processes are happening in an isolated lab. Think about the design process: there is always a testing phase, and when we put something new in the world we are testing on humanity - that's because technological innovation can inhabitate the core of our everyday life, it's not something that stays on the superficial layers of the world's skin.

values? First step of every implementation is going back to the pre-existing processes, translating them in a way that can bridge the previous and the new technologies.
So, the ethical implementation will ask us: which language are you using to 'write the instructions'? What effort are you making to leave a hyperspecialized language behind and avoid epistemic injustice, providing the common ground of terminology and competence to be aware of the circumstances in which the metaverse has been (and is being) developed?

For example, onboarding of the tool is a implementation problem: access to the metaverse has already been recognised as a counterargument to the metaverse itself. On the one hand because, no matter how well (and ethically) the immersive environment is designed, if implemented in a process where users are not trained to use VR headsets, or not used to moving their avatars, it can generate inequality of fruition not only of the tool itself, but also of the content of the immersive experience, let's imagine a VR meeting where some participants lose sight of the area where the meeting is taking place, hence losing part of the information shared, because of spatial audio or other sharing settings.
Another implementation issue concerns the lack of affordability and/or accessibility of VR headsets and other accessories[13]that can affect user participation and experience, because it can generate inequality whenever what is conveyed in the immersive experience has value or can generate value beyond the experience itself (e.g. training). Even when a metaverse is perfectly designed, it ought to be carefully decided under which conditions to activate its functions and in which specific contexts it is appropriate to use this tool or not, thus designing the implementation process in an ethical and just way. The same must be said of the choice of processes where metaverse could be implemented.

Furthermore, implementation refers to the improvement of a previous process: the ethical innovator will need to focus on the pre-existing needs and activities that can be actually improved thanks to the space-tool of metaverse and start out implementation from there, abandoning the progress for progress' sake mindset and thus freeing the process from itself, making it fully a part of our (virtual, digital, physical, hybrid) world.

And finally I want to mention one last extremely important point of ethical implementation, which I do not have the opportunity to examine in depth here: in the implementation preliminary analysis, it may become evident that not all processes need to be replaced by the new tools, and that implementation should therefore be avoided at some points. Only if we have the courage to ask ourselves how much metaverse we really need, and to answer the question even negatively, will we be able to make good use of it.

Conclusions

Only by taking into account both concepts, ethics in the design process and ethics in the implementation process, can we avoid a counterproductive "watchdog ethics" that risks arriving late, not grasping and thus not acting on critical issues as they arise, if not anticipating them by altering the very affordance of the metaverse space-tool. The risk here is also that a 'outsider ethics' can ultimately act as an antagonist to the process of the technological innovation itself, which must instead be accompanied and guided not only by the curiosity of the explorer and the business logic of the trader, but also by the values of the human being who inhabits, shapes and allows himself to evolve from the new world they have created.

---

[13] Pospiech, 2022

Acknowledgments

Many heartfelt thanks to Tommaso W. Bertolotti, PhD, for the proficous exchange of ideas and perspectives, and for the invaluable support during all the research progression. My gratitude also goes to Isabel De Peuter-Rutten, president of Euromersive, for her support and encouragement and for giving me the stage to discuss my ideass publicly for the first time in front of AWE Europe public.

References

1. VV. AA. *The internet does not exist*. London: Sternberg press, 2015
2. Bellini, Mauro, Della Mura, Maria Teresa. *Umano digitale. Verso un'etica dell'innovazione.* Milano: In Dialogo, 2023
3. Brey, P., Dainow, B. *Ethics by design for artificial intelligence. AI Ethics*, 2023. https://doi.org/10.1007/s43681-023-00330-4
4. Chalmers, David. *Reality+: Virtual Worlds and the Problems of Philosophy*. London: Allen Lane 2022
5. De Monticelli, Roberta, Conni, Carlo. *Ontologia del nuovo Ontologia del nuovo: la rivoluzione fenomenologica e la ricerca oggi.* Milano: Bruno Mondadori, 2008
6. Fasoli, Marco, Piredda, Giulia. *Filosofia, Tecnologia e scienze della mente.* Bologna: Il Mulino, 2023
7. Gibson, James J.
    a. *The Theory of Affordances. The Ecological Approach to Visual Perception.* Boulder: Taylor & Francis, 1979
    b. *The ecological approach to visual perception.* New York: Psychology Press, 1986
8. Husserl, Edmund, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Erstes Buch: Allgemeine Einführung in die reine Phänomenologie*, in *Jahrbuch für Philosophie und phänomenologische Forschung* Max Niemeyer, Halle a. d. S., 1913
9. Latour, Bruno, *La clef de Berlin et autres leçons d'un amateur de sciences*, La Découverte, Paris 1993
10. Latour, Bruno, *A Cautious Prometheus? A few steps towards a philosophy of design*, Keynote Lecture, History of Design Society, Falmouth, 3rd September 2008
11. Merleau-Ponty, Maurice, Landes, Donald A. (translated by). *Phenomenology of Perception.* London: Routledge, 2013
12. Notte, Riccardo. *La condizione connettiva. Filosofia e antropologia del metaverso.* Roma: Bulzoni, 2002.

frontiers

13. Pew research Center. *The state of online harassment.* Updated January 2021. https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/

14. Pfänder, Alexander. *Phänomenologie des Wollens: eine psychologische Analyse; Motive und Motivation*. Leipzig: Johann Ambrosius Barth, 1900

15. Pospiech, L. (2022). *10 Metaverse marketing challenges for affiliate marketers and brands.* Retrieved 19 August 2022, from https://zeropark.com/blog/metaverse-marketing-challenges/

16. Rosa, Nick, Furness, Tom (Foreword by). *Understanding the Metaverse: A Business and Ethical Guide.* Hoboken: Wiley, 2022

17. Walkme. *What is software implementation?* Updated October 2023. https://www.walkme.com/glossary/software-implementation/

What do policymakers need to know about harassment in the metaverse?

Verity McIntosh[1,2]*

[1] Bristol VR Lab, University of the West of England, Bristol, United Kingdom
[2] Digital Cultures Research Centre, University of the West of England, Bristol, United Kingdom

*** Correspondence:** Corresponding Author Verity McIntosh, verity.mcintosh@uwe.ac.uk

Keywords: metaverse, governance, abuse, harassment, policymaking.

Abstract

As immersive technologies and spatial computing paradigms move into the mainstream, public and political interest in the metaverse is growing. In some respects, the metaverse offers an exciting view of the future, one in which a global community can meaningfully connect regardless of where they are in the world. In contrast, however, early instances of 'proto-metaverse' spaces have been plagued by reports of harassment and abuse.

Policymakers around the world are now considering the role that governments might play in the regulation and governance of metaverse spaces, seeking to secure protections for citizens, and criminal accountability for offenders in this fast-evolving space.

This paper introduces some of the key issues for governments engaging with this topic, including the suitability of existing legislative frameworks, and consideration of a new category of harm that seeks to recognise the distinctive impact of 'conduct' abuses in metaverse environments.

1 Definition of terms

There is yet to form one coherent definition of the term 'metaverse' and there remains much debate about what is inferred by the term and how it is applied. This paper utilises the X Reality Safety Intelligence (XRSI) definition of the metaverse as:

"A network of interconnected virtual worlds with the following key characteristics: Presence, Persistence, Immersion and Interoperability" (XRSI, 2023)

Forms of harassment discussed in this paper are generally limited to behavioural activity occurring in real time in what might be considered a metaverse, or proto-metaverse environment. This may take the form of verbal or gestural abuse, and/or the use of embodied avatars and virtual objects against other users to enact behaviours experienced as aggressive, violating, offensive or demeaning. An understanding of harassment and abuse could reasonably be extended to include areas such as data and privacy abuses, identity cloning, social and political manipulation, fraud, theft and exploitation. For the purpose of clarity, this paper will focus solely on real time, peer-to-peer encounters involving one or more natural persons in a virtual environment. It should be noted, however that instances of harassment in metaverse contexts may form part of a wider pattern of abuse, taking place both on- and offline and should be considered in such a context when abuses are reported.

2 Harassment and abuse

In recent years, occurrences of harassment and abuse within proto-metaverse platforms, sometimes referred to as social VR platforms, have been well documented in the media, perhaps less thoroughly explored in a scholarly context, and in relation to the role and obligations that governments may have to intervene in this space.

Evidence suggests that instances of harassment tend to increase in virtual environments devoid of managed hosting or a clear purpose, with female users and minoritized people most likely to be targeted (Limina Immersive, 2018). A survey of over 600+ users in 2018 suggested that 49% of regular female VR users reported experiences of sexual harassment or abuse in virtual social spaces (Outlaw, 2018). Since then, with the rise in public adoption of VR headsets, the issue appears to have persisted and perhaps escalated. In 2021 the Center for Countering Digital Hate asserted that users of popular social VR platform, VRChat were exposed to abusive behaviour once every seven minutes (Center for Countering Digital Hate, 2021). Numerous reports of sexual harassment and abuse within the metaverse have been reported in the media (Eccles, 2022; Patel, 2021; Rifkind, 2022).

3 Impact

Although the nature of harassment and abuse in VR differs from real-world instances, the impact on individuals can be significant. Slater calls attention to the confluence of psychologically convincing Place Illusion (PI) and Plausibility Illusion (Psi) in virtual reality, giving users a strong sense of presence, and implicating their body in the virtual space. "If you are there (PI) and what appears to be happening is really happening (Psi) then *this is happening to you!* Hence you are likely to respond as if it were real. We call this 'response-as-if-real' RAIR. (Slater, 2009)

Several researchers have pointed to the compounding impact of "social presence" (Lee, 2004; Ratan, 2012) i.e. the awareness of being co-present with other users, conversing and taking consequential action in a shared, virtual environment. This attribute is often understood in combination with "self-presence" and "environmental presence", cumulatively forming a powerful sense of "being there" that has been identified as particular to virtual reality (Bailenson, 2018). Ratan has suggested that social presence might be considered to be the most impactful of the three, as the participation of other natural persons in a virtual space adds complex social cuing to the simulative environment, further convincing users of the veracity, immediacy and embodied nature of their experience (Ratan, 2012).

The UK's Cyberpsychology Research Group call attention to the contiguous emotional impact of negative experiences in metaverse environments "Just because these events happen online rather than offline doesn't mean they are not being experienced as real" (Askham, 2022). Madary & Metzinger take it a step further, introducing the possibility that "[t]orture in a virtual environment is still torture. The fact that one's suffering occurs while one is immersed in a virtual environment does not mitigate the suffering itself" (Madary & Metzinger, 2016).

In the context of all of the above it seems likely that the immersive and embodied nature of social, metaverse environments will significantly intensify the impact of harassment and abuse such as physical threats or simulated violence. In metaverse environments, non-consensual instances of touching, verbal harassment or invasion of personal space may put users at particular risk of psychological and emotional distress. Future developments such as haptic technology clothing may further heighten this affect by adding a physical sensation to abuse enacted in metaverse contexts in the future.

Even without the use of specific haptic technology, there is evidence to suggest that some people, using only a headset and controllers, experience uncanny physical sensations upon being touched in virtual environments. Some hypothesise that the psychologically convincing nature of metaverse environments can lead users to partially transfer their phenomenal self model (PSM) into that of an avatar , an effect akin to the Rubber Hand Illusion (Botvinick & Cohen, 1998). Consequently, they may report feeling pronounced physical sensations when they observe their avatar being touched or harmed, even though their physical body remains uncontacted (Desnoyers-Stewart et al., 2024; Madary & Metzinger, 2016; McIntosh & Allen, 2023).

A comparable phenomenon, 'phantom touch', is frequently discussed by users of social VR. Although largely under-researched in a formal setting, this sensation appears to involve users perceiving a touch sensation on their bodies that directly corresponds to a simulated act of touch in VR. Some users actively cultivate this sensation using virtual mirrors in order to associate avatar touch with tactile sensation. There would appear to be enhanced likelihood that those who experience a form of 'phantom touch' could be at greater risk of traumatic impact in the event of harassment and abuse (McIntosh & Allen, 2023).

One often-posed question in regard to VR abuse, from those not familiar with the technology is, 'why didn't you just take the headset off?'. Preliminary research suggests that rapid disengagement from VR, particularly under stress or anxiety, can provoke panic or dissociative episodes, therefore, the solution may not be as simple as disconnecting (Allen & McIntosh, 2022). This question also signals a tendency towards victim blaming, failing to account for well understood trauma-response behaviours such as freeze and appeasement in response to high stress, high risk encounters (Cantor & Price, 2007)

4 Design responses

In response to apparent abuses in proto-metaverse spaces, many app developers and platform owners have sought design solutions to mitigate the risk or severity of potential harms. Some have turned to social science research that may not have been initially conceived in relation to technology paradigms, drawing on research exploring physical and relational behaviour as a route into understanding the needs of social, virtual spaces.

Hideaki Matsui, a design lead at Google has publicly discussed their use of Proxemics (Hall, 1966) Hall's theory of Proxemics suggests that people will maintain differing amounts of distance from one another depending on the social setting and their cultural backgrounds. Google use this framework as a schematic, encouraging designers to construct virtual environments that conserve distances between users that are appropriate to the social context and levels of intimacy that might be anticipated in a particular encounter. As per Hall's design, they distinguish between public, social, personal and intimate space and design experiences accordingly. Their approach notably does not incorporate Hall's framing of such boundaries being informed by background and cultural context.

Michelle Cortese, Design Lead Manager at Meta extends the use of Proxemics to incorporate consent frameworks inspired by the BDSM community. She writes about the significant number of people, particularly women, who reported being sexually harassed or assaulted in multi-person virtual reality spaces in the late 2010s, and calls for an approach to personal space management that involves explicit and informed mutual consent.

"we suggest designers build granular controls that are easy to access and surface before intimate interactions begin. It's important that people can customize and control the types of experiences they're willing to have with other people in these close quarters before they happen" (Cortese & Zeller, 2019)

In the intervening years, many of these recommendations have been adopted, with features such as 'personal space bubbles' now available in most social VR apps. Personal space bubbles enforce an invisible boundary around the user, keeping other avatars at a designated distance, or rendering them invisible and inaudible when the allotted space is impinged. In some instances, users can choose only to be perceptible to those pre-designated as 'friends' to minimise the risk of harassment.

Whist such design features may prove useful, they can also create an imbalance of power that favours the aggressor. The onus is on the victim to apply extreme caution entering into a metaverses space, configuring complex safety features and limiting their own experience prior to entry, or attempting to do so in the moment whilst experiencing harassing behaviours. Those persistently harassing other users, notionally violating the terms of use of the platform, encounter no such barriers.

For victims of abuse; block, mute and report tools may be available, and are designed be deployed ad hoq in the event of unwanted attention or abusive behaviour. Such reporting features can be difficult to navigate in the moment, especially when abuse is ongoing. It is also generally unclear what responses or punitive measures might follow from the reporting of such instances. To date platforms are not obliged by any regulatory authority to make transparent their internal monitoring, evidentiary and justice systems, to disclose actions taken to investigate or remediate reports of abuse, or to notify the complainant of any actions taken (Allen & McIntosh, 2022; Center for Countering Digital Hate, 2021).

5 Regulation and governance

5.1 Suitability of existing laws

Around the world, governments are seeking advice on whether existing legislature is sufficient to ensure that their citizens are afforded the same rights, protections and freedoms in metaverse spaces as they might expect in comparable physical and digital spaces.

One key, anticipated challenge to efficacy, is that many legal frameworks related to abuse and harassment make clear distinctions between 'content' abuses which can include the posting and sharing of abusive materials such as text, imagery and video, and physical 'contact' abuses, which generally involve unwanted physical touch.

Several governments have sought to improve protections for citizens in 2D online platforms in recent years. New criminal designations are being written onto the statute books for online criminal behaviour such as 'cyber-flashing' and the posting of 'revenge porn'(Online Safety Act 2023, 2023). In the relatively new field of multi-person, metaverse environments, there is currently little legislative provision to account for abuses that might take place in psychologically convincing, simulative environments where multiple natural persons are co-present and interacting with one another.

Given what is understood about the immersive, embodied and relational qualities of metaverse environments, governments may need to specify a new category of harm. Perhaps one that recognises certain forms of user 'conduct' as harassment and abuse, even where there is no physical contact, or associated production or proliferation of content.

5.1.1 Case study

As an early test of the suitability of existing legislature, UK police announced in January 2024 that they were investigating an alleged instance of 'sexual attack' of a girl who is under 16, and was abused by a group of men in a social VR setting. (Camber, 2024)

In an interview with LBC News, The UK's Home Secretary, James Cleverly said "I know it is easy to dismiss this as being not real, but the whole point of these virtual environments is they are incredibly immersive. We're talking about a child here, and a child has gone through sexual trauma. It will have had a very significant psychological effect and we should be very, very careful about being dismissive of this." (Taylor, 2024)

In response to this case, the chairman of the UK's Association of Police and Crime Commissioners, Donna Jones was reported as saying "We need to update our laws because they have not kept pace with the risks of harm that are developing from artificial intelligence and offending on platforms like the metaverse." (Taylor, 2024)

The statements of two such prominent public figures signals an appetite at policy level to apply some of the principles discussed in this paper at the highest levels of governance. This specific case is understood to be ongoing at the time of publish. It will be interesting to see how existing legislation is applied and reconciled in this seemingly unprecedented case.

5.2 Accountability

Policymakers may wish to consider creating stronger links between activity in the metaverse and national law enforcement agencies. This would ensure that serious crimes committed in metaverse worlds don't remain under the exclusive jurisdiction of the platform's internal justice system, which is arguably better suited to technology-related issues than serious criminal offences. Public confidence will also need to be built such that anyone reporting abuses to civic authorities can expect to be understood, believed, and for their complaint to be acted upon.

Criminal prosecution of individuals for abusive conduct in the metaverse is one area that governments certainly need to consider. Another is the relative culpability and accountability of the companies providing metaverse apps, platforms and services. Where frequent instances of criminal activity, such as abusive behaviour are evidently taking place in a particular app or platform, regulators may wish to consider holding providers wholly or partially accountable. Particularly if they are failing to uphold terms of use, and encouraging or turning a blind eye to abusive behaviour.

In the US, holding platforms to account is likely to prove challenging. Section 230 of the Communications Act affords legal immunity for providers of interactive computer services with respect to the actions of their users (Section 230, 1934) In the UK, the new Online Safety Act (Online Safety Act 2023, 2023) has some provision for this, extending a “duty of care” to platform owners regarding what content users, particularly children should be able to encounter online. The challenge of ‘content’ versus ‘conduct’ and ‘contact’ is largely unaddressed in the Act, however the metaverse has been deemed explicitly in scope (Local Government Association, 2022). The Institute for Engineering and Technology recently called on UK government to ensure that new legislation is made fit for purpose in relation to social, spatial environments (Almond et al., 2024). The EU’s new Digital Services Act (European Union, 2023) goes further still, holding very large tech companies legally accountable for the content posted on their platforms. Again, the Act sets out a framework for addressing illegal ‘content’ online, however there is no direct provision for metaverse contexts, and it remains unclear how the more behavioural, conduct-based forms of abuse and harassment might be addressed by this new legal framework.

5.3 Jurisdiction

In most legislative frameworks, sovereign jurisdiction is determined by the geography of where an alleged crime has taken place. For many exponents of the metaverse, the promise of this new paradigm lies in its potential to be borderless and decentralised. Just as cryptocurrency could be conceived as an alternative to centralised banking systems, so the metaverse might be imagined as an alternative to state-based territoriality for interpersonal encounters. What then for state-based authorities looking to respond to reports of criminal activity, including reports of harassment and abuse in the metaverse?

As with the internet before it, questions of jurisdiction in metaverse contexts are proving challenging. Users of such spaces may be encountering one another in what experientially is a common metaverse environment, but connecting from very different territories, each with their own particular legal contexts. To further complicate matters, the metaverse environment visited might be provided by a company in another territory, with the underpinning technology stack hosted across multiple territories. What legal frameworks should then apply when abuses are detected? And which nation(s) should have the jurisdiction to prosecute criminal behaviour?

Laws governing interpersonal behaviour vary considerably between territories, and jurisdictional ambiguity can create a vacuum of legal accountability, a lag in governmental response to evident harms, and a gulf of support for victims of criminal behaviour.

Even in instances when jurisdiction is relatively unambiguous, or where laws can be expected to be common across territories, challenges can remain. For instance, most legal systems descended from English law e.g. Australia, Canada, New Zealand, Singapore and the United States, conform to similar systems of Tort law (civil laws pertaining to interpersonal wrongdoing between private persons). However, it remains unclear whether such laws would be legally applicable in metaverse contexts as the legal 'personhood' of an avatar is yet to be determined. Questions remain regarding whether the actions of an avatar in a virtual world should be considered directly analogous to the action of the embodied 'natural person' controlling it. Or whether avatar behaviour would be better understood as akin to a playable video game character (Cheong, 2022). Each approach would attract a very distinct legal response, particularly in relation to acceptable levels of interpersonal violence.

In the absence of legal certainty there is concern that cases of abuse and harassment may become entrenched in costly, intractable disputes regarding which legal jurisdiction applies, risking a drain on resources in multiple territories and lessening the likelihood of successful conviction (Europol, 2022; Kalyvaji, 2023).

One approach would be to make platforms responsible for ensuring that the legal protections of each user are implemented in the design of the space before they are granted access to a given metaverse environment. Where legal frameworks in different jurisdictions prove incompatible, this may lead to citizens from certain territories being excluded, or companies running multiple instantiations of metaverse environments, the user being directed to the space that is compliant with their domestic legal system. An alternative, or addition perhaps, is to encourage closer working with international agencies such as Interpol to ensure the complementarity of different governmental approaches, and to enhance international cooperation agreements to support cross-jurisdictional prevention and response to crimes involving metaverse technologies and environments.

5.4 Stakeholder literacy

Among the most immediately actionable opportunities for government agencies engaging with this topic, is to improve stakeholder literacy. This could be achieved by training programmes, giving stakeholders direct experience of embodied metaverse platforms, providing insight into the current trajectory and pace of technological developments, and the manner in which the affordances of this medium relate to issues of abuse and harassment. Governments may wish to consider prioritising the literacy of responsible bodies such as legislators, police and the judiciary. Public literacy campaigns may also be valuable in supporting citizens to understand their rights, and empowering them to make informed and empowered choices about their own engagement with the metaverse.

6 Conclusion

Although the metaverse is often positioned as a ‘future horizon’ technology, it is evident that early versions of the metaverse are already here, and that instances of harassment and abuse are taking place with potentially significant consequences for citizens. Governments have an opportunity to urgently consider the suitability and efficacy of existing legislature, and to assess whether new legal instruments are needed to reflect the distinctive experience of embodied, immersive, multi-person environments. Policymakers may also wish to consider prevention, reporting and prosecution strategies, as well as the accountability of both individuals and platforms/service providers in relation to abusive behaviours in metaverse environments. Programmes of metaverse literacy now could equip stakeholders and the wider public with the information they need to collectively design and advocate for more positive futures for the metaverse.

7 References

Allen, C., & McIntosh, V. (2022). *Safeguarding the metaverse*. https://www.theiet.org/media/9836/safeguarding-the-metaverse.pdf

Almond, E., McIntosh, V., & Allen, C. (2024, January 3). *An open letter to Ofcom on the need to urgently review how VR spaces are governed*. The IET. https://www.theiet.org/media/press-releases/press-releases-2024/press-releases-2024-january-march/3-january-2024-an-open-letter-to-ofcom-on-the-need-to-urgently-review-how-vr-spaces-are-governed

Askham, G. (2022, April 26). Metaverse: New Documentary Exposes Racial & Sexual Abuse. *Glamour*. https://www.glamourmagazine.co.uk/article/metaverse-misogyny

Bailenson, J. (2018). *Experience on demand : what virtual reality is, how it works, and what it can do* (First edit). W.W. Norton & Company,.

Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, *391*(6669). https://doi.org/10.1038/35784

Camber, R. (2024, January 1). British police probe VIRTUAL rape in metaverse. *Daily Mail*.

Cantor, C., & Price, J. (2007). Traumatic Entrapment, Appeasement and Complex Post-Traumatic Stress Disorder: Evolutionary Perspectives of Hostage Reactions, Domestic Abuse and the Stockholm Syndrome. *Australian & New Zealand Journal of Psychiatry*, *41*(5), 377–384. https://doi.org/10.1080/00048670701261178

Center for Countering Digital Hate. (2021, December 30). *New research shows Metaverse is not safe for kids*. Center for Countering Digital Hate (CCDH). https://counterhate.com/blog/new-research-shows-metaverse-is-not-safe-for-kids/

Cheong, B. C. (2022). Avatars in the metaverse: potential legal issues and remedies. *International Cybersecurity Law Review*, *3*(2), 467–494. https://doi.org/10.1365/s43439-022-00056-9

Cortese, M., & Zeller, A. (2019). *Designing Safer Social VR Using the ideology of sexual consent to make social VR a better place*. https://immerse.news/designing-safer-social-vr-76f99f0be82e

Desnoyers-Stewart, J., Bergamo Meneghini, M., Stepanova, E. R., & Riecke, B. E. (2024). Real human touch: performer-facilitated touch enhances presence and embodiment in immersive performance. *Frontiers in Virtual Reality*, *4*. https://doi.org/10.3389/frvir.2023.1336581

Eccles, L. (2022, January 22). My journey into the metaverse — already a home to sex predators. *The Sunday Times*.

European Union. (2023). *The Digital Services Act*. https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package

Europol. (2022). *Policing in the metaverse: what law enforcement needs to know, an observatory report from the Europol Innovation Lab*. https://doi.org/10.2813/81062

Hall, E. T. (1966). *The Hidden Dimension*. NY Doubleday.

Kalyvaji, M. (2023). Navigating the Metaverse Business and Legal Challenges: Intellectual Property, Privacy, and Jurisdiction. *Journal of Metaverse*, *3*(1), 87–92. https://doi.org/10.57019/jmv.1238344

Lee, K. M. (2004). Presence, Explicated. *Communication Theory*, *1*(1), 27–50. https://onlinelibrary-wiley-com.ezproxy.uwe.ac.uk/doi/epdf/10.1111/j.1468-2885.2004.tb00302.x

Limina Immersive. (2018). *Immersive Content Formats for Future Audiences*. www.digicatapult.org.uk

Local Government Association. (2022, April 19). *Online Safety Bill, Second Reading, House of Commons*. https://www.local.gov.uk/parliament/briefings-and-responses/online-safety-bill-second-reading-house-commons-19-april-2022

Madary, M., & Metzinger, T. K. (2016). Recommendations for Good Scientific Practice and the Consumers of VR-Technology. *Frontiers in Robotics and AI*, *3*. https://doi.org/10.3389/frobt.2016.00003

McIntosh, V., & Allen, C. (2023). *Child Safeguarding and Immersive Technologies: Key Concepts*. https://learning.nspcc.org.uk/research-resources/2023/child-safeguarding-immersive-technologies

Online Safety Act 2023 (2023). https://www.legislation.gov.uk/ukpga/2023/50/enacted

Outlaw, J. (2018). *Survey of Social VR Users*. The Extended Mind. https://www.extendedmind.io/2018-survey-of-social-vr-users

Patel, N. J. (2021, December 21). *Reality or Fiction?* Medium. https://medium.com/kabuni/fiction-vs-non-fiction-98aa0098f3b0

Ratan, R. (2012). Self-presence, explicated: Body, emotion, and identity extension into the virtual self. In *Handbook of Research on Technoself: Identity in a Technological Society* (pp. 321–335). IGI Global. https://doi.org/10.4018/978-1-4666-2211-1.ch018

Rifkind, H. (2022, February 21). The metaverse will be an abuser's paradise. *The Sunday Times*.

Section 230 (1934). https://crsreports.congress.gov

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions. Biological Sciences*, *364*(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Taylor, W. (2024, January 2). Police investigate "rape" in metaverse after group of men attack girl in virtual reality room. *LBC News*.

XRSI. (2023). *The Metaverse - X Reality Safety Intelligence (XRSI)*. XRSI. https://xrsi.org/definition/the-metaverse

Extended Reality is a Fundamentally Mediated Reality

Emmie Hine[1,2*]

[1]Department of Legal Studies, University of Bologna, Bologna, Italy

[2]Center for IT & IP Studies, KU Leuven, Leuven, Belgium

*** Correspondence:**
Emmie Hine
emma.hine@studio.unibo.it

Keywords: extended reality, mediation theory, technological mediation, online self-presentation, online experience, technology ethics, virtual reality, augmented reality.

Abstract

This article advances the idea that extended reality is a *fundamentally mediated reality*. Drawing on postphenomenological theory and real-world examples, it explores three levels of inherent mediation—devices, corporations, and self-presentation—arguing that these mediators implicate fundamental rights and risk making extended reality a fundamentally unequal reality.

1 Introduction

Extended reality (XR) technologies are increasingly entering the consumer domain, and with them questions about what they mean for our lives on- and offline. XR encompasses the inherently mediating technologies of virtual reality (VR) and augmented reality (AR), and as intermediaries between us and the physical and virtual spheres, they impact our experiences, wellbeing, and even fundamental rights.

Technologies already mediate our worlds, from phones reshaping interactions with friends to algorithms curating our news. Technological mediation has been the subject of philosophical inquiry both generic and specific. I draw on Ihde's theory of postphenomenology, which takes a pragmatic approach towards investigating the relations between humans and technological artifacts (Verbeek 2016; Ihde 2009). Verbeek (2016) applies this to a theory of technological mediation, arguing that we need to center the "mediated subject" and "the ways in which human beings give meaning to [technologies'] mediating roles," which encompass the epistemological, ethical, and metaphysical (Verbeek 2012). Van Den Eede (2011) presents a transparency-opacity mediation framework, looking at "design" (or "use") and "context" perspectives. I combine these approaches, looking primarily at the epistemological and ethical levels of abstraction. XR technology is an epistemological mediator that influences our perceptions of the virtual and physical and the information we can access within each, and an ethical mediator through its impacts on our experiences. Historically, not all groups have had equal experiences in cyberspace, and racist and sexist harassment are rife on social VR platforms (Hine 2024). XR's immersive, embodied realism creates experiences similar to those in the physical world (Hine 2024) and thus nondiscrimination, safety, privacy, and other fundamental rights are a concern in XR as well.

In this paper, I will explore how XR technologies shape our realities through three levels of abstraction: the device, the corporation, and self-presentation. Because they are intrinsic to XR, I argue it is a *fundamentally mediated reality,* which impacts experiences and fundamental rights.

2 Three Levels of Mediation

2.1 Devices

The most immediately apparent level of technological mediation is that of the device, which primarily mediates at the epistemological level. Unlike physical reality, extended reality is only accessible through devices. These devices mediate both the physical world by impacting user perception of their surroundings, and also the virtual experiences that devices grant access to. Regarding the physical world, XR devices mediate how much you perceive and how you perceive it. “See-through” AR devices provide a transparent (both physically and in terms of awareness of mediation) view of the physical world, while “pass-through” devices use video projected onto opaque interior screens to mimic physical transparency and attempt to achieve transparency of use. This creates pernicious visual impacts; field of vision and color range are restricted (Patel 2024), changing how the physical world is perceived on a fundamental level.

The virtual sphere is even more mediated by devices, which are responsible for creating and projecting virtual AR and VR content. The quality of this content is determined by the quality of the device (and its Internet connection, if a connected experience): the color spectrum, resolution, frame rate, latency, and tracking accuracy vary across devices. For instance, the $3,499 Apple Vision Pro has a screen resolution of 3,386 pixels per inch (PPI), nearly three times as many as the $499 Meta Quest 3 (Sorrel 2024), contributing to a superior user experience (Patel 2024). Additionally, while most consumer VR products only support visual and audio input, haptics and other sensory inputs like scent generators could eventually contribute to greater sensory fidelity, adding to the amount of mediation—but at a cost. Device mediation thus creates a fundamental level of inequality: those with the money to purchase higher-quality devices and higher-speed broadband will have superior experiences. While in physical reality, one cannot pay to have more or better senses than others,[1] it is highly likely that an extended reality will be a fundamentally unequal one, with tiered experience quality that embeds a discriminatory digital divide.

2.2 Corporations

While hardware mediates the individual’s sensory experiences, corporations decide how it will do so. Hardware manufacturers define the type and quality of sensory experiences that devices permit but also the type and quantity of data they can collect through sensors, acting as an epistemological mediator for both users and platforms. Platform companies then define the parameters of XR platforms, including platform moderation policies. Thus, they are essentially the governments of physical worlds, but more powerful because they define the fabric of extended reality itself. This gives them a corresponding amount of power over fundamental rights by, for instance, making decisions about behavioral policies and how they are enforced. This is a mediation of the ethical dimension, as it lays out what behaviors are acceptable and even what behaviors are possible through code. For instance, sexual harassment plagues many social VR platforms; after a beta tester was groped in Meta Horizon

[1] A possible exception is paying for corrective medical procedures such as cataract surgery or cochlear implants. However, while these elevate one’s experience of a specific sense above their pre-procedure baseline—and that of those who cannot access those procedures—it does not augment them compared to the general population like XR devices can.

Worlds, the company rolled out a "personal boundary" feature (Basu 2021), but platforms still struggle with real-time behavioral moderation to address harassment, especially of already marginalized groups (Hine 2024). Platform policies that are not well-crafted and equally applied will create an unequal reality, disparately impacting rights like safety, privacy, nondiscrimination, and expression.

2.3 Self-Presentation

Self-presentation is a form of cooperative mediation that is affected by hardware, platforms, and the individual. On the hardware level, this happens with the physical world, where an XR device becomes part of an individual's self-presentation. Some devices are more contextually transparent than others. AR glasses such as the Meta Ray-Ban blend in like normal glasses, while the Apple Vision Pro's "EyeSight" feature, which projects the user's eyes onto the front of the device when another person tries to interact with them, has been criticized as overtly uncanny (Patel 2024). However, mediating devices can become more contextually transparent over time; Apple's AirPods were initially mocked for looking strange, but eventually became a status symbol (Tinari 2023). XR devices may cause an off-putting self-presentation at first, but could change social perceptions as they become prevalent.

Virtual self-presentation also mediates interactions. This happens even in text form (Hine 2024), but VR experiences with avatars more directly mediate self-presentation via avatar customization. Early VR platforms had extremely limited options, excluding racial minorities and gender-diverse users from fully expressing their identities, but modern platforms are taking conscious steps towards inclusivity (Hine 2024). This can have broadly beneficial effects, allowing users to explore certain identities or present themselves in advantageous ways. However, it can also allow users to conceal their identity, which could encourage abuse without consequences, especially of users who present marginalized identities. This mediates both the epistemological dimension in terms of information available about users, but also the ethics dimension by facilitating different behaviors depending on self-presentation options. Again, it potentially creates disparate and discriminatory experiences for marginalized users, emphasizing the unequal impacts of fundamentally mediated reality on fundamental rights.

3 Conclusion

XR is inherently reliant on mediating technologies and corporations, and implicates both virtual and physical self-expression. Because of this, it is not just a case of technological mediation, but a fundamentally mediated reality. Combined with its psychological and physiological realism, this creates significant implications for fundamental rights, especially as XR becomes more prevalent and more sensorily equivalent to physical reality. This short paper is a first step towards exploring the full impacts of creating a fundamentally mediated reality, but it is clear that as XR develops, we must consider how the experiences it enables may not be equal or equitable for all. In creating a fundamentally mediated reality, we may have also created a fundamentally unequal one.

4 References

Basu, Tanya. 2021. "The Metaverse Has a Groping Problem Already." MIT Technology Review. December 21, 2021. https://www.technologyreview.com/2021/12/16/1042516/the-metaverse-has-a-groping-problem/.

Hine, Emmie. 2024. "Virtual Reality, Cyberspace, and Embodiment: A Historical Debate with Contemporary Resonance." SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.4714432.

Ihde, Don. 2009. *Postphenomenology and Technoscience: The Peking University Lectures*. State University of New York Press. https://philarchive.org/rec/IHDPAT-2.

Patel, Nilay. 2024. "Apple Vision Pro Review: Magic, until It's Not." The Verge. January 30, 2024. https://www.theverge.com/24054862/apple-vision-pro-review-vr-ar-headset-features-price.

Sorrel, Charlie. 2024. "Vision Pro Teardown Part 2: What's the Display Resolution?" iFixit. March 23, 2024. https://www.ifixit.com/News/90409/vision-pro-teardown-part-2-whats-the-display-resolution.

Tinari, George. 2023. "A History of Apple AirPods: From Controversial Launch to Universal Triumph." MUO. June 7, 2023. https://www.makeuseof.com/history-of-apple-airpods/.

Van Den Eede, Yoni. 2011. "In Between Us: On the Transparency and Opacity of Technological Mediation." *Foundations of Science* 16 (2): 139–59. https://doi.org/10.1007/s10699-010-9190-y.

Verbeek, Peter-Paul. 2012. "Expanding Mediation Theory." *Foundations of Science* 17 (4): 391–95. https://doi.org/10.1007/s10699-011-9253-8.

———. 2016. "Toward a Theory of Technological Mediation: A Program for Postphenomenological Research." In *Technoscience and Postphenomenology: The Manhattan Papers*. Lexington Books.

Considering Psychotherapy in the Metaverse

Lucie Kruse [1,*], **Catharina Rudschies** [2], **Sebastian Rings** [1], **and Frank Steinicke** [1]

[1] *Human-Computer Interaction, Universität Hamburg, Hamburg, Germany*
[2] *Department of Informatics, Universität Hamburg, Hamburg, Germany*

Correspondence*:
Lucie Kruse
lucie.kruse@uni-hamburg.de

ABSTRACT

Virtual environments provide many possibilities for immersive psychotherapy, e.g., exposure therapy, practicing difficult situations before encountering them in reality, or through relaxation exercises (Cieślik et al. (2020)). However, in many cases, these are single applications that can be experienced in the clinical context, with a physically co-located psychotherapist, who is constantly available on-site during the therapy. While remote therapy potentially provides several benefits and advantages, such as availability and transferability, it also leaves room for ethical and technical considerations, justifying the question: Can - and should - virtual therapy also be available online through the metaverse? In this paper, we provide a first glance at three important design considerations of integrating psychotherapy into the metaverse: uncertainties in emergency situations, transmission of non-verbal factors, and privacy and data protection requirements.

Keywords: metaverse, virtual reality, psychotherapy, digital health, emotion recognition, ethics in IT

1 INTRODUCTION

The concept of the metaverse is getting more tangible with each advancement in XR and networking technologies, providing a social, immersive environment where users can meet, enjoy entertainment and perform business operations. This also opens opportunities for providing immersive psychotherapy online, combining a variety of effective digital therapy offers, such as diagnostics, cognitive behavioral therapy, or online support groups, in one application, potentially with human support available. This could mean that instead of visiting a health expert in person, they could meet online in VR, represented by their respective avatars, in a chosen virtual environment. The meeting characteristics, as well as the representation of the users could be customized to the patient's needs and in the best case, this could improve the therapy outcome. Additionally, through the international connections in the metaverse, it would be possible to find a health expert at any time. In this paper, we identify current technical and ethical limitations from interviews with health experts and patients, and underline these statements with literature to point out design considerations for future metaverse applications in the psychotherapeutical domain.

2 INTERVIEWS AND FINDINGS

For a project on hybrid psychotherapy, we interviewed sixteen medical experts (M_x) and ten (former) patients and relatives (P_x). From these extensive interviews, this paper only considers certain aspects that can be transferred to applications in the metaverse. The relevant interview questions deal with the importance of personal contact between a patient and a medical expert, possible use cases for digital interventions, and prerequisites for (digital) diagnostics and therapy. All interviews were audio-recorded and transcribed. Consequently, they were evaluated with respect to our research question, identifying design considerations. Three larger themes emerged, which will be presented in the following parts.

