Immersive Virtual Reality Games for Societal Impact

Physical Activity, Active Aging, and Social Interaction

Sukran Karaosmanoglu

born in İzmir, Türkiye

an der Universität Hamburg eingereichte Dissertation

A cumulative doctoral dissertation submitted in fulfillment of the requirements for
the degree of Doctor rerum naturalium
(*Dr. rer. nat.*)
to the Faculty of Mathematics, Informatics, and Natural Sciences
Department of Informatics
Universität Hamburg
Germany

2025

Supervisor & First Examiner:	Prof. Dr. Frank Steinicke
Second Examiner:	Prof. Dr. Regan Mandryk
Third Examiner:	Prof. Dr. Kathrin Gerling
Chair of Examination Committee:	Prof. Dr. Janick Edinger
Deputy Chair of Examination Committee:	Prof. Dr. Eva Bittner
Date of Dissertation Defense:	21.05.2025

Note: This dissertation adopts the scientific pronoun "we" to acknowledge the valuable contributions of fellow researchers and students who were involved in the research papers of this dissertation.

Dedicated to my family.

ACKNOWLEDGMENTS

As I reach one of the milestones of my academic journey, and reflecting on the meaning of my name (Sukran means "thankfulness" in Turkish), I would like to express my gratitude to those who supported me and were there for me throughout this journey.

Firstly, I am thankful to *my doctoral father*, Prof. Dr. Frank Steinicke, for welcoming me into his research group, giving me the freedom to explore the scientific challenges I wanted to tackle, trusting and supporting me at every step, creating an environment where I could thrive, and for all the thumbs-up emojis that helped keep me motivated. Prof. Dr. Regan Mandryk, thank you for reviewing my dissertation and for your encouraging talks; you have always been a role model to me as a successful and incredibly supportive researcher. Prof. Dr. Kathrin Gerling, thank you for reviewing my dissertation; your work has always been inspirational, and I feel privileged to have you as one of my examiners. Prof. Dr. Janick Edinger and Prof. Dr. Eva Bittner, thank you for being part of my examination committee and for chairing my oral defense.

Along the way, I was fortunate to meet people who inspired me and made my academic life great: Prof. Dr. Lennart E. Nacke, thank you for your insightful advices, friendly conversations, successful collaborations, and for welcoming me into your group for a research visit. Dr. Katja Rogers, thank you not only for being academically critical, inspiring, and supportive but also for being a friend with whom I can laugh and cry. Dr. Sebastian Cmentowski, thank you for the digital co-working sessions, fruitful research discussions, and wonderful collaborations, but most importantly, thank you for being such a caring and supportive friend. Dr. Maximilian Altmeyer and Dr. Philipp Sykownik, thank you for the pleasant collaborations and the wonderful conversations we shared about academia and life. I also want to express my gratitude to all my other collaborators; without you, this dissertation would not have been possible.

I would like to extend my thanks to Universität Hamburg, the Bundesministerium für Bildung und Forschung (BMBF), the Deutscher Akademischer Austauschdienst (DAAD), and the Excellence Strategy of the Federal Government and the Länder for their generous financial support for my research. I would also like to thank the Hospital zum Heiligen Geist Hamburg for the nice collaborations on several projects. Lastly, I would like to thank all of my participants; without your participation, none of the research projects would have been possible.

To all my wonderful colleagues in the Human-Computer Interaction research group, thank you for creating such a friendly environment, for the motivating speeches, and for the coffee breaks; you made me feel at home. I must also thank all of my students who took my courses, worked with me as teaching and research assistants, and those who chose me to supervise their projects and theses; together, we accomplished many great things! My friends, Tom, Aslı, Nati, Iwona, and Toby, thank you for being there to cheer me up and for letting me rant when something was bothering me. Mary, Teresa, Susanne, and Jana, thank you for our coffee talks about life and academia. Thank you, volleyball, for helping me relax and take my mind off work.

Lastly and most importantly, thank you to my lovely family. I would never have achieved this without your unconditional support! Mom, thank you for always being there for me, no matter what I need. I am deeply sorry for worrying you from time to time with my extensive working hours, especially during deadline periods. Dad, I feel like the luckiest girl in the world to have you as my father. Thank you for everything, and especially for discussing academic topics with me, even though you are not part of that world. My little sister, you are the most precious thing to me! I have a large, supportive family that would be difficult to mention fully here, but a heartfelt thank you goes to my little Selen Nur, cousins, aunts, uncles, and grandparents. And yes, grandpa, I made it and hope I have made you proud! Finally, my partner, thank you for always being there for me, whether I was celebrating my successes or feeling overwhelmed by my challenges.

Son olarak ve en önemlisi, sevgili aileme teşekkür ederim. Bu başarıyı, koşulsuz destekleriniz olmadan asla elde edemezdim! Anneciğim, her zaman, neye ihtiyacım olursa olsun, yanımda olduğun için teşekkür ederim. Özellikle son teslim tarihleri sırasında yoğun çalışma saatlerimle seni zaman zaman endişelendirdiğim için özür dilerim. Babacığım, benim babam olduğun için kendimi dünyanın en şanslı kızı hissediyorum. Her şey için, özellikle de akademik dünyadan olmasan bile akademik konuları benimle tartıştığın için teşekkür ederim. Küçük kız kardeşim, sen benim için en değerli şeysin! Burada ismini tek tek anmak zor olsa da, benim büyük ve destekleyici bir ailem var. Onlara, yani benim küçük Selen Nur'uma, kuzenlerime, (her zaman kalbimde olan) teyzeme, yengeme, amcama, büyükannelerime, ve dedelerime içten bir teşekkür gönderiyorum. Ve evet, dede, başardım ve umarım seni gururlandırmışımdır!

Canım ailem, sizi çok seviyorum!

ABSTRACT

Many of today's challenges are tightly connected to a rapid digitalization, which is accelerating societal developments significantly. However, while technological advancements may bear the risk of amplifying societal challenges, they can also help to solve these challenges. With the goal of tackling societal challenges, this dissertation targets the United Nations' third sustainable development goal of "*good health and well-being - ensure healthy lives and promote well-being for all at all ages*". To this end, this dissertation takes a Human-Computer Interaction (HCI) perspective and focuses on designing, developing, and evaluating technology-based solutions, specifically immersive virtual reality (VR) games, to help address three societal challenges of today's world: (i) *physical activity*, (ii) *active aging*, and (iii) *social interaction*.

In the last decades, VR-head-mounted displays (HMDs) have experienced rapid improvements in performance, accessibility, and affordability. Especially, games have become a major application area of this technology since its immersive nature is an ideal fit for delivering engaging and playful experiences. Immersive VR games do not only entertain users but also have enormous potential for addressing pressing societal challenges due to their immersive, spatial, and motivating characteristics. However, our understanding of how to design and develop these games for such purposes is still limited. Accordingly, this dissertation presents research contributions on how immersive VR games can help address the three outlined societal challenges.

Overall, in this dissertation, we examine the spectrum of immersive VR games, from single-user experiences to multi-user environments, and address diverse target groups, including younger and older adults, as well as older adults with different degrees of dementia. As a first societal challenge, we focus on how immersive VR games can support physical activity. We provide an up-to-date overview, envision new training opportunities, and offer empirical insights into the benefits of using real-time continuous cues and various game elements. Following this, we show how VR games can support active aging by designing age-appropriate cognitive and physical exercises, considering requirements of older adults with dementia, evaluating the impact of augmented interactions, and analyzing advantages and disadvantages of different exergaming technologies. Lastly, we showcase how immersive VR games can support social interaction. We demonstrate their potential for creating social gaming environments, provide a theoretical understanding and uncover research gaps of asymmetric multiplayer immersive VR games, use asymmetric game design and resulting interdependences to bridge the gap between players

of different mediums, and explore how existing social ties influence the experience of asymmetric multiplayer immersive VR games. The dissertation concludes with a discussion of the main contributions, the key lessons learned, and the future research directions. Our work establishes a foundation for future research in the field of HCI and promotes the use of immersive VR games to foster a positive societal impact for individuals of all ages and conditions.