2.1 Negative Impacts of Remote Therapy

With all the positive aspects of the metaverse, such as independence of the user's location and a certain doctor's availability, virtual therapy could also pose risks in emergency situations. For example, one medical expert mentioned that during online therapy sessions, if there is an emergency, it can be difficult to directly reach the patient (M_{15}). Similarly, a patient mentioned that being alone when an online session ends would leave them without support when dealing with the aftermath of the session (P_7). This implies that if patients are in an acute crisis during or after using the metaverse for psychotherapy, they could possibly not receive the same, immediate help they would get if they were in a clinical setting. In contrast, when virtual therapy is offered 24/7 through the metaverse it may also contradict common practice in conventional (face-to-face) psychotherapy. Therapists usually aim to prevent patients from becoming dependent on them; rather, they work to improve patients' self-esteem and agency (Clemens (2010)). Hence, while the possibility for constant access to psychotherapy may sound beneficial on paper, excessive dependency on a (virtual) psychotherapist may even be counterproductive for the therapy's success and increase social isolation. Some studies also indicate that the metaverse has the potential for addiction (Bojic (2022)).

2.2 Non-Verbal Factors

One of the doctors said that the correct diagnosis of a patient, which is crucial to choose the right form of therapy, not only depends on what they verbally say but also on non-verbal factors such as gait patterns, facial expressions, focus and attention (M_{13}). A patient also mentioned difficulties in the communication with their psychotherapists when they met via video call, e.g. because the medical experts had limited perception of the patient's facial expressions or gestures, which can lead to misunderstandings and fatigue (P_1). In the context of the metaverse, these issues are also relevant. While current VR headsets provide facial tracking, classification of emotions and transferring emotions to a virtual avatar still remain areas of research (Khare et al. (2024)). For example, facial expressions with a Meta Quest Pro[1] are estimated using infrared images of the user's face and then mapped to generic facial expressions of virtual avatars. In this process, it is possible to lose accuracy or personal expression characteristics, deteriorating the emotion fidelity. Similarly, body tracking in standalone consumer headsets is often only performed by tracking the head and hands of users and interpolating between these positions. This can in turn lead to missing important information, such as distress, which can consequently result in a worse diagnosis or therapy experience. To counteract this gap of knowledge, it is possible to integrate biosensors, e.g. to measure the participant's skin conductance or heart rate. This would provide information about physiological stressors and potentially help patients' and medical experts' mutual understanding. However, research has shown that the assumption that emotions can be inferred from facial movements alone currently lacks scientific evidence and is "at best incomplete and at worst entirely lack[s] validity" (Barrett et al. (2019), p. 48). Further research is thus needed on how to improve non-verbal communication in the metaverse.

[1] https://www.meta.com/de/quest/quest-pro/

2.3 Privacy and Data Protection

In our interviews, most participants raised privacy concerns, especially in regard to personal data and sensitive information being collected and analysed. Interviewees required that data is securely stored and not shared with third parties (P_3). Particularly data revealed in conversations and collected through body tracking (Bailenson (2018)) or biosensors poses a high risk for the user's privacy. The amount and nature of the data also poses legal challenges. Current legislation only protects "personal data", meaning information through which a person is identifiable. However, some data collected via VR applications like aggregated or irreversibly anonymised data are not considered "personal data" and therefore lack legal protection, despite the potential to infer highly sensitive information from them (Wachter and Mittelstadt (2019)). For example, the inference of sensible attributes such as gender, age, and even income is possible with relatively high accuracy from seemingly innocuous information (Nair et al. (2022)). Given that mostly large companies control the infrastructure, software and state-of-the-art hardware necessary for the metaverse, concerns can also be raised about the control private companies have over the practices surrounding design and data processing.

3 IMPLICATIONS AND FUTURE WORK

This paper only presents a short glance into the psychotherapeutical requirements. Further research is still needed to study the effects of facial and body tracking, but improving the realism of avatar representation could foster a better understanding between users. Further, it has to be clarified what safety measures need to be in place for emergency situations. Finally, it is crucial to ensure privacy and data safety. Especially for medically sensitive data such as diagnoses, medical history of a patient or emotional states, the patient has to remain in control of their data which cannot be revealed to third parties. Future research in this direction should be considered to assess the benefits and risks of psychotherapeutic care in the metaverse.

Funding: Funded by the BMBF within the scope of the HIVAM project, Grant Number: 16SV8878.

REFERENCES

Bailenson, J. (2018). Protecting nonverbal data tracked in virtual reality. *JAMA pediatrics* 172, 905–906

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., and Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest* 20, 1–68

Bojic, L. (2022). Metaverse through the prism of power and addiction: what will happen when the virtual world becomes more attractive than reality? *European Journal of Futures Research* 10, 1–24

Cieślik, B., Mazurek, J., Rutkowski, S., Kiper, P., Turolla, A., and Szczepańska-Gieracha, J. (2020). Virtual reality in psychiatric disorders: A systematic review of reviews. *Complementary Therapies in Medicine* 52, 102480. doi:https://doi.org/10.1016/j.ctim.2020.102480

Clemens, N. A. (2010). Dependency on the psychotherapist. *Journal of Psychiatric Practice®* 16, 50–53

Khare, S. K., Blanes-Vidal, V., Nadimi, E. S., and Acharya, U. R. (2024). Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Information Fusion* 102, 102019. doi:https://doi.org/10.1016/j.inffus.2023.102019

Nair, V., Garrido, G. M., and Song, D. (2022). Exploring the unprecedented privacy risks of the metaverse. *arXiv preprint arXiv:2207.13176*

Wachter, S. and Mittelstadt, B. (2019). A right to reasonable inferences: re-thinking data protection law in the age of big data and ai. *Colum. Bus. L. Rev.* , 494

Biometrics and Inferred data in the metaverse

***Giovanni Sorrentino**[1,2]

[1]*CIRSFID, University of Bologna, Via Galliera 3, 40121 Bologna, Italy*

[2] *International School of Advanced Study, University of Camerino, Via Gentile III Da Varano, 62032, Camerino MC, Italy*

***Correspondence:** giovanni.sorrentino5@unibo.it

Keywords: metaverse, biometrics, data, GDPR, privacy, avatar.

Abstract

The arrival of the metaverse involves reassessing concepts related to data processing and privacy, particularly for biometric and inferred data. The European legal framework regulates biometric data processing, and it is essential to evaluate its suitability for the metaverse. Key aspects include the relationship between avatars and biometric data, especially when avatars closely resemble their users, and understanding how to handle inferred data—information that can indirectly reveal sensitive data. This analysis aims to review the current European framework and platform privacy policies to ensure compliance with the metaverse's unique challenges.

Introduction

The rapid advancement of technology outpaces legislative processes, posing a significant challenge for legislators, including those in Europe. Although European regulations in the tech field are high-quality, there is uncertainty if they can keep up with evolving technology, such as the metaverse. With no specific European regulation for the metaverse, the GDPR will provide guidance on data protection, particularly for avatars and biometric/inferred data use. It is crucial to determine where the GDPR is directly applicable and where it requires interpretation for the metaverse.

1 Biometric Data

Within the metaverse, a vast amount of data is collected, including not only non-sensitive personal data, but also particularly sensitive data, such as biometric data.

Biometrics refers to the measurement of physical aspects of the human body. This can include patterns of the skin or blood vessel networks under the skin; facial appearance, and behavioural traits (Smith, Mann and Urbas, 2018). This highlights the most crucial aspect of biometrics: their uniqueness. The use of biometrics has many potential applications in the metaverse. Specifically, with regard to avatars, the biometric data (or at least the inferred data) that can potentially be extracted and processed requires greater attention.

The processing of "non-sensitive" personal data does not entail any specific risks or obligations for the processor and falls within the scope of Article 4. On the other hand, the processing of biometric data, due to the inherent sensitivity of these data, is subject to a distinct discipline, contained in Article 9.

In EU, the GDPR defines biometric data as “personal data resulting from specific technical processing relating to the physical, physiological or behavioral characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data” (Art. 4(14)). The processing of this type of data is strongly restricted by the European legislator. As a result, Article 9 paragraph 1 of the GDPR establishes a general ban on the processing of biometric data for the purpose of uniquely identifying a natural person.

This general rule is subject to a derogation in the cases specified in the same Article 9 paragraph 2. The main derogations include cases where the data subject has given explicit consent or where the processing concerns personal data that has been “manifestly made public” by the data subject.

Consequently, when personal data is not technically processed for the specific purpose of uniquely identifying an individual, it should not be considered biometric data as defined by Art. 4, point 14) and processed as a non-sensitive data. Nevertheless, cases of biometric data processing in the metaverse are limited and are mainly problematic when used for the unique identification of a person. This is not occurring at present, although it cannot be ruled out for the future. However, a different discussion should be made regarding so-called “inferred data” (Bolognini, 2022).

2 Inferred Data

Although the GDPR does not provide a definition, “inferred data” refers to information derived from the analysis or interpretation of other data, such as human characteristic data. In the context of the metaverse, if data—including those generated by avatars—reveal sensitive information, the GDPR's Article 9 guidelines will apply. These guidelines impose restrictions and conditions on processing such sensitive data. The massive use of avatars, in fact, raises several potential legal issues. As a matter of practice, during registration on a platform, users are typically asked to create their own avatar. There are no specific requirements or guidelines for avatar design, and users may choose to use an avatar that looks very different from their real-life appearance. However, in practice, most users tend to create an avatar that resembles their real-life appearance (Gavrilova and Yampolskiy, 2011), which makes it possible to use the results of successful avatar recognition for human recognition, and vice versa.

The metaverse is likely to become increasingly realistic, so avatars will also become highly similar to the real user. For instance, considering the level of realism achieved by video games such as FIFA or Call of Duty or still the Meta’s Codec Avatar, can serve as an example. Excluding, for now, the numerous legal differences, this comparison is useful for understanding the level of quality and similarity that avatars could reach.

Based on experience to date, within the main platforms of the metaverse, there is no mention of biometric data, nor is consent required under Article 9 of the GDPR. However, this is still an unexplored area, as until now there has been no digital environment capable of simultaneously involving so much personal data as in the metaverse, and it will certainly be a critical issue in the future.
Excluding, for the time being, the challenges related to biometric data, it is important to note that a vast amount of inferred data can be deduced and processed by the platform itself and by third parties, without an adequate basis of consent.

In summary, an analysis of metaverse platforms Decentraland, Roblox, and Epic Games found that none of them specifically collect consent for biometric data, nor do they mention biometric data in their privacy policies. This lack of transparency raises questions about how these platforms process and utilize data related to user avatars. If the data associated with avatars is used, it could potentially be classified as personal data or special categories of personal data as per Article 9 of the GDPR, which could have legal implications. Above all, it should take into account not only biometric data, but also inferred data, which instead seem to be totally ignored. This classification could have significant implications for the handling and processing of such data. In fact, if not informed correctly, the user may not foresee the consequences of configuring and using their avatar. Sensitive data, such as racial or ethnic origin, health data, or other physical characteristics, could be inferred from this.

Although biometric data is not being processed with the aim of identifying individuals, it seems that platforms are not adequately considering the data that can be inferred during user activities. Therefore, even if biometric data were not actually processed for specific purposes, taking into account inferred data it would be appropriate to inform users of the consequences of their avatar configurations, as the principle of transparency requires.

Conclusions

The considerations in this article could also open up a discussion on the adequacy of the current European legal framework with regard to the metaverse. Until now, the most widely used technologies had not manifested the capacity to collect and process data that the metaverse is highlighting.
In summary, the integration of real-world elements into digital contexts in Web 3.0 and the imminent 4.0 brings the need for protecting users from potential negative consequences. Currently, avatars lack a unified legal definition, specific regulation, and sufficient attention from platforms and users. This lack of regulation can lead to various issues, which users may encounter, and on which platforms need to provide clear guidance. The situation becomes even more critical with the rise of artificial intelligence and the need for coordination with the AI Act, particularly focusing on biometric data. Therefore, it is essential to establish clear guidelines and regulations for avatars and other digital elements to ensure user protection and safe use of digital spaces.

References

Ball, M. (2022). The Metaverse: And How It Will Revolutionize Everything. Liveright Publishing Corporation, a division of W.W. Norton & Company.

Bolognini, L., & Carpenelli, F. (2022). Il futuro dei dati personali nel metaverso. Diritto, Economia e Tecnologie della Privacy.

European Data Protection Board (EDPB). (2019). Guidelines 3/2019 on processing of personal data through video devices.

Gavrilova, M. L., & Yampolskiy, R. (2011). Applying Biometric Principles to Avatar Recognition. In: Transactions on Computational Science XII. Lecture Notes in Computer Science, vol 6670. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-22336-5_8

Smith, M., Mann, M., & Urbas, G. (2018). Biometrics, crime and security. Routledge.

Creating the Metaverse: Technical Challenges and Solutions

Some Technical Challenges of Scaling from Social Virtual Reality to Metaverse(s)

Anthony Steed [1]

[1] *University College London, Computer Science, London, UK*

Correspondence*:
Corresponding Author
a.steed@ucl.ac.uk

1 INTRODUCTION

There are many visions for what metaverse systems will comprise: from explorations of technical affordances of next-generation communication systems (19, 1) through to rights and principles that should underpin such systems in use (21, 5, 3). If there is a consensus, it is that we don't quite know what the technical architecture or scope of capabilities of such as a system will be: Will it consistent of purely simulated virtual worlds or will it be grounded in mixed-reality spaces? What is the role of emergent network technologies such as peer to peer databases or blockchains? Who will control the systems and how will rights be enforced? How will the system be equitable to users with different backgrounds and access to interface devices?

We can see the early prototypes of future metaverse systems in today's social virtual reality (SVR) systems (e.g. see (20, 16). These allow multiple users to congregate in diverse virtual places, share fantastic or realistic representations of themselves and undertake a wide variety of social activities. To some extent, a metaverse system should allow any existing SVR environment and activity to be experienced within a more unified framework where one didn't need multiple accounts, multiple devices, multiple pieces of client software and experience with multiple interfaces. In a metaverse we would expect to be able to reach all of the virtual places supported and we would expect fluid interactions with other users. Today, while some users are already spending a lot of time in prototypical metaverse experiences, hardly any of the capital developed in one system (e.g. assets, but also social capital) are portable to other systems. Each is effectively a walled garden. Discoverability is low, with different experiences needing different software. Users are often not able to stay as a group within a single software client, never mind when switching to different software.

Today, designing a whole system that alleviates these frustrations, is an almost impossible task. Even the most evolved walled gardens of today lack significant features. Indeed some of the challenges about ownership, naming and scalability have already been identified in research going back over 30 years (e.g. see (22, 25)). There is a need for more systems to be prototyped that bring us closer to understanding the infrastructure components needed. One could point at other successful distributed engineering efforts such as the Internet and the World-Wide Web and claim that system will emerge by building on simple standards, and that relevant architectures will emerge from clear separation of responsibilities between clients and servers. But this belies that there are fundamentally difficult problems that client software that connects to metaverse(s) will need to solve, no matter their architecture. This essay highlighta three such problems that affect the client software that a user engages with: what interoperability means; how to scale awareness levels; and how to ensure accessibility.

2 INTEROPERABILITY

A metaverse system would necessarily be made up of client software that interact with server systems that provide critical features. While some of communication might be peer to peer between clients, as is done with some current game systems and research prototypes (e.g. see (25, 10)), within the services of metaverses we will find the distribution of description of virtual spaces, services for identifying and rendezvousing with other users, platforms for secure transactions, etc. Additionally, as we discuss in the next section, it is highly likely that dense social environments will require server infrastructure for simulation and message distribution. When we talk of interoperability, we thus identify that clear interfaces must exist for describing virtual spaces (i.e. are descriptive of scenes, containing geometry, materials, animations, etc.) but also for interaction between different processes across the Internet. Previous work on inter-operation has tended to focus on the latter, plotting architectures that build upon existing web services (e.g. (14)) or proposing new server simulation models[1]. There is an implicit assumption in a lot of this work: there will be a single integrated client that accesses these services. That is, if only these services could be described the client would take on the complete role of creating the user experience. Systems that have been built over the past couple of decades, have thus been largely based around this model of a large client that integrates all the interfaces necessary to connect to the back-end services. Second Life illustrates the advantages and disadvantages of this approach: while originally a closed system, the client was made open along with its protocols, so servers could be implemented to complement the official servers (for an overview see (2, 3)). Modern SVR systems are still walled gardens though their architectures will have some similarities.

While standardisation of services is certainly necessary, we highlight that almost all efforts are making this assumption that a single client is being supported. We note that with the rise of smartphones, consoles and tablets, the model of a single application taking control of the user experience has dominated. That is, while multi-tasking might be possible, certainly when it comes to immersive experiences a single client is responsible for creating the environments. Exceptions to this include the ‘holograms’ on Windows Mixed Reality/Hololens[2] or Volumes in VisionOS[3]. Both of these are examples of code from different sources that is integrated into a shared experience for the user. Notably both of these systems are primarily used for augmented reality situations, where the analogy of multiple applications matches a conceptual model of different objects in the world having different functions.

An interesting question is thus how far inter-operation between applications can be pushed. Rather than put all the behaviour code into a client that would then have to support code extensibility or interpreted code, why are we not treating the user experience more like an operating system or window manager, where different applications are responsible for different aspects of the experience? That is, how could we assemble an immersive experience from multiple applications?

Currently when we switch between different fully immersive systems each system takes full control of the user experience and is thus responsible for everything from input device interpretation, through simulation to rendering. Of course applications rely on run-time services for device interfacing, but there is little interaction between applications. This can be contrasted with mature desktop operating systems, where we not only have multiple applications and interfaces operating at once, but the screen is composited from multiple applications. Similarly, a modern web browser provides a variety of services over the operating

[1] M^2 Morpheus Platform `https://codex.msquared.io/technology/m2-morpheus-platform`

[2] `https://learn.microsoft.com/en-us/windows/mixed-reality/discover/hologram`

[3] `https://developer.apple.com/visionos/`

system, with complex application being written within distinct pages in a browser that supports multiple windows.

There are precedents for this in research prototypes. The DIVE system was an early peer to peer multi-user system that supported a variety of immersive interfaces (10). The user interface was written in TCL/TK, so the client was dynamically extensible at run-time, but also applications could connect to a running client and insert new 3D objects as interfaces. This was an analogy of how X11 remote applications worked. A slightly different approach is the concept of scene-graph as bus, where multiple application interact using a common scene graph as a shared resource. The scene-graph acts as an interface between applications rather than using any other inter-process communication (28). We can imagine a variety of ways that different spaces could be constructed from volumetrically constrained, or through the window, metaphors for layering and composition. We can also see some hints of this in recent discussions about constructing mixed reality systems that compose real and virtual scenes (e.g. VRCeption (13)).

We therefore suggest that there are more ways that metaverse system could be constructed than simply constructing a large client browser. While efforts to extend web browsers to support larger collaborative experience are very promising demonstration of certain facilities, we believe that the key step is to re-imagine how an immersive interface is built. In Section 4 we will elaborate on one specific goal for client-side interoperability: supporting accessibility.

3 SCALE

Fundamentally, interaction between N users requires $O(N^2)$ interactions between the clients supporting those users. Scalability to large numbers has long been an interest of developers from early distributed simulation days (e.g. see (22)). Early SVR systems used a variety of ways to determine whether or not to enabled communication between pairs of users or entities based on cellular partitions of spaces (17), reasoning about spatial overlap of users (6), by partitioning a graph of nearest neighbours(4) or fixing group size of interested players (7). The general area is sometimes referred to as interest management (18, 9).

Many current systems use a partitioning of users based on a shard or instance models. That is, users are partitioned up into groups of users that can be supported on a simple server. Current SVR services supports small numbers of users (< 50), as do many online games. Network measurement of existing systems (8) or open source examples (11) show that the server is a bottleneck: it has to ingest messages, simulate the world state and then distribute state changes.

There is thus a mismatch between the common vision of the metaverse as a massive online space, and the capabilities of the server and message passing infrastructure. But scalability techniques already note that user interest can be scoped down to the most proximate other users. Three levels of awareness have been identified (see also (25), chapter 12):

- Primary awareness: users are co-located and are proximate to each other. We would expect detailed animation, voice and joint interaction.
- Secondary awareness: awareness that a user is somehow in a reachable space, e.g. within the scope of a server or server network. This now means that rendezvous is possible and can be achieved easily, for example, by virtual travel.
- Tertiary awareness: awareness of the activity of activity anywhere in the population of users with which one is connected (e.g. friend networks), scoped by privacy policies. This to facilitate coordination and rendezvous of users.

Different systems might have additional levels, but it can be a useful exercise to map this to existing systems. Tertiary awareness is more like a social network or it might be labelled ambient awareness. The facilities provided are more about finding if a user is connected and available. The secondary awareness highlights that users are only reachable for primary awareness if they are in the same system. One can imagine tertiary awareness being an intra-system service of awareness, whereas to rendezvous one might need to install more software or download specific resources. Primary awareness is the full bandwidth experience. A distinction to secondary awareness is that in constrained situations, bandwidth and capacity is prioritised to support more relevant (i.e. proximate and important) users rather than users that are more in the background (e.g. far away or not engaged in the same activity).

These distinctions of awareness also highlight that in order to support travel in an apparently large-scale open world, multi-server support will be necessary. This then returns us to the question of partitioning load across servers, but now attempting to support seamless handover between servers. Again, research prototypes exist, (12, 15), but this is definitely an area that needs more research and development.

Finally, we will note that scalability will also have user interface aspects. Users want features to manage privacy and security, and these have to match with the practical routing of messages and distribution of assets. Users also want to travel as groups and maintain cohesion. These present significant challenges to the network infrastructure as primary awareness should reflect user preferences, especially as collaboration is necessary to support certain accessibility features.

4 ACCESSIBILITY

Accessibility is an increasingly important topic in virtual and augmented reality[4] and thus it will be for metaverse systems. Previous efforts have covered aspects such as supporting users with low-vision (29) through to having one user act in support of a second user with restricted movement (26). These considerations pose enormous challenges to the future developers of metaverse spaces: how to support the broadest range of users without all developers having to support all features.

We again highlight that in other domains, such as windowing or smartphone interfaces, there is a variety of support at operating system and interface management layers that help make applications more accessible: from magnifiers to screen readers. Users will want tools that allow them to modify their experiences within the metaverse. Either as plugins to tools (as is done with web browsers) or through some other interoperability mechanism. While VRML97/X3D are no longer used to describe full systems, we would note that within the specification there is a clear responsibility for the browser to provide a set of interaction techniques for certain actions (walk, fly, etc.) but without specifying how these should be implemented[5]. Thus browsers have a lot of flexibility in implementation and can support customisation. While OpenXR[6] provides some functionality to abstract specific of mixed-reality devices, application developers still need to implement the full interpretation of device signals and generate a full-screen interface. Some decoupling of world description from basic interaction would allow users to use the devices that they need or prefer (23).

Finally, we note that accessibility is facilitated by sharing of content. In recent years there has been a significant interest in tools that share immersive experiences, either through live streaming video, or more recently through re-sharing immersive experiences (e.g. see (27)). This touches on the other two themes of

[4] `https://xraccess.org/`
[5] `https://doc.x3dom.org/author/Navigation/NavigationInfo.html`
[6] `https://www.khronos.org/openxr/`

this essay: sharing content is an issue of supporting inter-operation between clients and sharing is one way of achieving scale by re-broadcasting. We should not think of metaverses as simple broadcasting systems, but as interfaces composed from multiple components. For example our recent prototype AccompliceVR (24) adds an overlay network for sharing single user VR experiences running on SteamVR [7] to remote users, and having those remote users appear as avatars superimposed in the original VR application. This was inspired by Vermillion[8] , which allows users to paint other SteamVR applications. Both applications exploit features of the SteamVR/OpenXR compositors and overlay systems which were originally designed to facilitate management of the device environment, application launching and immersive system controls. Such functionality hints at a potential interoperability route.

5 CONCLUSION

In this essay we have highlighted three challenges for future metaverse software. While current SVR software provides many delightful and useful collaborative experiences, there is no clear route to a system that could subsume the wide variety of features and capabilities that are expected. The challenges for any one developer of client software seem almost impossible to overcome with a single client: dealing with scale, interoperating with different services that will need to be implemented by different providers, while supporting accessibility to the broadest population. Our proposal is that we should rethink the role of client software, so that developers can focus on creating small components that are integrated by users or on behalf of users. By integrating lessons learned in other areas of software engineering for distributed applications, interoperability should be possible, which then enables scale and accessibility through allowing composition of systems and interfaces.

REFERENCES

1 .Abilkaiyrkyzy, A., Elhagry, A., Laamarti, F., and Saddik, A. E. (2023). Metaverse Key Requirements and Platforms Survey. *IEEE Access* 11, 117765–117787. doi:10.1109/ACCESS.2023.3325844

2 .Au, W. J. (2008). *The Making of Second Life: Notes from the New World* (New York: Harper Business)

3 .Au, W. J. (2023). *Making a Metaverse That Matters* (Wiley)

4 .Backhaus, H. and Krause, S. (2007). Voronoi-based adaptive scalable transfer revisited. *Proceedings of the 6th ACM SIGCOMM workshop on Network and system support for games - NetGames '07* , 49–54doi:10.1145/1326257.1326266

5 .Ball, M. (2022). *The Metaverse: And How It Will Revolutionize Everything* (Liveright)

6 .Benford, S. and Fahlén, L. (1993). A Spatial Model of Interaction in Large Virtual Environments. In *Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993, Milan, Italy ECSCW '93*, eds. G. de Michelis, C. Simone, and K. Schmidt (Dordrecht: Springer Netherlands). 109–124. doi:10.1007/978-94-011-2094-4_8

7 .Bharambe, A., Douceur, J. R., Lorch, J. R., Moscibroda, T., Pang, J., Seshan, S., et al. (2008). Donnybrook: enabling large-scale, high-speed, peer-to-peer games. *ACM SIGCOMM Computer Communication Review* 38, 389–400. doi:10.1145/1402946.1403002

8 .Cheng, R., Wu, N., Varvello, M., Chen, S., and Han, B. (2022). Are we ready for metaverse? a measurement study of social virtual reality platforms. In *Proceedings of the 22nd ACM Internet Measurement Conference* (New York, NY, USA: Association for Computing Machinery), IMC '22, 504–518. doi:10.1145/3517745.3561417

9 .Delaney, D., Ward, T., and McLoone, S. (2006). On Consistency and Network Latency in Distributed Interactive Applications: A Survey—Part I. *Presence: Teleoperators and Virtual Environments* 15, 218–234. doi:10.1162/pres.2006.15.2.218

10 .Frécon, E., Smith, G., Steed, A., Stenius, M., and Ståhl, O. (2001). An overview of the coven platform. *Presence* 10, 109–127

[7] Valve Corporation, `https://store.steampowered.com/steamvr`

[8] Mountainborn Studios OÜ, `https://vermillion-vr.com/`

11 .Friston, S., Olkkonen, O., Congdon, B., and Steed, A. (2023). Exploring Server-Centric Scalability for Social VR. In *2023 IEEE/ACM 27th International Symposium on Distributed Simulation and Real Time Applications (DS-RT)*. 56–65. doi:10.1109/DS-RT58998.2023.00016

12 .Funkhouser, T. A. (1995). RING: a client-server system for multi-user virtual environments. In *Proceedings of the 1995 symposium on Interactive 3D graphics* (New York, NY, USA: Association for Computing Machinery), I3D '95, 85–ff. doi:10.1145/199404.199418

13 .Gruenefeld, U., Auda, J., Mathis, F., Schneegass, S., Khamis, M., Gugenheimer, J., et al. (2022). VRception: Rapid Prototyping of Cross-Reality Systems in Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22, 1–15. doi:10.1145/3491102.3501821

14 .Havele, A., Polys, N., Benman, W., and Brutzman, D. (2022). The Keys to an Open, Interoperable Metaverse. In *Proceedings of the 27th International Conference on 3D Web Technology* (New York, NY, USA: Association for Computing Machinery), Web3D '22, 1–7. doi:10.1145/3564533.3564575

15 .Iimura, T., Hazeyama, H., and Kadobayashi, Y. (2004). Zoned federation of game servers: A peer-to-peer approach to scalable multi-player online games. In *Proceedings of the ACM SIGCOMM Workshop on Network and System Support for Games, NetGames'04*. 116–120. doi:10.1145/1016540.1016549

16 .Liu, Q. and Steed, A. (2021). Social Virtual Reality Platform Comparison and Evaluation Using a Guided Group Walkthrough Method. *Frontiers in Virtual Reality* 2

17 .Macedonia, M. R., Zyda, M. J., Pratt, D. R., Brutzman, D. P., and Barham, P. T. (1995). Exploiting reality with multicast groups. *IEEE Computer Graphics and Applications* 15, 38–45. doi:10.1109/38.403826

18 .Morse, K. L., Bic, L., and Dillencourt, M. (2000). Interest Management in Large-Scale Virtual Environments. *Presence: Teleoperators and Virtual Environments* 9, 52–68. doi:10.1162/105474600566619

19 .Oliveira, M., Jordan, J., Pereira, J., Jorge, J., and Steed, A. (2009). Analysis Domain Model for Shared Virtual Environments. *International Journal of Virtual Reality* 8, 1–30. doi:10.20870/IJVR.2009.8.4.2745

20 .Osborne, A., Fielder, S., Mcveigh-Schultz, J., Lang, T., Kreminski, M., Butler, G., et al. (2023). Being Social in VR Meetings: A Landscape Analysis of Current Tools. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (New York, NY, USA: Association for Computing Machinery), DIS '23, 1789–1809. doi:10.1145/3563657.3595959

21 .[Dataset] Parisi, T. (2021). The Seven Rules of the Metaverse. `https://medium.com/meta-verses/the-seven-rules-of-the-metaverse-7d4e06fa864c`

22 .Singhal, S. and Zyda, M. (1999). *Networked Virtual Environments: Design and Implementation* (Addison Wesley)

23 .Steed, A. (2019). The Vehicle Pattern for Simplifying Cross-Platform Virtual Reality Development. In *VR Developer Gems* (A K Peters/CRC Press). Num Pages: 13

24 .Steed, A. (2024). Accomplicevr: Lending assistance to immersed users by adding a generic collaborative layer. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (IEEE)

25 .Steed, A. and Oliveira, M. F. (2009). *Networked Graphics: Building Networked Games and Virtual Environments* (Elsevier)

26 .Thiel, F. J. and Steed, A. (2021). " lend me a hand"–extending the reach of seated vr players in unmodified games through remote co-piloting. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (IEEE), 214–219

27 .Thoravi Kumaravel, B. and Wilson, A. D. (2022). Dreamstream: Immersive and interactive spectating in vr. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17

28 .Zeleznik, B., Holden, L., Capps, M., Abrams, H., and Miller, T. (2000). Scene-Graph-As-Bus: Collaboration between Heterogeneous Stand-alone 3-D Graphical Applications. *Computer Graphics Forum* 19, 91–98. doi:10.1111/1467-8659.00401

29 .Zhao, Y., Cutrell, E., Holz, C., Morris, M. R., Ofek, E., and Wilson, A. D. (2019). SeeingVR: A Set of Tools to Make Virtual Reality More Accessible to People with Low Vision. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19* (Glasgow, Scotland Uk: ACM Press), 1–14. doi:10.1145/3290605.3300341

A Brief Review of Asymmetric Social VR

Sebastian Friston [1,*] **and Anthony Steed** [1]

[1] *University College London, Computer Science, London, UK*

Correspondence*:
Corresponding Author
sebastian.friston@ucl.ac.uk

1 INTRODUCTION

A commonality among metaverse definitions, is of a persistent platform that breaks through the constraints of time and space to expand into the physical world (36). Distinctively, this breakthrough can also go the other way. There is the promise of inhabiting the metaverse as a space of its own, and moving freely between the real and virtual. As a communication medium, the metaverse necessarily has social attributes. This blending of the real and virtual in a social context means the metaverse must support *asymmetry* - in physical modality, and social interaction.

Contemporary Social VR systems (SVRs) are sometimes likened to the embodied-virtual part of the metaverse, but could be more accurately seen as an archipelago of incompatible "microsverses". While many systems succeed in one or more ways, they are still too limited to act as a standalone embodied 'universe', let alone the metaverse. Specifically, systems are limited in the range of social situations they support. Social situations or interactions can be quantified, and categorised into discrete types (4). This categorisation can be applied to computer mediated interactions (12).

Day-to-day, people move between many types without conscious effort. Within these types, individuals take on different roles: customer, teller, moderator, entertainer, etc. These roles are enabled or mediated by social conventions, but also by space, and by technology. For example, a teller performs their function using point-of-sale equipment. An event manager has access to a guest list and radio, and spaces unavailable to the general audience. Contemporary SVR however makes few concessions to these roles. Examples of important requirements include: (1) user privileges and (2) multi-modal interaction or interoperability.

The real-world is a collection of inter-operating systems to which users have differing levels of control. The metaverse must be equally inter-operative if it is to play a role. This can be as simple as integrating a virtual meeting space with an online ticketing service - however this is still a problem for the current generation of SVRs. Most SVRs offer only one type of interaction metaphor - embodied 3D. However, embodied 3D is not necessarily the best interface for a task. In the real world, people often use specially designed interfaces, such as mixing desks. Further, if the metaverse is going to encompass more of people's

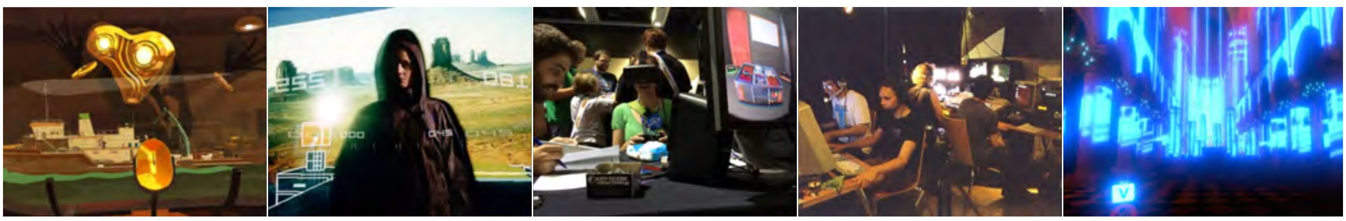


Figure 1. SVRs using tools to facilitate different social roles (left-to-right): The Under Presents (6), Desert Rain (19), Keep Talking and Nobody Explodes (18), Out Of This World (11) and Welcome To The Other Side (33)

daily lives, it must also encompass more of the mixed-reality continuum, so users can share a social situation across devices or platforms. SVRs also limit users' control over the virtual world or other users, prohibiting them from creating the pre-conditions for a large range of social situations. There are many reasons SVRs don't offer this yet, from commercial to security considerations. However, so long as SVRs remain walled-gardens, users will be continually forced outside of them, and they will not form part of the metaverse.

While the above is true of many contemporary SVRs, it is not true of all. In the history of social VR, there are many examples of asymmetric SVRs: those that are multi-modal, cross-platform and support different social roles. These systems demonstrate the importance of asymmetry and technical openness in building the metaverse, and they illustrate some the technical challenges involved - either through successful solutions, or the anecdotes of those who have tried.

2 ASYMMETRIC SOCIAL VR

In this section we briefly describe a number of asymmetric SVR platforms, applications and tool-kits as a starting point. They show how roles are enabled through different immersive and non-immersive tools.

2.1 Commercial

The Under Presents (6) is an immersive VR theatre experience. The same client is used by audience members and performers, however performers have access to a virtual dressing room where they can outfit their avatar with different appearances and abilities. For example, the ability to teleport audience members, with the goal of enhancing their performance. *Out Of This World* (11), was a VR-based game show. The producers combined dedicated virtual camera controls with real-world analogue mixers to live-broadcast television from inside an SVR, so production staff could perform their jobs in a familiar way and integrate with the studios' existing technology. *Desert Rain* (19) was a location-based mixed reality theatre experience. Thematically, the breakdown between the real and virtual world was used for dramatic effect. To this end, audience members would interact with functional props such as swipe-cards and a variety of displays. The producers orchestrated it with control systems that could observe both the physical and digital experiences simultaneously. In the VR game *Keep Talking and Nobody Explodes* (30), asymmetry is embraced as a mechanic. Non-embodied users participate entirely in the real-world, using printed materials to help an immersed user complete a task.

Vitillo (33) staged a virtual concert in VRChat. VRChat allows building custom environments in the Unity Editor, and has a basic scripting language as well. However VRChat was limited in terms of capacity, leaving most attendees to view pre-recorded versions of the concert out-of-sync. Additionally, due to limitations in permissions and scripting, the team had difficulties with synchronisation, and handling disruptive users. Saffo et al (25) replicated lab-based VR studies in VRChat. They similarly noted how VRChat's graphical programming language allowed them to re-create experiment logic, but a major limitation was the inability to communicate data outside of the client. Screenshots from these various systems are shown in Figure 1.

Cross-platform support generally is rare. The most common form is cross-play between an immersive client and a web-browser, predominantly in productivity tools. For example, Spatial (28) and Glue (10). Support is less common, though not non-existent, in games (e.g. Roblox (26)). Data exchange remains a problem; for example, organisers of a virtual innovation event were unable to marry-up Virbela attendees with ticket holders managed on another platform (17).

2.2 Platforms & Toolkits

Systems such as SCIVE (20) and FlowVR (1) aimed to allow building distributed VR systems as an arrangement of heterogenous nodes, and defined protocols through which new components could be added. These were followed by systems such as Simulator-X (21). This took an actor-based approach where the distributed system was expressed as a shared state common to all nodes. Heterogenous processes (the actors) could manipulate the shared state through events.

Nowadays, most VR applications are built using engines such as Unity or Unreal. Unreal has a built-in networking model. For Unity there are a number of networking SDKs (some of which are also available to Unreal developers) including Photon (8), DarkRift (7) & MLAPI (31). These integrate with any user code within Unity, but none explicitly support cross-platform connectivity outside Unity binaries. The Unity-compatible SVR SDK Ubiq includes a first-party Javascript client library explicitly for supporting heterogenous cross-platform behaviour in SVR (9). Ubiq operates by sending arbitrary messages between individual components in an Entity Component System, allowing single entities to be shared between platforms and to distribute logic between them.

2.3 Standards

The origins of networked VR began with distributed graphics for cluster rendering, supported by APIs such as Blue-c (23) and Wolverine (5). As a low-level API example, OpenGL has proved popular for distributed rendering. Standardisation has made it amenable to transparent distributed implementations (e.g. (15)) as well as cross-platform support (e.g. WebGL). At a higher level, Zeleznik et al. (37) proposed the scene graph as a common representation for exchanging data between different 3D applications. Hesina et al (14) created a distributed implementation of Open Inventor. For the web, the array of standards is now deep enough to support a federation of complete 3D XR worlds running entirely in the browser. These include XR device interaction (WebXR) as well as interactive 3D scene descriptions (X3DOM) and interchange of video (MP4), images (JPEG) and 3D models (glTF) (13). Mozilla Hubs is a proof-of-concept SVR application built on these technologies (22).

For distributed computing, the Virtual Reality Peripheral Network (VRPN) provides a device and network independent standard for combining VR peripherals (16). A more recent example is the Robot Operating System (ROS) (29), a set of standards and tools for building graphs of message passing components. ROS is designed for robotics but could build any distributed computing system.