ZUSAMMENFASSUNG

Viele der heutigen Herausforderungen sind eng mit einer rasanten Digitalisierung verbunden, die gesellschaftliche Entwicklungen erheblich beschleunigt. Technologische Fortschritte bergen zwar das Risiko, gesellschaftliche Herausforderungen zu verstärken, können aber auch zu deren Lösungen beitragen. Mit dem Ziel, gesellschaftliche Herausforderungen anzugehen, adressiert diese Dissertation das dritte Ziel für nachhaltige Entwicklung der Vereinten Nationen: "*Gute Gesundheit und Wohlbefinden – ein gesundes Leben gewährleisten und das Wohlbefinden für alle in jedem Alter fördern*". Diese Dissertation adoptiert eine HCI-Perspektive und konzentriert sich auf die Gestaltung, Entwicklung und Evaluierung technologiebasierter Lösungen, insbesondere immersiver VR-Spiele, um drei gesellschaftliche Herausforderungen der heutigen Welt anzugehen: (i) *körperliche Aktivität*, (ii) *aktives Altern* und (iii) *soziale Interaktion*.

In den letzten Jahrzehnten haben sich VR-HMDs rasant in Bezug auf Leistung, Zugänglichkeit und Erschwinglichkeit entwickelt. Insbesondere Spiele sind zu einem wichtigen Anwendungsbereich dieser Technologie geworden, da ihre immersive Natur ideal ist, um fesselnde und spielerische Erfahrungen zu vermitteln. Immersive VR-Spiele dienen jedoch nicht nur der Unterhaltung der Nutzer, sondern besitzen aufgrund ihrer immersiven, räumlichen und motivierenden Eigenschaften auch ein enormes Potenzial, um drängende gesellschaftliche Herausforderungen anzugehen. Dennoch ist unser Verständnis dafür, wie diese Spiele für solche Zwecke gestaltet und entwickelt werden können, sehr begrenzt. Diese Dissertation präsentiert Forschungsbeiträge, die zeigen, wie immersive VR-Spiele dazu beitragen können, die drei skizzierten gesellschaftlichen Herausforderungen anzugehen.

Insgesamt untersuchen wir in dieser Dissertation das Spektrum immersiver VR-Spiele, von Einzelspieler-Erfahrungen bis hin zu Mehrbenutzerumgebungen, und sprechen verschiedene Zielgruppen an, darunter jüngere und ältere Erwachsene sowie ältere Erwachsene mit unterschiedlichen Demenzgraden. Zunächst konzentrieren wir uns darauf, wie immersive VR-Spiele körperliche Aktivität unterstützen können. Wir geben einen aktuellen Überblick, stellen neue Trainingsmöglichkeiten vor und bieten empirische Einblicke in die Vorteile von kontinuierlichem Feedback in Echtzeit und anderen Spielelementen. Anschließend zeigen wir, wie immersive VR-Spiele das aktive Altern unterstützen können, indem wir altersgerechte kognitive und körperliche Übungen entwickeln, die Anforderungen älterer Menschen mit Demenz erforschen, die Auswirkungen augmentierter Interaktionen untersuchen und die Vor- und Nachteile verschiedener Exergaming-

Technologien analysieren. Schließlich zeigen wir, wie immersive VR-Spiele die soziale Interaktion unterstützen können. Wir demonstrieren ihr Potenzial für soziale Spielumgebungen, vermitteln ein theoretisches Verständnis für und decken Forschungslücken bei asymmetrischen immersiven VR-Mehrspieler-Spielen auf, nutzen asymmetrisches Spieldesign und daraus resultierende Interdependenzen für das gemeinsame Spielen über verschiedene Medien, und untersuchen, wie bestehende soziale Bindungen die Erfahrung asymmetrischer immersiver VR-Mehrspieler-Spiele beeinflussen. Die Dissertation endet mit einer Diskussion der wichtigsten gewonnenen Erkenntnisse und der offenen Fragestellungen für zukünftige Forschung. Unsere Arbeit stellt eine Grundlage für zukünftige Forschung im Bereich der Mensch-Computer-Interaktion dar und zeigt das Potenzial des Einsatzes immersiver VR-Spiele, um eine positive gesellschaftliche Wirkung für Menschen jeden Alters und jeder Lebenslage zu erzielen.

CORE PUBLICATIONS

This cumulative dissertation is based on the following peer-reviewed core publications. Each of these core publications is included in Part II of this dissertation.

The ✱ icon indicates shared first-authorship, while the 🎗 icon denotes honorable mentions.

CORE PUBLICATIONS

[CorePub1] **Sukran Karaosmanoglu**, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. Born to run, programmed to play: Mapping the extended reality exergames landscape. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, (CHI '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 979-8-4007-0330-0/24/05. DOI 10.1145/3613904.3642124. URL `https://doi.org/10.1145/3613904.3642124`.

[CorePub2] ✱🎗 Sebastian Cmentowski, **Sukran Karaosmanoglu**, Lennart E. Nacke, Frank Steinicke, and Jens Harald Krüger. Never skip leg day again: Training the lower body with vertical jumps in a virtual reality exergame. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. DOI 10.1145/3544548.3580973. URL `https://doi.org/10.1145/3544548.3580973`.

[CorePub3] **Sukran Karaosmanoglu**, Moritz Wegner, Sebastian Cmentowski, and Frank Steinicke. Move, react, repeat! the role of continuous cues in immersive exergames. *Proceedings of the ACM on Human-Computer Interaction*, 8 (CHI PLAY), 2024. DOI 10.1145/3677091. URL `https://doi.org/10.1145/3677091`.

[CorePub4] **Sukran Karaosmanoglu**, Lucie Kruse, Sebastian Rings, and Frank Steinicke. Canoe vr: An immersive exergame to support cognitive and physical exercises of older adults. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, (CHI EA '22), New York,

NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391566. DOI 10.1145/3491101.3519736. URL https://doi.org/10.1145/3491101.3519736.

[CorePub5] **Sukran Karaosmanoglu**, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. Lessons learned from a human-centered design of an immersive exergame for people with dementia. *Proceedings of the ACM on Human-Computer Interaction*, 5 (CHI PLAY), 2021. DOI 10.1145/3474679. URL https://doi.org/10.1145/3474679.

[CorePub6] ✱ Philipp Sykownik, **Sukran Karaosmanoglu**, Katharina Emmerich, Frank Steinicke, and Maic Masuch. Vr almost there: Simulating co-located multiplayer experiences in social virtual reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. DOI 10.1145/3544548.3581230. URL https://doi.org/10.1145/3544548.3581230.

[CorePub7] Katja Rogers, **Sukran Karaosmanoglu**, Dennis Wolf, Frank Steinicke, and Lennart E. Nacke. A best-fit framework and systematic review of asymmetric gameplay in multiplayer virtual reality games. *Frontiers in Virtual Reality*, 2, 2021. ISSN 2673-4192. DOI 10.3389/frvir.2021.694660. URL https://doi.org/10.3389/frvir.2021.694660.

[CorePub8] **Sukran Karaosmanoglu**, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. Feels like team spirit: Biometric and strategic interdependence in asymmetric multiplayer vr games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, (CHI '21), New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380966. DOI 10.1145/3411764.3445492. URL https://doi.org/10.1145/3411764.3445492.

[CorePub9] **Sukran Karaosmanoglu**, Tom Schmolzi, and Frank Steinicke. Playing with friends or strangers? the effects of familiarity between players in an asymmetric multiplayer virtual reality game. In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY Companion '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400700293. DOI 10.1145/3573382.3616079. URL https://doi.org/10.1145/3573382.3616079.

Author Contributions

I am the main author of the following peer-reviewed publications: [CorePub1, CorePub3, CorePub4, CorePub5, CorePub8, CorePub9]. As the main author, I led these papers and contributed significantly to their major research steps, including the ideation, conceptualization, study design, design and implementation of the games, data collection, analysis, and writing the papers. The data collection and game implementation steps of [CorePub8] were still part of my master's thesis. In all listed publications, the other co-authors contributed to various parts of the papers, such as the conceptualization, implementation, data collection, analysis, writing, or feedback.