On the web, the W3C maintains standards such as Simple Object Access Protocol (SOAP) for arbitrary object passing (34). WebSockets and HTTP, for transport, are also standardised. The existence of non-standard, if open, solutions such as Socket.io (27) suggests there may still be a standards gap at the messaging level however.

3 DISCUSSION AND TECHNICAL CHALLENGES

3.1 Authority and Roles

In the real world, roles are enabled through both social convention and technology, such as technology to mediate access to spaces like key-cards. As space in the metaverse is defined digitally, there are many ways to control access: Who one is vs. What one has vs. Where one is. In *The Under Presents*, actors gain abilities in a virtual dressing room (Where one is). In RecRoom, users can unlock abilities through ‘room keys’ - digital tokens that they possess (What one has). Virbella adopts a traditional user-admin-owner model where user accounts are assigned a fixed role (32) that follows them regardless of avatar or location (Who one is). In VRChat, the first person to enter a 'room' is the administrator, and has elevated privileges. Vitillo, for example, pre-created multiple rooms to ensure their team had control over all of them. With regards to standards specifically for authority, the only ecosystem with mature open standards is the web (35). In addition to standards such as Single Sign On (3), it is particularly interesting to consider the security model around code. Browsers sandbox untrusted code, a different approach taken to typical operating systems which use signing to execute trusted code. It is more than just a sandbox of one program however. Using certificates, browsers are able to create barriers between *domains*, allowing the server to be confident that the browser is only running the code the server itself delivered - there is trust between components of the distributed application, but not between the host system, browser, and application itself. This is important, because even if platforms like VRChat create local scripting languages, they will still be limited unless they can interact with a wider ecosystem in a trusted way.

3.2 Shared Representations vs Messaging

Roles are also mediated through tools. Certain roles are best performed using dedicated interfaces. For example as in *Out Of This World*. While Vitillo used the embodied VRChat client for everything, this was not by choice (33). Seamless movement along the mixed reality continuum should come naturally to the metaverse, however there are a number of implementation considerations. One of those is the distinction between shared representations and messaging systems.

Specifications such as X3DOM or Open Inventor effectively declare virtual machines. Developers can rely on any feature being available on any platform, but all platforms must implement the entire scope. In SVR, we see this in platforms such as VRChat, where users can use the Unity Editor to build environments, but only utilising existing Unity Components via the Scene file or scripting language. This is in contrast to distributed compute systems such as VRPN or ROS, which rely on shared message structures. Any component can create or receive messages, allowing a function to be implemented on only one platform.

Both approaches have their advantages. Arguably however any dichotomy between representation-based and message-based sharing is false. For example, running in Js, there is nothing to stop an X3DOM tab opening a WebSocket to a tracker, or synchronising this data with other peers using the SGAAB approach (37). While it is in theory straightforward to open the messaging layer of an SVR, the transport layer is only one part of building distributed applications. Systems must also agree on an addressing scheme, so messages are exchanged between the correct components in user code. This requires a number of functions to operate the same way between platforms and languages. An addressing scheme means defining a number of concepts that must be shared. For example, if addressing entities, there must the concept of an entity. If a language were not object oriented, for example, it would have to be given this concept in order to interact with the system. Some systems such as Simulator-X used semantic annotation on events, increasing the abstraction but also the potential for platforms to interact. In this way, the space between nothing but a shared transport layer, and a complete standard such as X3DOM, can be seen less as a delineation between two construction methods, and more of a continuum of abstractions. It is simply there are no popular examples that sit halfway to illustrate this yet.

Another representation consideration is how to move digital assets between platforms with different abilities. For example, it is easy to conceive of a system to transfer ownership of the materials, geometry and perhaps even the short-lived deterministic animations (such as bubbling) of a 'health potion' between two applications, but not so to transfer the function of that object, e.g. its effect on your health or personal capabilities in the game, if any, in a generalisable way.

3.3 Feature-as-a-service

As systems open up, there are many opportunities created by federating different black-box systems. In Ubiq-Genie (24), the SDK was used to connect text-to-voice, voice-to-text & a Large Language Model service to build an NPC which users could speak with. A contemporary commercial example is Ready Player Me (RPM) (2). Users design avatars in browsers which can then be loaded into a variety of SVR platforms. These examples illustrate a concept that could be called *feature-as-a-service*, where closed platforms use open standards to provide a feature or subsystem to multiple metaverse applications at once. There are many advantages of this model, which make it an exciting development. It frees developers from re-implementing the same subsystems, reducing the barrier to entry. It provides a more consistent experience between applications. Finally, commodification of certain features will bring down their costs overall, resulting in more metaverse content in sum.

4 CONCLUSION

In this paper we have briefly reviewed a number of SVR systems to derive lessons on the importance of asymmetry in the metaverse. Some of these are platforms in their own right, and others are applications created on platforms. What they have in common is that whether by choice or necessity, they have embraced asymmetry through a combination of immersive and non-immersive tools. In doing so, these systems have been able to explore social situations outside the scope of many contemporary SVRs. Arguably, the metaverse cannot fulfil its potential as a walled garden, because users will continually be forced outside it. Interoperability is necessary for expansion. These systems can offer some lessons on what this requires, and how it might one day be achieved.

REFERENCES

1 .Allard, J., Gouranton, V., Lecointre, L., Limet, S., Melin, E., Raffin, B., et al. (2004). FlowVR: A Middleware for Large Scale Virtual Reality Applications. In *Parallel Processing. 10th International Euro-Par Conference* (Springer Berlin Heidelberg). 497–505. doi:10.1007/978-3-540-27866-5_65. ISSN: 03029743

2 .Altundas, S. and Karaarslan, E. (2023). Cross-platform and personalized avatars in the metaverse: Ready player me case. In *Digital Twin Driven Intelligent Systems and Emerging Metaverse* (Singapore: Springer Nature Singapore). 317–330. doi:10.1007/978-981-99-0252-1_16

3 .Auth0 (2024). Single sign on. https://auth0.com/docs/authenticate/single-sign-on

4 .Bassetti, C. (2017). Social Interaction in Temporary Gatherings. In *Group and Crowd Behavior for Computer Vision* (Elsevier). 15–28. doi:10.1016/B978-0-12-809276-7.00003-5

5 .Chardavoine, F., Ageneau, S., and Ozell, B. (2005). Wolverine: A Distributed Scene-Graph Library. *Presence: Teleoperators and Virtual Environments* 14, 20–30. doi:10.1162/1054746053890297

6 .Coulombe, A. (2020). Let's dive into the under presents: Tempest. https://medium.com/alive-in-plasticland/lets-dive-into-the-under-presents-tempest-pt-1-2d1ef2168c5f

7 .Dark Rift (2021). DarkRift Networking. https://www.darkriftnetworking.com/

8 .Exit Games (2021). Photon. https://www.photonengine.com/

9 .Friston, S., Congdon, B., Numan, N., Brandstätter, K., Izzouzi, L., Thiel, F., et al. (2023). Extending the open source social virtual reality ecosystem to the browser in ubiq. In *Proceedings of the 28th International ACM Conference on 3D Web Technology* (New York, NY, USA: Association for Computing Machinery), Web3D '23. doi:10.1145/3611314.3615903

10 .Glue (2023). Glue. https://www.glue.work/

11 .Greenhalgh, C., Bowers, J., Walker, G., Wyver, J., Benford, S., and Taylor, I. (1999). Creating a live broadcast from a virtual environment. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques - SIGGRAPH '99* (ACM Press), 375–384. doi:10.1145/311535.311591

12 .Groh, G., Lehmann, A., Wang, T., Huber, S., and Hammer, F. (2010). Applications for social situation models. In *Proc. of the IADIS Int. Confs. Informatics 2010, Part of the MCCSIS 2010*. 75–82

13 .Havele, A., Polys, N., Benman, W., and Brutzman, D. (2022). The Keys to an Open, Interoperable Metaverse. In *The 27th International Conference on 3D Web Technology* (Evry-Courcouronnes France: ACM), 1–7. doi:10.1145/3564533.3564575

14 .Hesina, G., Schmalstieg, D., Furhmann, A., and Purgathofer, W. (1999). Distributed Open Inventor. In *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '99* (New York, New York, USA: ACM Press), 74–81. doi:10.1145/323663.323675

15 .Humphreys, G., Eldridge, M., Buck, I., Stoll, G., Everett, M., and Hanrahan, P. (2001). WireGL: A scalable graphics system for clusters. In *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*. 129–140

16 .Ii, R. M. T., Hudson, T. C., Seeger, A., Weber, H., Juliano, J., and Helser, A. T. (2001). VRPN: A Device-Independent, Network-Transparent VR Peripheral System. In *Proceedings of the ACM symposium on Virtual reality software and technology*. doi:10.1145/505008.505019

17 .Jauhiainen, J. S. (2021). Entrepreneurship and Innovation Events during the COVID-19 Pandemic: The User Preferences of VirBELA Virtual 3D Platform at the SHIFT Event Organized in Finland. *Sustainability* 13, 3802. doi:10.3390/su13073802

18 .Kohler, C. (2014). Defuse a bomb with friends in this brilliant oculus game. https://www.wired.com/2014/09/keep-talking-nobody-explodes/

19 .Koleva, B., Taylor, I., Benford, S., Fraser, M., Greenhalgh, C., Schnädelbach, H., et al. (2001). Orchestrating a mixed reality performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle Washington USA: ACM), 38–45. doi:10.1145/365024.365033

20 .Latoschik, M. E., Fröhlich, C., and Wendler, A. (2006). Scene Synchronization in Close Coupled World Representations using SCIVE. *The International Journal of Virtual Reality* 5, 47–52

21 .Latoschik, M. E. and Tramberend, H. (2011). Simulator x: A scalable and concurrent architecture for intelligent realtime interactive systems. In *2011 IEEE Virtual Reality Conference*. 171–174. doi:10.1109/VR.2011.5759457

22 .Mozilla Corporation (2021). Mozilla Hubs. https://hubs.mozilla.com

23 .Naef, M., Lamboray, E., Staadt, O., and Gross, M. (2003). The blue-c distributed scene graph. In *Proceedings of the 2003 IEEE Virtual Reality Conference* (IEEE Comput. Soc), 275–276. doi:10.1109/VR.2003.1191157. ISSN: 1087-8270

24 .Numan, N., Giunchi, D., Congdon, B., and Steed, A. (2023). Ubiq-Genie: Leveraging External Frameworks for Enhanced Social VR Experiences. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (Shanghai, China: IEEE), 497–501. doi:10.1109/VRW58643.2023.00108

25 .Saffo, D., Di Bartolomeo, S., Yildirim, C., and Dunne, C. (2021). Remote and Collaborative Virtual Reality Experiments via Social VR Platforms. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama Japan: ACM), 1–15. doi:10.1145/3411764.3445426

26 .Sammy, K. (2022). Play roblox on web browser. https://alvarotrigo.com/blog/roblox-web-browser

27 .SocketIO (2023). Socketio. https://socket.io/docs/v4/how-it-works/

28 .Spatial (2023). Spatial.io. https://spatial.io/

29 .Stanford Artificial Intelligence Laboratory et al. (2018). Robotic operating system. https://www.ros.org

30 .Steel Crate Games (2023). The co-op bomb defusing party game. https://keeptalkinggame.com/

31 .Unity Technologies (2021). MLAPI - Unity Multiplayer Networking. https://docs-multiplayer.unity3d.com/

32 .Virbela (2022). Campus roles. https://support.virbela.com/s/article/Campus-Roles

33 .Vitillo, A. (2021). Behind the scenes of the vr concert welcome to the other side. https://skarredghost.com/2021/01/05/jean-michel-jarre-vr-concert-postmortem/

34 .W3C (2023). Simple object access protocol. https://www.w3.org/TR/soap/

35 .W3C (2023). W3c security activity. https://www.w3.org/Security/

36 .Wang, H., Ning, H., Lin, Y., Wang, W., Dhelim, S., Farha, F., et al. (2023). A Survey on the Metaverse: The State-of-the-Art, Technologies, Applications, and Challenges. *IEEE Internet of Things Journal* 10, 14671–14688. doi:10.1109/JIOT.2023.3278329

37 .Zeleznik, B., Holden, L., Capps, M., Abrams, H., and Miller, T. (2000). Scene-Graph-As-Bus: Collaboration between Heterogeneous Stand-alone 3-D Graphical Applications. *Computer Graphics Forum* 19, 91–98. doi:10.1111/1467-8659.00401

Acoustic coherence in extended reality: challenges and strategies.

Joanna Luberadzka [1,*], **Enric Guso Muñoz** [1,2], **Umut Sayin** [1], **Adan Garriga Torres** [1]

[1] *Eurecat, Centre Tecnològic de Catalunya, Tecnologies Multimèdia, Barcelona*
[2] *Universitat Pompeu Fabra, Music Technology Group, Barcelona*

Correspondence*:
joanna.luberadzka@eurecat.org

ABSTRACT

In recent years, extended reality (XR) has gained interest as a platform for human communication, with the emergence of the ”Metaverse” promising to reshape interactions. At the same time, concerns about harmful behavior and criminal activities in virtual environments have increased. This paper explores the potential of technology to support social harmony within XR, focusing specifically on audio aspects. We introduce the concept of acoustic (in)coherence and discuss why it is crucial for smooth interaction. We further explain the challenges of speech communication in XR, including noise and reverberation, and review sound processing methods to enhance the auditory experience. Finally, we present a VR demonstration in which users can perceive differences between coherent and incoherent auditory scenes and listen to several audio enhancement techniques while being immersed in the virtual scene.

Keywords: virtual acoustic simulation, binaural audio, acoustic matching, signal enhancement, deep learning

1 INTRODUCTION

The idea of XR as a platform for human communication has become increasingly popular in recent years. It has led to the emergence of the concept of the ”Metaverse,” which is speculated to become the leading medium of communication in the future (Dzardanova et al. (2022); Dwivedi et al. (2022)). The primary aim of such virtual worlds is to enable individuals from different locations to interact within a shared audiovisual environment. In parallel to the widespread appreciation for the technological potential and the promise of fostering social connections, there is also a consensus that the Metaverse could facilitate toxic behavior and pose threats of criminal activity (Gómez-Quintero et al. (2024)). In response to this, there has been a growing interest in studying ways in which technology could promote social harmony and inclusive behavior within virtual worlds. In this brief paper, we take a closer look at this issue from the perspective of audio technology. We identify challenges related to speech communication in XR and discuss sound processing methods that could improve the overall auditory experience in XR. We support this discussion with a VR demonstration featuring spatial sound and selected audio enhancement methods.

2 CHALLENGES

Human interaction heavily relies on listening, making high-quality sound crucial for effective communication. Various studies illustrate how factors such as noise, reverberation, and the inaccuracy of binaural cues can impair our ability to understand speech, localize sound sources in space, or concentrate on specific sounds. Even when the linguistic message is comprehended, reduced audio quality escalates the effort required for listening, resulting in measurable physiological changes in the body, typically associated with psychosocial stressors (Francis and Love (2020)). Consequently, social interactions are likely to be less harmonious in environments with poor acoustics.

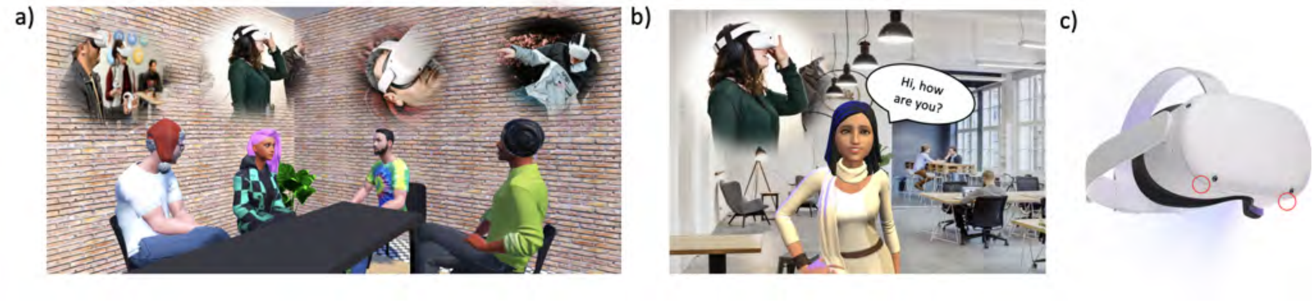


Figure 1. Common auditory scenes in XR. a) Meeting in a virtual room, b) Meeting with an augmented avatar, c) microphone location in the Quest 2 VR Headset.

Figure 1.a illustrates a common VR meeting scenario where four users, depicted as avatars, gather around a virtual table to engage in a conversation. Because the users are physically located in different spaces, their headset microphones, apart from speech, capture additional environment-specific disturbances e.g. noise, background sounds, wind or movement disturbances, and reverberation. Additionally, microphone positioning may not be optimal for capturing voice with high quality (See Fig.1.c). The final mixture of sounds delivered to the user originates from several different environments and may create a perceptual mismatch which we refer to as *acoustic incoherence*. Acoustic incoherence may arise from the evident contrast between the acoustics of individual sound sources or from the discrepancies between perceived sounds and their visual representations within the virtual environment. In augmented reality scenarios (Fig.1.b), where remotely recorded sounds must seamlessly integrate with the user's real soundscape, such incoherence may be even more noticeable.

3 STRATEGIES

Binaural technology is a foundation for most audio experiences in virtual and augmented reality. Through a rigorous preparation of the signal in the left and right audio channels, an auditory illusion of sound being placed at a specific location inside a defined environment can be created. In an XR meeting scenario, the voice captured by the headset on one user's end becomes a virtual audio source for the other participants. Figure 2 illustrates a possible audio processing pipeline applied to such a recording before reaching the listener's ears: Initially, the sound emitted by User 1 in Environment 1 – s_1e_1 – is captured by the microphones of the headset. Before transmitting the signal to User 2, environmental disturbances degrading the recording must be eliminated through *denoising* and *dereverberation* to obtain the clean source signal s_1. In parallel, the originally captured signal serves as a source of environmental information e_1 extracted in the *acoustic characterization* step, essential for later adapting the audio received from User 2. Once the recording is cleaned of unwanted disturbances, it can be passed as input to the binaural rendering stage. Binaural rendering involves a) *spatialization*, which creates a perception of a specific direction according to the relative position and orientation between the source and receiver, and b) *acoustic matching*, which applies the acoustic properties of the target room. The exact order of these operations depends on the chosen binaural rendering method (See Gari et al. (2022) for a review). Finally, the spatialized sound with modified room acoustic properties – $s_1 e_2$ – is delivered to the user's ears. It's noteworthy that in VR, as the user moves their head, the relative direction of sound arrival changes, requiring real-time updates of the entire binaural rendering block.

The depicted pipeline specifically addresses augmented reality, where the target acoustic space is defined by the user's actual location. Consequently, the properties of the target environment must be estimated through acoustic characterization. In the graph, the user's own voice recording serves as the source of this information. However, alternative methods exist for estimating room acoustic properties, such as using images as a source of information (Chen et al. (2022)). In VR setups, this step is typically unnecessary since the reverberation can be simulated from scratch using geometric methods based on polygon meshes in the scene (Välimäki et al. (2016)).

Task	**Traditional approaches**	**Deep-learning approaches**
Denoising	Virag (1999); Krishnamoorthy and Prasanna (2009)	Yuliani et al. (2021); Pascual et al. (2017)
De-reverberation	Li and Deng (2021); Nakatani et al. (2006)	Ochieng (2023); Su et al. (2020b)
Acoustic characterization	Ratnam et al. (2003); Kendrick et al. (2007); Hua (2002)	Martin et al. (2023); Steinmetz et al. (2021)
Binaural synthesis	Cuevas Rodriguez et al. (2022); Rafaely et al. (2022)	Huang et al. (2022); Lluís et al. (2022); Zhu et al. (2024)
Acoustic matching	Välimäki et al. (2016); Peters et al. (2012)	Koo et al. (2021); Im and Nam (2024); Su et al. (2020a)

Table 1. Selected papers discussing either traditional or deep-learning-based approaches for the sub-tasks of virtual acoustic simulation.

All the aforementioned sub-tasks have been subjects of research in audio and acoustics, leading to diverse approaches utilized in a broad range of audio processing devices. Notably, recent advancements in artificial intelligence are beginning to change the landscape of audio technology. Besides numerous deep learning solutions for individual signal processing tasks, there is a trend of replacing entire conventional audio processing pipelines with end-to-end neural approaches. Table 1 outlines a non-exhaustive list of examples comparing traditional signal-processing-based solutions to their deep-learning-based counterparts.

4 DEMONSTRATION

Building on our previous discussion, we designed a simple audiovisual environment in **Unity**, simulating a typical meeting in VR, where four participants engage in a conversation. Users of our demo[1] are virtually placed at the same table as the avatars and have the opportunity to observe and listen to their conversation, while also being able to modify audio processing via the user interface.

The signal flow of our application is depicted in Figure 4: Each avatar is assigned to an audio source. The clean audio sources are first corrupted with a variety of noises and reverberations to mimic a real scenario (i.e. each participant located in a different noisy space)[2]. Next, the audio is processed by one (or none) of the signal enhancement methods listed in the GUI (See Fig.3 for details). The enhanced audio is spatialized based on the relative position between the avatar and the listener and placed in a simulated acoustic environment i.e. a shoebox room corresponding to the visual scene. We use the **Steam**

[1] Video available at `https://www.youtube.com/watch?v=SaXMYn8b3eg&t=124s`

[2] We utilized **DiPCo** dataset for the conversations (Van Segbroeck et al. (2019)), **ACE** database for room impulse responses (Eaton et al. (2015)), **AVAD-VR** for anechoic recordings of instruments (Thery and Katz (2019)), and **Freesound** for other background sounds.

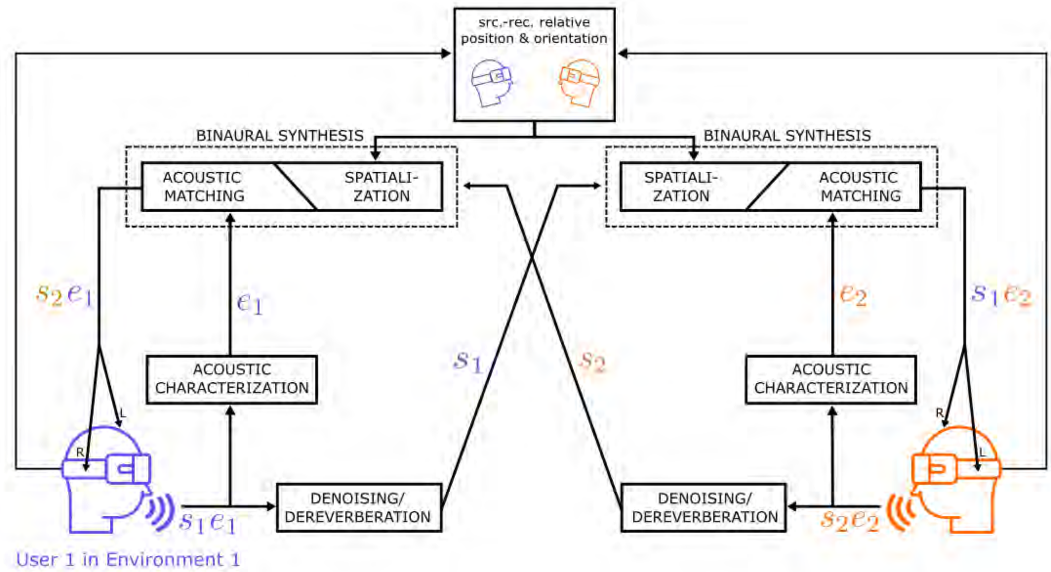


Figure 2. Diagram depicting a possible audio processing pathway for augmented reality conversations.

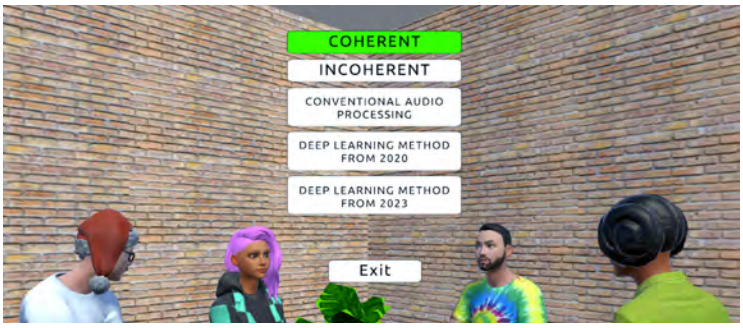


Figure 3. User interface for selecting one of the signal enhancement options: COHERENT: clean original without signal enhancement; INCOHERENT: noisy, reverberant audio without signal enhancement; CONVENTIONAL AUDIO PROCESSING: spectral noise gating (Sudheer Kumar et al. (2023)); DEEP LEARNING METHOD FROM 2020: Facebook Denoiser (Defossez et al. (2020)); DEEP LEARNING METHOD FROM 2023: DeepFilterNet (Schröter et al. (2023))

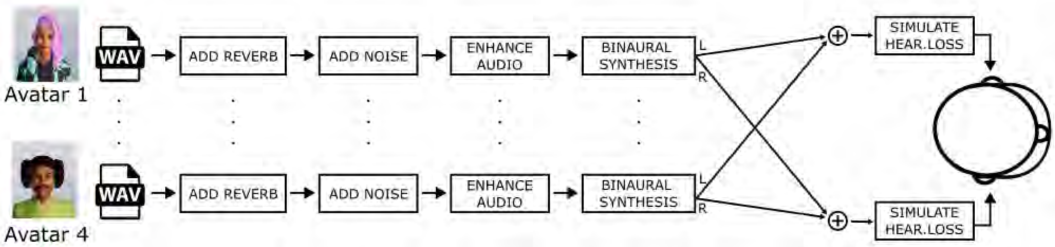


Figure 4. Signal flow in the proposed VR demo.

Audio spatializer plugin. Additionally, a hearing loss simulator feature is included in the application to showcase how the acoustic environment would be perceived by individuals with hearing impairments. We use **3DTuneIn** hearing loss simulator plugin (Cuevas-Rodríguez et al. (2019)). Apart from spatialization and hearing loss simulation, the signal is processed offline. In the future, we plan to implement a fully real-time version which will require adapting the deep-learning models to operate on a frame-by-frame basis.

Informal interviews with participants of the demo confirm that acoustic coherence has a crucial impact on the willingness to take part in the VR meeting. Users reported that when no signal enhancement was applied, they had a hard time following the conversation and associating individual voices with the avatars. While it was clear that all speech enhancement methods improve the auditory experience to some extent, participants had distinct preferences when rating the methods. This observation indicates that there are important differences between the enhancement strategies in terms of the introduced distortions. Further studies are necessary to quantify how such distortions interact with the binaural synthesis and to what extent various aspects of acoustic coherence effectively influence the social interactions in XR.

5 CONCLUSIONS

In this brief overview article, we introduced the concept of acoustic coherence in XR. We provided a concise explanation of the importance of coherent auditory scenes in speech communication and highlighted the challenges of achieving ideal audio signals in remote XR meeting scenarios. Through our scientific discussion and VR demo, we aim to motivate further advancement in audio processing methods for XR, as well as the investigation of their perceptual and social implications.

REFERENCES

Chen, C., Gao, R., Calamia, P., and Grauman, K. (2022). Visual acoustic matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18858–18868

Cuevas-Rodríguez, M., Picinali, L., González-Toledo, D., Garre, C., de la Rubia-Cuestas, E., Molina-Tanco, L., et al. (2019). 3d tune-in toolkit: An open-source library for real-time binaural spatialisation. *PloS one* 14, e0211899

Cuevas Rodriguez, M. et al. (2022). 3d binaural spatialisation for virtual reality and psychoacoustics

Defossez, A., Synnaeve, G., and Adi, Y. (2020). Real time speech enhancement in the waveform domain. *arXiv preprint arXiv:2006.12847*

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., et al. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management* 66, 102542

Dzardanova, E., Kasapakis, V., Gavalas, D., and Sylaiou, S. (2022). Virtual reality as a communication medium: a comparative study of forced compliance in virtual reality versus physical world. *Virtual Reality* 26, 737–757

Eaton, J., Gaubitch, N. D., Moore, A. H., and Naylor, P. A. (2015). The ace challenge—corpus description and performance evaluation. In *2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)* (IEEE), 1–5

Francis, A. L. and Love, J. (2020). Listening effort: Are we measuring cognition or affect, or both? *Wiley Interdisciplinary Reviews: Cognitive Science* 11, e1514

Gari, S. V. A., Robinson, P. W., and Calamia, P. T. (2022). Room acoustic characterization for binaural rendering: From spatial room impulse responses to deep learning. In *International Congress on Acoustics*

Gómez-Quintero, J., Johnson, S. D., Borrion, H., and Lundrigan, S. (2024). A scoping study of crime facilitated by the metaverse. *Futures* , 103338

Hua, Y. (2002). Blind methods of system identification. *Circuits, Systems and Signal Processing* 21, 91–108

Huang, W. C., Markovic, D., Richard, A., Gebru, I. D., and Menon, A. (2022). End-to-end binaural speech synthesis. *arXiv preprint arXiv:2207.03697*

Im, J. and Nam, J. (2024). Diffrent: A diffusion model for recording environment transfer of speech. *arXiv preprint arXiv:2401.08102*

Kendrick, P., Li, F. F., Cox, T. J., Zhang, Y., and Chambers, J. A. (2007). Blind estimation of reverberation parameters for non-diffuse rooms. *Acta Acustica united with Acustica* 93, 760–770

Koo, J., Paik, S., and Lee, K. (2021). Reverb conversion of mixed vocal tracks using an end-to-end convolutional deep neural network. In *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE), 81–85

Krishnamoorthy, P. and Prasanna, S. M. (2009). Temporal and spectral processing methods for processing of degraded speech: a review. *IETE Technical Review* 26, 137–148

Li, Y. and Deng, L. (2021). An overview of speech dereverberation. In *Proceedings of the 8th Conference on Sound and Music Technology: Selected Papers from CSMT* (Springer), 134–146

Lluís, F., Chatziioannou, V., and Hofmann, A. (2022). Points2sound: from mono to binaural audio using 3d point cloud scenes. *EURASIP Journal on Audio, Speech, and Music Processing* 2022, 1–15

Martin, I., Pastor, F., Fuentes-Hurtado, F., Belloch, J., Azpicueta-Ruiz, L., Naranjo, V., et al. (2023). Predicting room impulse responses through encoder-decoder convolutional neural networks. In *2023 IEEE 33rd International Workshop on Machine Learning for Signal Processing (MLSP)* (IEEE), 1–6

Nakatani, T., Kinoshita, K., and Miyoshi, M. (2006). Harmonicity-based blind dereverberation for single-channel speech signals. *IEEE Transactions on Audio, Speech, and Language Processing* 15, 80–95

Ochieng, P. (2023). Deep neural network techniques for monaural speech enhancement and separation: state of the art analysis. *Artificial Intelligence Review* 56, 3651–3703

Pascual, S., Bonafonte, A., and Serra, J. (2017). Segan: Speech enhancement generative adversarial network. *arXiv preprint arXiv:1703.09452*

Peters, N., Choi, J., and Lei, H. (2012). Matching artificial reverb settings to unknown room recordings: A recommendation system for reverb plugins. In *Audio Engineering Society Convention 133* (Audio Engineering Society)

Rafaely, B., Tourbabin, V., Habets, E., Ben-Hur, Z., Lee, H., Gamper, H., et al. (2022). Spatial audio signal processing for binaural reproduction of recorded acoustic scenes–review and challenges. *Acta Acustica* 6, 47

Ratnam, R., Jones, D. L., Wheeler, B. C., O'Brien Jr, W. D., Lansing, C. R., and Feng, A. S. (2003). Blind estimation of reverberation time. *The Journal of the Acoustical Society of America* 114, 2877–2892

Schröter, H., Rosenkranz, T., Maier, A., et al. (2023). Deepfilternet: Perceptually motivated real-time speech enhancement. *arXiv preprint arXiv:2305.08227*

Steinmetz, C. J., Ithapu, V. K., and Calamia, P. (2021). Filtered noise shaping for time domain room impulse response estimation from reverberant speech. In *2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)* (IEEE), 221–225

Su, J., Jin, Z., and Finkelstein, A. (2020a). Acoustic matching by embedding impulse responses. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE), 426–430

Su, J., Jin, Z., and Finkelstein, A. (2020b). Hifi-gan: High-fidelity denoising and dereverberation based on speech deep features in adversarial networks. *arXiv preprint arXiv:2006.05694*

Sudheer Kumar, E., Jai Surya, K., Yaswanth Varma, K., Akash, A., and Nithish Reddy, K. (2023). Noise reduction in audio file using spectral gatting and fft by python modules. In *Recent Developments in Electronics and Communication Systems* (IOS Press). 510–515

Thery, D. and Katz, B. F. (2019). Anechoic audio and 3d-video content database of small ensemble performances for virtual concerts. In *Intl Cong on Acoustics (ICA)*

Välimäki, V., Parker, J., Savioja, L., Smith, J. O., and Abel, J. (2016). More than 50 years of artificial reverberation. In *Audio engineering society conference: 60th international conference: dreams (dereverberation and reverberation of audio, music, and speech)* (Audio Engineering Society)

Van Segbroeck, M., Zaid, A., Kutsenko, K., Huerta, C., Nguyen, T., Luo, X., et al. (2019). Dipco–dinner party corpus. *arXiv preprint arXiv:1909.13447*

Virag, N. (1999). Single channel speech enhancement based on masking properties of the human auditory system. *IEEE Transactions on speech and audio processing* 7, 126–137

Yuliani, A. R., Amri, M. F., Suryawati, E., Ramdan, A., and Pardede, H. F. (2021). Speech enhancement using deep learning methods: A review. *Jurnal Elektronika dan Telekomunikasi* 21, 19–26

Zhu, Y., Kong, Q., Shi, J., Liu, S., Ye, X., Wang, J.-C., et al. (2024). End-to-end paired ambisonic-binaural audio rendering. *IEEE/CAA Journal of Automatica Sinica* 11, 502–513

Digital Twin Embodied Interactions Design: Synchronized and Aligned Physical Sensation in Location-Based Social VR

Ju Zhang [1,*], Yuke Pi [1], Haoyang Du [1], Fang Ma [1] and Xueni Pan [1]

[1] *SeeVR Lab, Department of Computing, Goldsmiths, University of London*

Correspondence*:
Ju Zhang
j.zhang@gold.ac.uk

ABSTRACT

Using VR HMD body tracking, we propose a ‘cross-interaction’ design that bridges the digital and physical realms within a location-based VR environment. This design (or ’dual-realm’ setup) merges a physical laboratory with its digital twin, enabling users to collaborate and interact across geographic boundaries. The use of full-body avatars, hand tracking, and audio-driven facial expressions facilitates a mixed reality experience, featuring *digital* body non-verbal communication and synchronized *physical* sensation. Our implementation highlights the seamless fusion of the physical and virtual worlds through precise environment alignment, avatar representations, and enhanced social interactions. Furthermore, the introduction of custom Wi-Fi-based VR trackers offers the potential to refine the system’s motion capture capabilities and interaction with movable objects, providing a cost-effective alternative to conventional motion capture systems. The project demonstrates significant potential for future research in social interaction, communication, and entertainment within VR settings.

1 INTRODUCTION

This project explores a multiplayer experience set within both a physical laboratory and its 1:1 mapped digital twin environment in VR. Central to this initiative is the creation of a dynamic virtual environment where individuals, especially students, can gather, collaborate, and engage in shared activities without the constraints of geographical boundaries. Our vision aligns with and extends the trend toward digitalization, evident in the burgeoning utilization of social media and video conferencing tools like Facebook, WhatsApp, Zoom, and Microsoft Teams (Skulmowski and Rey, 2020). Our project pushes the boundary of the very concept of “metaverse”, and underscores the potential of VR in reshaping digital and physical social interactions.

This project addresses a significant research gap in the field of human social interaction, by tracking user behaviour (Pan and Antonia, 2018), more advanced non-verbal cues (Ma and Pan, 2022) and introducing physical haptic feedback within a social VR setting, leading a novel approach that extends beyond the predominant focus on visual and auditory feedback in VR (Waltemate et al., 2018). This project also contributes to the ongoing exploration of avatar-based interaction psychology, offering a unique platform for studying identity, belonging, and representation in VR. By allowing users to embody avatars capable of physical interactions, the project opens new ways for understanding how users perceive and construct their identities within virtual spaces (Freeman et al., 2017).

The subsequent sections provide a thorough thinking of avatar design, the technical implementation, cross-interactions with physical sensations, and the potential for further study.

2 AVATAR AND INTERACTION DESIGN

In order to strike a balance between natural interactions that mirror real-world movements and magical interactions (Kolesnichenko et al., 2019) which transcend physical limitations, we delved into several aspects in our design.

Embodiment Enabling users to embody themselves within the avatar's body enhances user immersion by allowing users to become one with their avatar's body (Kilteni et al., 2012). In this multiplayer project, we chose embodied avatars. One of the reasons to use full-body avatars to ensure safety in a dual-realm multiplayer experiment is to avoid the potential risk of players accidentally colliding with each other.

Locomotion Our dual-realm setup allows users to Navigate the physical space naturally, without removing their VR headsets. This could enhance users' comfort and safety while reducing motion sickness (Cherni et al., 2020). This approach significantly improves the sense of presence and immersion, making virtual interactions more authentic.

Hand Control For hand interaction includes: poke physics button, complex hand grab, gesture recognition by intuitive hand use in the real world, which is natural interaction, and hand UI ray for magic interaction. The interactable objects such as the VR headset have been designed to be held in different ways.

3 IMPLEMENTATION AND TECHNIQUES

The technical implementation of this project focuses on achieving precise alignment between the virtual world and the physical room, enabling multiplayer interactive experiences with full-body self-avatars, and exploring the potential of bespoke VR tracker devices (see Figure 1).

3.1 Accuracy align the virtual environment

The alignment of the virtual and physical rooms is the key to this location-based VR experience. Thus, to make this function work, two main challenges need to be solved in this project: one is to ensure the physical room and objects match perfectly, and the other is to accurately move the headset's relative location in both virtual and real spaces.

For the first challenge, three distinct LiDAR camera-based 3D scan software were tested to conduct 3D scans of the room for accuracy. Remodel the space within Blender with precise measurement, taking care not to modify the dimensions or arrangement of any objects. Cross-reference the outcomes from each software package to guarantee absolute alignment.

The second challenge is to address the synchronisation between the actual laboratory and the virtual one. The temporary solution at that time was to maintain its precise location alignment by the origin anchor point. The point is the starting angle and position of the VR headset, tested in advance by developers. We marked the point on the ground of the real-world scene with tape, so that each time the experience begins, the VR device enters the virtual scene at the same location and angle. This enables the device to automatically calibrate the virtual world, achieving a mapping that matches reality.