I am the shared-first author of the following peer-reviewed publications: [CorePub2, CorePub6]. The other shared-first authors also used the respective papers in their dissertations. In [CorePub2], while my focus was on planning and preparing the study design and evaluating both the quantitative and qualitative data, Dr. Sebastian Cmentowski's focus was on conducting the expert interviews and study. In [CorePub6], while my focus was on the qualitative data analysis, Dr. Philipp Sykownik's focus was on executing the study. All other parts of both papers were done in close collaboration between the respective main authors. The remaining co-authors of both papers contributed to various parts of the papers, such as data analysis, writing, or feedback.

I am the second author of the following peer-reviewed publication: [CorePub7]. As the second author, I contributed significantly to the following steps in this research: the conceptualization, methodology, data extraction, analysis, and paper writing. Dr. Katja Rogers is the main author and led the paper in all steps of the research. The remaining co-authors contributed to various parts of the paper, such as the data analysis, or feedback.

CORE PREPRINTS

In addition to the core publications, this cumulative dissertation is also based on the following non peer-reviewed core preprints. Each of these core preprints is included in Part II of this dissertation.

CORE PREPRINTS

[CorePre1] **Sukran Karaosmanoglu**, Bastian Kordyaka, and Frank Steinicke. Unlocking the potential: The role of game elements on player motivation and intention of long-term use in virtual reality exergames. 2025. DOI 10.25592/uhhfdm.16793. URL https://doi.org/10.25592/uhhfdm.16793.

[CorePre2] **Sukran Karaosmanoglu**, Sebastian Finnern, Frank Steinicke, and Katja Rogers. Evaluating augmented locomotion and range of reachable objects for older adults in a virtual reality exergame. 2025. DOI 10.25592/uhhfdm.16795. URL https://doi.org/10.25592/uhhfdm.16795.

[CorePre3] **Sukran Karaosmanoglu**, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. Exercube vs. virtual reality: A comparative study of exergame technologies for older adults. 2025. DOI 10.25592/uhhfdm.16797. URL https://doi.org/10.25592/uhhfdm.16797.

Author Contributions

I am the main author of the following preprints: [CorePre1, CorePre2, CorePre3]. As the main author, I led these papers and contributed significantly to their major research steps, including the ideation, conceptualization, study design, design and implementation of the games, data collection, analysis, and writing the papers. In all listed preprints, the other co-authors contributed to various parts of the papers, such as the conceptualization, implementation, data collection, analysis, writing, or feedback.

OTHER PUBLICATIONS

Moreover, I was also involved in the following papers, which also influenced this dissertation, however, they are not part of this cumulative dissertation:

The ✱ icon indicates shared first-authorship, while the 🎗 icon denotes honorable mentions.

OTHER PUBLICATIONS

[OtherPub1] Sebastian Rings, **Sukran Karaosmanoglu**, Lucie Kruse, Daniel Apken, Tobias Picker, and Frank Steinicke. Using exergames to train patients with dementia to accomplish daily routines. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY '20), New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450375870. DOI 10.1145/3383668.3419883. URL `https://doi.org/10.1145/3383668.3419883`.

[OtherPub2] ✱ Jenny Gabel, **Sukran Karaosmanoglu**, Celeste Mason, Sebastian Rings, and Frank Steinicke. Corona beat - kicking the sedentary habit induced by prolonged social distancing. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, (IEEE VR VRW '21), 2021. DOI 10.1109/VRW52623.2021.00226. URL `https://doi.org/10.1109/VRW52623.2021.00226`.

[OtherPub3] Julia Hertel, **Sukran Karaosmanoglu**, Susanne Schmidt, Julia Bräker, Martin Semmann, and Frank Steinicke. A taxonomy of interaction techniques for immersive augmented reality based on an iterative literature review. In *2021 IEEE International Symposium on Mixed and Augmented Reality*, (ISMAR '21), 2021. DOI 10.1109/ISMAR52148.2021.00060. URL `https://doi.org/10.1109/ISMAR52148.2021.00060`.

[OtherPub4] Lucie Kruse, **Sukran Karaosmanoglu**, Sebastian Rings, Benedikt Ellinger, and Frank Steinicke. Enabling immersive exercise activities for older adults: A comparison of virtual reality exergames and traditional video exercises. *Societies*, 11(4), 2021. ISSN 2075-4698. DOI

10.3390/soc11040134. URL https://doi.org/10.3390/soc11040134.

[OtherPub5] Lucie Kruse, **Sukran Karaosmanoglu**, Sebastian Rings, Benedikt Ellinger, Daniel Apken, Thandiwe Feziwe Mangana, and Frank Steinicke. A Long-Term User Study of an Immersive Exergame for Older Adults with Mild Dementia during the COVID-19 Pandemic. In Jason Orlosky, Dirk Reiners, and Benjamin Weyers, editors, *International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, (ICAT-EGVE '21). The Eurographics Association, 2021. ISBN 978-3-03868-142-7. DOI 10.2312/egve.20211322. URL https://doi.org/10.2312/egve.20211322.

[OtherPub6] Lucie Kruse, **Sukran Karaosmanoglu**, Sebastian Rings, and Frank Steinicke. Evaluating difficulty adjustments in a vr exergame for younger and older adults: Transferabilities and differences. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*, (SUI '22), New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450399487. DOI 10.1145/3565970.3567684. URL https://doi.org/10.1145/3565970.3567684.

[OtherPub7] 🎖 Katja Rogers, **Sukran Karaosmanoglu**, Maximilian Altmeyer, Ally Suarez, and Lennart E. Nacke. Much realistic, such wow! a systematic literature review of realism in digital games. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, (CHI '22), New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391573. DOI 10.1145/3491102.3501875. URL https://doi.org/10.1145/3491102.3501875.

[OtherPub8] ✱ Sebastian Cmentowski, **Sukran Karaosmanoglu**, Fabian Kievelitz, Frank Steinicke, and Jens Krüger. A matter of perspective: Designing immersive character transitions for virtual reality games. *Proceedings of the ACM on Human-Computer Interaction*, 7 (CHI PLAY), 2023. DOI 10.1145/3611023. URL https://doi.org/10.1145/3611023.

[OtherPub9] Bastian Dewitz, **Sukran Karaosmanoglu**, Robert W. Lindeman, and Frank Steinicke. Magic, superpowers, or empowerment? a conceptual framework for magic interaction techniques. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, (IEEE VR VRW '23), 2023. DOI 10.1109/VRW58643.2023.00246. URL https://doi.org/10.1109/VRW58643.2023.00246.

[OtherPub10] Jana Franceska Funke, Anja Schikorr, **Sukran Karaosmanoglu**, Teresa Hirzle, Frank Steinicke, and Enrico Rukzio. Tiles to move: Investigating tile-based locomotion for virtual reality. *Proceedings of the ACM on Human-Computer Interaction*, 7 (CHI PLAY), 2023. DOI 10.1145/3611060. URL https://doi.org/10.1145/3611060.

[OtherPub11] Thereza Schmelter, Lucie Kruse, **Sukran Karaosmanoglu**, Sebastian Rings, Frank Steinicke, and Kristian Hildebrand. Towards more inclusive and accessible virtual reality: Conducting large-scale studies in the wild. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI EA '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394222. DOI 10.1145/3544549.3583888. URL https://doi.org/10.1145/3544549.3583888.

[OtherPub12] Sebastian Cmentowski, **Sukran Karaosmanoglu**, and Frank Steinicke. From virtual gains to real pains: Potential harms of immersive exergames, 2024. URL https://arxiv.org/abs/2405.05915.

[OtherPub13] **Sukran Karaosmanoglu**, Elisabeth L Fittschen, Hande Eyicalis, David Kraus, Henrik Nickelmann, Anna Tomko, and Frank Steinicke. Language of zelda: Facilitating language learning practices using chatgpt. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, (CHI EA '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703317. DOI 10.1145/3613905.3648107. URL https://doi.org/10.1145/3613905.3648107.

[OtherPub14] Katja Rogers, Teresa Hirzle, **Sukran Karaosmanoglu**, Paula Toledo Palomino, Ekaterina Durmanova, Seiji Isotani, and Lennart E. Nacke. An umbrella review of reporting quality in chi systematic reviews: Guiding questions and best practices for hci. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 31(5), 2024. ISSN 1073-0516. DOI 10.1145/3685266. URL https://doi.org/10.1145/3685266.

[OtherPub15] Samuli Laato, **Sukran Karaosmanoglu**, Juho Hamari, Marlies Brunnhofer, and Bastian Kordyaka. Character preferences in competitive games: A study among league of legend players. In *Proceedings of the 58th Hawaii International Conference on System Sciences*, (HICSS '25), 2025. DOI 10.24251/HICSS.2025.514. URL https://doi.org/10.24251/HICSS.2025.514.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

3D	Three Dimensional
6-DOF	Six Degrees of Freedom
ACM	Association for Computing Machinery
AI	Artificial Intelligence
ANOVA	Analysis of Variance
ART	Aligned Rank Transform
AR	Augmented Reality
AV	Augmented Virtuality
CAVE	CAVE Automatic Virtual Environment Display
COVID-19	Coronavirus Disease 2019
ESA	Entertainment Software Association
GPEDT	Goals, People, Exercises, Design, Technologies
HCD	Human-Centered Design
HCI	Human-Computer Interaction
HMD	Head-Mounted Display
HIIT	High-Intensity Interval Training
IPQ	Igroup Presence Questionnaire
ISO	International Organization for Standardization
MR	Mixed Reality
NMMSP	Networked Minds Measure of Social Presence
PACT	People, Activities, Contexts, Technologies
PCI	Player-Computer Interaction
RQ	Research Question
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR	Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
PXI	Player Experience Inventory
RVC	Reality-Virtuality Continuum

SPGQ	Social Presence in Gaming Questionnaire
SSQ	Simulator Sickness Questionnaire
SUS	Slater-Usoh-Steed Questionnaire
VE	Virtual Environment
VR	Virtual Reality
WHO	World Health Organization
XR	Extended Reality

Part I

DISSERTATION

1 INTRODUCTION

In today's digitalized world, we experience a multitude of benefits and challenges that affect and transform our society as a whole. Foremost, we see enormous benefits of technology. We interact with technology in every aspect of every day life—at home, at work, and in public spaces. We communicate in real-time via messenger applications, relax using video games, and even turn on the lights in our homes using artificial intelligence (AI)-integrated smart home devices. This interaction with computing systems has been shaping our past, present, and will continue to shape our future profoundly. However, we also see the societal challenges amplified by digitalization, inspiring us—as HCI experts—to actively engage in addressing them. Accordingly, we believe that HCI "*has a lot to contribute to the process of solving societal challenges*" [212].

1.1 SOCIETAL CHALLENGES

Societal challenges are complex issues that hinder sustainable development and prosperity of future generations [214]. Within the Horizon 2020 framework for research and innovation, the European Commission highlighted several of these pressing challenges, including "*health, demographic change and wellbeing*" [63]. Similarly, the United Nations' 2030 agenda for sustainable development has outlined 17 sustainable development goals to achieve a better future globally [214]. In this dissertation, we focus on the third sustainable development goal of the United Nations: "*good health and well-being - ensure healthy lives and promote well-being for all at all ages*". Specifically, we explore how technology-based solutions, particularly immersive VR games, can make a positive societal impact by contributing to reaching this development goal. Exemplary, we target three pressing societal challenges that are influenced but can also be supported by technological advances: (i) *physical activity*, (ii) *active aging*, and (iii) *social interaction*. We derived these aspects from the United Nations' third development goal, forming the foundation for good health (i.e., physical activity [233]), particularly for older generations (i.e., active aging [231]), and well-being (i.e., social interaction [155, 156]).

PHYSICAL ACTIVITY Nowadays, many jobs involve working from home in front of a computer. While this development has many advantages, such as efficiency and accessibility, it also introduces a new challenge: Insufficient physical activity as well as sedentary lifestyles have greatly increased [135, 229] and evolved into one of the major problems of societies [232]. According to recent statistics, 1.8 billion adults do not meet daily physical activity recommendations—they are physically inactive [200, 233]. Since this ultimately increases the risk of severe diseases (e.g., cardiovascular conditions) and places a big burden on our economies (e.g., health systems), the World Health Organization (WHO) recommends ensuring "*access to programmes, services and equipment that provide affordable, enjoyable and inclusive opportunities for all people to be active*" [232].

ACTIVE AGING Technological advancements are contributing to a rapidly aging population [229]: We are experiencing a "*longevity revolution*" [215]. The United Nations predicts that the population of older adults (65+) will grow to 1.5 billion people by 2050 [215]. This has a profound impact not only on the health of individuals, but also on government policies and financial resources [216]. Therefore, it is crucial to support older adults to stay active as long as possible to retain an independent life and autonomy [231].

SOCIAL INTERACTION Technology is also transforming our social interactions significantly by allowing us to connect with others remotely and instantly. According to the WHO's Geneva Charter for well-being agenda, "*addressing the impacts of digital transformation*" is one of the five target areas for "*creating sustainable well-being societies*" [61]. While technology can have negative effects, such as contributing to isolation [61] and weakening close connections [114], it can also increase our potential social circles dramatically and support well-being (e.g., [72, 78, 221]). Accordingly, we need to "*strengthen the benefits*" of digital transformation [155] and develop technology-driven opportunities for improved social interactions in our increasingly digital world.

Focusing on these three key societal challenges—(i) *physical activity*, (ii) *active aging*, and (iii) *social interaction*—, this dissertation explores how we can design, develop, and evaluate immersive VR games that contribute to their solution. While the connection between immersive VR games and the three selected societal challenges may not be immediately apparent, we believe in the potential of immersive VR games for creating positive societal impact. In the following, we explain how the unique capabilities of immersive VR games merit their exploration for the benefit of society.

1.2 IMMERSIVE VR GAMES

Advances in technological developments and affordability have evolved immersive VR technology from specialized, niche equipment to a widely available hardware [195]. Nowadays, VR hardware and its applications are being increasingly used in many areas, such as education (e.g., [73, 85]) and health (e.g., [169, 170]). According to a report on VR hardware [195], this increase is expected to continue and the market revenue is forecast to reach $17.89 billion USD in 2029.

The most common form of immersive VR technology is HMD-based setups. These systems typically consist of two high-level components, i.e., (i) a head-worn display (i.e., VR-HMD) and (ii) two controllers. While a VR-HMD delivers a stereoscopic and real-time head-dependent view, combining VR-HMD with tracked controllers enables users to interact with virtual environments (VEs) [198]. Due to this technology's high visual and auditory fidelity, users get fully immersed in computer-generated three dimensional (3D) environments [198]. Compared to other traditional media, VR technology offers embodied interactions (e.g., [69]), allows to explore new locations virtually without the need to travel (e.g., [205]), and enables to engage in safe scenarios (e.g., exposure to virtual spiders for people with spider phobia [74]) as well as unrealistic scenarios (e.g., flying [115, 241]). These distinctive features of immersive VR make it an ideal technology for a variety of applications. One of these application areas, that has received a lot of attention, are *games*.

Games are interactive rule-structured play systems [207]. Since ancient times, they have been in our lives, serving various purposes [132]. Games are a sign of innovation and creativity; "*games do not just adopt*" [59] the established standards of computer science (e.g., presenting a notification in a new dialog window), but actively create novel solutions (e.g., using calm messaging style for texts). They have been instrumental in pushing the boundaries of computer science and served as "*a perfect test-bed for HCI explorations*" [107]. For example, many AI algorithms (e.g., Minimax for chess [213]) were created or improved through games [240].

Nowadays, games are used by a wide range of users and have become a part of many people's daily lives. According to the 2024 report of the Entertainment Software Association (ESA) [62], 61% of US-based people across all ages play video games at least one hour a week. The widespread adoption of games is also apparent in the financial size of the gaming industry; its market volume has been expected to reach $398.20 billion USD in 2029 [194]. Games offer enjoyment and relaxation for players [62], but can also have serious benefits and target serious issues. They stimulate the brain [116], improve well-being [OtherPub5], and increase awareness about societal issues [113].