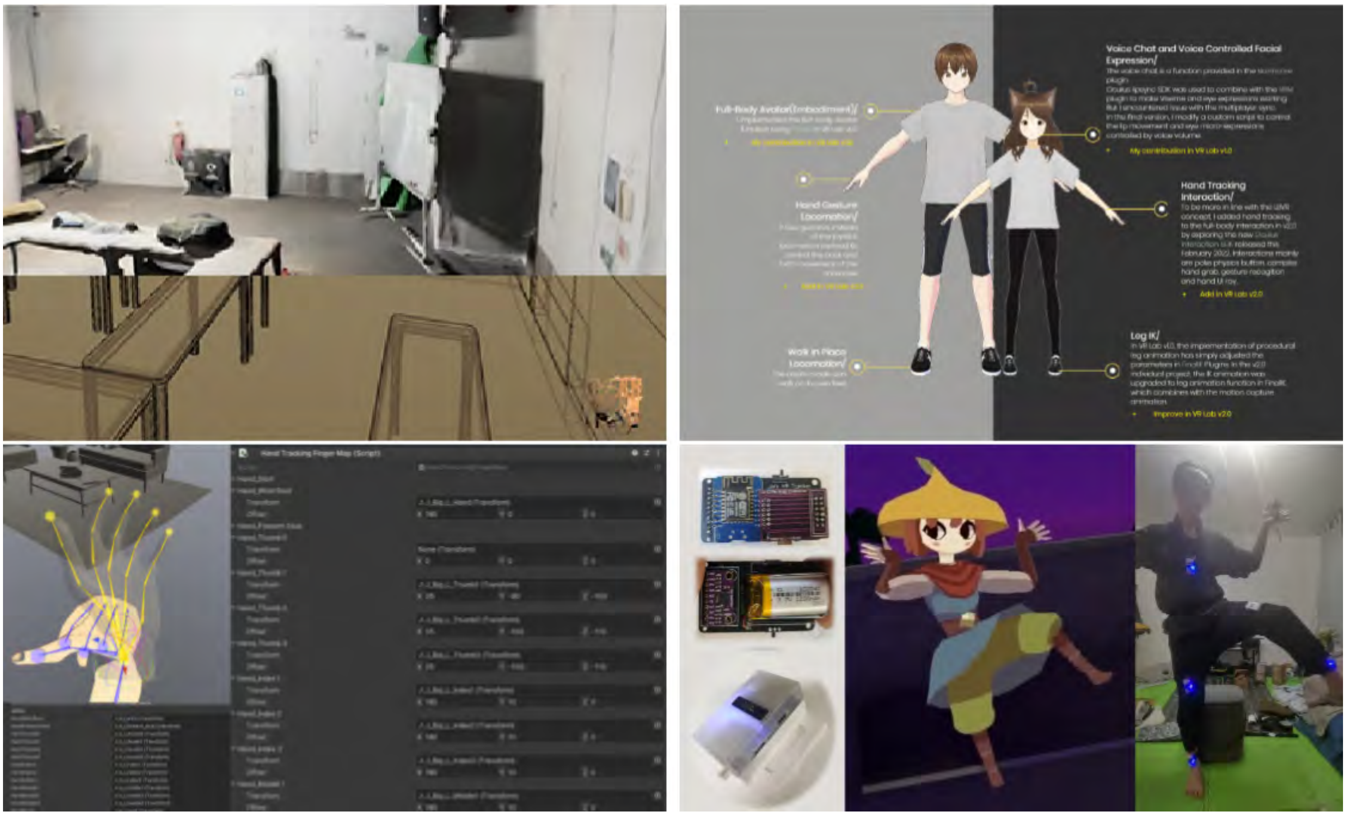


Figure 1. Top Left: 3D scanned data and the model from Blender; Top Right: Full-Body Avatar creation pipeline; Bottom Left: Hand tracking bone remapping; Bottom Right: Modular VR tracker on the left, tested in VRChat with full body tracking in the middle, with a live camera image on the right.

3.2 Full-Body Self-Avatar Representation for Multiplayer Interactions

The creation of avatars is facilitated through the use of VRoid Studio Figure 1, with the Unity game engine laying the foundation for the development of a sophisticated multiplayer platform. This platform is distinguished by several key components: the integration of the Normcore API, which enriches the multiplayer experience; the Oculus Interaction SDK, renowned for its precision in head and hand tracking; and Final IK, which brings to life realistic hand movements, comprehensive body dynamics, and leg actions through Leg IK. In this project, 3-point full-body tracking is employed.

The system's capabilities are further augmented by the incorporation of voice chat, achieved through the Normcore plugin, providing a seamless communication channel within the multiplayer environment. To deepen the level of immersion, the Oculus Lipsync SDK works with the VRM plugin. This collaboration is pivotal in achieving the synchronisation of Viseme, enhancing the realism of spoken words, and enabling expressive eye movements and facial expressions, thus enriching the user's immersive experience.

3.3 9-Point Full-body Motion Capture and Movable Object Tracking

An issue identified in this project is that easily movable objects (e.g., chairs) can pose challenges to the sustained operation of this precisely aligned digital twin experience, as well as lead to some deficiencies (e.g., gates and cabinet doors) in cross-interaction. To address this problem, a Wi-Fi-based VR tracker prototype that operates independently of any lighthouse system is explored. These VR trackers can also improve the full-body avatar to 9-point tracking.

This bespoke device draws inspiration from the SlimeVR [1] open-source initiative. 9-axis motion sensor has been utilised to combine with Wi-Fi to transmit information. Thus, this offers a cost-effective alternative to optical tracking methods. In an extension of the foundational VR tracking capabilities, the unit has been enhanced to become a modular VR tracker with magnetic connectivity. Figure 1 Moreover, a 9-pin magnet port has been integrated onto the tracker. This port enables the possibility of incorporating additional functionalities such as supplementary batteries, triggers, buttons, rolling wheels, lights, and various sensors. There is also potential for further development, given the availability of a 9-pin magnet port, which includes two pins designated for power.

4 CROSS-INTERACTIVE: DIGITAL INTERACTION SYNCED WITH PHYSICAL SENSATION

4.1 Object Interactions

Users could interact with various objects in our digital twin room. For instance, they could pick up pre-synced objects that exist in both the virtual and real worlds (e.g. VR headset in Figure 3). This is facilitated by hand tracking rather than controller tracking, to make the act of picking up objects a more natural interaction. Below are two examples of the designed individual cross-interactions in Figure 2:

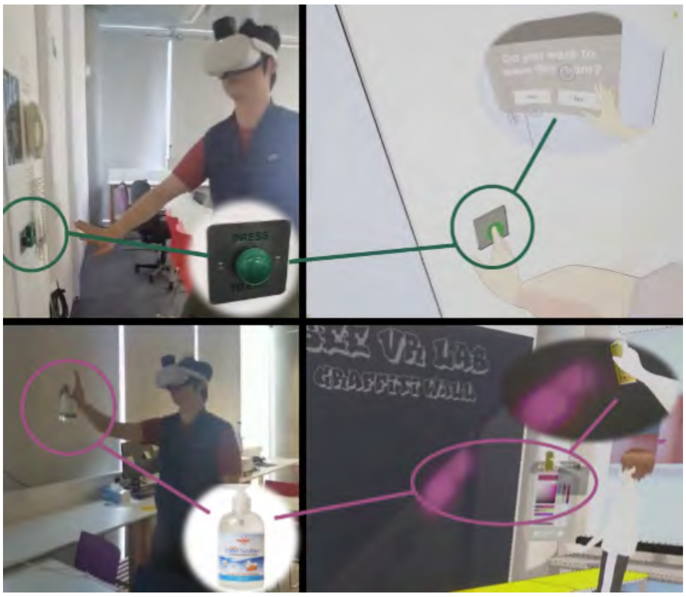


Figure 2. The top images depict interaction with the real exit button, while the bottom images show the hand sanitizer being used as a graffiti spray paint .

- Press the real button This project integrated the use of existing physical door exit buttons as the means to exit the virtual experience. When a player presses the real-world button, a virtual menu appears, confirming their decision to exit the virtual world. Another similar button interaction is the light swift in the real room can change the virtual room wall colour.

[1] https://github.com/SlimeVR

- Spray the graffiti bottle Users can interact with virtual objects, such as spraying graffiti spray paint, on the location, we placed the hand sanitizer that users can pick up, which provides tactile feedback and further blurs the boundary between the real and virtual worlds.

Figure 3. Left: Multiplayer guess and drawing game; Right: Recording a specific pose for the interactable object

4.2 Socially Involved Interactions

"Socially Involved Interaction" refers to interactions between individuals, using the simple example of picking up objects. If, during this interaction, real objects are passed on to others, it becomes a form of social cross-interaction. This kind of social cross-interactions bridges the gap between the digital and physical realms and transcends the limitations of traditional VR with more meaningful non-verbal communication.

- Hand tracking: shaking hands and high five Users can feel the sensation of a real handshake in the physical world simultaneously, also the high five is used when people achieve some gaming goals together. Or just simply, show numbers by using fingers to others.

- Body tracking: body language and hugs Full-body IK tracking ensures that users can experience physical sensations associated with hugging within the virtual environment. In this application, a drawing game Figure 3 is designed, and users communicate with each other through non-verbal cues and body language to convey ideas and guess each other's drawings. This social cross-interaction adds dimension to this body-tracking VR experience, where words are replaced by expressive gestures and movements.

5 USER FEEDBACK

The project has been showcased during industry event days. Two example quotes are:

'Wow, this is an amazing experience; I have never tried this type of experience before. Being so strongly embodied in the avatar and being able to shake hands with another avatar while physically feeling a real person is a very unique sensation.'

'At the beginning, I was afraid I would hit something, but soon after I touched the wall and the table around me, I felt much more confident and had a sense of safety due to the accurate mapping. I can walk freely within the 100 per cent cartoon style VR environment. Interacting with some physical objects that react inside the VR is an exciting experience.'

6 CONTRIBUTION AND FUTURE WORK

The project can facilitate virtual social gatherings and act as a valuable resource for researchers to examine social interactions and the interplay between virtual environments and reality. Technically, the VR trackers would be a good way to improve this project in both full-body and object tracking. One is through a cost-effective and convenient solution for alternative motion capture, these trackers streamline the process of capturing full-body movements. Another is the trackers' compact and lightweight design does not necessitate any lighthouse systems. This function makes them exceptionally suitable for tracking doors, drawers, and other objects in environments where lighting is limited. The VR trackers serve as an invaluable tool for collecting data on the movement of specific objects. Future efforts could also leverage this pipeline as a controlled setting for conducting experiments and collecting behavioural data, including body movement, facial expressions, eye gaze, and more. This is particularly relevant for studies in social science, communication, and entertainment.

ACKNOWLEDGMENTS

This project is funded by the AHRC Project "Immersive, Innovative, and Interactive Experience" AH/T011416/1.

REFERENCES

Cherni, H., Métayer, N., and Souliman, N. (2020). Literature review of locomotion techniques in virtual reality. *International Journal of Virtual Reality* 20, 1–20. doi:10.20870/IJVR.2020.20.1.3183

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D. M., Spanlang, B., et al. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine* 47, 2393–2400. doi:https://doi.org/10.1017/s003329171700040x

Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence Teleoperators amp Virtual Environments* 21. doi:10.1162/PRES_a_00124

Kolesnichenko, A., McVeigh-Schultz, J., and Isbister, K. (2019). Understanding emerging design practices for avatar systems in the commercial social vr ecology. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (New York, NY, USA: Association for Computing Machinery), DIS '19, 241–252. doi:10.1145/3322276.3322352

Ma, F. and Pan, X. (2022). Visual fidelity effects on expressive self-avatar in virtual reality: First impressions matter. *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* doi:https://doi.org/10.1109/vr51125.2022.00023

Pan, X. and Antonia (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology* 109, 395–417. doi:https://doi.org/10.1111/bjop.12290

Skulmowski, A. and Rey, G. D. (2020). Covid -19 as an accelerator for digitalization at a german university: Establishing hybrid campuses in times of crisis. *Human behavior and emerging technologies* 2, 212–216. doi:https://doi.org/10.1002/hbe2.201

Waltemate, T., Gall, D., Roth, D., Botsch, M., and Latoschik, M. E. (2018). The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Transactions on Visualization and Computer Graphics* 24, 1643–1652. doi:https://doi.org/10.1109/tvcg.2018.2794629

Digitizing Fashion: 3D Garment Reconstruction for Photorealistic Virtual Try-On and Avatar Creation

Hossein Javidnia [1,*], Diego Gonzalez [1], Arman Najari [1], Cai Felip ,[1], Yolanda Gómez,[1]

[1] *Union Avatars, C/ de la Llacuna, 162, Sant Martí, 08018 Barcelona, Spain*

Correspondence*:
Hossein Javidnia
hossein@unionavatars.com

ABSTRACT

This research paper investigates 3D garment reconstruction, crucial for photorealistic human digitization and virtual try-ons, yet challenged by the lack of extensive clothing models. It explores template-free methods, particularly focusing on Signed Distance Functions (SDF) and the Marching Cubes algorithm for 3D mesh creation, offering lower processing demands and better adaptability to complex garments. The study employs the Pixel-Aligned Implicit Function for High-Resolution Clothed Human Digitization (PiFU) model for high-resolution clothed human digitization, including color estimation. It discusses the potential impact of these advancements on avatar creation and the metaverse, suggesting significant contributions to digital identities and the internet's future.

Keywords: 3D Garment, Avatar, Metaverse, Virtual Try-on, Virtual clothes

1 INTRODUCTION

The introduction of 3D garment reconstruction has revolutionized computer vision and graphics, impacting fashion, gaming, and the emerging metaverse significantly. This technology's capability to model and recover clothed figures and garments from images or point clouds is crucial for achieving photorealism. While there have been advances in modeling unclothed human shapes and poses, clothed human and garment reconstruction presents distinct challenges. This paper examines the complexities of 3D garment reconstruction, highlighting both template-based and template-free approaches. It specifically focuses on the use of Signed Distance Functions (SDF) and the Marching Cubes algorithm Lorensen and Cline (1998) for 3D mesh reconstruction, addressing the critical aspects of creating realistic 3D garments and figures.

1.1 Relevant Work

Recent studies highlight the transformative impact of 3D garment reconstruction on the fashion industry. Techniques for virtual try-on and 3D clothing construction, as noted in Tuan et al. (2021), offer new avenues for virtual garment fitting and sustainable fashion. 3D printing's role in garment production, discussed in Kang and Kim (2019), introduces innovative approaches to meet sustainable fashion demands. Research by Lee (2021) and Sun and Valtas (2016) showcases how 3D technologies enhance supply chain efficiency and integrate with traditional manufacturing, respectively. The advantages of 3D software for zero-waste design and its broad impact across the fashion value chain are emphasized in studies by McQuillan (2020) and Arribas and Alfaro (2018). Additionally, Hwang and Hahn (2017) examines the significant role of 3D technologies in fashion education, highlighting their contribution to curriculum development and skill enhancement. These advancements across design, production, sustainability, and education sectors underline 3D garment reconstruction's pivotal role in evolving fashion industry practices.

2 METHODOLOGY

2.1 Base Architecture

Our network's architecture, grounded in PiFU (Pixel-Aligned Implicit Function for High-Resolution Clothed Human Digitization) Saito et al. (2019), features two parallel processes for 3D garment generation and vertex color estimation. It employs a convolutional encoder to extract features from input images, which, alongside 3D points, are processed by a Multi-Layer Perceptron (MLP) Rumelhart et al. (1986). The MLP assesses whether points are inside or outside the garment mesh and estimates the RGB colors for points within the 3D space. This dual-process approach ensures detailed and realistic 3D garment reconstruction, leveraging continuous implicit functions for both mesh reconstruction and color estimation, with a requirement for ”watertight” meshes that are volume-enclosed and perfectly sealed. This architecture as illustrated in figure 1 forms the foundation of our approach to 3D garment reconstruction, enabling us to generate detailed and realistic 3D garments.

2.2 3D Space Evaluation

During the training phase, N points are strategically sampled from the 3D space's target mesh M, to enhance model robustness. This is achieved by initially selecting points on the surface of M, and then adding random white noise ϵ_i to each point, where ϵ_i follows a normal distribution with mean 0 and variance σ^2, resulting in the set of sampled points $P_{\text{sampled}} = \{p_1 + \epsilon_1, p_2 + \epsilon_2, \ldots, p_N + \epsilon_N\}$.

2.2.1 Color Estimation Branch

For each point p_i in P_{sampled}, the MLP estimates the RGB color values $C(p_i)$. The estimated color $C(p_i)$ is compared against the target color $C_{\text{target}}(p_i)$ to compute the color loss:

$$\text{Loss}_{\text{color}} = \frac{1}{N} \sum_{i=1}^{N} ||C(p_i) - C_{\text{target}}(p_i)||^2 \tag{1}$$

2.2.2 3D Reconstruction Branch

For each point p_i in P_{sampled}, the MLP predicts whether the point is inside (1) or outside (0) the mesh, denoted as $V(p_i)$. The binary prediction $V(p_i)$ is compared to the actual inside/outside status $V_{\text{target}}(p_i)$, to compute the reconstruction loss:

$$\text{Loss}_{\text{reconstruction}} = -\frac{1}{N} \sum_{i=1}^{N} \left[V_{\text{target}}(p_i) \log(V(p_i)) + (1 - V_{\text{target}}(p_i)) \log(1 - V(p_i)) \right] \tag{2}$$

2.2.3 Overall Training Objective

The overall training objective is to minimize the combined loss $\text{Loss}_{\text{total}}$, which aggregates the losses from both branches:

$$\text{Loss}_{\text{total}} = \lambda_{\text{color}} \cdot \text{Loss}_{\text{color}} + \lambda_{\text{reconstruction}} \cdot \text{Loss}_{\text{reconstruction}} \quad (3)$$

where λ_{color} and $\lambda_{\text{reconstruction}}$ are weighting coefficients that balance the importance of color accuracy and geometric fidelity in the training process.

By optimizing $\text{Loss}_{\text{total}}$, the network is trained to predict the color and geometric structure of 3D garments accurately, leading to high-fidelity digital human representations. This process involves strategic sampling of points around the target mesh and a robust training methodology that facilitates the effective learning of garment morphology and color, significantly enhancing the realism of the generated 3D models. During the inference phase, as presented in figure 2, points covering the designated 3D space are systematically evaluated in batches at a specified resolution. In the reconstruction branch, the point cloud is categorized based on its spatial relationship (internal or external to the target mesh), enabling the reconstruction of the 3D mesh through the Marching Cubes algorithm Lorensen and Cline (1998). Concurrently, in the color branch, RGB color values are meticulously estimated for points deemed to be part of the mesh.

2.2.4 Training Adjustment

Adapting PiFU Saito et al. (2019) for garment reconstruction necessitates changes in the training phase, especially in point sampling strategies, to accommodate garments' unique morphology compared to human bodies. Figure 3 highlights the differences in point sampling between full human reconstruction and garment reconstruction. To ensure points remain within the garment mesh for accurate reconstruction, the sigma value of the applied noise is adjusted. Moreover, to capture the garment's detailed morphology accurately, noise sigma values are varied across point subsets, with half at standard levels and the other half increased threefold. This approach helps balance point distribution, addressing imbalances and favoring points inside the mesh, which in turn enhances the network's ability to learn the garment's structure effectively.

3 DATASET SUMMARY: DEEPFASHION3D

The DeepFashion3D dataset Zhu et al. (2020) offers a comprehensive platform for initial testing due to its variety in data formats, including point clouds and 3D meshes, and its high variance and detail across 590 distinct clothing types totaling 1212 items. The dataset provides meshes in OBJ format with accompanying MTL files for textures and point clouds in PLY format. To adapt DeepFashion3D for use with PiFU Saito et al. (2019), modifications are necessary due to the original meshes being single-layered and not ”watertight.” The adaptation process involves using Blender Blender Foundation (2023) to thicken the meshes, thereby adding volume and converting them into ”watertight” versions, as illustrated in figure 4 which compares the original and modified meshes.

Additionally, the dataset undergoes scaling adjustments to fit PiFU's requirement for reconstructing entire bodies rather than individual garments. This entails applying a uniform scale adjustment of 450 and recentering the meshes based on their volumetric center of mass. Further preprocessing includes the application of Precomputed Radiance Transfer (PRT) Villberg (2003) to enhance mesh quality for rendering. Initially offering 360-degree views, the rendering perspective is now narrowed to frontal views ranging from $-25°$ to $+25°$, as showcased in figure 5. This step is critical for generating high-quality renders from specific angles, and it is complemented by the production of normal maps, textures, and segmentation masks to meet PiFU's training data specifications.

4 TRAINING

The dataset was segmented into training and validation sets, with 85% of the samples allocated for training and the remaining 15% set aside for validation and testing. Adjustments to the data loader enhanced memory efficiency and data loading speed.

To minimize memory demands, training occurred in single-batch increments. Considerations for future enhancements include utilizing more robust GPUs to facilitate larger batch sizes. The initial learning rate was maintained at $1e-3$, and each iteration produced $10,000$ points with a sigma value of 8. Over two days of training on an RTX 3060, the network completed 6 epochs.

5 RESULTS AND DISCUSSION

Our research on 3D garment reconstruction has shown notable progress, particularly with template-free techniques using SDF and the Marching Cubes algorithm Lorensen and Cline (1998). Utilizing the PiFU model Saito et al. (2019), we've achieved highly detailed and realistic garment reconstructions from single images. The results demonstrate remarkable accuracy in garment shape and texture, closely mirroring actual garments. This accuracy is further enhanced by precise vertex color estimation, adding to the visual realism by capturing the original colors and patterns accurately.

A significant achievement is the creation of detailed 3D models suitable for applications like virtual try-on and avatar customization in the metaverse. The models maintain high fidelity across various poses and body types, proving the effectiveness and flexibility of our reconstruction approach.

The improvement in automatic skinning weights estimation, crucial for garment fitting on avatars, represents a considerable advancement, overcoming initial challenges through mesh decimation, alignment, and weight cleaning. This development is pivotal for dressing virtual characters realistically, facilitating their use in gaming, fashion, and virtual reality.

In conclusion, our work demonstrates that template-free 3D garment reconstruction, alongside sophisticated skinning weight estimation methods, significantly enhances the realism and utility of virtual garments. This underscores the transformative potential of our techniques in digital fashion creation and interaction within immersive environments.

For visual results of the method, see the supplementary material.

FUNDING

This research was supported by the Ministry of Economic Affairs and Digital Transformation of the Spanish Government, whose funding and support have been essential to the success of this project.

REFERENCES

Arribas, V. and Alfaro, J. A. (2018). 3d technology in fashion: from concept to consumer. *Journal of Fashion Marketing and Management: An International Journal* 22, 240–251

Blender Foundation (2023). Blender - a 3d modelling and rendering package. `https://www.blender.org`. Accessed: 2023-11-26

Hwang, J. and Hahn, K. (2017). A case study of 2d/3d cad virtual prototype simulation programs to enhance student performance in student-centered fashion design education. *J Textile Eng Fashion Technol* 3, 578–584

Kang, M. and Kim, S. (2019). Fabrication of 3d printed garments using flat patterns and motifs. *International journal of clothing science and technology* 31, 653–662

Lee, Y. K. (2021). Transformation of the innovative and sustainable supply chain with upcoming real-time fashion systems. *Sustainability* 13, 1081

Lorensen, W. E. and Cline, H. E. (1998). Marching cubes: A high resolution 3d surface construction algorithm. In *Seminal graphics: pioneering efforts that shaped the field*. 347–353

McQuillan, H. (2020). Digital 3d design as a tool for augmenting zero-waste fashion design practice. *International Journal of Fashion Design, Technology and Education* 13, 89–100

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature* 323, 533–536

Saito, S., Huang, Z., Natsume, R., Morishima, S., Kanazawa, A., and Li, H. (2019). Pifu: Pixel-aligned implicit function for high-resolution clothed human digitization. In *Proceedings of the IEEE/CVF international conference on computer vision*. 2304–2314

Sun, D. and Valtas, A. (2016). 3d printing for garments production: An exploratory study. *Journal of Fashion Technology and Textile Engineering* 4, 1000139

Tuan, T. T., Minar, M. R., Ahn, H., and Wainwright, J. (2021). Multiple pose virtual try-on based on 3d clothing reconstruction. *IEEE Access* 9, 114367–114380

Villberg, A. (2003). *Precomputed Radiance Transfer*. T-111.500 technical report, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory. Spring 2003

Zhu, H., Cao, Y., Jin, H., Chen, W., Du, D., Wang, Z., et al. (2020). Deep fashion3d: A dataset and benchmark for 3d garment reconstruction from single images. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16* (Springer), 512–530

FIGURE CAPTIONS

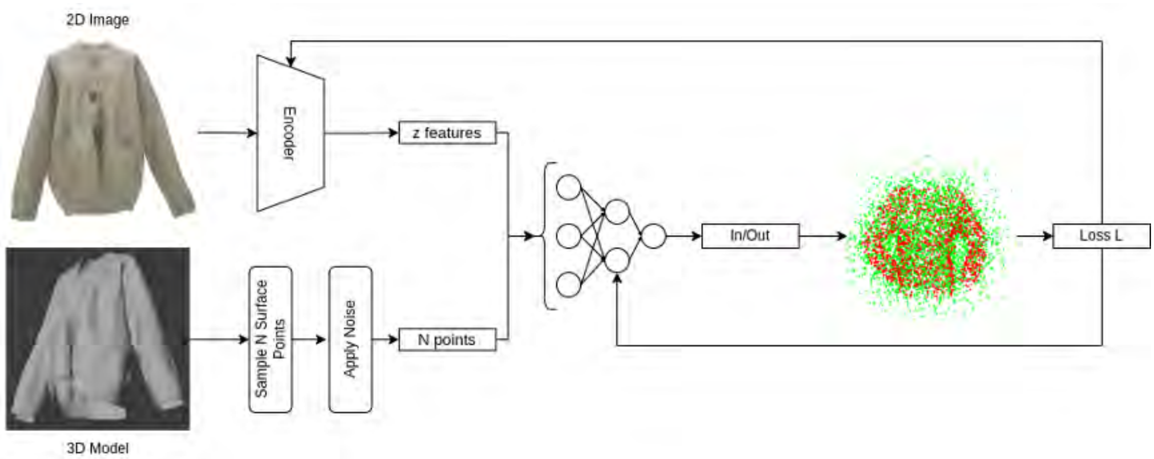


Figure 1. The foundational network architecture of our approach to 3D garment reconstruction

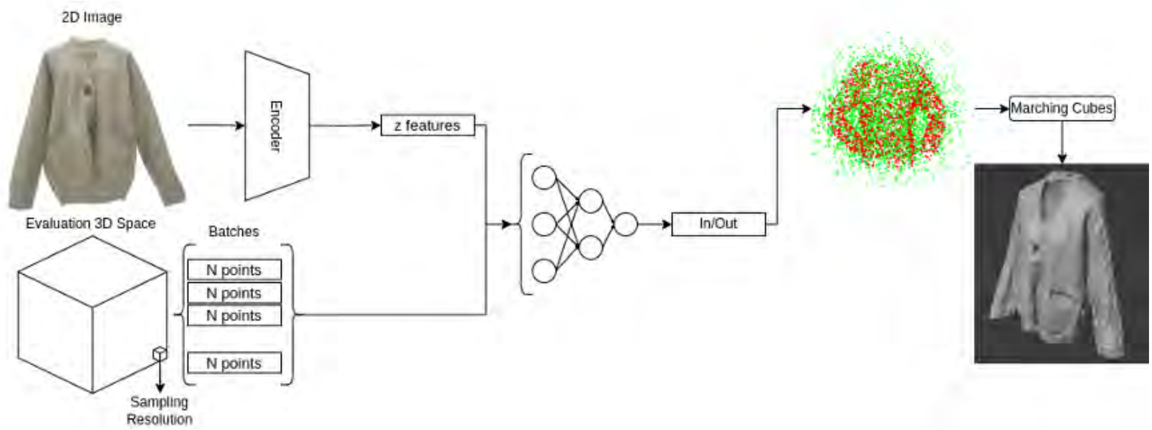


Figure 2. Inference stage - from 2D image to 3D garment model

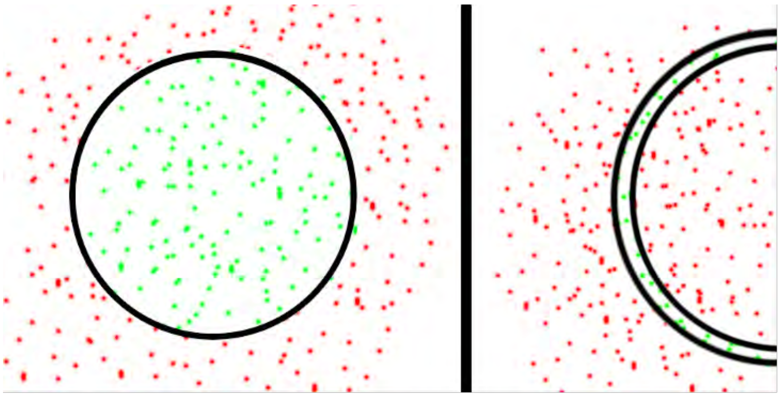


Figure 3. Adjustments in point sampling due to the different morphology of garments compared to human bodies

Figure 4. The original mesh (left) with its watertight counterpart (right) processed using Blender

Figure 5. DeepFashion3D prepreocessing

Medicine, Neurotechnology and the Metaverse

Immersive Virtual Reality: a paradigm shift and a disruptive tool for visual neurorehabilitation.

Eduardo Garcia-Giler[1,2], Bill Zhang[1], Mariana Misawa[3], Yulia Pyatova[3], Monica Daibert-Nido[3], Atri Sen[1], Samuel N. Markowitz[1,3], Kyle Cheung[1], Danielle Tchao[4], Pi Nasir[1], Inci Yamana Bajin[5], Eric Bouffet[5], Lora Appel[2,4,6] and Michael Reber[1,2,3,*].

[1]Donald K Johson Eye Institute, Krembil Research Institute, University Health Network, Toronto, ON, Canada.
[2]UMR-S 1329 INSERM, Strasbourg, France.
[3]Department of Ophthalmology and Vision Science, University of Toronto, Toronto, ON, Canada.
[4]OpenLab, University Health Network, Toronto, ON, Canada.
[5]Department of Neuro-Oncology, The Hospital for Sick Children, Toronto, ON, Canada.
[6]Department of Health Informatics, York University, Toronto, Canada.

*** Correspondence:**
michael.reber@inserm.fr

Keywords: ophthalmology, low vision, neurorehabilitation, perceptual learning, immersive virtual reality.

Abstract
Introduction: Visual neurorehabilitation in Ophthalmology addresses the challenges faced by individuals with incurable low vision. The World Health Organization defines low vision as persistent impairment despite standard treatments, affecting daily activities. Approximately 206 million people globally suffer from moderate to severe visual impairment. Existing neurorehabilitation programs often require expensive clinical setups, limiting accessibility for underserved communities. The use of standalone immersive virtual reality (IVR) headsets, like the Meta Quest Pro, emerges as a revolutionary approach. These IVR systems facilitate remote development of rehabilitation programs, incorporating real-world scenarios and personalized protocols based on individual performance. This innovation enhances efficiency and offers diagnostic and prognostic tools, making neurorehabilitation more accessible and effective.
Methods: Developed on the Unity3D platform, immersive virtual reality applications include Re:Vision for the stimulation of spatial perception, Re:SaccD and V(o)R for oculo-motor control and vestibular-ocular reflex assessments. Two clinical trials, evaluating the feasibility and effectiveness of Re:Vision as an interventional tool, were performed using the Meta Quest 2/Pro: Study 1, a single-arm phase II trial, occurred from July 2022 to October 2023. Study 2, a 4-week randomized trial, took place from February 2021 to October 2023, partially during Covid-19 restrictions. Re:SaccD and V(o)R were tested among healthy volunteers and individuals with oculo-motor deficits.
Results: The visual neurorehabilitation application Re:Vision shows effectiveness in patients with visual field defects by increasing visual perception. In addition, IVR is a meaningful platform to successfully deliver exploratory and investigational measures, including head/eye tracking concomitant to audiovisual performance, in clinical evaluation tasks (Re:SaccD and V(o)R).
Discussion: The development of IVR as a tool for interventional and exploratory purposes in ophthalmology and visual neurorehabilitation is a potential game changer in the field, evolving towards a holistic approach, emphasizing personalized care and independent home-based therapy. The use of IVR could ensure equitable access to specialized treatments, particularly benefiting patients in rural areas. It addresses challenges posed by the pandemic, such as travel restrictions and staff shortages, allowing patients to receive timely care without delays. The model promotes self-sufficiency and empowers individuals to manage their health independently, aligning with the current need for flexible and accessible healthcare solutions.

Introduction.

Visual neurorehabilitation is a discipline in Ophthalmology aimed at improving quality of life of individuals with low vision. Low vision, as defined by the World health Organization, refers to "*A person who has impairment of visual functioning even after treatment and/or standard refractive correction, and has a visual acuity of less than 6/18 to light perception, or a visual field of less than 10 degree from the point of fixation, but who uses, or is potentially able to use, vision for planning and/or execution of a task*'' (World Health Organization, 1980; ICO, 2002). In other words, low vision cannot be corrected or cured with glasses, medication or surgery therefore affecting activities of daily living such as reading or driving and strongly restricting learning and development in children and adolescents. Individuals with low vision show increased risks of falls, anxiety/depression and social dependence. It is estimated that 206 million people worldwide have moderate to severe visual impairment which includes retinal diseases (glaucoma, macular degeneration, retinopathies), oculo-motor defects (amblyopia, nystagmus, strabismus) and visual field defects (hemianopia). Modern visual neurorehabilitation programs are available for some of these defects with variable efficacy/effectiveness. However, they often require expensive equipment operated by highly qualified personnel in clinical settings therefore limiting access of underserved communities where geographical barriers and financial constraints limit accessibility to such supportive therapies.

The utilization of stand-alone, mobile, immersive virtual reality (IVR) headsets like the Meta Quest Pro as a platform for delivering supportive interventions and diagnostic/prognostic tools holds the potential to revolutionize the field of rehabilitation, particularly in visual neurorehabilitation. The standalone IVR technology allows the development of remotely controlled, home-based rehabilitation programs incorporating real-world situations, enhancing efficiency through patient-tailored protocols using algorithms that refine tasks based on the individual's previous performance. These properties also favor the implementation of existing standardized diagnostic tests in Ophthalmology in VR and opens avenues for developing new testing paradigms. This innovative approach aligns seamlessly with the trajectory of modern, personalized, and precision medicine. Here, we will provide a summary of clinical and feasibility results derived from clinical trials and case series that underscore the usefulness, reliability, and potency of VR. Detailed clinical outcomes are detailed elsewhere(1).

Methods.

Studies design.

Study 1 was designed as a single-arm phase II feasibility and proof-of-concept trial with pre/post-intervention analysis(1) (Research Ethics Boards - REB# 1000076413 Hospital for Sick Children, Toronto, Canada and REB# 21-5978 - University Health Network - UHN, Toronto, Canada, and registered on clinicaltrials.gov registration NCT05065268). The study took place from July 2022 to October 2023. Study 2 was a 4-week, open label, reference-controlled, phase 2 randomized, feasibility and proof-od-concept trial assessing feasibility and potential effectiveness of 3D-MOT-IVR versus a reference - biofeedback training (UHN REB# 20-6143- NCT04685824). The study took place from February 2021 to October 2023, partially during the Covid-19 restrictions. Case analyses involving either healthy volunteers (< 3) or individuals with nystagmus (< 3) were performed at the low vision centre, ophthalmology clinic, Toronto Western Hospital, UHN, Toronto, ON, Canada under supervision of an ophthalmologist. All studies followed the principles outlined in the World Medical Association Declaration of Helsinki. All participants provided written informed consent.

Immersive virtual reality applications.

Using Unity3D platform, we have developed and encoded several interventional and diagnostic/investigational applications in the Meta Quest Pro:

- Re:Vision is an audiovisual stimulation developed to restore/enhance spatial detection of elements in the peripheral environment using both vision and audition. The application is based on the multiple-object paradigm (MOT) designed in the late 80's to study visual attention in humans(2) and implemented in 3D (3D-MOT) with correlated spatial sound. The central features of the 3D-MOT task closely match attentionally demanding and complex real-world situations(3).
- Re:SaccD is a visual application evaluating oculo-motor control in VR (saccades and anti-saccades) based on the Interleaved Pro/Anti-Saccade Test (IPAST)(4) test. These outcome measures are used

in research and in clinical settings to study eye movements and gaze control or to evaluate the effects of a visual rehabilitation program on oculo-motor control.

- V(o)R is an application assessing the vestibular-ocular reflex (VOR – gaze stabilisation) in VR. Participants focus on a blue fixation point localized in their central visual field and move their heads left to right for 30 seconds at a pace of 1Hz while keeping their eyes fixed on the central point.

Outcome measures, data collection and analysis.

In Re:Vision, Re:SaccD and V(o)R, data are collected remotely from the clinic or from the participants' home. After each trial or session, an encrypted .csv file is sent by the Meta Quest 2/Pro through WiFi to the laboratory computers. Remotely collected measures include safety (assessment of cybersickness using virtual reality- induced symptoms and effects – VRISE questionnaire(5) uploaded in the application), adherence and compliance to the intervention/test (# of trials/session performed, # of interruptions, date/time, duration of the trial/session) and task-specific performance (hit/miss rates, reaction time, latency, speed of target). Outcomes measured in clinic include ophthalmological evaluation (visual fields, reading speed), adoption and usability (qualitative questionnaires). Head tracking and left/right eye tracking (for Quest Pro) data were collected for exploratory purposes (understanding the tracking strategy of a participant) or for clinical evaluation (impaired oculo-motor or gaze control). Data were compiled and processed using in-house python codes, and analysed using descriptive statistics (mean, SD or median, min, q1, q3, max). Statistical comparisons between pre- and post-intervention data were performed using paired samples t-test in JASP (v0.17.2.1)

Results

Clinical outcomes – restoration of visual perception after Re:Vision intervention.

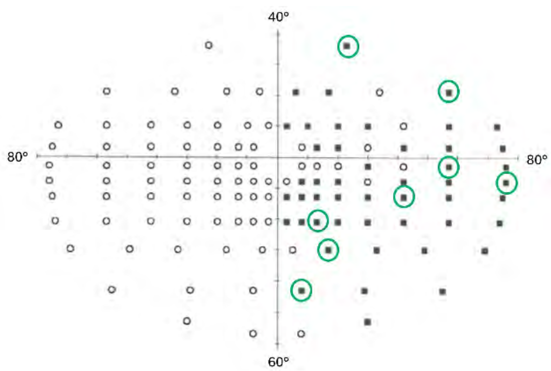


Fig. 1: Example of an Esterman binocular field testing. Green circles show perceived points after Re:Vision intervention. From (1).

The effectiveness of Re:Vision delivered using the stand-alone Meta Quest2/Pro as an interventional tool in individuals with permanent visual field defects (hemianopia) due to brain tumour has been tested clinically in a case series(6,7) and a pilot study(1) (Study 1). Summary of the results indicate a significant and clinically meaningful increase in the number of points perceived at the Esterman binocular field test (average Δ = +5.6 ± 2.0 points, corresponding to an average of 20.8% ±8.9 increase in visual field restoration, n = 7/12 responders) after using Re:Vision every other day for 15 minutes for 6 weeks at home. Fig. 1 shows an example of pre/post-intervention results at the Esterman test for 1 participant. Eye/head tracking was recorded to analyze the tracking strategy developed by the participants when performing covert tracking (following a target without moving eyes/head) (Fig. 2).