Games played with immersive VR technology (i.e., *immersive VR games*) have enormous potential to tackle today's societal challenges. Immersive VR games benefit from the spatial nature of VR along with the inherent entertaining nature of gaming. In these games, players can be fully immersed in VEs, where they can interact with virtual objects and have fun with entertaining quests in the meantime. Compared to traditional games (e.g., console games), immersive VR games allow players to take a first-person perspective and move around in the VE themselves, rather than just controlling their characters. Similarly, players can engage in full-body movement in a unique way. Instead of making a character jump in a computer game or jumping in front of a TV screen in a console game, players can actually perform a jump to overcome the obstacles approaching them. These games also enable distinct opportunities for designers and developers. They can have a full control over VEs and customize them. Accordingly, research on immersive VR games presents promising results and opportunities, indicating that these games can motivate users to move (e.g., [22]), enable virtual visits to fully immersive and realistic locations for older adults (e.g., [OtherPub5]), and offer possibilities for embodied social interaction (e.g., [CorePub6]).

1.3 RESEARCH QUESTIONS AND CONTRIBUTIONS

Despite the potential of immersive VR games to address societal challenges, we have only limited knowledge of how to design and build them optimally to be effective tools to counter these challenges. Therefore, this dissertation targets this research gap and formulates the following overarching research question (RQ):

> ? **RQ:** *How can immersive VR games support addressing contemporary societal challenges?*

As our core contribution, we present 12 HCI research papers (nine peer-reviewed publications and three preprints) covering literature reviews, design and implementation of various immersive VR games, empirical studies, and design and theoretical implications. Following Wobbrock and Kientz [227]'s categorization in the field of HCI, we specifically list our contribution types, but we note that some publications contain more than one contribution type (for a full overview of the contributions in this dissertation, see Part II):

- two survey contributions present an overview of the landscape of immersive VR games.
- nine artifact contributions offer insights into the design and implementation of immersive VR games.
- ten empirical research contributions provide empirical evidence on the use of immersive VR games to support addressing societal challenges.
- twelve theoretical and opinion contributions present implications for the design and theory of immersive VR games.

To address our overarching RQ, we break it down into three RQs, each addressing a key societal challenge: (i) *physical activity*, (ii) *active aging*, and (iii) *social interaction*. Through our work (see Part II for our research papers), we demonstrate how immersive VR games can support addressing these challenges and ultimately contribute to the United Nations' goal of "*good health and well-being*".

1.3.1 *Physical Activity*

As the first societal challenge, we focus on *physical activity*. While technology is one of the contributing factors to the lack of physical activity (e.g., easy access to many products online), it can also help address this issue by providing engaging ways to motivate people to move more. Inspired by WHO's recommendation on "*provid[ing] affordable, enjoyable and inclusive opportunities*", we look into immersive VR exercise games—in short exergames [124][1]—that blend fun, immersion, and physical activity together. Therefore, RQ_1 asks:

> **?** **RQ1:** *How can immersive VR games support physical activity?*

As a first step to answer RQ_1, we perform a scoping review. Consequently, we provide an overview of immersive exergame research, develop a taxonomy, establish reporting guidelines, and uncover research gaps [CorePub1]. After providing an in-depth overview of the design and research of immersive exergames, we present the results of three empirical studies. Our first study involves the movement-experts-driven design of a VR exergame targeting lower-body-focused training [CorePub2]. With this research, we show the potential of VR exergames that incorporate personalized feedback and expert-driven design to provide a targeted training for challenging scenarios. Next, we focus more on the feedback aspect and explore how to provide real-time continuous cues through various modalities to help players correctly execute exercises in immersive VR exergames [CorePub3]. As a last step, we target the visionary goals of VR exergames: to do physical exercise and long-term use. We investigate the building blocks of the most successful commercial VR exergame to understand how game elements influence motivation (i.e., inspiration) to do physical exercise and long-term use [CorePre1]. With our contributions (see Figure 1.1), we show that immersive VR games can support physical activity by providing training opportunities, featuring continuous cues that guide players in performing movements, and including game elements that inspire people to engage in physical exercise and sustain long-term use of immersive VR exergames.

1 We note that exergames can be referred to as exertion (e.g., [140, 141]), motion-based (e.g., [10, 191]), or movement-based games (e.g., [12, 128]).

Immersive VR Games for Physical Activity

Immersive VR Exergames	**Born to Run, Programmed to Play: Mapping the Extended Reality Exergames Landscape (CHI'24) [CorePub1]**
Training in Immersive VR Exergames	**Never Skip Leg Day Again: Training the Lower Body with Vertical Jumps in a Virtual Reality Exergame (CHI'23) [CorePub2]**
Continuous Cues in Immersive VR Exergames	**Move, React, Repeat! The Role of Continuous Cues in Immersive Exergames (CHI PLAY'24) [CorePub3]**
Game Elements in Immersive VR Exergames	**Unlocking the Potential: The Role of Game Elements on Player Motivation and Intention of Long-Term Use in Virtual Reality Exergames [PREPRINT] [CorePre1]**

Figure 1.1: Immersive VR Games for Physical Activity research papers. For legibility, the left side lists the main focus, while the right side lists the publications and preprints. We outline our papers following Wobbrock and Kientz [227]'s categorization in the field of HCI research: survey, artifact, empirical, theoretical, and opinion contribution. The icon indicates honorable mentions.

1.3.2 *Active Aging*

Although immersive VR games have become available for older adults (e.g., [29, 60, 143, 236]), we still know very little about designing immersive VR games for this demographic. Nevertheless, given that these games involve cognitive and physical elements, they hold significant potential for promoting *active aging*. Yet, there are several challenges to consider. Among others, this generation has largely grown up in a non-digitalized world [133], may be more hesitant to try new technologies [89], and experiences age-related changes in their abilities [76, 98]. Therefore, there is a clear need to gain an empirical understanding of how immersive VR games can support older adults. Thus, we formulate RQ2 as the following:

> ? **RQ2:** *How can immersive VR games support active aging?*

To answer RQ2, we particularly focus on immersive VR exergames since physical activity is crucial for active aging in older adults [231, 233]. Accordingly, we first focus on designing age-appropriate exercises for older adults. By involving end-users (i.e., older adults with and without dementia) and other stakeholders (e.g., physiotherapist) in the design process, we design and develop an immersive VR exergame featuring cognitive and physical exercises [CorePub4]. While

Immersive VR Games for Active Aging

Cognitive and Physical Exercises for Older Adults	Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults (CHI EA'22) [CorePub4]
Inclusivity for Older Adults	Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia (CHI PLAY'21) [CorePub5]
Augmentation for Older Adults	Evaluating Augmented Locomotion and Range of Reachable Objects for Older Adults in a Virtual Reality Exergame [PREPRINT] [CorePre2]
Exergaming Technologies for Older Adults	ExerCube vs. Virtual Reality: A Comparative Study of Exergame Technologies for Older Adults [PREPRINT] [CorePre3]

Figure 1.2: Immersive VR Games for Active Aging research papers. For legibility, the left side lists the main focus, while the right side lists the publications and preprints. We outline our papers following Wobbrock and Kientz [227]'s categorization in the field of HCI research: survey , artifact , empirical , theoretical , and opinion contribution .

our results indicate promising results, we also find points for improvements (e.g., tutorial), particularly for older adults with dementia. In [CorePub5], we specifically target older adults with dementia. We design and develop an immersive VR exergame that presents realistic scenarios and utilizes real-life-like interaction concepts to offer active aging opportunities for this demographic. After studying realistic scenarios, we explore unrealistic interactions in immersive VR exergames as a potentially effective approach to improve player motivation and physical activity [CorePre2]. As we introduce players to enhanced virtual abilities within VR, in this research, we focus exclusively on older adults without dementia. Our research indicates that, in contrast to younger generations, older adults do not experience a significant increase in motivation due to augmented interactions, and may even decrease their physical activity. Finally, by comparing the two emerging exergame technologies, we reveal the advantages and disadvantages of using immersive (VR-HMD) and non-immersive (i.e., projection-based exergame system, ExerCube) setups for older adults [CorePre3]. With our work answering to RQ2 (see Figure 1.2), we demonstrate how immersive VR games can support active aging by integrating cognitive and physical exercises, incorporating realistic scenarios as well as interactions for older adults with dementia, considering the differences between younger and older adults when designing enhanced virtual abilities, and highlighting the advantages and disadvantages of immersive and non-immersive exergaming technologies for older adults.