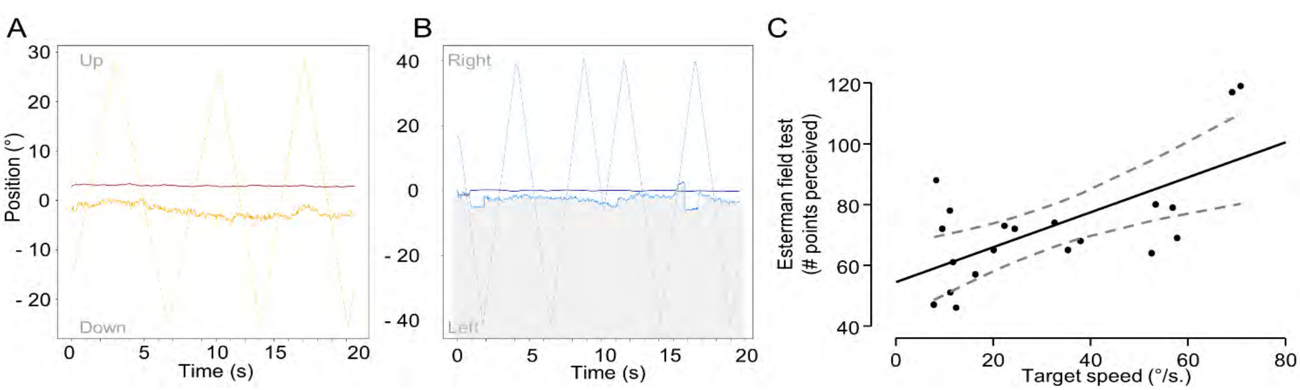


Fig. 2: Example of head/eye/tracking results using Meta Quest Pro during Re:Vision. Vertical axis (**A**) and horizontal (**B**) axes show target (dotted orange/blue line), head (red/blue line) and eye (orange/blue line) movements. The grey area in B indicates the blind visual field in this participant. This trial was successful (correct target choice) suggesting that the participant could track the target when entering the blind field without moving the head/eyes. **C**. Correlation between # of points perceived at Esterman and performance (speed of target for correct choice) at Re:Vision (Pearson's r = 0.65, p = 0.003) suggesting that the better covert tracking (without head/eye movement) led to more points perceived in the Esterman binocular field test. From (1).

Adoption, safety, usability, compliance and adherence of Re:Vision program in immersive VR. The potential use of immersive VR as a home-based interventional tool in visual neurorehabilitation is innovative, therefore data about its use in real-world conditions is still lacking. In our finalized case series(6,7) and clinical studies on young individuals (NCT05065268 – Study 1, < 30 group) and older adults (NCT04685824 – Study 2, > 65 group) we measured usability and adoption of the VR technology, adherence and compliance to the home-based Re:Vision program as feasibility criteria defined based on studies using virtual-reality as an interventional tool(8,9) and on CONSORT guidelines for feasibility trials(10). Results are summarized in Tables 1-3.

Group (n)	**Average age (years)**	**Adoption (%, ratio)**	**Adherence/Compliance (%, ratio) (protocol completion criteria ≥ 80%)**	**Remote data collection through Wifi (%, ratio)**
< 30 (12)	16.8 ±6.5	100 (12/12)	ITT 92.8 (182/196)	100 (182/182)
< 30 ♀ (4)	*18.5 ±9.2*	*100 (4/4)*	*ITT 95.6 (44/46)*	*100 (44/44)*
< 30 ♂ (8)	*16.0 ±5.3*	*100 (8/8)*	*ITT 92.1 (129/140)*	*100 (129/129)*
> 65 (7)	83.2 ±5.3	70 (7/10)	ITT: 75 (96/128) PP: 83 (93/112)	100 (132/132)
> 65 ♀ (5)	*84.4±4.7*	*71 (5/7)*	*ITT: 67.7 (65/96) - PP: 77.5 (62/80)*	*100 (62/62)*
> 65 ♂ (2)	*80.5 ±7.7*	*67 (2/3)*	*ITT/PP: 96.9 (31/32)*	*100 (31/31)*

Table 1: Demographics and feasibility results (ITT: intention-to-treat, PP: per protocol).

Safety, assessed using the VRISE questionnaire scores evaluating cybersickness, showed an overall average difference after/before of – 1.3 points (after = 33.3/35, before = 34.6/35, and 1 score < 25 points) in the < 30 group and – 0.9 points (after = 33.8/35, before = 34.7/35) in the > 65 group indicating no adverse event (cybersickness symptoms scores criteria < 25) related to the use of Re:Vision (details in Table 2). One drop-out in the > 65 group was due to the intervention: The weight of the Meta Quest 2 headset was too heavy for the participant's neck muscles which became sore after the 15-minute session.

Group (n)	**Safety (average difference score before/after session)**				
	Nausea	**Disorientation**	**Dizziness**	**Fatigue**	**Instability**
< 30 (10)	- 0.18 ±0.19	- 0.15 ±0.19	- 0.22 ±0.25	- 0.44 ±0.23	- 0.08 ±0.14
< 30 ♀ (3)	*- 0.17 ±0.33*	*- 0.13 ±0.25*	*- 0.27 ±0.49*	*- 0.58 ±0.60*	*0.14 ±0.33*
< 30 ♂ (7)	*- 0.17 ±0.16*	*- 0.16 ±0.21*	*- 0.20 ±0.24*	*- 0.37 ±0.31*	*- 0.05 ±0.11*
> 65 (7)	- 0.03 ±0.07	- 0.09 ±0.19	- 0. 25 ±0.19	- 0. 37 ±0.18	- 0.02 ±0.10
> 65 ♀ (5)	*- 0.08 ±0.25*	*- 0.09 ±0.17*	*- 0.34 ±0.28*	*- 0.45 ±0.21*	*- 0.21 ±0.80*
> 65 ♂ (2)	*0.00 ±0.00*	*0.00 ±0.00*	*- 0.03 ±0.13*	*- 0.03 ±0.13*	*0.00 ±0.00*

Table 2: Safety measures.

Qualitative feedback was performed at the end of the intervention period using an adapted version of the virtual-reality neuroscience questionnaire (VRNQ)(5). Median score, 47.5 [33, 39.3, 51.8, 55] indicated usability of Re:Vision in The Meta Quest 2/Pro (total cutoff threshold = 40). Detailed analysis among 8 items (scored 1: extremely difficult/low to 7: extremely easy/high, cutoff threshold = 5) is shown in Table 3.

Group (n)	**Usability (VRNQ)**							
	Putting device on	**Adjusting device**	**Quality of graphics**	**Quality of sound**	**Quality of VR technology**	**Use of the pointer**	**Selection using the pointer**	**Device recharging process**
< 30 (10)	6.0 ±1.2	6.0 ±1.2	4.9 ±1.2	4.9 ±0.7	5.6 ±1.0	6.2 ±0.9	5.5 ±1.4	6.4 ±0.8
< 30 ♀ (3)	*5.3 ±1.5*	*5.3 ±0.6*	*5.3 ±1.5*	*5.0 ±0.0*	*5.3 ±1.5*	*5.7 ±1.2*	*5.3 ±1.5*	*5.7 ±1.2*
< 30 ♂ (7)	*6.3 ±1.1*	*6.3 ±1.3*	*4.7 ±1.1*	*4.9 ±0.9*	*5.7 ±0.8*	*6.4 ±0.8*	*5.6 ±1.4*	*6.7 ±0.5*
> 65 (7)	5.00 ±1.07	4.75 ±1.04	5.63 ±0.92	5.38 ±0.92	5.75 ±0.71	4.38 ±1.69	4.63 ±1.19	6.13 ±0.64
> 65 ♀ (5)	*4.67 ±1.03*	*4.50 ±1.05*	*5.67 ±1.03*	*5.50 ±0.84*	*5.83 ±0.75*	*4.33 ±1.75*	*4.33 ±1.21*	*6.17 ±0.75*
> 65 ♂ (2)	*6.00 ±0.00*	*5.50 ±0.71*	*5.50 ±0.71*	*5.00 ±1.41*	*5.50 ±0.71*	*4.50 ±2.12*	*5.50 ±0.71*	*6.00 ±0.00*

Table 3: Usability measures.

We recorded the time of the day when participants were performing Re:Vision at home (fig. 3). The < 30 group performed most of the sessions (55%, n = 98) between 15:00 and 21:00 taking an average time of 30 ±4 min. to complete a session (including pre/post-intervention VRISE questionnaire and 3 blocks of 15 trials of 20 s each). The > 65 group performed 71% of the sessions between 9:00 and 15:00 taking an average time of 28 ±7 min. There was no effect of the time of day of the intervention on performance in neither group.

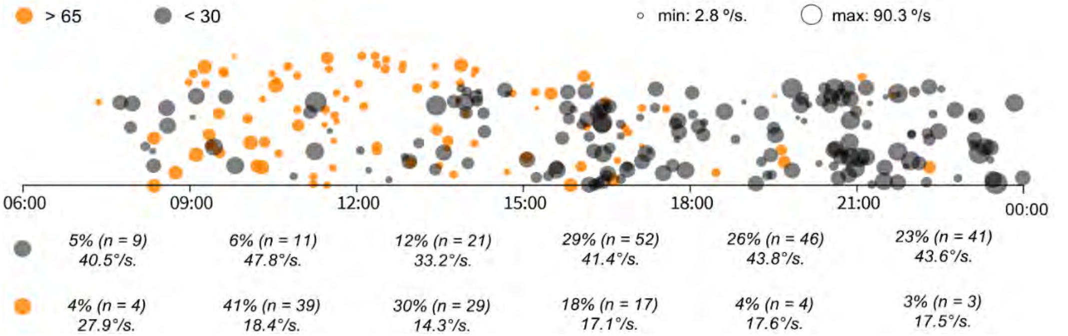


Fig. 3: Time of day of the intervention in < 30 group (orange) and > 65 group (grey).

Exploratory outcomes: Performance at Re:Vision

Because of the paradigm underlying Re:Vision using a 1:1 staircase dynamic adaptative method for target speed regulation, the correct choice rate stably oscillated around 60% (< 30: 62.2 ±4.2 %, > 65: 60.0 ±2.5 %). Therefore, the speed of the target for a correct choice is an appropriate measure of performance at Re:Vision (fig. 4). Speed of target for correct choice in the < 30 group is (17.2 ±2.2 °/s) is on average 2x higher compared to > 65 group (9.5 ±2.4 °/s, $t_{(30)}$ = 9.66, p = 1.01E-10) and stable over time in both groups, indicating no improvement at the Re:Vision task.

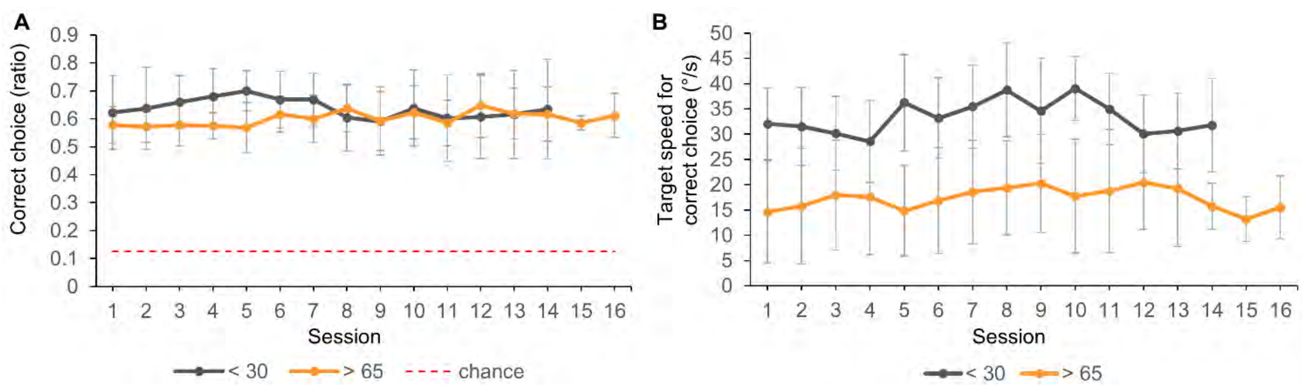


Fig. 4: Comparison of performance at Re:Vision between < 30 and > 65 group. **A**. Correct choice rates. Red dashed line show chance level (1/8). **B**. Average speed of target for correct choice.

Exploratory outcomes: Head/Eye tracking.

We compared target tracking strategy between the 2 groups of patients with different visual impairments (< 30 group: hemianopia, > 65 group: macular degeneration) by recording head movements during overt tracking (when eyes and head movement are allowed) in Re:Vision sessions performed at home. Individual with macular degeneration > 65 (> 65 group) tracked the moving target using head movement (r = 0.46 ±0.11 - Fig. 5) whereas individuals with hemianopia < 30 (< 30 group) do not (r = 0.11 ±0.20 - fig. 5).

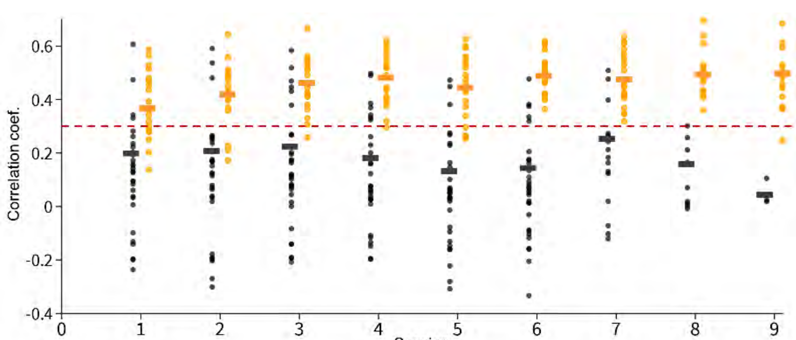


Fig. 5: Head/target tracking correlation during Re:Vision stimulation sessions in <30 and > 65 groups. Red dashed line = correlation threshold of head tracking. Correlation coefficients > 0.32 indicate target tracking using head movements.

Diagnostic/Investigational tools in VR: the power of the Meta Quest Pro.

Examples of the use of Meta Quest Pro as a platform for diagnostic applications in visual neurorehabilitation. Recordings of head/eye movement are performed remotely from lab. computers.

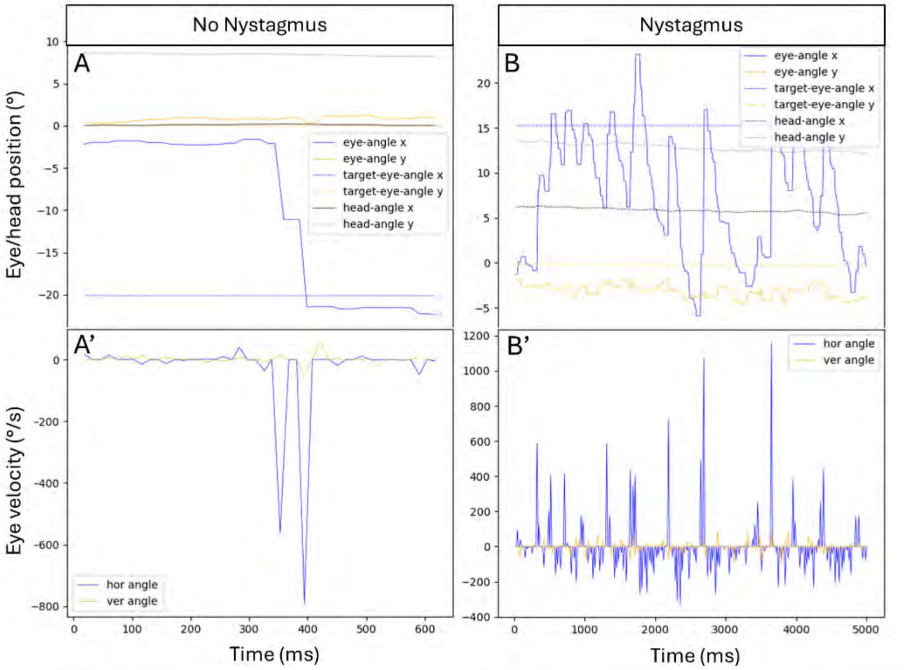


Fig. 6: Eye/head movements analysis using Meta Quest Pro during saccade tests (Re:SaccD application) comparing individuals with no nystagmus (A, A') and nystagmus (B, B'). A shows horizontal target position (blue dotted line) at - 20° and eye movement (saccade - blue line) starting with a 300ms and aligning with target position with a speed of 600-800°/s (blue lines in A'). There is no head movement (dark line) nor vertical movements.
B shows horizontal target position (blue dotted line) at 15° and jittered eye movement (saccade - blue line) unable to align with target position at 15°. The task stops after 5 s. due to the inability to fixate the target position. Note small jittering saccades in the vertical axis in B (orange line) B'' shows the eye velocity with jittering saccades typical of nystagmus. There is no head movement

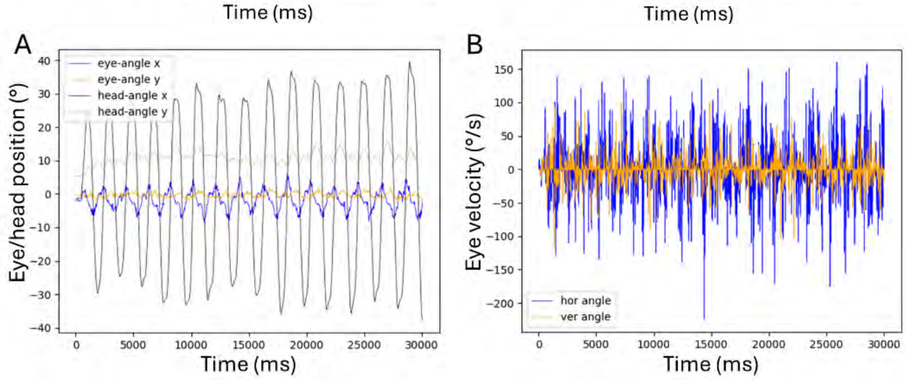


Fig. 7: Eye/head movements analysis using Meta Quest Pro during vestibulo-ocular reflex (VOR) in healthy individuals. A shows head/eye positions with oscillatory movement of the head (black line) and compensatory eye movements (blue line). B shows eye velocity with moderate speed (< 200°/s) distinctive of compensatory eye movements in VOR.

Discussion/Conclusion:

Our results indicate that IVR using the Meta Quest 2/Pro is a reliable tool to provide audiovisual neurorehabilitation programs remotely to the participants' home with a control of the device/application and a safe collection of data from the lab computers through WiFi. Adoption, usability, and safety meet the criteria for deployment in individuals aged 10 to 96 years old with visual impairment holding strong promises for the development of IVR as a platform to provide tele/e-health in different patient populations. Neurorehabilitation using IVR represents a paradigmatic shift in the conceptualization of healthcare, embracing a more holistic vision and facilitating the delivery of personalized care. It ensures equitable access, offering the capability to receive supportive therapy at home independently, without direct supervision by healthcare professionals. This empowers patients residing outside urban areas to access specialized therapies. These advantages gain heightened significance in the recent context of the pandemic, given the challenges posed by travel restrictions, staff shortages, and related constraints, which have resulted in treatment delays for patients.

References

1. Misawa et al., 2024, medRxiv, 2024.01.25.24301666
2. Pylyshyn & Storm. Spatial Vision. 1988;3:179–97
3. Meyerhoff et al., Atten Percep. Psy. 2017, 79:1255–74.
4. Yep et al., Front Aging Neurosci. 2022;14:842549.
5. Kourtesis et al., Front Hum Neurosci. 2019;13:417.
6. Daibert-Nido et al., Am. J. . Case Rep. 2021.
7. Daibert-Nido et al., Front Neurol. 2021;12:680211.
8. Huygelier et al., Appl Neuropsychol Adult. 2020 18;1–21.
9. Shim L et al., Clin. 2023 Apr 17;17:17.
10. Eldridge et al., Pilot Feasibility Stud. 2016;2:64.

Towards immersive teleclinics in the Metaverse: the case of neuropsychological consultation

Toinon Vigier[1], **Didier Acier**[2], **Samuel Bulteau**[3,4], **Solène Thevenet**[4] **and Yannick Prié**[1,*]

[1] *Nantes Université, École Centrale Nantes, CNRS, LS2N, UMR 6004, F-44000 Nantes, France*
[2] *Nantes Université, Centre de recherche en éducation de Nantes, CREN, UR 2661, F-44000 Nantes, France*
[3] *Nantes Université, CHU Nantes, INSERM, MethodS in Patients-centered outcomes and HEalth Research, SPHERE, F-44000 Nantes, France*
[4] *CHU Nantes, F-44000 Nantes, France*

Correspondence*:
Corresponding Author: Département Informatique, Polytech Nantes, Rue Christian Pauc, BP 50609, F-44306 Nantes cedex 3
yannick.prie@univ-nantes.fr

ABSTRACT

The Metaverse is often deemed very important for the future of medicine and healthcare, yet there exist very few real health applications making use of social VR, its core feature. Here we first present our vision of what could be a virtual clinic in the Metaverse that would allow patients and practitioners to meet in VR and carry out various activities. We then introduce Tecnis, our implementation of such a clinic for psychotherapeutic consultation and neuropsychological testing. Our system features several spaces such as a waiting room, a consultation room, and standalone activities, as well as data-based dashboads for clinicians to lead the sessions.

Keywords: Cognitive testing, Medical consultation, Metaverse, Psychotherapy, Teleclinic

1 INTRODUCTION

Immersive Virtual Reality (VR) has long been studied and used in medicine, for a large variety of use cases in care (e.g. pain relief Chan et al. (2018)), rehabilitation (e.g. low-vision rehabilitation treatment Ehrlich et al. (2017)), diagnosis (e.g. early assessment of Alzheimer's disease Howett et al. (2019)), or education (e.g. nursing education Cieslowski et al. (2023)) and training (e.g. surgical training Mao et al. (2021)). Beyond VR, the Metaverse concept can be defined as a 3D virtual space for social interactions and activities. It encompasses XR technologies, avatars, and persistent 3D worlds, offering diverse applications in education, leisure, gaming, work, or business Park and Kim (2022). It differs from VR, however, in its focus on services, its ability to be accessed with non-immersive devices, and its necessary scalability. Current Metaverse tools and practices are mostly related to social VR, with digital avatars meeting in virtual worlds, such as those offered by VRChat, RecRoom or Workrooms.

Metaverse applications in healthcare and medicine can be categorized in five main topics: medical education and training (e.g. surgical simulations, patient education for professionals); medical visualization (immersive and collaborative analysis of patient data, 3D models ad AR-assistant for surgery, distant monitoring. . .); telerehabilitation (e.g. immersive technologies for mental or physical therapy at home); virtual consultation and telemedicine (e.g. meetings between practitioners and patients in virtual environments); and social support (e.g immersive group therapy sessions). Although the metaverse in medicine is a trending topic, many of its featured applications can appear as rebranded versions of existing technologies such as autonomous VR medical applications or AR-based assisted surgery, without taking into account its fundamental social aspects. Due to the Covid pandemics, telehealth care services have gained significant attention and acceptance Wong et al. (2021), and Metaverse-based telehealth is on the agenda Sampaio et al. (2021). However if support groups between patients have used social VR technology for a long time, virtual consultations that would allow the medical meeting of patients and practitioners are still lacking, and most of the proposals have remained mere ideas, with a lack of implementation, let alone assessment. Therefore, examining how the Metaverse and social VR can contribute to teleconsultation is needed Dwivedi et al. (2022).

VR has been widely studied for the specific case of mental health applications, Freeman et al. (2017), showing promises for diagnosing and treating conditions like dementia, anxiety, and phobias through techniques like Cognitive Behavioral Therapy (CBT) and Virtual Reality Exposure Therapy (VRET) Cieślik et al. (2020). For neuropsychological testing and training, its ability to immerse people in ecological and controlled environments for stimulating cognitive and executive functions, while measuring patient's performance and behavior, enables more precise, valid and personalized results than paper-and-pencil tests Kim et al. (2021). Medical practitioners (psychiatrists, psychologists, psychotherapists...) are increasingly incorporating VR into patient care, follow-up and monitoring during consultations, often using Head-Mounted Displays (HMDs) for their patients and dedicated commercial applications that they can control on handheld tablets. More rarely, they can lend HMD to their patients who can exercise at home. However, very few studies have focused on the use of social VR technologies for patient-practitioner meeting in the Metaverse.

In this paper, we explore the potential of social VR for immersive clinical teleconsultation. First, we detail our vision and the main principles of immersive teleclinics, where patients meet practitioners and benefit from supervised or independent assessment or remediation activities. We then present a prototype designed and implemented as part of a concrete project with psychiatrists, psychologists and researchers, aimed at providing comprehensive remote patient management, including clinical interviews and neuropsychological testing activities with therapists. We conclude with some recommendations for future works.

2 PRINCIPLES FOR IMMERSIVE TELECONSULTATION

We propose our vision for teleconsultation in the Metaverse, which is based on a few principles (Fig. 1):

1. Immersive remote consultations between medical practitioners and patients should take place in *immersive teleclinics* run by socio-economic players, such as public or private hospitals or clinics, telemedicine operators, who make social VR consultation rooms available to their staff, or lend them to independant practitioners.
2. An immersive teleclinic is a set of *spaces*, e.g. welcoming spaces, social rooms, waiting rooms, consultation rooms, collective or individual exercice rooms, etc. It also provides web access for

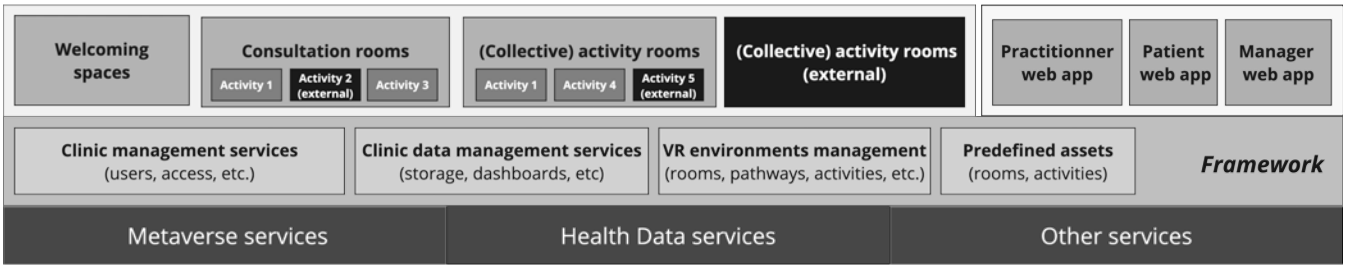


Figure 1. An immersive teleclinic as a set of social VR spaces and rooms where activities can take place, associated to web applications. The clinic makes use of external activity modules. It is built upon a framework featuring several services and predefined assets, based on various lower-level services (metaverse, web, health data, etc).

functionalities that do not need immersion, such as users and appointments management, offline access to data and dashboards, etc.

3. Depending on available technology, patients and practitioners in VR can perform various *activities* in the clinic, ranging from simple social interaction in consultation via avatars to remote examination with specific sensors, as well as a variety of exercises, either dedicated to assessment or remediation. Activities can be carried out alone (e.g. cognitive training) or collectively (e.g. elbow rehabilitation), and be accompanied or not by practitioners. Activities can take place in *consultation rooms* or dedicated *activity rooms*.
4. Activities can be already offered by the teleclinic software, but it is important that third party vendors can propose external modules that can be plugged into the clinic using an Activity API.
5. An immersive teleclinic is built on a *framework* offered by specific socio-economic players such as software companies, metaverse providers, etc. It can run on premise or in the cloud, and be connected to clinical information systems, secured health data management systems, etc. The framework is composed of a set of predefined modules for clinic and data management, as well as predefined rooms, activities, etc. that can be tailored.
6. The use of an immersive teleclinic and immersive teleconsultation should be *integrated within patients' healthcare paths* in a coherent way. The paths would be composed of real and virtual teleconsultations, depending on the activities that would be carried out. For instance, a first encounter, followed by distant testing, then a consultation associated to a prescription with distant monitoring related to performances in assessment exercises that would be carried out every week.
7. Patients at home can either possess their own HMD and be autonomous, or necessitate that a handover assistant be present, who would bring the HMD, set up the access, and accompany the patient during the session. Such profession needs to be considered in relation to existing paramedical professions, and integrated into the range of healthcare services.

This principles and the associated architecture described in Figure 1 are by no means definitive, but should give the main insights of our proposal.

3 THE CASE OF NEUROPSYCHOLOGICAL CONSULTATION: TECNIS PROJECT

The Tecnis (*Téléclinique Immersive*) project is a joint program between Nantes University and the University Hospital of Nantes. Its aim is to investigate how to take immersive cognitive testing to the next level, by integrating them into current practices. The interdisciplinary team is composed of one neuro-psychologist, one psychiatrist, and several clinical psychologists, XR/UX/HCI scientists, a UX researcher and a development engineer.

The methodology we followed is classical user-centered design. We first carried out several interviews and focus groups with various stakeholders outside of the team (neuro-psychologists, psychiatrists, neurologists, geriatrician, general practitioner), so as to describe what psychiatric patients' journey are, from detection to long-term follow-up or remission, and identify where and when immersive technologies could be useful. We also developed our general vision (section 2), and wrote numerous scenarios related to the various possibilities, both in VR and AR, before settling on the topic we would be focusing on, namely neuropsychological teleconsultation, for carrying out tests remotely. We made this choice because such testing is a time-consuming activity in a overloaded healthcare sector, that do not rely solely on dialog, that can be instrumented with data, and that can re-occur for longitudinal follow-up. We then held 8 interdisciplinary co-design workshops that helped us make our key design decisions, define our spaces, the types of activities that would be carried out, as well as the interactions. We also underwent used immersive prototyping, and had our prototypes tested and commented upon by a ”partner patient”, several nurses and hospital patients.

Our main design decisions were as follows.

- We should allow patients and clinicians to meet in VR during full clinical sessions composed of discussions, neuropsychological testing, and debriefings, as in conventional practice. The meeting would take place on a virtual desk, while each user would seat in front of a table/desk in reality.
- We should provide a waiting room for patients to wait for the practitioner, who could be late, as in real life, first because it acts as a transition space before a medical meeting, and second because waiting time is specifically important in psychotherapy, to help patients focus.
- Apart from psychotherapeutic interview, activities in immersive psychotherapy are mainly related to *cognitive performance testing and/or cognitive remediation or training*. Those activities can be carried out either on the practitioner's desk, or in ”standalone” independent VR environments. They can be performed alone or accompanied by the practitioner.
- As in a real psychoterapeutic consultation, the practitioner should be able to lead the session and decide of the activities beforehand, or during the session with a control panel. He should be provided with data-based tools for lived enriched observation during interviews and activities, as well as dashboards for post-activity or post-session debriefing.
- In both desk-based and standalone activities, behavioral data is to be automatically collected so as to produce indicators for the practitioner to access to in dedicated real-time, post-session or longitudinal dashboards.

The patient / practitioner workflow is described on Fig. 2. First both users are welcomed in their home space, where they can consult their agenda, specify their avatar, follow tutorials, etc. From there, patients teleport in the waiting room, where several waiting activities are proposed, to spend between 5 and 20 minutes, while practitioners teleport in the consultation room where they can prepare the upcoming session. When they are ready, practitioners teleport into the waiting room so as to welcome patients, engage in discussion, and invite them to join the consultation room. Once there, practitioners can lead the consultation, propose desk test activities, and discuss the results. At some point, they can invite patients to stand up and head towards the standalone activity space, where standalone testing can take place, before going back at the desk for other activities or data-based debriefing.

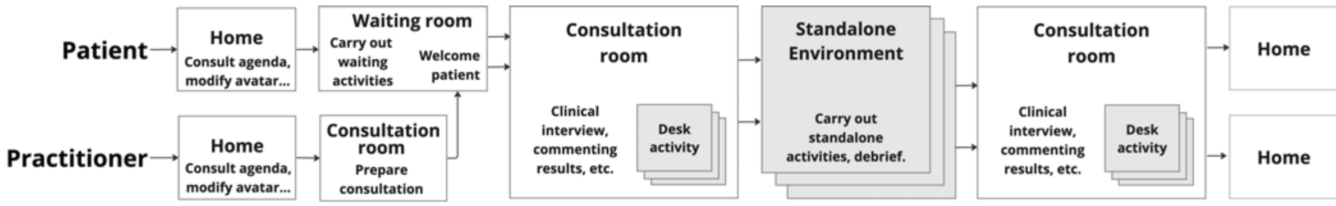


Figure 2. Simplified workflow for a consultation in Tecnis. The practitioner welcomes the patient in the waiting room before proceeding to the consultation room where a mix of discussions and activities can occur. Elements provided by the clinic are in white, activity modules plugged into the clinic in gray.

Figure 3. The clinic is composed of various environments. From left to right: Waiting room with different chairs and waiting activities, the female practitioner welcomes the patient; Consultation room with a desk-based activity (here a version of the *Tower of London* test); Both users in a standalone activity environment (here *Shelves* that we developed for executive function testing, see Ribeiro et al. (2024)).

We can now rapidly describe the various environments we have developed (Fig. 3). The *waiting room* is a environment that allows the user to teleport from chair to chair (or bench when outside in the garden) and carry out waiting activities. Amongst the ones we have designed with the medical members of our team, we implemented the following: an activity for controlling one's respiration; another for consulting medical information and newspapers on a table; a contemplation activity with large scale images accompanied with sound; and a gardening stand where patients can select and plant flowers creatively. The *consultation room* features a desk, where the practitioner has access to a control panel and several screens, one of them for sharing information with the patient. The control panel allows to launch tests, control their unfolding and access the resulting data. In our prototype e have implemented a fully functional version of the "Tower of London" test. Lastly, both the patient (by standing and moving) and the practitioner (staying in the chair) can go into *standalone tests environments*, where the patient carries out the activity while the practitioner controls and observes the task, getting information on the control tablet.

4 CONCLUSION

Beyond this, teleconsultation and teleclinics in the Metaverse should be studied for others domains than psychotherapy, and a lot of topics are waiting for careful studies: effects of avatars, integration of remote sensors technologies, experiential aspects of teleconsultation and visits to clinic in the Metaverse, business models, integration of immersive teleclinics into health systems, etc.

REFERENCES

Chan, E., Foster, S., Sambell, R., and Leong, P. (2018). Clinical efficacy of virtual reality for acute procedural pain management: A systematic review and meta-analysis. *PLOS ONE* 13, e0200987. doi:10.1371/journal.pone.0200987

Cieslowski, B., Haas, T., Oh, K. M., Chang, K., and Oetjen, C. A. (2023). The Development and Pilot Testing of Immersive Virtual Reality Simulation Training for Prelicensure Nursing Students: A Quasi-Experimental Study. *Clinical Simulation in Nursing* 77, 6–12. doi:10.1016/j.ecns.2023.02.001

Cieślik, B., Mazurek, J., Rutkowski, S., Kiper, P., Turolla, A., and Szczepańska-Gieracha, J. (2020). Virtual reality in psychiatric disorders: A systematic review of reviews. *Complementary Therapies in Medicine* 52, 102480. doi:10.1016/j.ctim.2020.102480

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., et al. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management* 66, 102542. doi:10.1016/j.ijinfomgt.2022.102542

Ehrlich, J. R., Ojeda, L. V., Wicker, D., Day, S., Howson, A., Lakshminarayanan, V., et al. (2017). Head-Mounted Display Technology for Low-Vision Rehabilitation and Vision Enhancement. *American Journal of Ophthalmology* 176, 26–32. doi:10.1016/j.ajo.2016.12.021

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., et al. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine* 47, 2393–2400. doi:10.1017/S003329171700040X

Howett, D., Castegnaro, A., Krzywicka, K., Hagman, J., Marchment, D., Henson, R., et al. (2019). Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation. *Brain* 142, 1751–1766. doi:10.1093/brain/awz116

Kim, E., Han, J., Choi, H., Prié, Y., Vigier, T., Bulteau, S., et al. (2021). Examining the Academic Trends in Neuropsychological Tests for Executive Functions Using Virtual Reality: Systematic Literature Review. *JMIR Serious Games* 9, e30249. doi:10.2196/30249

Mao, R. Q., Lan, L., Kay, J., Lohre, R., Ayeni, O. R., Goel, D. P., et al. (2021). Immersive Virtual Reality for Surgical Training: A Systematic Review. *J. Surgical Research* 268, 40–58. doi:10.1016/j.jss.2021.06.045

Park, S.-M. and Kim, Y.-G. (2022). A Metaverse: Taxonomy, Components, Applications, and Open Challenges. *IEEE Access* 10, 4209–4251. doi:10.1109/ACCESS.2021.3140175

Ribeiro, N., Vigier, T., Han, J., Kwon, G. H., Choi, H., Bulteau, S., et al. (2024). Three Virtual Reality Environments for the Assessment of Executive Functioning Using Performance Scores and Kinematics: An Embodied and Ecological Approach to Cognition. *Cyberpsychology, Behavior, and Social Networking* 27, 127–134. doi:10.1089/cyber.2023.0314

Sampaio, M., Navarro Haro, M. V., De Sousa, B., Vieira Melo, W., and Hoffman, H. G. (2021). Therapists Make the Switch to Telepsychology to Safely Continue Treating Their Patients During the COVID-19 Pandemic. Virtual Reality Telepsychology May Be Next. *Frontiers in Virtual Reality* 1. doi:10.3389/frvir.2020.576421

Wong, M. Y. Z., Gunasekeran, D. V., Nusinovici, S., Sabanayagam, C., Yeo, K. K., Cheng, C.-Y., et al. (2021). Telehealth Demand Trends During the COVID-19 Pandemic in the Top 50 Most Affected Countries: Infodemiological Evaluation. *JMIR Public Health and Surveillance* 7, e24445. doi:10.2196/24445

Virtual, Personalized Rehabilitation Environment

Paweł Śniatała[1*], Szymon Baliński[1], Mariusz Głąbowski[1],
Sławomir Michalak[2,3], Ewa Kozielewska – Zwierska[2,3] ,

[1]Faculty of Computing and Telecommunication, Poznan University of Technology, Poznan, Poland

[2]Stroke Unit, University Hospital in Poznan, Poland

[3]Department of Neurosurgery and Neurotraumatology, Poznan University of Medical Sciences, Poznan, Poland

*** Correspondence:**
Pawel Sniatala
pawel.sniatala@put.poznan.pl

Keywords: Metaverse, Neurology, Rehabilitation, Augmented Reality, Virtual Reality

Abstract

The paper presents a rehabilitation system using virtual (VR) and augmented (AR) reality. AR, implemented as a "Smart Mirror", is used as an intelligent interface that, once a person is recognized, personalizes further dialogue with the user. Next, VR offers, a dedicated, for the person set of rehabilitation exercises.

1 Introduction

Health care is one of the most important factors for the overall, physical, social and mental well-being of the entire world's people. Augmented and virtual realities in healthcare market size was valued at more than USD 2.5 bilion in 2022 and is predicted to register over 21% CAGR during the forest period 2023-2032 [1]. Emerging new technologies, if possible, are used in the Health Care area. Metaverse is one such technology, so it is used in many areas of Health Care [2, 3]. Virtual Reality has become a novel method for stroke rehabilitation within the past decade. Through the simulation of everyday activities, individuals recovering from strokes can enhance their self-care abilities in a manner that is typically unattainable in a hospital setting.

Among the many applications of Metaverse, in this article we would like to present the use of this technology in the rehabilitation process of the elderly and/or those in the rehabilitation process after various types of injuries or strokes. Currently, the elaborated system is used in the rehabilitation of neurological patients (e.g., after strokes). As is well known, the rehabilitation process in the case of neurological diseases, is a long process, and after a period of hospitalization, then, already at home, the patient must independently perform, often tedious, various exercises. Some of these exercises can be proposed to be performed in virtual reality. Performing these exercises in virtual reality helps to increase their attractiveness and, therefore, their effectiveness (the patient is more likely to perform the exercises in a properly designed, attractive virtual reality). Of course, assuming that the motor requirements appropriate for the exercises set are met.

2 Medical Background

The process of rehabilitation is initiated early in the acute phase of cerebral ischemia. Later, it continues during the patient's stay in the Stroke Unit. The disabilities caused by stroke vary; thus, there is a need for tailored post-stroke rehabilitation that will be appropriate for the needs of stroke patients and their everyday activities at home. Most stroke survivors are discharged into their homes. Thus, the question of how to design rehabilitation training appropriate for use in the home should be raised. To initiate the process of adaptation in some Stroke Centers (e.g., University Hospital in Poznan) a, "model apartments" (Fig. 1) are arranged for initial patient training. Using different objects like balls, iron, spoons, cups, etc., as physical therapy tools in a "model apartment" supports the patient's everyday activity at home. As technology becomes more pervasive and familiar, it can support rehabilitation in hospital and/or home environments.

The success of post-stroke rehabilitation depends on factors like patient pre-stroke activity, circulatory sufficiency, the severity and location of the stroke, and support from family and caregivers. Active participation of the stroke patient in the rehabilitation process is crucial for optimal recovery. Thus, the need for tailored rehabilitation in post-stroke patients is clearly expressed [4] and guidelines were already addressed as "evidence-based rehabilitation of mobility after stroke (ReMoS)" [5].

As a result of cooperation with Poznan University of Technology, the Virtual Reality (VR) Systems (Fig. 2) are currently used in everyday practice in the Stroke Unit at University Hospital in Poznan. Stroke patients are mainly older persons who are only sometimes familiar with computer technology. However, unexpectedly, their tolerance of VR is excellent. Currently, we perform a study on the tolerance of VR in stroke patients. A questionnaire was elaborated that included the presence of vertigo, nausea, diplopia, headache, chest pain, arrhythmia, anxiety, and sweating before and after training with VR. Moreover, the use of computer/smartphone/games at home, the educational level and profession are considered. Blood pressure and heart rate before and after VR training are also monitored. Patient evaluation with the questionnaire is performed at baseline and after 7 days of the training. We noticed an excellent tolerance of VR training in stroke patients. Interestingly, some symptoms, like intension tremor, that are present in real-world training, disappeared in VR training.

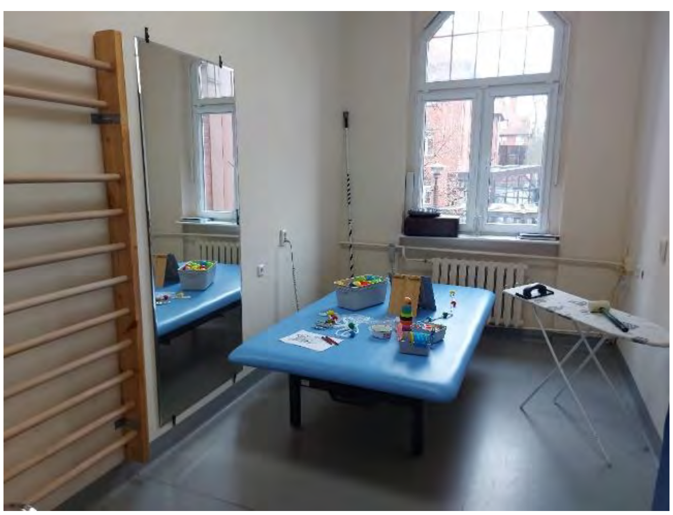


Figure 1. “Disability Escape-Room” as a “model apartment” in Stroke Unit at University Hospital in Poznan.

Figure 2. VR scenario for neuro-rehabilitation.

2.1 Disseminate neurorehabilitation

The use of technology and VR enables the dissemination of neurorehabilitation, which enables reaching such goals like:

- physical and cognitive restoration
- rehabilitation techniques enrichment
- increased effectiveness
- wider access to society
- Tele-neurorehabilitation.

Neurorehabilitation supported with simple, everyday activity tools and advanced technology, opens the possibilities that improve motor training and cognitive functions and enables the treatment of disabling symptoms like neglect syndrome.

3 Virtual Rehabilitation System

The general idea of the system is presented in Fig. 3. In the proposed solution, we use elements of augmented reality (AR) and Virtual Reality (VR). AR is implemented in the form of a "Smart Mirror", which in our system is used to recognize the patient/person being rehabilitated in order to then suggest personalized, for that person, exercises in the virtual world. In the presented sample scenario, after recognizing the person ("John") according to the planned course of the rehabilitation process, he has, for example, proposed origami exercises in the virtual world. Of course, the range of proposed exercises depends on the suggestions of the doctor and rehabilitator and, as was mentioned, dedicated to a specific person.

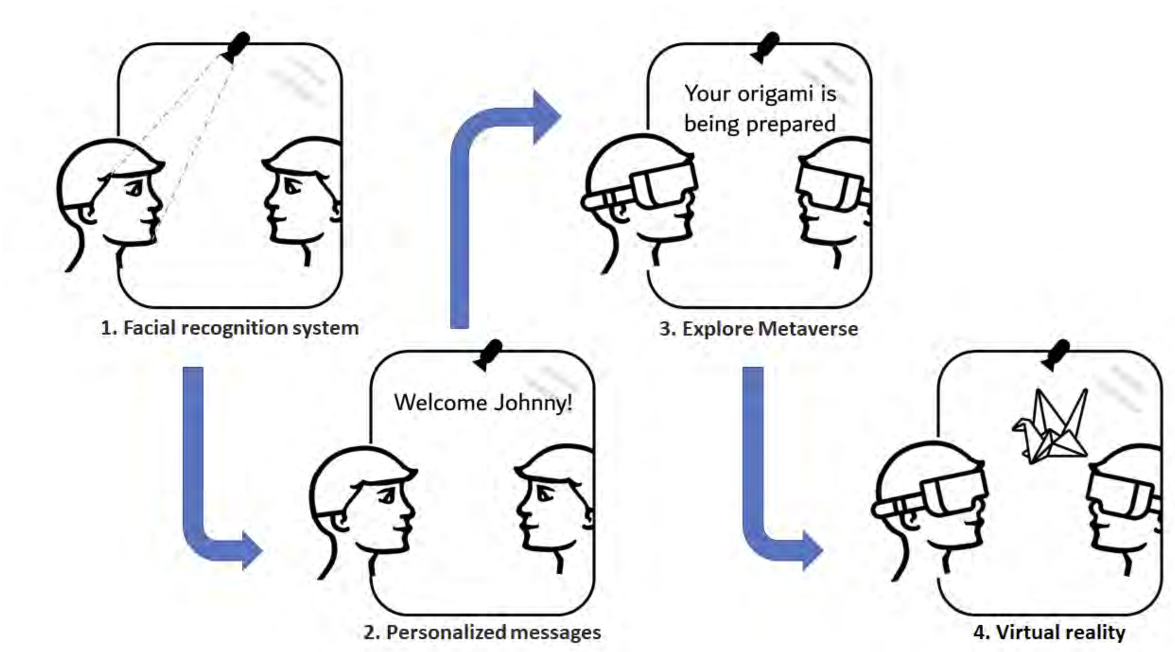


Figure 3. General idea of the Virtual Rehabilitation System

3.1 Smart Mirror

Human-machine interaction becomes an important issue in the digitalization era. We want a machine to interact with us in a personalized way. It means that the machine should recognize us and next offer personalized actions. This idea was utilized in the aforementioned “Virtual Rehabilitation System” by a design a “Smart Mirror” presented in Fig. 4. One of the key features of this smart mirror is its ability to recognize individual faces and personalize the information displayed on the mirror. The mirror interacts with the user, learning their interests and preferences, and displaying personalized content that caters to their specific needs. Additionally, users are able to communicate with the mirror using voice commands, making the interaction more intuitive and user-friendly. In our system, which is dedicated to the rehabilitation of patients, the person interacting with the mirror is offered activities that best suit his/her current state of recovery/rehabilitation.

Figure 4. Implementation of the Smart Mirror

3.2 Smart Miror – behind the scene

The system is running on a Raspberry Pi platform. However, the platform communicates with more powerful computer where the core application is running. It is a web based application implemented in Django environment. Facial recognition is implemented with the help of the PyTorch and OpenCV libraries. Required data are stored in a SQL database (in our case MariaDB). The software structure of the system is presented in Fig. 5. The facial recognition feature of the Smart Mirror is powered by AI-based algorithms that can recognize faces and compare them to faces in the system's database. The system's ability to differentiate between authorized and unauthorized persons is based on this technology, which can detect even subtle differences between faces. With this technology, the Smart Mirror provides enhanced security by monitoring who enters the room and alerting users when an unauthorized person is detected.

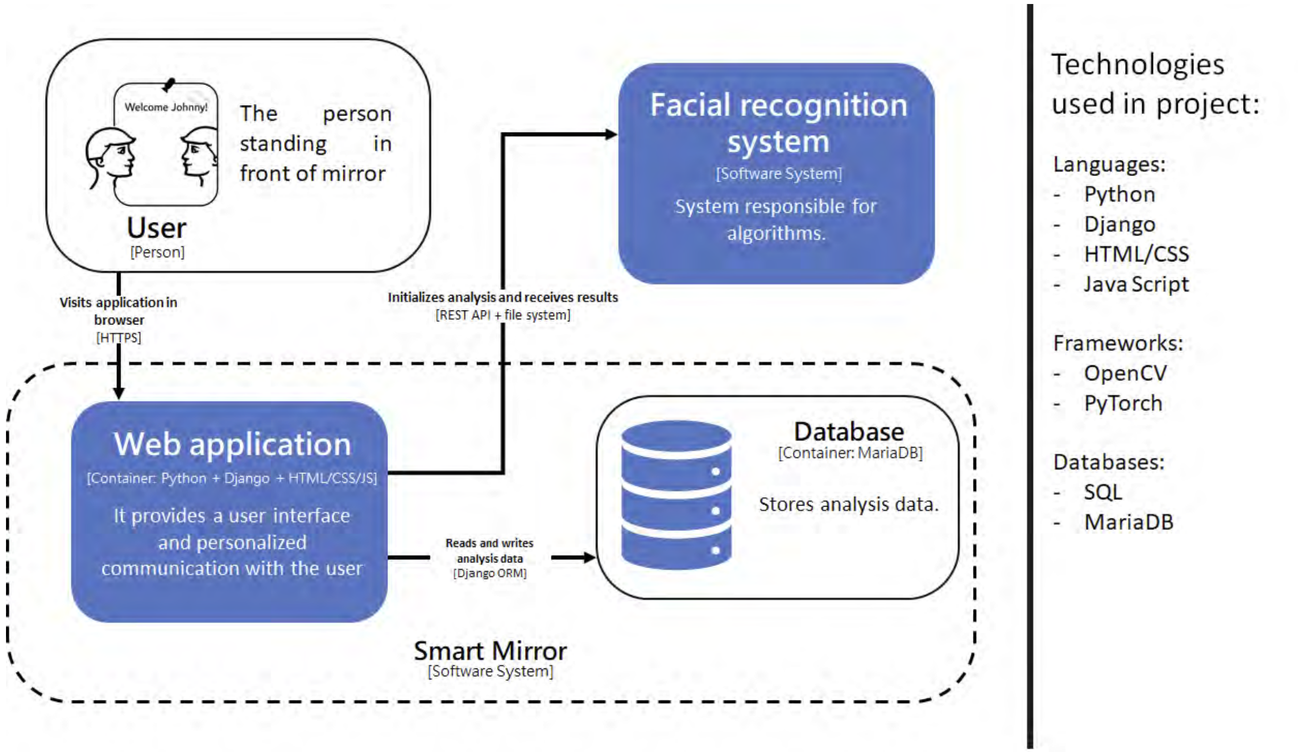


Figure 5. Software implemented "behind" the Smart Mirror.

The facial recognition feature of the Smart Mirror is powered by AI-based algorithms that can recognize faces and compare them to faces in the system's database. The system's ability to differentiate between authorized and unauthorized persons is based on this technology, which can detect even subtle differences between faces. With this technology, the Smart Mirror provides enhanced security by monitoring who enters the room and alerting users when an unauthorized person is detected.

3.3 Virtual exercises

As was illustrated in the Fig. 3, the Virtual Rehabilitation System will propose the most appropriate and/or expected virtual interaction with the recognized person. In our case, it can be an origami exercise. In fact, origami is a good example proposal since it requires hand-eye coordination, develops fine-motor skills and supports mental concentration – all of which stimulate the brain). The origami exercise is done in a virtual reality. Of course, there are many others existing application, which can be used for the rehabilitation purposes [6].

4 Conclusion

Over the last decade, Virtual Reality has emerged as a novel approach to stroke rehabilitation, offering a unique treatment method. By replicating real-world activities, individuals recovering from strokes can engage in practicing self-care tasks within an environment that is typically unfeasible to recreate within a hospital setting. The utilization of virtual reality in this context is becoming increasingly prevalent, with its potential medical applications far from being fully explored. Its profound impact on stroke survivors is evident, as they leverage VR technology to rehearse essential daily activities, foster new neural connections, and enhance their self-assurance. As more and more survivors utilize this technology to retrain their limbs, the future of Virtual Reality in stroke recovery appears promising. VR has recently emerged as a valid addition to conventional therapy by incorporating rehabilitation strategies in a novel and low-cost approach [7]. VR-based therapy can provide a positive learning experience, and be engaging and motivating. The system presented in this paper was prepared and implemented for a patients under the care of doctors from the Stroke Unit of University Hospital in Poznan, Poland. Exercises carried out with patients using the described system produce positive results. Further expansion of the palette of exercises available in virtual reality is planned. In the future, we plan to install such systems in elderly homes, which will allow remote rehabilitation of people staying there.

5 References

[1] Mariam Faizullabhoy, Gauri Wani, "Augmented and Virtual Reality in Healthcare Market", Report ID:GMI3144, Dec. 2022. www.gminsights.com, .

[2] Al-Ghaili AM, Kasim H, Al-Hada NM, Hassan ZB, Othman M, Tharik JH, et al. A Review of Metaverse's Definitions, Architecture, Applications, Challenges, Issues, Solutions, and Future Trends. IEEE Access. 2022;10:125835–66.

[3] Chengoden R, Victor N, Huynh-The T, Yenduri G, H.Jhaveri R, Alazab M, et al. Metaverse for Healthcare: A Survey on Potential Applications, Challenges and Future Directions. IEEE Access. 2023;11:12764–94.

[4] Scheffler B, Schimböck F, Schöler A, Rösner K, Spallek J, Kopkow C. Tailored GuideLine Implementation in STrokE Rehabilitation (GLISTER) in Germany. Protocol of a Mixed Methods Study Using the Behavior Change Wheel and the Theoretical Domains Framework. Front Neurol. 2022 Jul 27;13:828521. doi: 10.3389/fneur.2022.828521

[5] Pogrebnoy D, Dennett A. Exercise programs delivered according to guidelines improve mobility in people with stroke: a systematic review and meta-analysis. Arch Phys Med Rehabil 2019. 10.1016/j.apmr.2019.06.015

[6] Vyas, P., Harris, M.C., Brown, D.J., Shires, L. (2023). Games for Stroke Rehabilitation: An Overview. In: Brooks, E., Brown, D.J. (eds) Virtual Reality Games for Rehabilitation. Virtual Reality Technologies for Health and Clinical Applications. Springer, New York, NY. https://doi.org/10.1007/978-1-0716-3371-7_1

[7] Laver K.E., Lange B., George S., Deutsch J.E., Saposnik G., Crotty M. (2017) Virtual reality for stroke rehabilitation. Cochrane Database of Systematic Reviews. Issue 11. Art. No.: CD008349. doi: 10.1002/14651858.CD008349.pub4.

Identity and the body – metaverse applications

Avatars in mixed reality meetings: A field study of avatar facial realism on felt and perceived emotion

Georgiana Cristina Dobre [1,2,*], **Marta Wilczkowiak** [3], **Marco Gillies** [2], **Xueni Pan** [2] **and Sean Rintel** [3]

[1] *Human-Centered Artificial Intelligence, University of Augsburg, Augsburg, Germany*
[2] *SEEVR Lab, Department of Computing, Goldsmiths, University of London, London, United Kingdom*
[3] *Microsoft Research, Cambridge, United Kingdom*

Correspondence*:
Georgiana Cristina Dobre
cristina.dobre@uni-a.de

ABSTRACT

We conducted a within-subjects study to examine how realistic faces and cartoon faces on avatars effect communication, task satisfaction, sense of presence, and emotional state perception in mixed reality meetings. Over the course of two weeks, six groups of co-workers (14 people) held recurring meetings using Microsoft HoloLens2 devices, each person embodying a personal full-body avatar with either a realistic face or cartoon face. Half of the groups started with the realistic face avatar and switched to the cartoon face version halfway through (RC condition), and the other half with the cartoon-face avatar first (CR condition). For this paper, we focus on our results of emotion felt and perceived. We found that participants reported more positive emotions with the realistic face avatar but only in the RC condition. This positive emotion also seemed to have reduced over time. The realistic face avatar also caused more error in judgement of their colleagues emotions, and again this was worse in the RC condition.

Keywords: avatar facial realism, mixed reality, meetings, work, field study, longitudinal, communication, presence, tasks, emotion

1 INTRODUCTION

As remote meetings increase, there has been a growing demand for $3D$ immersive systems that address the limitations of traditional $2D$ formats. The aim of immersive systems is to connect remote users as if they were in the same location, allowing people to work more effectively on shared tasks by preserving spatial relationships Otto et al. (2006) and social behaviours such as proximity or gaze Bailenson et al. (2001). Avatar representations in immersive systems may range from floating spheres with hands to full or partial humanoid bodies with different appearance styles (e.g., cartoon, realistic). There are positive and negative aspects to different avatar styles. Realistic avatars may make people feel uncomfortable of affinity Shin et al. (2019). Cartoon styling may lead to anxiousness about the appropriateness in a professional context Bailenson and Beall (2006).

Most of what we know about avatar appearance in meeting-style settings comes from one-off lab studies in virtual reality environments, in which research participants often look at only short animations or still images of avatars Waltemate et al. (2018); Lugrin et al. (2015). However, real-life collaborative work involves users who know each other and interact regularly trying to get real work done. Verbally, this is served well by the spatial audio common to most immersive environments. Visually, many questions remain around how different avatar styles enable identification, recognition of facial expressions and gestures Burgoon et al. (2016), negotiate proxemics Hall et al. (1968), and authenticity Oh et al. (2018). Little is known about how these findings apply to MR, the effects in real-world contexts, and the longitudinal effects on avatar acceptance. In this paper, we report our results of six groups of co-workers (14 people) from a global technology company conducting a series of virtual meetings using Microsoft HoloLens2 (HL2) devices for $2-3$ weeks. Each participant used a personalised avatar with either a realistic (R) face or cartoon (C) face. Half the groups began with R avatars and the other half with C avatars; all groups switched halfway through the study period. Our main focus was to determine whether the acceptance ratings for both avatar styles would change as novelty waned. We have reported our previous findings in Dobre et al. (2022). Here, we focus on our findings in self-reported and perceived emotional states of individuals during immersive virtual meetings. Our research questions focuse on the ability to recognise others' emotional states. **RQ1**: Does the avatar representation change the self-reported emotional states (a) overall and (b) over time? **RQ2**: Does the avatar representation effect how accurately people perceive emotional states (a) overall and (b) do they improve over time?

2 METHODOLOGY

A summary of our method is provided below. Further details can be found in Dobre et al. (2022).

Participants and Tasks. Following ethics authorization (provided by Microsoft Research's Institutional Review Board (IORG0008066, IRB00009672)), we recruited Microsoft employees in groups of 2 or 3 via emails. All participants worked together and were willing to conduct one of their regular daily meetings in mixed reality using HL2 for a period of 2 to 3 weeks (10 meetings). 14 participants (7 female, 6 male, 1 non-binary; aged $21-45$) completed the study, forming 6 groups. The members of each group remained the same throughout the study. To maintain a high level of ecological validity, participants conducted a regular work meeting for at least $10-15$ minutes (often stand-ups, status reports, or catch-ups).

Avatars. Participants used full-body avatars in C and R style,with personalised faces. There were two male and two female bodies, each with a C and R appearance. The avatars were animated in real-time using inverse kinematics: the hands moved when the HL2 detected hand movement using its external cameras, and the legs moved when the headset detected location movement based on the headset's position. Facial animation was generated using a blink animation and simple lip-flapping based on voice amplitude.

Procedure. Participants' pictures were used to create their C and R avatars. For each of the actual experiment meeting sessions, the participants logged on via HL2 app menu, changed their avatar to the corresponding one for that week (C or R), and appeared around a virtual table, standing for the duration of their meeting. After the meeting, participants were reminded to complete the corresponding questionnaires.

Avatar Order. We controlled the order in which the avatar styles were used. Half of the participants used the C avatar in Week 1 (W1) and then R in Week 2 (W2). The other half swapped the avatars order. We coded the data from participants who used the C avatar in W1 followed by R in W2 the *Cartoon—Realistic* **condition (CR)**, and the data from those who had the R avatar first *Realistic—Cartoon* **condition (RC)**.

Data. We collected data from questionnaires and daily meetings. The participants completed consent forms and the following questionnaires: demographic and on-boarding, and after each meeting a daily questionnaire. Each group alternated between using one avatar type for half of their meetings and the other avatar type for the other half. In total 54 meetings were held, resulting in 124 daily questionnaire responses (although one was missed due to a technical issue). This means there were 63 questionnaire responses from meetings with R avatars and 61 responses from meetings with C avatars. Half of the groups used C avatars first, while the other half used R avatars first. Unfortunately, two groups were unable to complete all 10 sessions due to circumstances beyond our control: one triad had 8 sessions (with R avatars first) and one dyad had 6 sessions (with C avatars first). Both triads also balanced the order of avatar use, with one starting with C avatars and the other starting with R avatars. The daily questionnaire was divided into two parts. The first part focused on communicative effects (which were reported previously in Dobre et al. (2022)). The second part focused on the recognition of emotional states and the usefulness of cues for perceiving these emotional states. Four emotional states were selected based on the UWIST mood checklist Matthews et al. (1990): optimistic, focused, annoyed, and stressed. Participants were asked to rate their own emotional states and the perceived emotional states of their colleagues on a $1-7$ Likert scale ranging from "strongly disagree" to "strongly agree." Participants in triads rated the perceived emotional states of the other two participants.

Data Analysis. For calculating average emotion state scores, the negative emotions (Annoyed and Stressed) were *reverse coded* to represent an overall degree of positive emotion. We calculated the accuracy of the perceived emotional state by computing the error that participants had when perceiving the emotional states of their colleagues. We determined the error by mapping the absolute value of the difference between the self-reported emotional state and the perceived rating of the emotional state onto a scale of $[0,1]$. With a maximum rating of 7 and a minimum rating of 1, the largest possible error was 6 $(7-1)$, which was mapped to a value of 1. The smallest error, which occurred when the self-reported rating was the same as the perceived rating of the emotional state, was mapped to a value of 0. For example, if the self-reported rating was 6 and the perceived rating was 2, then the error was 4 $(6-2)$; this was then mapped between $[0,1]$, gaining the value of 0.667. We calculated this error for each pair of participants in a group. In dyads, we considered the error for each participant in perceiving the emotional state of their colleague: P1's error in perceiving P2's emotional state ($P1\_to\_P2$) and P2's error in perceiving P1's emotional state ($P2\_to\_P1$). In triads, we considered each participant's error in perceiving the emotional states of all other participants in the triad. For example, in a group with participants P1, P2, and P3, we took into account all six possible combinations: $P1\_to\_P2$, $P1\_to\_P3$, $P2\_to\_P1$, $P2\_to\_P3$, $P3\_to\_P1$, and $P3\_to\_P2$.

3 RESULTS

RQ1a: Self-Reported Emotions Overall. Participants self-reported their emotional states daily after each meeting. We were interested in the self-reported emotional state while participants embodied different avatar styles and how this changes over time. We calculated the self-reported score for each avatar for each participant, averaging the data from Week 1 and Week 2. A paired t-test was conducted to compare the scores from C and R avatars, regardless of the Order or Emotional State, and no significant difference was found (p= 0.98; mean$\pm$ variance - C: 4.27 ± 1.28; R: 4.26 ± 1.26).

As the order in which participants experience an avatar could influence their subjective experience, we also conducted separate paired t-tests between R and C avatars for each condition. We found no significant difference in self-reported emotional state for the CR order (p= $.65$ C W1: 4.75 ± 2.07; R W2: 4.69 ± 1.83). However, for the RC order, participants reported more positive emotions when using R avatars in Week 1 compared to C avatars in Week 2 (**p=.009**, R W1: 4.9 ± 1.29; C W2: 4.23 ± 1.81), see Fig. 1A. We also conducted paired t-tests for each Emotional State and for each order, but no significant results were found.

For each participant we calculated the average score for each avatar and each Emotional State (ES) separately. Here we did not reverse the ratings for Annoyed and Stressed (i.e., a high in annoyed would indicate negative emotion). A paired t-test was conducted to compare the self-reported ES ratings for C and R avatars, regardless of condition. Participants self-reported feeling more Optimistic in meetings using R avatars compared to C (**p=.025**; C: 4.6 ± 1.3; R: 5.06 ± 1.68), see Fig. 1B. There was no significant difference for the other emotions (Focused p=.92, Annoyed p=.58, or Stressed p=.71).

RQ1b: Self-Reported Emotions overtime. The data was split by avatar style and weekly session of avatar use, and regression analyses were conducted on the self-reported emotional states over time. We found a significant result for Realistic avatars used in the first week. Specifically, while using R avatars in week 1, participants self-reported feeling less Optimistic (**p=.034**) and more Stressed over time (**p=.024**). There were no significant results for the Focused and Annoyed emotional states or for the C avatars.

RQ2a: Accuracy of perceived emotional states overall. We were interested in the effect of avatar type on the accuracy of people's perception of their co-workers' emotional states. A repeated measure 2×4 ANOVA on the Normalised Error using avatar style (R, C) and emotional state (Optimistic, Focused, Annoyed, Stressed) as within-subjects factors, and order (CR and RC) as a between-subjects factor showed a significant difference in the error of perceived emotional states when using C versus R avatars ($F(1, 18) = 5.13$, **p=.036**, $\eta^2 = .22$). Specifically, participants perceived their colleagues' emotional states with fewer errors when using C (M=.22) compared to R (M=.25).

We found a significant interaction effect between the order and the type of avatar ($F(1, 18) = 5.91$, **p=.026**, $\eta^2 = .247$). This is evident of more errors for R avatars when they were used in the RC order compared to CR. Specifically, when participants used R avatars in week 1, their mean error rate was .270, but .226 when R avatars were used in week 2.

For the interaction effect between avatar style and order, we conducted a post-hoc analysis using a paired t-test to compare errors made by participants in the CR and RC orders. No significant difference was found in the CR order (p= .91; C W1:$.228 \pm .006$; R W2: $.226 \pm .003$). However, in the RC order, participants made more errors in perceiving their colleagues' emotional states while using the R avatar in the first week and then the C avatar in the second week (**p=.004**; R W1: $.27 \pm .004$; C W2: $.21 \pm .001$), see Fig. 1C.

There was also an interaction effect between the order, the avatar style, and the emotional state of the participant ($F(3, 18) = 3.71$, **p=.017**, $\eta^2 = .171$), see Fig. 1D. When examining the data by emotional state, we found that the error rates varied depending on the order and style of avatar used. For C avatars, the error rate was generally higher for the positive emotional states (Optimistic and Focused) in the RC order compared to the CR order. However, for the negative emotional states (Annoyed and Stressed), the pattern was reversed, with the error rate being higher for the CR order. For Realistic avatars, the error rate was higher for almost all emotional states except for the Optimistic state.

A two-factors ANOVA with avatar style (C and R) as the dependent variable and order as the between-subjects factor revealed significant differences between C and R avatars for the Annoyed (F=4.41, **p=.05**) and Stressed (F=5.32, **p=.033**) emotional states, but no significant differences were found for the Optimistic (F=2.77, p= .11) or Focused (F=.11, p= .74) emotional states.

RQ2b: Accuracy of perceived emotional states over time. We calculated the regression for each avatar style considering the order they were used in the weekly meetings, each of which had data from 5

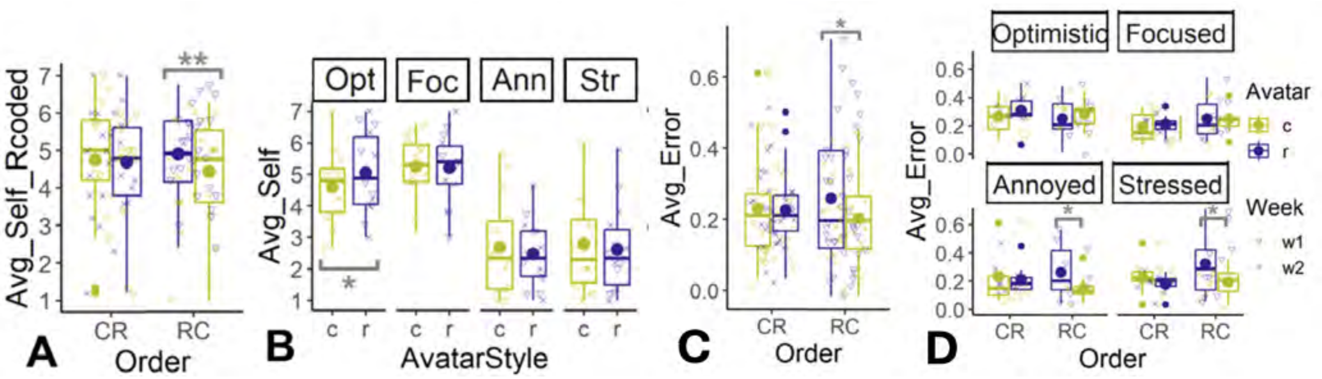


Figure 1. Averaged self-reported emotional state (reverse coded for Stressed & Annoyed) (**A & B**); Averaged normalised error of the perceived emotional state (**C & D**).

meetings. We found significant results for the R avatar only. For those starting with R avatars in W1, there was an increase in the error over time of perceiving others as Optimistic (**p=.05**) and Annoyed (**p=.02**). However, for those participants who did not use R avatars until their second week (CR condition), there was a decrease in the error for Annoyed (**p=.01**) and Focused (**p<.001**).

4 DISCUSSION

Participants reported more positive emotions with the Realistic face avatar, but only when they had it first (RC condition)(**RQ1a**). However, over time, this positive emotion seemed to diminish (**RQ1b**). For **RQ2a**, it is the Realistic face avatar which seems to be causing more error in judging their colleague's emotions; again, this was worse when they had the R avatar first (RC condition). Over time (**RQ2b**), those in the RC condition seemed to have performed worse with the realistic avatar. However, in the CR condition there was an improvement in their performance with the same realistic face avatar.

Realism of appearance can create confusion about the realism of expressions. The realism of avatar appearance seems to have an impact on our emotions and quality of our judgements. Overall, when using Realistic face avatars, participant self-reported more optimism while making more emotion judgement errors. Here, the order in which the avatars were introduced played a role in how participants felt and perceived emotions: when starting with Realistic face avatars, participants reported feeling more positive and made more errors in judging others' emotions compared to when they used Cartoon face avatars in the subsequent week or compared to when they used realistic avatars after cartoon avatars.

Using Realistic face avatars first had impact on errors made when using Cartoon face avatars, too: when using Cartoon face avatars in week 2, emotion judgement errors where higher for positive emotions and lower for negative emotions than when using Cartoon face avatars first. **This suggests that starting with more Realistic face avatars may lead to an over-reliance on expressions and subsequent judgement errors.** This might be related to findings from our previous work Dobre et al. (2022) that reported that participants rated Realistic face avatars as more appropriate and useful for understanding others, despite using the same animation technology as Cartoon face avatars.

Shifting Assessment Over Time. Examining temporal effects sheds new light on users' experience when utilizing Realistic face avatars. Over time, the optimism when using Realistic face avatars (RC order) decreased and stress increased. Interestingly, the errors in emotion judgement increased for Realistic face avatars in RC order, while they decrease in the CR order. The decrease in judgement errors in CR order could be linked to the findings from Dobre et al. (2022): the importance of avatar appearance decreased over time for both Cartoon face and Realistic face avatars in that order. Meanwhile, the increased errors in RC order could be attributed to an over-reliance on expressions discussed previously and a lack of feedback loop. More in-depth analysis would require study over longer period of time and analysis of other factors, such as correlation of emotion judgement errors with emotional state and analysis of animation cues.

Implications

First impressions set expectations for mixed reality meetings, but these change over time, and realism's communicative value is fragile and can lead to errors in emotional perception. Greater ecological validity in evaluating avatar use may provide more value than meeting thresholds of accuracy in realistic depiction (Wilczkowiak et al. (2024)).

REFERENCES

Bailenson, J. N. and Beall, A. C. (2006). Transformed Social Interaction: Exploring the Digital Plasticity of Avatars. In *Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments*, eds. R. Schroeder and A.-S. Axelsson (Springer Netherlands). 1–16. doi:10.1007/1-4020-3898-4_1

Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2001). Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments* 10, 583–598

Burgoon, J. K., Guerrero, L. K., and Floyd, K. (2016). *Nonverbal communication* (Routledge)

Dobre, G. C., Wilczkowiak, M., Gillies, M., Pan, X., and Rintel, S. (2022). Nice is different than good: Longitudinal communicative effects of realistic and cartoon avatars in real mixed reality work meetings. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–7

Hall, E. T., Birdwhistell, R. L., Bock, B., Bohannan, P., Diebold Jr, A. R., Durbin, M., et al. (1968). Proxemics [and comments and replies]. *Current anthropology* 9, 83–108

Lugrin, J.-L., Latt, J., and Latoschik, M. E. (2015). Anthropomorphism and illusion of virtual body ownership. In *ICAT-EGVE*. 1–8

Matthews, G., Jones, D. M., and Chamberlain, A. G. (1990). enRefining the measurement of mood: The UWIST Mood Adjective Checklist. *British Journal of Psychology* 81, 17–42. doi:10.1111/j.2044-8295.1990.tb02343.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2044-8295.1990.tb02343.x

Oh, C. S., Bailenson, J. N., and Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI* 5. doi:10.3389/frobt.2018.00114

Otto, O., Roberts, D., and Wolff, R. (2006). A review on effective closely-coupled collaboration using immersive cve's. In *Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications*. 145–154

Shin, M., Kim, S. J., and Biocca, F. (2019). The uncanny valley: No need for any further judgments when an avatar looks eerie. *Computers in Human Behavior* 94, 100–109

Waltemate, T., Gall, D., Roth, D., Botsch, M., and Latoschik, M. E. (2018). The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE transactions on visualization and computer graphics* 24, 1643–1652

Wilczkowiak, M., Jakubzak, K., Clemoes, J., Treptow, C., Read, K., Porubanova, M., et al. (2024). Ecological validity and the evaluation of avatar facial animation noise. In *Workshop Series on Animation in Virtual and Augmented Environments (ANIVAE 2024)*

Focus groups in the metaverse: shared virtual spaces for patients, clinicians, and researchers

Amestoy, Beñat.[1*], Donegan, Tony. [1*], Calvis, Isaac[1]., Swidrak, Justyna[1]., Rodriguez, Edgar[1]., Combalia, Andrés. [2,3,4], Oliva, Ramon. [5], and Sanchez-Vives, Maria V.[1,6]

[1] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), c. Villarroel, 170, 08036 Barcelona, Spain.

[2] Servei de Cirurgia Ortopèdica i Traumatologia, Hospital Clínic de Barcelona, Universitat de Barcelona (UB), c. Villarroel, 170, 08036 Barcelona, Spain

[3] Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona (UB), c. Casanova, 143, 08036 Barcelona, Spain.

[4] Departament de Cirurgia i Especialitats Medicoquirúrgiques, Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona (UB), c. Casanova, 143, 08036 Barcelona, Spain.

[5] Event Lab, Faculty of Psychology, University of Barcelona, Barcelona, Spain

[6] Institución Catalana de Investigación y Estudios Avanzados (ICREA), Passeig de Lluís Companys, 23, 08010 Barcelona, Spain

*** Correspondence:**
Corresponding Author
msanche3@recerca.clinic.cat

Keywords: metaverse, virtual reality, applications, rehabilitation, shared spaces, clinical VR

Perspective

Abstract

Patient focus groups can be invaluable for facilitating user-centered design of medical devices and new technologies, effectively capturing patient experiences to help thoroughly assess feasibility, tolerance, and usability. While the metaverse holds promise for healthcare applications, its use in patient focus groups remains unexplored. In this Perspective we discuss the potential of the metaverse for conducting focus groups with patients. The theme of the focus group was the design and development of a therapeutic virtual reality application for patients with low back pain. We carried out a pilot study comparing a focus group in a shared virtual space versus a physical location. This experience was positively received by patients, researchers, and clinicians, suggesting the metaverse is a viable medium for conducting these meetings and has potential advantages for remotely located participants, opening the doors for future expansion beyond focus groups to encompass all kinds of patient support

and information groups. This approach fosters patient-centered healthcare by helping to incorporate patient voices directly into the design process, ultimately leading to improved healthcare delivery, patient satisfaction, and treatment outcomes.

Introduction

The metaverse is envisioned as a consistent virtual environment and shared space where people can interact and collaborate in real time while represented by avatars (Mystakidis, 2022; Shoa et al., 2023). These avatars can be customized to the users' real aspect, but they may choose other representations. While the metaverse as a space where different systems seamlessly interoperate is not yet a reality, current shared virtual spaces can provide the experience of being with others and are an experimental space for investigating different aspects of virtual social interactions such as rules, feelings, emotions, behaviours, or restrictions (Schroeder, 2001; Giannopoulos et al., 2008; Gottschalk, 2010; Pan et al., 2012; Pan and Hamilton, 2018).

The healthcare landscape is undergoing a digital revolution, with advances in technology continuously reshaping how patients interact with the medical community. Patient engagement in the co-design of studies, medical devices, and intervention strategies has emerged as a critical aspect of modern healthcare (Mummah et al., 2016). This paper explores the potential applications of the metaverse for conducting meetings with patients, in particular focus groups, examining its potential benefits and addressing any existing challenges. We do that based on our own experience using a custom shared virtual space called VRUnited (Oliva et al., 2023). VRUnited is a virtual reality application aimed at supporting multiple people interacting simultaneously in the same virtual environment. Each participant is represented by an avatar, facilitating collective virtual experiences. We used VRUnited to conduct a focus group with patients in the context of the creation of a therapeutic virtual rehabilitation program for people with low back pain. We take this as a departing point in this Perspective so that we can discuss the potential of the metaverse for conducting focus groups with patients and in a wider sense, for meetings of researchers and clinicians with patients.

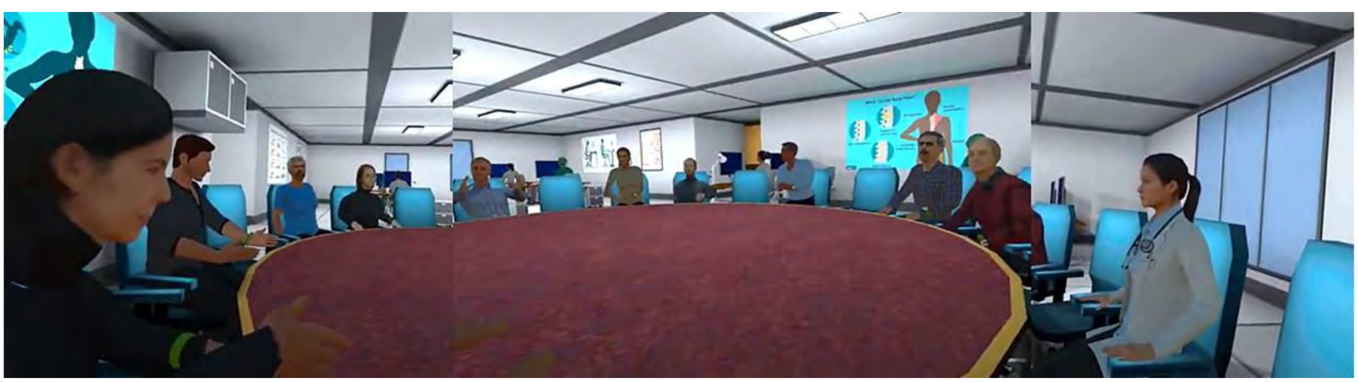


Figure 1. Patients, clinicians and researchers in a shared space in immersive VR United in order to carry out a focus group.