1.3.3 *Social Interaction*

As our final societal challenge, we target *social interaction* by following the WHO's recommendation to "*strengthen the benefits*" of digital transformation [155]. Immersive VR games hold great potential to improve player and social experiences [35]. In these games, unlike traditional digital media, players can interact with one another in real-time within a shared VE. However, there are also challenges to face when designing multiplayer immersive VR games. These include high expectations for realism that mirror real-life social scenarios, limited access to hardware among players, and varying levels of existing social ties that can benefit or hinder social interaction. Hence, we pose RQ3 as follows:

> ❓ **RQ3:** *How can immersive VR games support social interaction?*

To address RQ3, we start by investigating the potential of immersive VR in multiplayer gaming scenarios. Accordingly, in [CorePub6], we develop a VR application that enables players to engage in a couch-coop gaming experience virtually, regardless of their physical locations. Our findings, which compare the VR scenario with real-life co-located gameplay, indicate promising results in terms of player and social experiences and suggest technical improvements in VR hardware. Having demonstrated the potential of multiplayer VR games, which naturally involve multiple players, we focus on understanding the differences—or asymmetries—in game design and between players. To understand what asymmetric features have been explored and how asymmetries in multiplayer immersive VR games affect players' experiences, we conduct a systematic literature review [CorePub7]. As a result, we provide an overview and develop a framework to guide the design of multiplayer immersive asymmetric VR games. After providing a theoretical understanding of these games, we conduct two empirical studies targeting research gaps in this area. In [CorePub8], we investigate whether and what types of interdependence between players can lead to comparably similar player experiences in asymmetric VR games where only one player wears a VR-HMD. Our results suggest that primarily strategic interdependence can deliver comparable experiences for players of different mediums. After seeing the potential of asymmetric multiplayer immersive VR games, we question if these games are inherently social or can be affected by existing social ties (i.e., friends vs. strangers) between players [CorePub9]. We reveal that, regardless of familiarity, these games provide high levels of social and player experience, even in VEs. By answering RQ3, we present how

Immersive VR Games for Social Interaction

Social Multiplayer Immersive VR Games	**VR Almost There: Simulating Co-located Multiplayer Experiences in Social Virtual Reality (CHI'23) [CorePub6]**
Asymmetric Multiplayer Immersive VR Games	**A Best-Fit Framework and Systematic Review of Asymmetric Gameplay in Multiplayer Virtual Reality Games (Frontiers in VR'21) [CorePub7]**
Interdependence in Asymmetric Multiplayer Immersive VR Games	**Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games (CHI'21) [CorePub8]**
Familiarity in Asymmetric Multiplayer Immersive VR Games	**Playing with Friends or Strangers? The Effects of Familiarity between Players in an Asymmetric Multiplayer Virtual Reality Game (CHI PLAY Companion'23) [CorePub9]**

Figure 1.3: Immersive VR Games for Social Interaction research papers. For legibility, the left side lists the main focus, while the right side lists the publications and preprints. We outline our papers following Wobbrock and Kientz [227]'s categorization in the field of HCI research: survey, artifact, empirical, theoretical, and opinion contribution.

immersive VR games can support social interaction (see Figure 1.3) by offering an alternative to real-world social gaming scenarios, providing a theoretical framework for understanding asymmetric multiplayer immersive VR games, featuring interdependences between players to overcome the hardware divide, and using asymmetric game design to offer positive experiences, regardless of familiarity between players.

1.4 DISSERTATION OVERVIEW

This dissertation consists of two parts: the synopsis (Part I) and a publication and preprint collection (Part II). The synopsis, which is currently being read, provides a brief overview of the dissertation framework, articulates the motivation behind the dissertation, introduces the RQs and research contributions, provides a guide through the collection of contributions, and offers an overarching discussion and future work considerations. Part II contains the collection of publications and preprints.

The first eight chapters of this dissertation forms the synopsis (Part I). Chapter 1 provides a motivation to explore immersive VR games as a means to overcome societal challenges. The second chapter, Chapter 2, briefly introduces basic terminologies and foundational concepts needed to understand the contributions of the dissertation. Chapter 3 introduces the HCI methods used in the research collection of this dissertation. The central part of the dissertation comprises the following chapters: Chapter 4 summarizes the contributions of this dissertation on immersive VR games for physical activity to answer RQ1. Chapter 5 answers RQ2 and includes a summary of our contributions targeting active aging. Chapter 6 provides an overview of our publications addressing social interaction and thus answers RQ3. Chapter 7 presents a general discussion on the findings of the research collection. We conclude our synopsis with a final chapter (Chapter 8).

The final three chapters are the publication and preprint collection of this dissertation (Part II). We list the published version of each publication and preprint sorted by RQs: immersive VR games for physical activity (Chapter 9), for active aging (Chapter 10), and for social interaction (Chapter 11).

2 THEORETICAL FOUNDATION

This chapter introduces the basic concepts and theoretical foundations of immersive VR games. We limit the scope of this chapter to the foundational work that is necessary to understand and to discuss the research covered in the subsequent chapters and its overarching contribution. For a more detailed description of the previous work of each publication and preprint, we refer to the individual papers of this dissertation (see Chapter 9, 10, and 11).

2.1 IMMERSIVE VR

According to Slater and Sanchez-Vives [188], "*virtual reality is 'reality' that is 'virtual'*" [188]. However, how do we classify the "realities" and define VR? The reality-virtuality continuum (RVC) was introduced by Milgram et al. [130] and classifies realities based on two endpoints defined as the real and VE. According to this classification, while the real environment contains only real-world objects, the VE consists entirely of computer-generated, artificial objects, creating a VR [129]. Milgram and Kishino [129] explains mixed reality (MR) as the blending of real and virtual realities. In RVC, when we move away from the real world towards the VE, we encounter augmented reality (AR), which can be defined as the integration of virtual, artificial objects within the real-world environment. In contrast, augmented virtuality (AV) describes the opposite concept: the integration of real-world objects in the computer-generated VEs. Today, VR-HMDs are capable of delivering both VR and MR (including AR and AV) content. We refer to all these concepts together (excluding only the real environment) under the umbrella term as extended reality (XR) and adhere to these definitions throughout this dissertation (see Figure 2.1). However, we also acknowledge the ongoing discussions regarding the classification of realities [186, 193]. For example, Skarbez et al. [186] revisited the RVC. The authors argue that current immersive VR systems in fact fall under the term of MR since these could only control users' exteroceptive senses (e.g., sight), but not the interoceptive senses (e.g., vestibular), indicating that VR technology is still far away from providing fully immersive experiences.