The need for patient focus groups in the design of digital technologies

Within the context of evidence-based design of medical devices and digital technologies, it is critical to work closely with patients throughout design and development, for developing solutions that are not only technically innovative but also highly relevant, safe, and effective for the intended users (Birckhead et al., 2019). Focus groups can offer unique insights that might not be readily apparent to engineers, designers, or even healthcare professionals (Leung and Savithiri, 2009; Dil et al., 2024), and these insights are more likely to be developed in a shared discussion setting. This approach ensures that the development of medical devices is not only technically viable but also aligned with the actual needs, preferences, and experiences of the end-users — the patients.

Patients can identify potential safety issues or risks associated with the use of the device that might not be evident in laboratory settings or through theoretical analysis. Understanding how devices will be used in real-world settings helps in designing products that maintain their efficacy outside of controlled environments, thereby reducing the likelihood of misuse or errors. Devices designed with input from patients are more likely to be embraced and used correctly, gaining faster acceptance and market penetration, and thus benefiting both the manufacturer and the healthcare community (Garmer et al., 2004). Furthermore, ensuring that devices address clinically relevant issues as identified by patients helps in aligning product development with healthcare priorities and outcomes.

Focus groups in the metaverse

There are some considerable potential advantages to conducting focus group meetings in shared virtual settings. While initial setup costs for VR equipment and software may be significant, virtual meetings can ultimately lead to cost savings by reducing the need for in-person appointments, travel expenses, and associated overhead costs for healthcare facilities. Since patients with chronic pain often have reduced mobility, allowing them to attend a meeting from home may be beneficial, and unlike other forms of videoconferencing such as Zoom, in immersive VR the user has a strong sense of presence and *really being there* with other people in the shared environment. In addition, the novelty of VR can provide a heightened engagement, and a potential distraction effect from pain, potentially leading to increased levels of interest and participation in the shared discussion.

There may also be some drawbacks or disadvantages. Technical glitches such as hardware malfunction, software bugs or poor internet connection speed, poor usability or discomfort/cybersickness using VR headsets could hinder the meeting for some patients, particularly so with longer meetings. There is also the potential for misinterpretation or miscommunication. While the virtual avatars have mouth animations triggered by the microphone in the HMD, and full upper body and head tracking providing visuomotor congruence with real life movements, other more subtle forms of non-verbal communication such as facial expression and eye gaze are currently not captured (although this is likely to change in future iterations thanks to recent technological advances in VR hardware). Additionally, the security and privacy of patient data within VR platforms would need to be carefully addressed.

To explore some of these issues, we conducted a pilot study with patients with chronic low back pain that had been independently testing a virtual reality rehabilitation program at home for five days. The program consisted of a set of therapeutic experiences, games and exercises designed for the rehabilitation of the low back pain, a development within the project "XR-PAIN: eXtended Reality-Assisted Therapy for Chronic Pain Management" (see Funding section). The system is based on embodiment of virtual bodies (Slater et al., 2009; Sanchez-Vives et al., 2010; Blanke, 2012; Maselli and Slater, 2013) and realization of a variety of movements and strategies in order to reduce pain and disability and improve range of motion and movement confidence (Matamala-Gomez et al., 2019;

Donegan et al., 2022; Álvarez de la Campa Crespo et al., 2023). Four volunteer low-back pain patients were given a VR system (Quest 3, Meta, California) with the program installed. They were instructed to use the program at home for 20 minutes daily for one week in order to pilot the contents and get their feedback. These patients also agreed to provide a front-facing photograph to have a customized avatar made for when they attended the virtual focus group.

We organized a focus group session in a shared virtual space using VR United with four patients, one clinician (traumatologist), two physiotherapists, two developers and three researchers. A structured session was organized, where in a highly interactive dialogue guided by the researchers, the patients actively reported on their experience with the rehabilitation program, content relevance and appropriateness of the contents, user experience, effectiveness, or suggestions for improvement. They also commented on different aspects of the contents, such as their experience with different games and exercises.

In a physical focus group meeting with the patients conducted 3 days after the virtual focus group meeting, the patients reported on their experience of the virtual focus group. They valued being able to be at their own homes without the time and physical effort of travelling. They found it comforting to know that they could meet in a common space. They also found it “weird the first two minutes and then it’s like a normal meeting”. At some point, they found it “very natural”. The experience was found to be less stressful and less intimidating than videoconferences, given that one does not need to prepare the environment nor their personal appearance, since the space is virtual and an avatar is used. Curiously, they also saw it as an advantage that the avatar does not show the use’s real emotions. The users found their embodiment to be “natural”. No serious adverse effects were reported after an 80 minute meeting, but one of the users found that by the end the head-mounted display (a Quest 3) felt heavy. They also made new suggestions, like having the name of each person by the avatar since after a round of introductions one normally forgets the names of people at the meeting, a clear demand for augmented virtual reality. Interestingly, in a physical meeting after the experience, users had the impression that they had been together previously, they knew and recognized each other not only by their physical aspect but also by their voices and body movements.

A look into the future

The potential applications of the metaverse in healthcare and medical training and practice are numerous. Several recent reviews and surveys have addressed this topic (Bansal et al., 2022; Usmani et al., 2022; Yang et al., 2022; Ahuja et al., 2023; Suh et al., 2023; Ullah et al., 2023). However, to our knowledge, its potential use for patient focus groups has not yet been discussed. Furthermore, there seem to be no reports on user experiences with remotely located patients in shared virtual environments. While the potential benefits of using the metaverse for focus groups compared to video conferencing (Kite et al., 2017; Greenspan et al., 2021) require further research and quantification, the shared virtual space offers a qualitatively different experience and memory for participants. Additionally, patients have expressed a preference for the privacy afforded by avatars and the virtual setting. Incorporating patient focus groups into the design and development of health technologies is crucial for developing solutions that are not only technically innovative but also highly relevant, safe, and effective for the intended users. This approach enhances the quality of healthcare delivery, patient satisfaction, and overall treatment outcomes, marking a significant shift towards more patient-centred healthcare solutions. Our early pilot studies comparing this experience in a shared virtual space versus a physical space have been positively valued by patients, researchers, and clinicians. This opens the door to future expansion, not only for focus groups but also for different types of support and information groups, both between patients themselves and with healthcare practitioners.

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.

Author Contributions

R.O. developed VRunited. The customized avatars were created by I.C., E.R. and T.D. B.A. carried out the work with the patients. A.C. recruited the patients. MVS-V coordinated the study, obtained the funding and wrote the first version of the paper. All authors contributed to structure the discussion, attended the virtual focus groups and revised the written paper.

Funding

Supported by XR-PAIN Grant Agreement No 101070533, funded by the European Media and Immersion Lab project financial support to third parties, which is funded by the European Union, and by Departament de Recerca i Universitats de la Generalitat de Catalunya (AGAUR 2021-SGR-01165 - NEUROVIRTUAL), supported by FEDER.

Acknowledgments

We want to thank the patients that participated in the experience.

References

Ahuja, A.S., Polascik, B.W., Doddapaneni, D., Byrnes, E.S., and Sridhar, J. (2023). The digital metaverse: Applications in artificial intelligence, medical education, and integrative health. *Integrative Medicine Research* 12(1)**,** 100917.

Álvarez de la Campa Crespo, M., Donegan, T., Amestoy-Alonso, B., Just, A., Combalía, A., and Sanchez-Vives, M.V. (2023). Virtual embodiment for improving range of motion in patients with movement-related shoulder pain: an experimental study. *Journal of Orthopaedic Surgery and Research* 18(1)**,** 729.

Bansal, G., Rajgopal, K., Chamola, V., Xiong, Z., and Niyato, D. (2022). Healthcare in metaverse: A survey on current metaverse applications in healthcare. *Ieee Access* 10**,** 119914-119946.

Birckhead, B., Khalil, C., Liu, X., Conovitz, S., Rizzo, A., Danovitch, I., et al. (2019). Recommendations for methodology of virtual reality clinical trials in health care by an international working group: Iterative study. *JMIR Mental Health* 6. doi: 10.2196/11973

Mummah, S. A., Robinson, T. N., King, A. C., Gardner, C. D., and Sutton, S. (2016). IDEAS (Integrate, Design, Assess, and Share): A Framework and Toolkit of Strategies for the Development of More Effective Digital Interventions to Change Health Behavior. *J Med Internet Res* 18, e317. doi: 10.2196/jmir.5927

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience* 13**,** 556-571.

Dil, N., Castiglioni, A., Kim, K., Aravind, N., and Torre, D. (2024). Use and implementation of focus groups. *Med Teach* 46(3)**,** 317-319. doi: 10.1080/0142159X.2024.2304451.

Donegan, T., Ryan, B., Sanchez-Vives, M.V., and Swidrak, J. (2022). Altered bodily perceptions in chronic neuropathic pain conditions and implications for treatment using immersive virtual reality. *Frontiers in Human Neuroscience*.

Garmer, K., Ylvén, J., and Karlsson, I.M. (2004). User participation in requirements elicitation comparing focus group interviews and usability tests for eliciting usability requirements for medical equipment: a case study. *International Journal of Industrial Ergonomics* 33(2)**,** 85-98.

Giannopoulos, E., Eslava, V., Oyarzabal, M., Hierro, T., Gonzalez, L., Ferre, M., and Slater, M. (2008). "The Effect of Haptic Feedback on Basic Social Interaction within Shared Virtual Environments", in: *LECTURE NOTES IN COMPUTER SCIENCE.* Springer).

Gottschalk, S. (2010). The presentation of avatars in second life: Self and interaction in social virtual spaces. *Symbolic interaction* 33(4)**,** 501-525.

Greenspan, S.B., Gordon, K.L., Whitcomb, S.A. and Lauterbach, A.A., (2021). Use of video conferencing to facilitate focus groups for qualitative data collection. American Journal of Qualitative Research, 5(1), pp.85-93.

Kite, J. and Phongsavan, P., (2017). Insights for conducting real-time focus groups online using a web conferencing service. F1000Research, 6.

Leung, F.-H., and Savithiri, R. (2009). Spotlight on focus groups. *Canadian Family Physician* 55(2)**,** 218-219.

Maselli, A., and Slater, M. (2013). The building blocks of the full body ownership illusion. *Frontiers in Human Neuroscience* 7. doi: 10.3389/fnhum.2013.00083.

Matamala-Gomez, M., Gonzalez, A.M.D., Slater, M., and Sanchez-Vives, M.V. (2019). Decreasing pain ratings in chronic arm pain through changing a virtual body: different strategies for different pain types. *The Journal of Pain* 20(6)**,** 685-697.

Mystakidis, S. (2022). "Metaverse". (Encyclopedia: MDPI).

Oliva, R., Beacco, A., Gallego, J., Abellan, R.G., and Slater, M. (2023). The Making of a Newspaper Interview in Virtual Reality: Realistic Avatars, Philosophy, and Sushi. *IEEE computer graphics and applications* 43(6)**,** 117-125. doi: 10.1109/MCG.2023.3315761.

Pan, X., Gillies, M., Barker, C., Clark, D.M., and Slater, M. (2012). Socially Anxious and Confident Men Interact with a Forward Virtual Woman: An Experimental Study. *PLoS ONE* 7**,** e32931.

Pan, X., and Hamilton, A.F.d.C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology* 109**,** 395-417. doi: 10.1111/bjop.12290.

Sanchez-Vives, M.V., Spanlang, B., Frisoli, A., Bergamasco, M., and Slater, M. (2010). Virtual hand illusion induced by visuomotor correlations. *PloS one* 5(4)**,** e10381.

Schroeder, R. (2001). *The social life of avatars: Presence and interaction in shared virtual environments.* Springer Science & Business Media.

Shoa, A., Oliva, R., Slater, M., and Friedman, D. (Year). "Sushi with Einstein: Enhancing Hybrid Live Events with LLM-Based Virtual Humans", in: *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*), 1-6.

Slater, M., Pérez Marcos, D., Ehrsson, H., and Sanchez-Vives, M.V. (2009). Inducing illusory ownership of a virtual body. *Frontiers in neuroscience* 3**,** 676.

Suh, I., McKinney, T., and Siu, K.-C. (Year). "Current perspective of metaverse application in medical education, research and patient care", in: *Virtual Worlds*: MDPI), 115-128.

Ullah, H., Manickam, S., Obaidat, M., Laghari, S.U.A., and Uddin, M. (2023). Exploring the potential of metaverse technology in healthcare: applications, challenges, and future directions. *IEEE Access*.

Usmani, S.S., Sharath, M., and Mehendale, M. (2022). Future of mental health in the metaverse. *General Psychiatry* 35(4)**,** e100825.

Yang, D., Zhou, J., Chen, R., Song, Y., Song, Z., Zhang, X., et al. (2022). Expert consensus on the metaverse in medicine. *Clinical eHealth* 5**,** 1-9.

Embodied Time Travel in VR: Witnessing Climate Change and Action to Overcome It

Yuke Pi[1], Xueni Pan[1], Mel Slater[2], Justyna Świdrak[3,4]

[1]SEEVR Lab, Department of Computing, Goldsmiths, University of London, London, UK
[2] Event Lab, Department of Clinical Psychology and Psychobiology, Institute of Neuroscience of the University of Barcelona, Barcelona, Spain
[3] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS) , Barcelona, Spain
[4] Institute of Psychology Polish Academy of Sciences, Warsaw, Poland

*** Correspondence:**
Yuke Pi[1]
y.pi@gold.ac.uk

Keywords: climate change, virtual reality, embodiment, environmental attitudes

Abstract

This study explores the impact of embodied experiences in Virtual Reality (VR) on individuals' attitudes and behaviour towards climate change. A total of 41 participants were divided into two groups: an embodied group that interacted with a virtual environment through full-body avatars, and a non-embodied group that observed the scenarios from an invisible observer's point of view. The VR experience simulated the progressive consequences of climate change across three generations within a family, aiming to make the abstract and distant concept of climate change a tangible and personal issue. The final scene presented an optimistic scenario of a future where humans had successfully combated climate change through collective action. The evidence suggests that there is some effect of Embodied compared to Observed on the carbon footprint response, but only after week 2 for carbon footprint. These findings suggest that immersive VR experiences that incorporate virtual embodiment can be an effective tool in enhancing awareness and motivating pro-environmental behaviour by providing a powerful and personal perspective on the impacts of climate change.

1 Introduction

In the future metaverse there will be many activities we can join - from entertainment to education. Some activities will be run by people connected in real-time (e.g., a drama teacher), others run by entirely virtual humans (non-player character, or NPC) in an automatic scenario (e.g., an immersive movie). Here we demonstrate a type of scenario that people can enter that could result in positive behavioural change. Our example is like an immersive movie except it includes embodiment into several generations of a virtual human character that experiences the effects of climate change first-hand - the past, present, a horrific future, and a possible positive alternate future if both individual and collective action is taken now (Lee, 2023).

In the face of escalating climate change events, such as wildfires, massive flooding, and extreme weather patterns, there is an increasing urgency to educate and engage the global populace. While evidence that these are caused by climate change induced by human activity is unequivocal, there remains a challenge: translating this complex and abstract proposition into a tangible, personal reality for individuals, who often feel powerless about it (Kenis & Mathijs, 2010).

There is growing evidence suggesting that Virtual Reality (VR) experiences can profoundly influence environmental attitudes and behaviour (Markowitz, 2018; Hofman, 2022). For example, VR-mediated natural experiences have been shown to increase connectedness to nature (Soliman, 2017; Spangenberger, 2022) and promote pro-environmental behaviour as effectively as real life (Deringer, 2021). There is also evidence that VR can increase climate change engagement, awareness, intentions, and behaviour (Queiroz, 2018). While VR's impact on environmental attitudes has been studied, the unique effect of embodied, first-person experience of climate change remains unexplored. With a wide field-of-view head-tracked and stereo head-mounted display, it is possible to substitute a person's real body with a life-sized virtual body, which moves with their movements, and which can be reflected in a virtual mirror. This typically gives rise to the illusion of ownership over the virtual body, where the participant feels that the virtual body is their own (Kilteni, 2012). Prior research has demonstrated that this kind of experience can lead to profound psychological changes in participants through implicit learning (Slater, 2017), including changes in attitudes and behaviour related to age (Banakou, 2013; Hershfield, 2011; Cowie, 2018).

We conducted an experimental study with 41 people, where we simulated various levels of climate deterioration in VR to provide an embodied experience of the consequences. The goal was to examine, by embodying a full-body avatar to experience the simulation of anthropogenic climate change consequences. Our VR experience includes a unique application of virtual embodiment: not only as a passive observer, but through the eyes of a child, and later an adult, witnessing the escalating climate change. We expected that such an embodied (*vs* non-embodied) experience of time travel could increase people's individual and collective climate action by making the *invisible* climate change *visible* and personal. In particular, our hypothesis was that, first, after the VR scenario, participants would reduce their carbon footprint as compared to before their VR scenario (**H1a**). We were also interested in whether this effect would still be detectable after 6 weeks (**H1b**). Our second hypothesis was that participants in the embodied condition would have a more significant decrease in their carbon as compared to the observer's group (**H2**).

2 Methods

The experiment used a between-subjects design with two conditions: embodied or observed experience. Participants were asked to choose the avatar aligned with their identified gender. The gender of all other virtual characters in the scenario aligned with that of the chosen avatar.

Participants. A total of 41 participants (21 female) were recruited from the campus of Goldsmiths, London University, 21 in the embodied group and 20 the observed (non-embodied). The VR experience was developed with the Unity 3D engine for Oculus Quest Pro. This study was approved by the University ethics board. Each participant received an Amazon voucher valued at £15. Ethics approval was obtained from the Department of Computing Ethics Committee at Goldsmiths.

Scenarios. This project spans three generations of a parent and children witnessing the impact of climate change from within their own home. In the embodied condition, participants see their virtual bodies when looking down at themselves and its reflections in a mirror. The movements of their head and upper body were mapped to their virtual body through real-time tracking from headsets and controllers. The participant, embodied as a child, interacted with a parent in the same room, while the TV portrayed information about climate change. The outside scene portrayed a green pasture with animals grazing. In the second phase, the participant was embodied as a teenager, with the older parent in the same, now dilapidated, room. The TV described the causes and consequences of climate change, and how it is being fought. The scene outside had deteriorated with fires in the distance. In the third phase, the participant was a parent and interacted with a child who pleaded for the future. The TV no longer works, and the outside scene is deserted. A final scene set 50 years in the future portrays a more optimistic scenario, including from that future standpoint a TV program about history, portraying events where worldwide mass collective action had forced politicians to take climate change seriously, instigating fundamental changes that led to the positively changed scenario. In the observed condition, the participant observed the same scenes but from an invisible third-person perspective, with no control over the avatar. In the observed group, participants did not have any avatar, and there were the two virtual characters (always parent and child) sitting on the sofa watching TV.

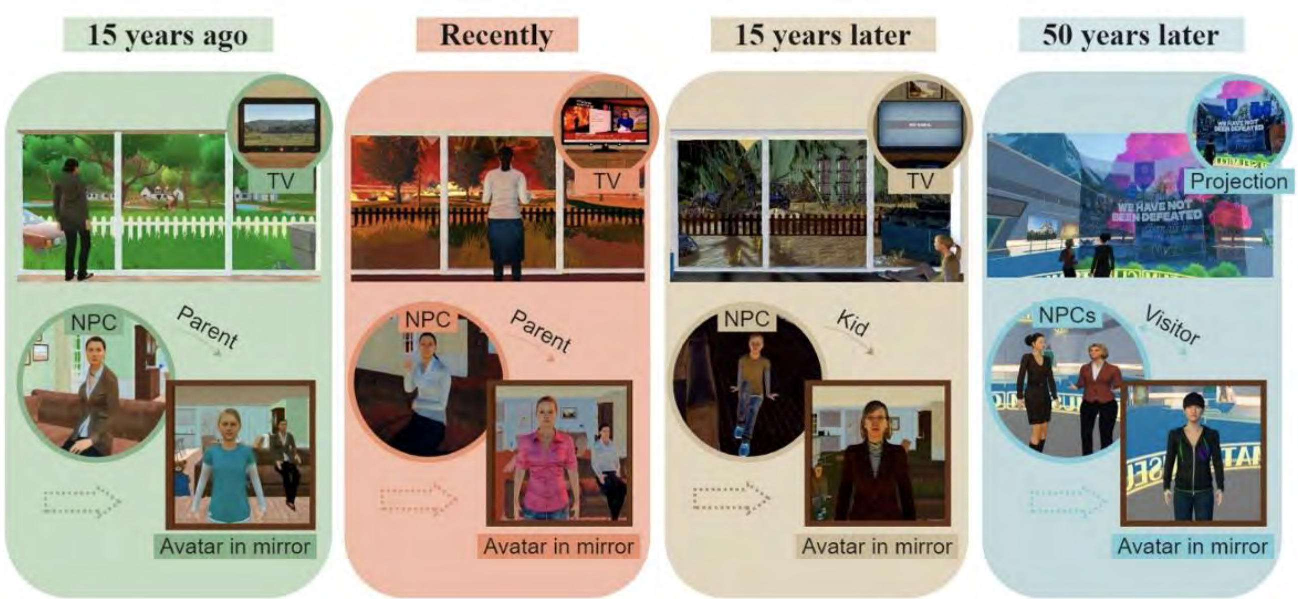


Figure 1 | Scenarios evolution of the interactive environment for the embodied group. A video walkthrough can be found here: https://youtu.be/q53ikT_pjh0

Procedure. Participants were first asked to read and sign a consent form. Then they filled in the pre-study questionnaires. Participants were reminded they could withdraw at any point. They were instructed on how to use the head-mounted display (HMD). The experimenter showed the participants in the embodied group how to calibrate their virtual bodies and perform avatar interactions such as switching the TV on and off and grabbing small objects. In contrast, the non-embodied group did not have any interactions. Both groups experienced the four phases mentioned above, each phase lasting about 150 seconds and automatically transitioning to the next one. Eye gaze and movement data were logged. After the experiment, they removed the HMD and answered the post-questionnaires. Finally, the participants were thanked for their participation and were asked to fill out follow-up questionnaires **after 2 weeks** and **6 weeks**.

Measurements. We measured real-world behavioural change after the VR experience with a carbon footprint calculator (CF) from World Wildlife Fund UK website (WWF)[1]. This was done three times:

[1] https://footprint.wwf.org.uk/questionnaire

before the experiment (carbon 1), 2 weeks after the experiment to measure any potential effect (carbon 2), and again 6 weeks later to measure any lasting effect (carbon 3). We also used a personal and political environmental willingness questionnaire (Lubell, 2007) to measure a change in the perceived personal and political influence on climate change and a shortened Discrete Emotions Questionnaire (Harmon-Jones, 2016) to measure the emotional reaction to the exposure. These results are not included, due to the limited size of this paper.

Statistical analysis. We carried out a Bayesian analysis of the data. In order to understand the effect of the experimental condition on the response variables, for each variable we used a linear predictor for each response variable of the form:

$$Cond + P1 + Cond \times P1$$

where,

$$Cond = \begin{cases} 0, Observed \\ 1, Embodied \end{cases}$$

$$P1 = period\ 1\ score \tag{1}$$

More formally, this can be expressed as:

$$\beta_1 + \beta_2 C_i + \beta_3 P_i + \beta_4 (C_i \cdot P_i)$$
$$i = 1,2, \ldots, n \tag{2}$$

where, n is the number of participants, C_i is the condition (Observed=0, Embodied=1) for the ith participant. This model indicates that the dependent variables, the scores for periods 2 and 3 $(P2, P3)$ depend on the condition, the score for period 1, and the interaction between the condition and the period 1 score. The prior distributions of the parameters were Student t with scale σ and degrees of freedom υ, each of which had a Cauchy distribution with median 0 and scale parameter 2.5.

3. Results

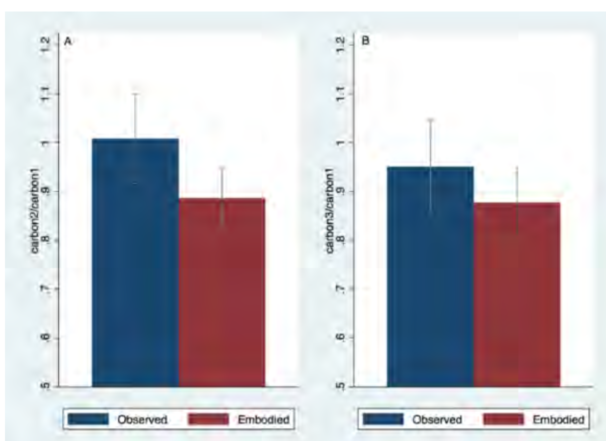


Figure 2 | Bar chart showing the means and standard errors for the ratios of the change of carbon (A) carbon2/carbon1 and (B) carbon3/carbon1.

We observed a reduction in Carbon Footprint in both conditions: for the embodied condition, Carbon footprint was lower after VR, and maintained its level 6 weeks onwards (Mean± SE – carbon1: 16.4

± 1.51; carbon2: 13.5 ± 1.16; carbon3: 13.1 ± 1.05). The same pattern was the same for the observed group (Mean± SE - carbon1: 16.2 ± 1.68; carbon2: 15.2 ± 1.66; carbon3: 14.0 ± 1.43). Figure 2 shows the changes in carbon footprint, suggesting that there were changes from period 1 (prior to the exposure) to periods 2 (2 weeks after the exposure) and period 3 (6 weeks after the exposure). However, the change is more pronounced for period 2. However, this graph still does not take into account the relationship between carbon2, carbon3 conditional on carbon1. This can be done with the analysis based on Eq. (2). We carried out a Bayesian analysis of these data with the following characteristics. For the distributions of the carbon footprint, the model (2) did not provide a good fit to the data. Instead, we used:

$$\frac{carbon_{ki}}{carbon_1} = \beta_1 + \beta_2 C_i \tag{3}$$

where k is 2 or 3 (meaning carbon2 or carbon3).

As shown in Table 1, taking into account the level of carbon1, the posterior probability that carbon2 in the experimental condition is less than in the control condition 2 weeks after the exposure is 1 – 0.070 = 0.930. However, this does not hold for carbon3. The posterior probability of that carbon3 in the experimental condition is less than in the control 1 – 0.358 = 0.642, indicating only a low probability of a slight.

Table 1| Summary posterior distributions of the parameters, showing for each distribution the Mean and SD, the 95% credible interval, and the probability of the parameter being positive.

Parameter	Coefficient of	Mean	SD	2.5%	97.5%	Prob>0
carbon2						
β_1		0.97	0.04	0.87	1.05	1.000
β_2	condition	-0.09	0.06	-0.21	0.04	0.070
σ		2.48	0.71	1.38	4.11	
υ		1.70	0.86	1.03	3.45	
carbon3						
β_1		0.85	0.06	0.72	0.97	1.000
β_2	condition	-0.03	0.09	-0.21	0.14	0.358
σ		3.67	0.97	2.10	5.87	
υ		4.51	44.29	1.10	8.73	

4. Discussion

Our result suggests that there is some effect of Embodied compared to Observed on the carbon footprint responses, but only after week 2 for carbon footprint. This is evidence that the embodiment experience in VR influenced individual behaviour, measured here as real-world carbon footprint, in the short term only. One single exposure may not be enough to lead to lasting change, instead a variety of similar experiences may be necessary.

Metaverse(s) may become an important source of entertainment and education, where users will experience reconstructions of past events and simulations of the future as active participants. The aim of the study was to investigate the influence of such a climate change consequences simulation on attitudes and behaviour. It would be interesting to study whether the same effects would be observed in a social metaverse, where several people could share such educational experience. We may hypothesize that participating in such simulation with others, especially those familiar or close to the user, could have potentially much stronger impact, especially with personalised, and not generic, avatars aging along the timeline, as users may find the experience much more personal, and personal responsibility to reduce climate change is associated with various climate actions (Bouman, 2020).

To summarize, the simulation of the climate change scenarios can be a useful tool for increasing climate action, but further studies are needed to maximize its effectiveness.

References

Banakou, D. a. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, 12846--12851.

Bouman, T. a. (2020). When worry about climate change leads to climate action: How values, worry and personal responsibility relate to various climate actions. *Global Environmental Change*, 102061.

Cowie, D. a. (2018). The development of bodily self-consciousness: changing responses to the Full Body Illusion in childhood. *Developmental science*, e12557.

Deringer, S. A. (2021). Virtual reality of nature can be as effective as actual nature in promoting ecological behavior. *Ecopsychology*, 219--226.

Harmon-Jones, C. a.-J. (2016). The discrete emotions questionnaire: A new tool for measuring state self-reported emotions. *PloS one*, e0159915.

Hershfield, H. E. (2011). Increasing saving behavior through age-progressed renderings of the future self. *Journal of marketing research*, S23--S37.

Hofman, K. a. (2022). The effectiveness of virtual vs real-life marine tourism experiences in encouraging conservation behaviour. *Journal of Sustainable Tourism*, 2022.

Kenis, A., & Mathijs, E. (2010). Beyond individual behaviour change: the role of power, knowledge and strategy in tackling climate change. *Environmental Education Research*, 45-65.

Kilteni, K. a. (2012). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 373--387.

Lee, H. a. (2023). *IPCC, 2023: Climate Change 2023: Synthesis Report, Summary for Policymakers. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPC.*

Lubell, M. a. (2007). Collective action and citizen responses to global warming. *Political Behavior*, 391--413.

Markowitz, D. M. (2018). Immersive virtual reality field trips facilitate learning about climate change. *Frontiers in psychology*, 2364.

Queiroz, A. C. (2018). Immersive virtual environments and climate change engagement. *Proceedings of the immersive learning research network*, 153e164.

Slater, M. (2017). Implicit learning through embodiment in immersive virtual reality. *Virtual, augmented, and mixed realities in education*, 19-33.

Soliman, M. a. (2017). The impact of immersive technology on nature relatedness and pro-environmental behavior. *Journal of Media Psychology*.

Spangenberger, P. a.-C. (2022). Becoming nature: Effects of embodying a tree in immersive virtual reality on nature relatedness. *Scientific reports*, 1311.

Towards the Incluverse: an Inclusive Metavarse (for People with Intellectual Disability)

Pablo Pérez [1,*], **Marta Orduna** [1], **Jesús Gutiérrez** [2], **and María Nava-Ruiz** [3]

[1] *Extended Reality Lab, Nokia, Spain*
[2] *Universidad Politécnica de Madrid, Madrid, Spain*
[3] *Fundación Juan XXIII, Madrid, Spain*

Correspondence*:
Corresponding Author
pablo.perez@nokia.com

ABSTRACT

We describe several research initiatives targeted to develop XR technologies that allow for immersive communication experiences applied to wellbeing, therapy, remote supervision, and vocational training use cases. The project aims to apply these technologies to people with intellectual disability and other vulnerable psychosocial situations as a fundamental step towards the development of a fully inclusive Metaverse or *Incluverse*.

Keywords: inclusive metaverse, extended reality, intellectual disabiliy, psychosocial vulnerability

1 INTRODUCTION

A common problem faced by the development of AI and XR technologies is their bias towards certain user profiles (Torralba and Efros, 2011). Including diverse populations from the beginning of the research process ensures universal use and accelerates their acceptance by the population. In the Metaverse, as in life, there is a risk of exclusion for people with disabilities. Therefore, it is important to work on the accessibility of extended reality technology. For instance, Mott et al. (2019) propose that XR technology should be accesible by design, adapting content, interacion, and devices to people with diverse physical limitations. The use of XR for people with intellectual disability has been mainly researched with relatively short VR-based interventions aimed at training some cognitive skills (Nabors et al., 2020). There is still limited experience on specifically desiging the XR technology to be accessed by them.

Focusing on this population has some interesting characteristics. Firstly, the specific needs of this population can clearly benefit from the potential of immersive technology. For example, there is a high comorbidity with some mental disorders (e.g. anxiety or depression), where virtual reality has proven to be very effective in therapeutic settings. Secondly, strong adaptation of technology is required, as well as the methods used to evaluate it. For instance, both experiences and interfaces must be usable while minimizing the cognitive load required. Similarly, evaluation tools (such as psychometric questionnaires) must be adapted and simplified to the fullest extent possible. And thirdly, because the two aforementioned characteristics make developing and evaluating technology for individuals with intellectual disabilities a powerful testing ground for extrapolation to the general population. Applications developed will have validated their utility beyond the “wow effect” of XR and can be used and evaluated while minimizing the cognitive effort required.

Figure 1. Images from the project experiences. The top row shows the immersive music and experience for the elderly: a live concert (left), a user (center), and the remote side of the telepresence session (right). The bottom row shows the user view of the therapy application (left), a trial of the telepresence system in the training apartment (center), and a 360-degree recording of a Basic Patry Operations class.

Considering these insights, we have initiated a series of research projects aimed at developing, adapting, and validating eXtended Reality (XR) technology for users with intellectual disabilities. In this paper, we describe the various initiatives underway, as well as the benefits observed during their implementation.

2 THE PROJECT

The project is targeted at the users of the occupational training center for opportunities and labor integration (COFOIL, in Spanish) of Fundación Juan XXIII. This Foundation is dedicated to aiding individuals with intellectual disabilities via educational programs, vocational training, and social inclusion efforts to improve their life quality and societal integration. The center caters to diverse user profiles, each with unique needs. In executing the project, target groups were pinpointed to potentially benefit from XR technology. For these groups, tailored actions were devised to align the technology's potential with the users' specific needs.

Although the realm of extended reality is vast, and the project often combines different aspects of it, the approach we take is that of Distributed Reality (Perez et al., 2022a). Our goal is to provide users with the ability to feel present in a remote location (either through recorded video or real-time telepresence) while being able to perceive their own nearby reality (their own body and objects of interest). To achieve this, our basic technological building blocks are video-based avatars for the representation of users within the immersive scene (Gonzalez Morin et al., 2023), natural interaction interfaces based on physical object manipulation (Cortés et al., 2022), real-time immersive telepresence based on 360-degree video (Kachach et al., 2021), and collection of patient behavior parameters and biomarkers, which can be used to identify mental states and cognitive processes, track patient progress, and develop tools for emotional and cognitive regulation (Itti, 2015; Kyriakou et al., 2019).

Each experience involves applying technology to meet the needs of a specific group of participants, with the direct user count ranging from 10 to 20 for each experience, and an approximate duration of 4 to 6 months. They are described below. Table 1 shows a summary.

Table 1. Summary of the experiences

Experience	Music	Therapy	Remote support	Training
Target	Wellbeing	Autonomy (mobility)	Autonomy (indepentent life)	Employment
Technology	360-degree video	Ego-avatars, biomarkers	Telepresence	All
Status	Complete	Started	Started	In design

2.1 Immersive music and experiences for the elderly

The life expectancy of individuals with intellectual disabilities has dramatically increased over the last 40 years, introducing new challenges as the first generation of older adults with disabilities emerges (considered from age 35 with Down syndrome and 45 without). This generation faces unique needs, including a lack of life experiences that make it difficult for them to experience and express complex emotions. In response to this challenge, several actions were designed to progressively assess the benefits of immersive video and telepresence technology to provide access to such life experiences.

The first part of the project was an immersive musical experience program, since virtual reality can facilitate the development of music therapy and sensory stimulation (Perez et al., 2022b). First, participants attended 10 live concerts of music bands performing songs significant to them. Concerts were recorded with 180-degree stereo cameras, and individual songs were extracted and edited separately. Afterwards, concerts were recreated in 15-minute sessions viewed with VR headset. The feedback from the users, 12 elderly individuals from the center, was overwhelmingly positive, with high scores on adapted social and spatial presence questionnaires indicating that participants truly felt present at the live concerts, even interacting with the musicians. Given the success of this initiative, the experience has been permanently installed at the occupational center of the Foundation and is beginning to be used in specific instances with users experiencing acute crises, which can be alleviated with immersive music.

A second phase focused on the difficulty of the users to travel to places that have been significant to them.Visits were made to the hometowns of some project participants, where the participant would present their town and its significant childhood places to other companions. The visit was recorded with a 360-degree camera, allowing other companions to later make the visit virtually.

Finally, real-time telepresence sessions were conducted. In this case, users employed a mobile immersive telepresence prototype, Nokia Owl (Kachach et al., 2021), to virtually visit someone significant to them: a former teacher they had not seen for several years. The sense of presence and the emotional connection achieved were significant and much higher than what is usually achieved with a standard video call.

2.2 Anxiety treatment therapy

This project applies Systematic Desensitization (SD), a cognitive-behavioral technique for addressing fears and phobias, as described by Wolpe (1961). SD aims to incrementally decrease sensitivity towards anxiety-inducing situations via mental visualization while performing an anxiety-incompatible response. Used in therapy under therapist supervision, it involves exposure to gradually increasing anxiety triggers. This method teaches patients to manage the emotions associated with these experiences effectively. Employing an incompatible response, like breathing exercises, can trigger an automatic reaction that mitigates the aversive stimulus's discomfort, thereby diminishing the resulting anxiety.

In a specific case at the occupational center (COFOIL) of the Juan XXIII Foundation, a common stress-inducing situation was identified: the fear of ascending or descending stairs, a daily challenge for some, potentially hindering autonomous attendance (e.g., fear of metro escalators). The approach involves applying SD combined with VR for COFOIL users with a fear of stairs, aiming to help them confront this situation and improve their quality of life. The method entails several phases, including creating an individualized hierarchy of anxiety-inducing situations using real photos of the feared stimuli (stairs), diaphragmatic breathing training to counteract anxiety, and desensitization by associating the anxiety-provoking situation with the learned incompatible response (relaxation through diaphragmatic breathing) using VR headsets. This process, expected to lead to counterconditioning through repetition, aims to reduce anxiety levels upon exposure to the feared stimulus, with the ultimate goal of generalizing these outcomes to real-life situations.

In this case, audiovisual stimuli for all phases are provided through a VR headset. The approach combines 360-degree videos of both neutral and anxiety-inducing environments (such as familiar stairs leading to the occupational center) with virtual scenarios offering six degrees of freedom. Full-body avatars based on pass-through video are used to enhance user embodiment in the virtual experience. Additionally, various markers (pulse, skin conductivity, electromyography) are monitored to assess the user's physiological state.

2.3 Immersive telepresence in supervised apartments

This use case focuses on Nokia Owl telepresence prototype (Kachach et al., 2021), which is based on 360-degree video communication. The video is captured with a 360-degree camera and streamed in real-time. Its distinctive feature is that the remote user participates in the conversation through VR headsets, which allow viewing a portion of the scene, called viewport, that changes with head movements. The main benefits are related to the immersion experienced by the remote user, offering an experience closely resembling being physically present where the 360-degree camera is located.

The system is implemented in the training apartment and supervised apartments of the Foundation. In the training apartment, individuals attending the Foundation prepare for independent and autonomous living. In the supervised apartments, 3-4 people live independently with professional support, each with specific needs. With this setup, the basic scenario involves a professional remotely using the VR headsets and the Owl device in the supervised apartment, thus exploring immersive communication technology from video quality necessary for safe remote supervision to additional tools that might enhance interaction and the system capabilities. Regularly using telepresence at home environments has also relevant considerations in terms of social behavior and privacy, which are considered in the project as well (Boudouraki et al., 2022).

Supervision planning involves defining daily tasks, some split into subtasks, for residents to practice, facilitating an assessment of progress, identifying arising issues, and determining if the level of support required changes over time. These activities, whether group-based or individual, depend on residents' characteristics, such as their experience and the dynamics of the supervised apartments. Specific tasks that can be performed under the remote support professional's supervision include doing laundry, hanging clothes, cleaning the toilet, and cooking simple dishes.

2.4 Teletraining for employment placement

The use of XR in training can improve the efficiency of training and therefore the employability of these individuals (Radhakrishnan et al., 2021). This project utilizes XR technology to overcome challenges in delivering vocational training courses to people with disabilities, addressing the limited availability of spots and the difficulties faced by students who cannot attend in person due to medical reasons or vacation breaks. By leveraging the immersive and presence-enhancing features of XR technology, this approach allows for remote participation in training sessions, such as cooking lessons, with a level of engagement and assessment comparable to physical attendance. This innovative use of XR aims to ensure that all interested individuals, especially those with greater support needs like COFOIL users, have access to valuable professional certification opportunities that can significantly enhance their employment prospects.

The pilot program targets individuals from COFOIL's socio-labor insertion area, focusing on those with fewer opportunities for accessing traditional training courses. It offers a combination of vocational training, specifically in basic pastry operations, facilitated through immersive technologies, and cognitive skill training aimed at improving attention, inhibition, and processing speed. This dual approach not only seeks to equip participants with practical skills and professional certifications but also aims to boost their cognitive functions, thereby enhancing their overall performance and efficiency in the workplace.

3 BENEFITS

Each experience within the project is accompanied by an evaluation, with the expectation of observing improvements in various assessed aspects. However, a key objective of the project is to integrate these experiences into the daily lives of its users, tailoring them to their specific needs. This approach can sometimes complicate isolating the effects of the treatments or using control groups to obtain clear results. Nevertheless, it also allows for the observation of the benefits of technology when it is adapted to the users' daily routines. These benefits arise not only from the technology itself but, more importantly, from how the use of technology and the project design are embedded in the users' everyday life.

In occupational training centers for opportunities and labor integration, a conventional educational offer (classes, workshops, etc.) is provided. Projects like ours present an alternative, high-quality form of training based on aspects significant to the participants, enabling them to interact with new technologies and professional profiles (technology researchers) usually beyond their reach. In terms of labor integration, this approach opens up new avenues beyond traditional occupations in logistics, gardening, or hospitality. Individuals with disabilities are often accustomed to engaging in the same types of activities and interacting with familiar technologies. Being disabled should not preclude them from working with innovative technologies. Therefore, it is crucial for technology companies to approach disabilities with an open mind and the willingness to test and adapt technologies to various circumstances.

The projects are designed to ensure successful participation by the individuals involved. Although the outcomes may vary in significance, the critical factor is that participants feel capable of engaging with the technology, which significantly boosts self-esteem in individuals who typically have fewer opportunities to pursue their desires. Participants feel a sense of social recognition by working with researchers who also promote and present these projects, a sentiment shared by their families.

Placing technology at the service of people is of utmost importance. In designing the various experiences, efforts were always made to address the concrete and real needs of the users (within the technological possibilities), iterating several times until finding the specific action that could have the most beneficial impact for them. This process of adaptation and personalization has been perceived as one of the most valuable aspects of the projects.

4 CONCLUSIONS

In this paper, we have described several ongoing projects aimed at designing, developing, and validating XR experiences that meet the needs of individuals with intellectual disabilities. Throughout the development of these experiences, we have recognized the importance of placing the individual at the center of the research. On one hand, designing experiences that address the specific needs of these individuals makes the technology much more effective. On the other hand, it's crucial to emphasize that the human aspects of the project itself (how users are treated and considered) are at least as valuable for the users as the technology itself. Adhering to these principles will allow us to develop a more inclusive Metaverse, not only for individuals with disabilities but for anyone who experiences it.

ACKNOWLEDGMENTS

This work has been partially supported by UNICO I+D 6G 2022 project TSI-064200-2022-009 (INCLUVERSO 5G) and Red.es project 2021/C005/00144164 (COMODIA) of the Spanish Government within the framework of the Recovery, Transformation and Resilience Plan, as well as by project PID2020-115132RB (SARAOS) funded by MCIN/AEI/10.13039/501100011033 of the Spanish Government.

REFERENCES

Boudouraki, A., Reeves, S., Fischer, J. E., and Rintel, S. (2022). Mediated visits: Longitudinal domestic dwelling with mobile robotic telepresence. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16

Cortés, C., Rubio, M., Pérez, P., Sánchez, B., and García, N. (2022). QoE study of natural interaction in extended reality environment for immersive training. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (IEEE), 363–368

Gonzalez Morin, D., Gonzalez-Sosa, E., Perez, P., and Villegas, A. (2023). Full body video-based self-avatars for mixed reality: from e2e system to user study. *Virtual Reality* , 1–19

Itti, L. (2015). New eye-tracking techniques may revolutionize mental health screening. *Neuron* 88, 442–444

Kachach, R., Morcuende, S., Gonzalez-Morin, D., Perez-Garcia, P., Gonzalez-Sosa, E., Pereira, F., et al. (2021). The owl: Immersive telepresence communication for hybrid conferences. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (IEEE), 451–452

Kyriakou, K., Resch, B., Sagl, G., Petutschnig, A., Werner, C., Niederseer, D., et al. (2019). Detecting moments of stress from measurements of wearable physiological sensors. *Sensors* 19, 3805

Mott, M., Cutrell, E., Franco, M. G., Holz, C., Ofek, E., Stoakley, R., et al. (2019). Accessible by design: An opportunity for virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (IEEE), 451–454

Nabors, L., Monnin, J., and Jimenez, S. (2020). A scoping review of studies on virtual reality for individuals with intellectual disabilities. *Advances in Neurodevelopmental Disorders* 4, 344–356

Perez, P., Gonzalez-Sosa, E., Gutiérrez, J., and García, N. (2022a). Emerging immersive communication systems: Overview, taxonomy, and good practices for qoe assessment. *Frontiers in Signal Processing* 2. doi:10.3389/frsip.2022.917684

Perez, P., Vallejo, E., Revuelta, M., Redondo Vega, M. V., Guervós Sánchez, E., and Ruiz, J. (2022b). Immersive music therapy for elderly patients. In *ACM International Conference on Interactive Media Experiences*. 47–52

Radhakrishnan, U., Koumaditis, K., and Chinello, F. (2021). A systematic review of immersive virtual reality for industrial skills training. *Behaviour & Information Technology* 40, 1310–1339

Torralba, A. and Efros, A. A. (2011). Unbiased look at dataset bias. In *CVPR 2011* (IEEE), 1521–1528

Wolpe, J. (1961). The systematic desensitization treatment of neuroses. *The Journal of nervous and mental disease* 132, 189–203

Construing the virtual self: The psychological functions of virtual identity in online videogames

Adrià Gabarnet[1,2*], Adrián Montesano[3], Guillem Feixas[1,2]
[1]Department of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain
[2]Institute of Neuroscience, University of Barcelona, Barcelona, Spain
[3]Department of Psychology and Educational Sciences Studies, Open University of Catalonia, Barcelona, Spain
*Corresponding Author
agabarnet@ub.edu

Abstract
In this study, we intended to put forward a typology of the different types of virtual identity that can be construed by online videogame players. We interviewed 202 players of Massive Multiplayer Online Role-Playing Games (MMORPGs) and administered them the Repertory Grid Technique (RGT). With this evaluation tool, we were able to explore how our participants perceived themselves in relation to their ideal self and their main characters from an MMORPG of their choice, and quantitatively measure their perceived discrepancies between those three aspects of their identity. We conducted a k-means cluster analysis to classify our participants according to their actual self–ideal self, actual self–virtual self, and virtual self–ideal self discrepancies. Three main types of virtual self construal arose: The projection type, characterized by players who viewed themselves as different from their ideal self and construed a virtual self that was similar to their ideal self; the exploration type, characterized by players whose actual self was closer to their ideal self and would create a virtual self that was different from both; and the proximal type, characterized by players whose virtual self was very similar to their actual self. This typology could be a useful tool for any future research that focuses on virtual environments to better understand its users or the predisposing or maintaining factors for any problematic usage of the platform.
Keywords: virtual self, repertory grid technique, avatars, videogames, identity construal.

1. Introduction

To many individuals, virtual environments represent a safe space where they can express aspects of their identity that they would not usually show in their offline lives. Even personality traits that the person sees as ideal can be inhibited due to social constraints or the need to preserve their already established identity within their usual social groups **(1,2)**. However, virtual environments offer their users the possibility to create an entirely new identity thanks to their anonymous nature and the customizability of the virtual avatars or profiles through which users interact with each other. This is especially true in MMORPGs (Massive Multiplayer Online Role-Playing Games), a genre of online videogames characterized by their highly customizable characters and focus on avatar-to-avatar interaction in different formats (competitive, cooperative, socialization...). MMORPG players could construe their characters' identities so that they possess traits that they would not usually associate with their offline self. Some of these traits could be associated with the players' ideal selves, in which case the avatar could have a compensatory function for any perceived lacking in their offline identity. This phenomenon has been observed especially in players with low self-esteem **(3,4)**. Other players could use their virtual identity within MMORPGs to explore alternative personas that are not necessarily desired ones, serving an exploratory or playful function in an act of "identity tourism" **(5,6)**.

The aim of this study was to explore the different ways in which MMOPRG players could construe their virtual identity through their characters according to how it relates to their actual and ideal selves, and what psychological function those virtual selves could have. Our main hypothesis was that the optimal clustering model for players would follow the typology proposed in our previous work, composed by four groups: Projection type, where players who see themselves as different from their ideal self would create a virtual self that was closer to it in order to compensate for their perceived offline deficiencies; Exploration type, where players who already see themselves as similar to their ideal self construe a virtual self that is different from both their actual and ideal self in an exercise of identity tourism; Proximal type, where players construe their virtual self as similar to their actual self, since they would see it more as an extension of their offline identity or a tool to interact with the virtual environment; and finally, the Unspecified type, where players would show an equal perceived distance between their actual, ideal and virtual selves, perhaps as a consequence of a lower level of cognitive differentiation **(7).**

2. Methods

To test our hypothesis, we contacted MMORPG players in different online forums and websites, asking them to participate. Our inclusion criteria for participation were to be at least 18 years old, being an active player of any MMORPG, not being a professional player, not playing under the influence of alcohol or other recreational drugs, and not being diagnosed with any severe mental disorder. Participants would answer an online survey to test their applicability for the study and gather sociodemographic and playing habits information. Those who met our criteria were later contacted through email to schedule a date for an online interview. Where they would be administered the Repertory Grid Technique (RGT), a psychological evaluation instrument developed from the framework of Personal Construct Theory. According to PCT, every individual construes their subjective perception of themselves and other people around them through a system of constantly evolving bipolar constructs (e.g., selfish vs. selfless), that they construe and update with each social interaction throughout their lives and that help them interpret and predict the world around them **(8,9)**. Based on that theory, The RGT was developed as a semi-structured interview that can be used to study an individual's identity by examining their construct system **(10,11)**. Based on how they apply those constructs to different elements in their lives (i.e. their actual self, ideal self, significant others, or even other aspects of themselves such as their virtual self), the RGT allows to quantitatively measure the perceived discrepancies between those elements. Because of that, we administrated the RGT and extracted the actual-ideal, actual-virtual, and virtual-ideal self discrepancy measures to analyze how similar or different these three elements were perceived. Based on those variables, we conducted a k-means cluster analysis to search for the optimal clustering model for our data and later tested the significance of the identified clusters' differences with a series of one-way ANOVAs.

3. Results and Discussion

Our analyses concluded that the clustering model with the best fit to our data was composed of three clusters. The first cluster was characterized by a low actual-ideal discrepancy (0.27) and high actual-virtual and virtual-ideal discrepancies (0.48 and 0.47, respectively). According to our theoretical typology was given the "Exploration type" label. The second cluster was characterized by low actual-ideal and virtual-ideal discrepancies (0.25 and 0.28, respectively) and a somewhat low actual-virtual discrepancy (0.32). Although the average actual-virtual discrepancy in that cluster would not be considered a particularly low score **(12)**, it was still substantially lower than the average actual-virtual discrepancy in the other two clusters, and for this it was given the "Proximal type" label. Finally, the third cluster was characterized by high actual-ideal and actual-virtual discrepancies (0.42 and 0.46, respectively) and a low virtual-ideal discrepancy (0.25). Because of that, it was given the "Projection type" label. In the supplementary material, Table 1 shows the average discrepancy variables found in all three clusters.

After conducting a series of one-way ANOVAs, significant differences were found between the three clusters in their actual-ideal discrepancy ($F(2,199) = 89.86$, $p < 0.001$), actual-virtual discrepancy ($F(2,199) = 94.88$, $p < 0.001$), and virtual-ideal discrepancy ($F(2,199) = 142.10$, $p < 0.001$). We conducted pairwise comparisons using the Bonferroni correction and observed that actual-ideal discrepancy was significantly higher in the cluster labeled "Projection type" than in the other two groups ($p < 0.001$ in both differences), showing a perception of themselves that was more distanced from their ideal self. Actual-virtual discrepancy was significantly lower in the cluster labeled "Proximal type" than in the other two groups ($p < 0.001$ in both differences). Lastly, virtual-ideal discrepancy was significantly higher in the cluster labeled "Exploration type" than in the other two groups ($p < 0.001$ in both differences). In the supplementary material, Figure 1 shows a boxplot with the differences between clusters for all three variables.

These results seem to validate, at least partially, the typology proposed in our previous work **(7)**. Although we did not observe any type of virtual identity construal that would fall under the "unspecified type", we were able to identify the other three. In the projection type, we found people who were not particularly satisfied with their perception of themselves and who resorted to construing an idealized version of themselves through their MMORPG characters, perhaps as a compensatory strategy, as contemplated by previous studies **(3,4)**. Secondly, we identified the exploration type, where players who were satisfied with themselves, as shown by their shorter distance with their ideal self, would construe a virtual identity that was removed from both their actual and ideal selves, following a more performative motivation, which can be a satisfying experience in and of itself, as evidenced by the phenomenon of identity tourism **(5,6)**. Finally, in the proximal type, we found people who did not see their virtual selves as different from their actual selves at all. This could be the case of people who do not see their characters as a separate aspect of their identity, but rather as a way to extend their offline self into the virtual environment so that they can interact with other players within as they already are offline. Another possible explanation for this type of virtual self construal would be that players see the character not as a manifestation of their identity at all but as a tool with which they can manipulate the game mechanics, like a bat for a baseball player.

This typology could represent a valuable tool for any future research that focuses on online videogame players and wishes to take on an identity-based perspective. It could shed some light on the motivations that drive players, and even into the underlying factors that could predispose or maintain problematic or excessive gaming. If validated in different types of virtual environments, such as social media, this typology could even be used in studies focused on virtual identity on a wider scale and contribute to the understanding of identity and how new technologies have affected our notion of how it is construed.

4. References

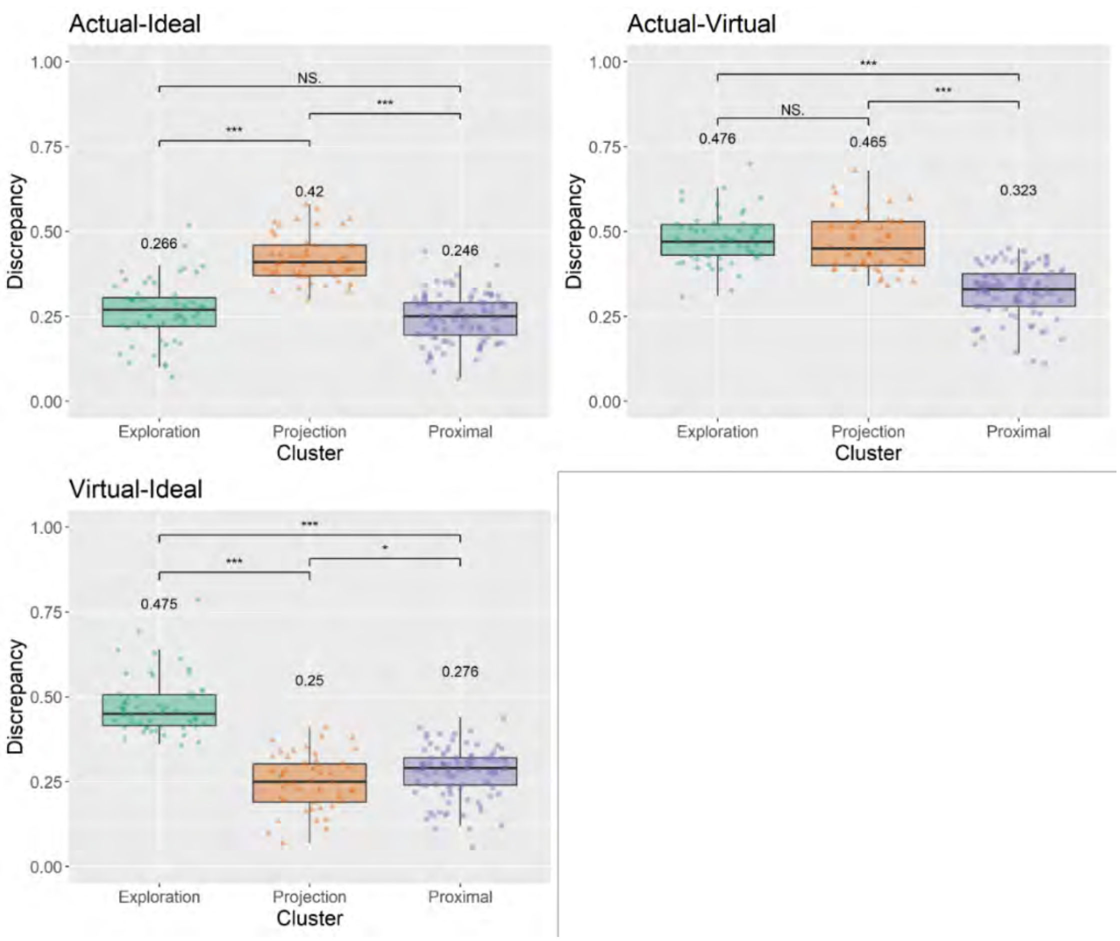
1. Lecky P. Self-consistency; a theory of personality. Washington, DC, US: Island Press; 1945. 154 p. (Self-consistency; a theory of personality).
2. Oyserman D, Elmore K, Smith. Self, self-concept, and identity. En: Handbook of self and identity. 2012. p. 69-104.
3. Dengah HJF, Snodgrass JG. Avatar Creation in Videogaming: Between Compensation and Constraint. Games Health J. agost 2020;9(4):265-72.
4. Michikyan M, Dennis J, Subrahmanyam K. Can You Guess Who I Am? Real, Ideal, and False Self-Presentation on Facebook Among Emerging Adults. Emerg Adulthood. 1 febrer 2015;3(1):55-64.
5. Gilbert R, Thadani V, Handy C, Andrews H, Sguigna T, Sasso A, et al. The psychological functions of avatars and alt(s): A qualitative study. Comput Hum Behav. març 2014;32:1-8.
6. Ribeiro J. The increase of the experiences of the self through the practice of multiple virtual identities. PsychNology J. 2009; Available at: https://www.semanticscholar.org/paper/The-increase-of-the-experiences-of-the-self-through-Ribeiro/b1d352131c716b0cce9b9838c88e9115b8adf575
7. Gabarnet A, Montesano A, Feixas G. Virtual-self Identity Construal in Online Video Games: A Repertory Grid Study Protocol. Aloma Rev Psicol Ciènc Educ Esport. 12 maig 2022;40(1):9-21.
8. Chiari G, Nuzzo ML. Personal construct theory within psychological constructivism: Precursor or avant-garde? In: Personal construct theory: A psychology for the future. Carlton South VIC, Australia: Australian Psychological Society; 1996. p. 25-54.
9. Walker BM, Winter DA. The Elaboration of Personal Construct Psychology. Annu Rev Psychol. 2007;58(1):453-77.
10.Feixas G, Cornejo, J. M. Manual de la Técnica de Rejilla mediante el Programa RECORD 2.0. 1996.
11. Fransella F, Bell R, Bannister D. A Manual for Repertory Grid Technique. John Wiley & Sons; 2004. 283 p.
12. Trujillo A. Técnica de la rejilla interpersonal: un estudio de validación en población comunitaria. Universitat de Barcelona; 2016. Available at: https://dialnet.unirioja.es/servlet/tesis?codigo=158549

Appendix: Supplementary Material

1 Supplementary Figures and Tables

1.1 Supplementary Figures

Supplementary Figure 1.



Boxplots with the mean differences between clusters in the three self-discrepancy variables (actual-ideal, actual-virtual and virtual-ideal). Over each box is shown the average score of the self-discrepancy variable within the indicated cluster. The significance of the mean differences is indicated by the brackets above, where three asterisks indicate a p-value <0.001, two asterisks indicate a p-value between 0.001 and 0.049, one asterisk indicates a p-value between 0.05 and 0.1 and “NS.” Indicates a p-value over 0.1.

1.2 Supplementary Tables

Supplementary Table 1: Mean scores for the Euclidian distances of the three self-discrepancy variables (actual-ideal, actual-virtual and virtual-ideal) grouped as per the resulting clusters from the k-means analysis.

	N	Actual-Ideal Discrepancy	Actual-Virtual Discrepancy	Virtual-Ideal Discrepancy	Label
Cluster 1	55	0.27	0.48	0.47	Exploration
Cluster 2	99	0.25	0.32	0.28	Proximal
Cluster 3	48	0.42	0.46	0.25	Projection

Navigating the Metaverse: Enhancing Non-Verbal Communication Through Externalized Interoceptive Signals

Jonathan Giron [1,*], Ori Chajuss [1] Emi, Schenkelbach Kliener [1] and Einat Geva [1]

[1] *Advanced Reality Lab. Reichman University, Sammy Ofer School of Communications, Herzliya, Israel*

Correspondence*:
Jonathan Giron
jonathan.giron@runi.ac.il

ABSTRACT

The Metaverse, a virtual environment where users interact with each other and digital entities, presents a novel frontier for human communication and interaction. However, one of its prominent challenges lies in the lack of non-verbal communication cues, which are integral to conveying emotions, intentions, and social nuances in face-to-face interactions. This challenge can lead to misunderstandings, misinterpretations, and a sense of disconnect among users. This paper examines the significance of non-verbal communication in human interaction and highlights the limitations of current Metaverse platforms in replicating these cues. Furthermore It explores the use of externalizing interoceptive signals using visual and auditory elements to provide a new channel of nonverbal communication pathway to accommodate these challanges. By offering this new communication pathway, interactions among individuals in the Metaverse could occur with reduced misunderstandings and foster improved, or even healthier, social dynamics.

Social virtual interaction: Interocpection, Psychophysiology, Sensory contingencies, Heart rate

1 INTRODUCTION

The past decade has brought significant advancements in virtual reality (VR) technology, with consumer-level head-mounted display (HMD) devices being widely used for gaming, education, work, and social interaction. Some believe that HMDs are the next technological step, similar to the impact of personal computers and smartphones in previous eras. This has led to the beginning of the transition of the Metaverse from fiction to reality. The Metaverse virtual world connects people in an immersive simulation, enabling enhanced collaboration and experiences Stephenson (1994) . The Metaverse would allow people to visit simulated locations, interact with others digitally, and experience high-fidelity simulations of extreme activities incorporating social interactions. To make this vision a reality, VR technology must be immersive, provide high-level graphical simulations, and facilitate natural interpersonal communication. Researchers have shown that VR experiences can trigger different emotions and awareness in users, highlighting the importance of understanding how users construct themselves and interact with others Knapp et al. (2013); Slater and Sanchez-Vives (2016). Amidst these possibilities, challenges also arise, such as maintaining positive social rapport, ensuring individual safety and overcoming limitations posed by virtual interactions.

In our work, we explore nonverbal communication in virtual social scenarios and propose a framework to overcome challenges in these realms. We suggest using visualization of affective indices in social virtual spaces to compensate for nonverbal cues that may not be available, thereby improving social interaction. The proposed framework entails capturing electrocardiogram (ECG) and respiratory signals, calculating the beats per minute for each signal, and subsequently visualizing them as orbs as part or separate from the avatar within the virtual space.

2 THE CHALLENGE AND OPPORTUNITY OF COMMUNICATING IN THE METAVERSE

Social communication and expression in VR is still developing. As with any new platform of communication, It often appropriates social and cultural built for past media. Biocca (1992) . While advancements in embodied interactions have significantly contributed to the development and progression of virtual interactions, the realm still lacks the intricate and nuanced nonverbal communication patterns found in face-to-face interactions.. This includes facial expressions, changes in skin tone such as blushing Thorstenson et al. (2020), and even chemical signals transmitted between individualsMishor et al. (2021); Ravreby et al. (2022). It is suggested that embodied avatars provide a high level of social presence with conversation patterns that are similar to face-to-face interactionsSmith and Neff (2018). Using full-body tracking to induce a sense of embodiment, we can dramatically help users depict their emotions and intentions, making non-verbal communication more powerful than verbal communication Smith and Neff (2018). The sense of embodiment refers to the feeling of owning, controlling, and being inside a body Banakou et al. (2016). To extend this, various studies have shown that a person can even achieve the sense of embodiment towards objects that are not part of the physical body Schettler et al. (2019), as well as virtual bodies in VR experiences Salomon et al. (2013). Various experiments have explored the effects of being embodied in an avatar employing full body ownership illusions on behavior, including body size, skin color, and gender swapping, which can affect feelings, thinking, bias, and empathy Slater et al. (2008); Slater (2009); Kilteni et al. (2013); Banakou et al. (2016); Neyret et al. (2020).

3 AFFECTIVE INDICES FOR NONVERBAL COMMUNICATION IN THE METAVERSE

Non-verbal communication is generally defined as any behavior that can be interpreted as informative but is not purely linguistic Knapp et al. (2013). Non verbal communication has been suggested to account for 65 percent of the meaning derived in an interaction Guerrero and Floyd (2006). Maloney et al. (2022) evaluated nonverbal communication in virtual reality. they describe that using full body tracking avatars dramatically improves users' ability to convey emotions and intentions through body language and subtle cues Maloney et al. (2020). Bio-responsive or affective systems, as defined by Stepanova et al. (2020), encompass interactive systems that utilize biosignals such as respiration, Heart Rate (HR), brainwaves, skin conductance, among others, as input. These systems generate responsive outputs that adapt to these signals, thereby enabling alternative modes of physiological interaction Stepanova et al. (2020). According to Lee et al. (2022), visually representing human biosignals can enhance bodily awareness Lee et al. (2022). In their study, they examined the integration of physiological feedback into virtual scenarios for entertainment purposes. Their findings underscore the growing recognition that digitally representing human biosignals can significantly enhance bodily awareness. Lee et al. (2022). Hassib et al. (2017) developed a chat mobile application called "HeartChat", which integrates HR as a cue to increase awareness Hassib et al. (2017) . Providing physiological feedback to users in a VR environment can make the user more aware of his/her own emotional state, and in a collaborative setup, it can help collaborators understand each other Tan et al.

(2014). In their research, Chen et al. (2017) have presented four different multi-sensory visualizations of HR data in immersive VR experiences. They found that "providing feedback of HR enhanced the user enjoyment more than not providing it at all" Chen et al. (2017) . Patibanda et al. (2017) measured how "Regular breathing exercises can be a beneficial part of leading a healthy life" Patibanda et al. (2017). They created a game called "life tree", which is a VR game in which a player controls the growth of a tree by practicing pursed-lip breathingPatibanda et al. (2017). In the "life tree" experiment participants enjoyed the visual feedback based on their breathing patterns by expanding and contracting its trunk on inhalation and exhalationPatibanda et al. (2017). Davies (1996) was a pioneer in the use of breathing as a control mechanismDavies and Harrison (1996). The project presents the use of breathing to assist players feel centered in their physical bodies during immersion in a way that is similar to the effect of practicing meditation Davies and Harrison (1996). Her project Osmose is "an immersive interactive virtual reality installation with a head-mounted display and real-time motion tracking based on breathing and balance"Davies and Harrison (1996). Desnoyers-Stewart et al. (2019,2020) presented the "JeL" as an interpersonal sincrony inducing immersive installation, they investigated the immersion of two users in a virtual underwater environment and examined how their breathing could influence the movement of a jellyfish. To correspond with the diaphragm movements of the user, the jellyfish's swimming animation progresses according to the Breathing Rate (BR). JeL's research aims to "create a stronger awareness of our bodily functions through a breath-based interaction" Desnoyers-Stewart et al. (2019); Stepanova et al. (2020). In addition, researchers have found that VR effectively incorporates breathing as a form of interaction that promotes mindfulness and relaxation Patibanda et al. (2017). Sra et al. (2019) proposed that "breathing is a directly controlled physiological signal that can facilitate unique and engaging play experiences through natural interaction in single and multiplayer virtual reality games Sra et al. (2019) .

3.1 The Developed System

To evaluate this concept, we developed a VR experince that includes a psychophysiological interface. We tested two prototypes, the first iteration included visualization of the user's HR and BR (as seen in 1). Most individuals are not consciously aware of the sensation of their heartbeat. Respiration visualization offers a somewhat clearer representation since it directly correlates with the user's breathing action. However, it necessitates users to breathe in a highly precise manner to fully grasp that what they are observing is indeed their own BR Davies and Harrison (1996). The system was designed to examine the emotional reaction of the users to the visual representation of their own affective bio signals. The data collected from this pilot study was analyzed and used as a basis for the second iteration and prototype. In the initial evaluation, the user was presented with two similar virtual objects in their field of view, one connected to the physiological data and the other to mock data simulating generic bio measurements. In addition, the voice of the user was also visualized in comparison to mocked voice input, in order to add another layer of immersion. The second iteration of the experience included new functionalities that enhance user embodiment and immersion. These include an avatar bound to the user's head, torso, and hands, a mirror avatar, and the option to visualize HR and BR data within the individual's avatar or next to it 2.

4 CONCLUSIONS

In the past decade, the rapid evolution of virtual reality (VR) technology, particularly with the widespread adoption of consumer-level head-mounted display (HMD) devices, has transformed various aspects of human interaction and communication. The concept of the Metaverse, once confined to the realms of science fiction, is gradually transitioning into reality, promising a virtual world where individuals can

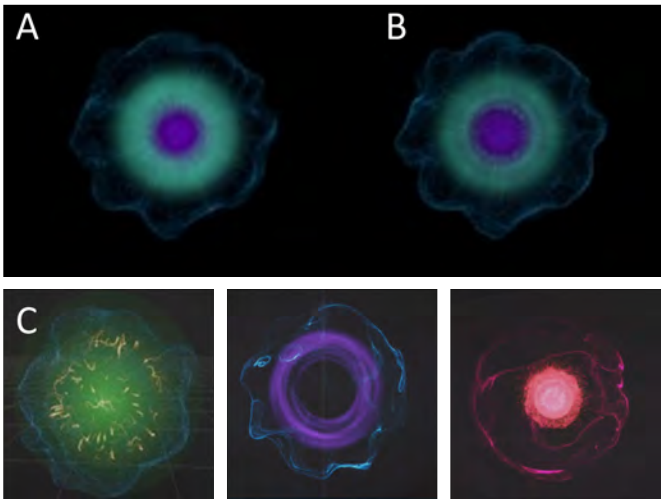


Figure 1. Virtual reality orb objects. A. Random generated visualization. B. Physiological signals live HR, and BR. C. Different bio-visualization objects tested in the first prototype

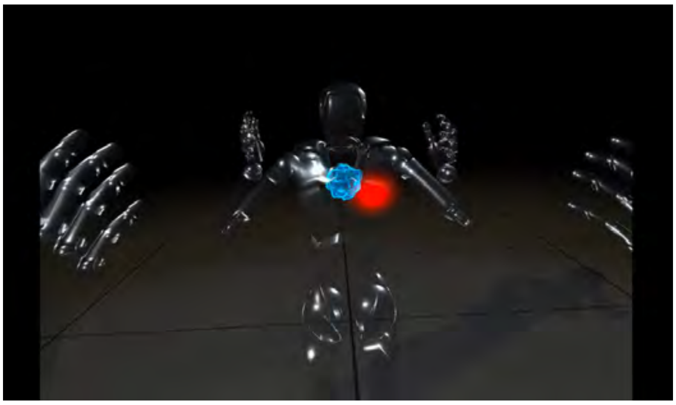


Figure 2. Embodied physiological visual representations, embedded within the users avatar

connect, collaborate, and experience immersive simulations. However, as this transition unfolds, one of the critical challenges that emerge is the replication of non-verbal communication cues within these virtual environments. Non-verbal communication, encompassing gestures, facial expressions, and social biological signals, plays a vital role in conveying emotions, intentions, and social nuances in face-to-face interactions. In virtual social scenarios facilitated by VR technology, the absence of these cues can lead to misunderstandings, misinterpretations, and a diminished sense of connection among users. To address this challenge, our work explores innovative approaches to integrate non-verbal communication cues into virtual environments, thereby enhancing the richness and authenticity of social interactions. We propose a framework that leverages the visualization of internal physiological measurements, such as Heart Rate and respiration, as novel channels of interaction. Externalizing interoceptive signals through visual and auditory elements offers users a pathway to develop a profound understanding of their emotional states. Moreover, it may serve as a novel means for individuals to authentically convey their true selves to others in the virtual space. This innovative communication channel holds the potential to foster mutual understanding and enrich collaborative settings within virtual environments. Moreover, our research demonstrates the potential of bio-responsive systems in VR to enhance immersion and embodiment, enabling users to establish a stronger connection with their virtual avatars and surroundings. By addressing the challenge of communicating in the Metaverse, we can unlock the full potential of VR technology to foster meaningful connections, collaboration, and experiences in virtual spaces.

REFERENCES

Banakou, D., Hanumanthu, P. D., and Slater, M. (2016). Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias. *Frontiers in Human Neuroscience* 10

Biocca, F. (1992). Communication Within Virtual Reality: Creating a Space for Research. *Journal of Communication* 42, 5–22. doi:10.1111/j.1460-2466.1992.tb00810.x

Chen, H., Dey, A., Billinghurst, M., and Lindeman, R. W. (2017). Exploring the design space for multi-sensory heart rate feedback in immersive virtual reality. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction* (New York, NY, USA: Association for Computing Machinery), OzCHI '17, 108–116. doi:10.1145/3152771.3152783

Davies, C. and Harrison, J. (1996). Osmose: towards broadening the aesthetics of virtual reality. *ACM SIGGRAPH Computer Graphics* 30, 25–28. doi:10.1145/240806.240808

Desnoyers-Stewart, J., Stepanova, E. R., Pasquier, P., and Riecke, B. E. (2019). JeL: Connecting Through Breath in Virtual Reality. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow Scotland Uk: ACM), 1–6. doi:10.1145/3290607.3312845

Guerrero, L. K. and Floyd, K. (2006). *Nonverbal communication in close relationships* (Routledge)

Hassib, M., Buschek, D., Wozniak, P. W., and Alt, F. (2017). HeartChat: Heart Rate Augmented Mobile Chat to Support Empathy and Awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA: Association for Computing Machinery), CHI '17, 2239–2251. doi:10.1145/3025453.3025758

Kilteni, K., Bergstrom, I., and Slater, M. (2013). Drumming in Immersive Virtual Reality: The Body Shapes the Way We Play. *IEEE Transactions on Visualization and Computer Graphics* 19, 597–605. doi:10.1109/TVCG.2013.29. Conference Name: IEEE Transactions on Visualization and Computer Graphics

Knapp, M. L., Hall, J. A., and Horgan, T. G. (2013). *Nonverbal Communication in Human Interaction* (Cengage Learning). Google-Books-ID: rWoWAAAAQBAJ

Lee, S., El Ali, A., Wijntjes, M., and Cesar, P. (2022). Understanding and Designing Avatar Biosignal Visualizations for Social Virtual Reality Entertainment. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15

Maloney, D., Freeman, G., and Wohn, D. Y. (2020). "Talking without a Voice": Understanding Non-verbal Communication in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction* 4, 1–25. doi:10.1145/3415246

Mishor, E., Amir, D., Weiss, T., Honigstein, D., Weissbrod, A., Livne, E., et al. (2021). Sniffing the human body volatile hexadecanal blocks aggression in men but triggers aggression in women. *Science Advances* 7, eabg1530. doi:10.1126/sciadv.abg1530. Publisher: American Association for the Advancement of Science

Neyret, S., Navarro, X., Beacco, A., Oliva, R., Bourdin, P., Valenzuela, J., et al. (2020). An Embodied Perspective as a Victim of Sexual Harassment in Virtual Reality Reduces Action Conformity in a Later Milgram Obedience Scenario. *Scientific Reports* 10, 1–18. doi:10.1038/s41598-020-62932-w. Publisher: Springer Nature

Patibanda, R., Mueller, F. F., Leskovsek, M., and Duckworth, J. (2017). Life Tree: Understanding the Design of Breathing Exercise Games. In *Proceedings of the Annual Symposium on Computer-Human*

Interaction in Play (New York, NY, USA: Association for Computing Machinery), CHI PLAY '17, 19–31. doi:10.1145/3116595.3116621

Ravreby, I., Snitz, K., and Sobel, N. (2022). There is chemistry in social chemistry. *Science Advances* 8, eabn0154. doi:10.1126/sciadv.abn0154. Publisher: American Association for the Advancement of Science

Salomon, R., Lim, M., Pfeiffer, C., Gassert, R., and Blanke, O. (2013). Full body illusion is associated with widespread skin temperature reduction. *Frontiers in Behavioral Neuroscience* 7, 65. doi:10.3389/fnbeh.2013.00065

Schettler, A., Raja, V., and Anderson, M. L. (2019). The Embodiment of Objects: Review, Analysis, and Future Directions. *Frontiers in Neuroscience* 13

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 3549–3557. Publisher: The Royal Society

Slater, M., Perez-Marcos, D., Ehrsson, H. H., and Sanchez-Vives, M. V. (2008). Towards a digital body: the virtual arm illusion. *Frontiers in Human Neuroscience* 2, 6. doi:10.3389/neuro.09.006.2008

Slater, M. and Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI* 3. doi:https://doi.org/10.3389/frobt.2016.00074

Smith, H. J. and Neff, M. (2018). Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12

Sra, M., Jain, A., and Maes, P. (2019). Adding proprioceptive feedback to virtual reality experiences using galvanic vestibular stimulation. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–14

Stepanova, E. R., Desnoyers-Stewart, J., Pasquier, P., and Riecke, B. E. (2020). JeL: Breathing Together to Connect with Others and Nature. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven Netherlands: ACM), 641–654. doi:10.1145/3357236.3395532

Stephenson, N. (1994). *Snow crash* (Penguin UK)

Tan, C. S. S., Luyten, K., Van Den Bergh, J., Schöning, J., and Coninx, K. (2014). The Role of Physiological Cues during Remote Collaboration. *Presence: Teleoperators and Virtual Environments* 23, 90–107. doi:10.1162/PRES_a_00168

Thorstenson, C. A., Pazda, A. D., and Lichtenfeld, S. (2020). Facial blushing influences perceived embarrassment and related social functional evaluations. *Cognition and Emotion* 34, 413–426. doi:10.1080/02699931.2019.1634004. Publisher: Routledge _eprint: https://doi.org/10.1080/02699931.2019.1634004