2.1.1 *Head-Mounted Displays*

In 1965, Sutherland [201] introduced his vision of an "*ultimate display*" that creates a seamless reality within virtuality; it "*can control*

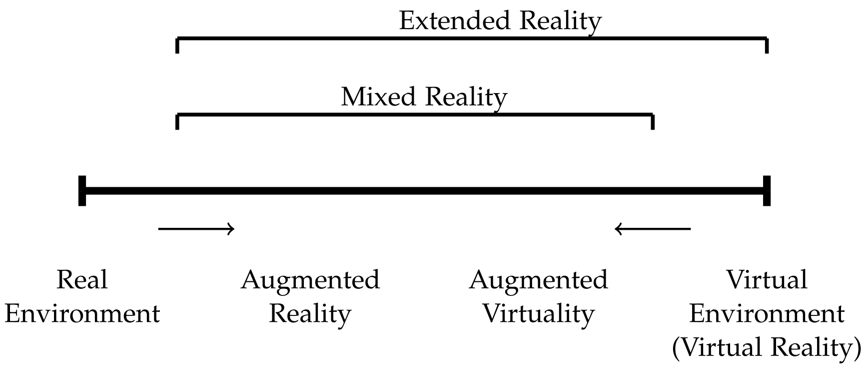

Figure 2.1: The reality-virtuality continuum. The figure is adapted from Milgram et al. [130] and adjusted to include the XR concept.

the existence of matter" and "*with appropriate programming such a display could literally be the Wonderland into which Alice walked*" [201]. A few years later, he presented the first step to his vision [202]: one of the first HMDs. Back then, the HMD was heavy and stationary, however, it laid the foundation for current VR-HMDs [188]: Users experienced a stereoscopic view of a 3D VE (i.e., wire-frame line drawings), which was rendered based on their head movements. Today, while VR hardware development still has a long way to go, now VR systems (e.g., standalone) can be used by people regardless of location and without in-depth technical knowledge, allowing people to freely create and design VEs. Given these developments, the VR industry is experiencing growth [195] and technology is making its way into our daily lives. However, like any technology, it brings challenges. For example, the most common discomfort issues brought by VR-HMDs are related to cybersickness [118], ergonomics, and digital eye strain [91]. According to the sensory conflict theory, cybersickness (typically measured using the simulator sickness questionnaire (SSQ), which we also employ in the research papers of this dissertation [108][1]) typically occurs due to a mismatch between the vestibular and visual senses and includes symptoms such as nausea and dizziness [118]. Ergonomics discomfort captures various symptoms such as neck pain and sweating [91]. Finally, digital eye strain includes several symptoms related to the eyes, such as burning and dry eyes [91]. To ensure a pleasant experience, VR researchers and practitioners should consider the associated challenges with using VR-HMDs.

2.1.2 *Immersion*

One of the main goals of immersive VR technology is to provide realistic perceptions, even though users are aware that they are ex-

1 We note the discussions about the use of the Simulator Sickness Questionnaire (SSQ) to measure cybersickness in VR research, since it was originally developed for measuring symptoms of flight simulator training [18, 91, 108].

periencing only a VE. According to Slater and Sanchez-Vives [188], immersion[2] refers to the technical ability of a VR system to deliver such experiences to provide natural "*sensorimotor contingencies*" [188]. For instance, when we turn our head while wearing a VR-HMD, we expect the VR-HMD to provide an update in our view corresponding to such an action.

Technical immersion is not a binary concept, but a continuum [188]. Slater and Sanchez-Vives [188] considers VR-HMDs more immersive than CAVE Automatic Virtual Environment Displays (CAVEs) (i.e., projection based immersive system) [48]. While a VR-HMD can simulate what a CAVE can do (e.g., the virtual representations of users' physical body), the reverse is not possible. There is also a large body of work considering different factors, such as field of view, head tracking, and high display resolution, to create more immersive VR experiences [188]. Although we are not aware of a taxonomy for classifying technologies according to their immersiveness, researchers generally agree that an immersive VR system should provide a head-dependent stereoscopic view that exceeds the user's field of view without requiring more than head-based input [71, 188]. We therefore follow this concept when we refer to immersive technologies in this dissertation, and consider technologies that do not fit this definition as non-immersive.

2.1.3 *Presence*

Imagine wearing a VR-HMD and being virtually immersed in one of the world's stunning beaches. As the sun sets, the sky comes alive with vibrant shades of red and blue, while birds soar overhead, chirping melodiously. Even though you know you are not physically on the beach and are just lying down on your couch, you still feel as if you are actually there [187].

According to Slater and Sanchez-Vives [188], presence is a subjective experience resulting from immersion and is mostly measured via questionnaires [182] (e.g., Igroup Presence Questionnaire (IPQ) [180], Slater-Usoh-Steed Questionnaire (SUS) [189]). Presence is typically linked to the system's technical capacity, since it can allow users to experience the VE as their real one. Slater [187] classifies two types of illusion of presence induced by immersive VEs: (i) place and (ii) plausibility illusion. The place illusion refers to the feeling of "*being there*" in the VE even though users know that it is a VE. The illusion of plausibility concerns events that are perceived as if they are actually happening [187, 188]. Consequently, the users can react to virtual events in the same way they would react in the real world (e.g., moving to the side to avoid a virtual apple falling from a virtual tree in VR). When we consider scenarios involving more than one person in immersive environments, we can have the illusion of

2 Immersion can be conceptualized differently in other domains (e.g., [117]).

co-presence: "*being there with the others, or virtual togetherness*" [190]. Similarly, Biocca et al. [20] refers to the concept of social presence, "*being with another*" [20], which considers psychological aspects (e.g., psychological involvement) in addition to co-presence. The perceived social presence can be measured using the Networked Minds Measure of Social Presence (NMMSP) [19]. Similar to other VR researchers, in this dissertation, we use the previously introduced questionnaires to assess presence and social presence where appropriate.

2.2 GAMES

Digital games, which can be played on a broad category of devices such as smartphones, computers, and VR-HMDs, are interactive rule-based computer systems. According to Dyck et al. [59], "*computer games are one of the most successful application domains in the history of interactive systems*" [59]. Tekinbas and Zimmerman [207] identifies six core components that define games: (i) system, (ii) players, (iii) artificiality, (iv) conflict, (v) rules, and (vi) quantifiable outcome. We explain this definition with an example using one of the most successful digital games of all times: Pac-Man [146]. In Pac-Man, there are various objects and the relationships between them within an environment that form a *system*: The walls create an *artificial* maze-like environment in which the *players* navigate the Pac-Man. *Rules* are the essence of games that tell players how to play; in Pac-Man, players need to avoid the ghosts, which represents the inherent *conflict*. As a result, players achieve a quantifiable *outcome*, a score based on the number of points they collect.

2.2.1 *Games Research*

Games research has grown to be a well-established subfield within the HCI domain [32]. An increasing number of papers have explored Player-Computer Interaction (PCI) [32]. While some research (e.g., [36]) focused on solely on game design and resulting outcomes (situated in "*ontological paradigm*" [32]), some game researchers used games as a platform for wider investigations. As such, games have played a key role in the advancement of computer science [107, 240]. For example, we saw how AI can approximate human intelligence through games: Deep Blue[3] won against the world chess champion in 1997 [240]. Similarly, games have been used as a platform to reach a comprehensive understanding of technologies or other areas of research, according to the "*epistemological paradigm*" [32]. For instance, researchers used games as a test-bed environment to explore collisions in VEs [39] or to evaluate the embodiment of multiple VR characters [OtherPub8]. In line with this, Dyck et al. [59] also suggested that the lessons learned

3 https://www.ibm.com/history/deep-blue

from game research (e.g., "*deep customizability*") feature of user interface elements [59]) can be applied to other interactive applications in the field of HCI.

Nevertheless, games also have characteristics that make them unique compared to other applications. For example, games are inherently designed for fun [107]. Even though they might target serious purposes (like suggested in "*operative paradigm*" [32]), they always benefit from the entertainment feature. Unlike other applications, games do not always target efficiency or ease of use, but might actually aim to frustrate or emotionalize players [107]. Given the distinct motivational characteristics of games, a large body of work (e.g., [5, 83, 84, 181]) has even used game elements in non-game contexts, a concept known as gamification [55], or motivational affordances [111].

2.2.2 *Playability and Player Experience*

Regardless of their research focus, research in games often aims to enhance playability and player experience. Corresponding to that, Nacke et al. [145] distinguished between the terms *playability* and *player experience*. Playability concerns the technological and design aspects of a game [145, 223] (e.g., consistent reaction from a game [54]). Playability has been specified as a prerequisite for fostering a good, positive player experience [223]. Many scholars have examined the "*psychological, cognitive, and emotional aspects*" [101] that players experience during and after their engagement with games [223], referring to this concept as *player experience*. Given that player experience encompasses a wide range of dimensions, game research evaluates various aspects related to it (e.g., interest / enjoyment [179], curiosity [1], empathy [50]). Similarly, game research uses various methods to measure these aspects [145], such as questionnaires (e.g., Player Experience Inventory (PXI) [1], Social Presence in Gaming Questionnaire (SPGQ) [50]), biometric measurements (e.g., heart rate), and gameplay data (e.g., time events of players' interaction with specific game objects). Like other game researchers, we benefit from all the aforementioned methods and assess various aspects of player experience in this dissertation to address our RQs.

2.3 IMMERSIVE VR GAMES

Immersive VR games are one of the most common applications of VR technology [110]. Since they support multi-modal sensory high-fidelity experiences, these games immerse players in game environments. Unlike traditional digital setups that rely on mouse or keyboard inputs, players can spatially interact with virtual objects through hand tracking. Players can embody a game character and experience the VE from a first-person perspective, creating a viewing experience similar to real life. Beyond realistic interactions, these games allow players to take on the role of fictional characters and experience enhanced virtual abilities (e.g., exaggerated jumping abilities), making players feel more motivated [100]. For example, a player can become a Spider-Man [45] and swing between buildings in a city using web lines to catch an enemy. Immersive VR games also offer "*access to inaccessible activit[ies]*" [75], for instance, by enabling players who use wheel-chairs to ski [75]. Accordingly, immersive VR games are known to provide more positive experiences [160] and enable players to feel more present [161]. However, like all immersive VR applications, these games are prone to the disadvantages of VR technology (e.g., cybersickness [118], limited physical tracking space [199], and usability issues [2]). As a result, immersive VR games also require attention in their design and implementation to ensure enjoyable, purposeful gaming experiences.

2.3.1 *Immersive VR Exergames*

When players complete tasks in games by performing physical activity, we consider these games as exergames. Accordingly, we follow Mueller et al. [141]'s definition: "*game[s] where the outcome [...] is predominantly determined by physical effort*". These applications offer a blend of gaming with exercise [124]. Notable examples of this genre include Dance Dance Revolution [14] and Wii Fit games [149].

Immersive VR technology has introduced distinct benefits to the exergames genre, so-called *immersive VR exergames*: (i) intuitive spatial interaction, (ii) presence, (iii) deep involvement, and (iv) control over VEs. Many VR hardware systems are equipped with built-in six degrees of freedom (6-DOF) tracking, which allows for natural spatial movements in the VEs. Consequently, players in VR exergames can perform intuitive physical movements without the need for artificial controls or additional hardware (e.g., Wii Fit balance board). Players in immersive VR exergames can feel presence in the game, rather than merely controlling it from the outside. Players become characters who play tennis instead of controlling a character on the screen by performing tennis movements. Similarly, the immersive nature of VR can enable players to become deeply involved in the game and distract them from the intensity of the movements being performed [65]. Additionally, immersive VR exergames provide designers with the opportunity to control the VEs, enabling the creation of supportive environments for physical activity (e.g., non-player characters cheering [237, 243]).

However, immersive VR exergames face specific challenges. While these games are designed to promote physical activity and induce exertion, players wear VR-HMDs that block their view of the real world, requiring additional care to ensure the safety of players. Similarly, since these games involve physical movements, their design and consideration of the target user group is crucial to avoid any potential harm. When exergames designed for older adults, the targeted intensity of physical effort may be entirely different from when considering younger adults. Although there are potential challenges with these games, research on immersive VR exergames also reports promising results. For instance, they provide more motivating [22] and performance inducing [235] gameplay compared to non-immersive versions of exergames (e.g., TV-based setup). However, research on immersive VR exergames is still in its infancy and requires empirical investigations to enhance their enjoyability, effectiveness, and inclusiveness.

2.3.2 *Multiplayer Immersive VR Games*

After introducing immersive VR exergames, we focus on another subgroup of games that are part of this dissertation: multiplayer immersive VR games. When more than one person is actively involved in an immersive VR game, we refer to it as a *multiplayer immersive VR game* [70]. These games create interaction, cooperation, or competition between multiple players. According to the ESA report in 2024 [62], multiplayer games play a key role in connecting people; 55% of the US population plays video games with others once a week. Multiplayer games also provide benefits beyond the gaming world, such as supporting well-being [52], forming new friendships [41, 224], and overcoming challenging times (e.g., Coronavirus Disease 2019 (COVID-19)) [38, 72, 131].

The benefits of multiplayer games in general can be extended to multiplayer immersive VR games as well. Unlike traditional digital game setups, in multiplayer immersive VR games, through body and hand tracking, players can virtually interact with each other in real time in VEs. These interactions can take place without being bound to physical location. Previous research shows that multiplayer immersive VR games provide socially rich experiences (e.g., empathy) and offer a higher player experience (e.g., flow) compared to a non-immersive version (i.e., monitor and gamepad controller) [35]. Nevertheless, creating multiplayer immersive VR games can also be challenging. For example, immersive VR could be potentially an isolating experience due to its immersion [21]. Hence, players may focus on VEs rather than the multiplayer gaming experience or social interactions. Alternatively, players may seek a high social experience that they might have in real life while playing multiplayer games (e.g., co-located couch coop games), which may be difficult to fully realize in VR.

2.3.3 *Asymmetric Multiplayer Immersive VR Games*

One of the main challenges of designing and building multiplayer immersive VR games is the involvement of multiple players in the gameplay. Each player is unique; they have different preferences, abilities, and access to hardware. Therefore, creating engaging multiplayer experiences that cater to each player can be quite difficult. We refer to these differences at the player and game design level as *asymmetry* [87]. For example, some people may prefer to take an active role in games (e.g., physical navigation), while others may prefer to take less active roles (e.g., guiding). Similarly, some players may be good at cognitive tasks (e.g., decision-making), while others may be good at tasks that require good reaction time. Alternatively, maybe a grandchild has a VR-HMD device at home that they want to play games on, while a grandparent may only have a mobile phone as a digital device. To account for and leverage all these asymmetries, we can design and build multiplayer immersive VR games, so-called *asymmetric games* [87].

Asymmetric video games have been shown to be more social (e.g., social presence) than the symmetric versions, where players have the same hardware and game mechanics to interact with the game [86]. The investigation of asymmetric games also gained traction in VR research: These investigations mainly targeted the single-owned VR-HMD scenarios where a VR player plays together with a player, who does not wear a VR-HMD [81, 82]. However, only limited immersive VR studies comprehensively engaged with the game design of asymmetric multiplayer games to enable multiplayer engagement of all players involved and enhance the resulting player experience.

3 RESEARCH DESIGNS, METHODS, AND ANALYSES

HCI is an interdisciplinary research field [172]. Therefore, HCI draws from various fields, in particular, computer science, psychology, as well as design, and applies and adapts their methods [119].

Foremost, this dissertation uses methods from computer science. Our research contributions mainly involve the development of interactive software applications, in particular nine immersive VR games. To design and implement these games, we used the multipurpose development engine, Unity[1], tailored towards real-time 3D applications. While this dissertation does not focus on hardware development, it uses several different pieces of hardware, including tethered (e.g., HTC Vive Pro in [CorePub2]) and untethered VR-HMDs (e.g., Meta Quest 2 in [CorePub3]), 3D-printed devices (e.g., a camera controller in [CorePub5]), physiological sensors (e.g., heart rate wristband in [CorePub8]), and the ExerCube (i.e., interactive projection-based display system in [CorePre3]) to enable users to experience the games. For more information on the details of each immersive VR game, we refer to the individual papers (see Chapter 9, 10, and 11).

Additionally, this dissertation utilizes design and psychology methods that may not be familiar to every computer science researcher. Therefore, before presenting the collection of papers, we introduce the research designs, methods, and analyses used in the core publications and preprints of this dissertation. Accordingly, we classify our contributions according to their designs, methods, and analyses to provide a brief overview: see Table 3.1.

3.1 RESEARCH DESIGNS

HCI research typically follows one of three research designs: (i) quantitative, (ii) qualitative, or (iii) mixed-methods. Creswell [46] sees quantitative and qualitative research designs at either end of a continuum, with a mixed-methods design in between. Thus, a study may be more inclined towards qualitative, quantitative, or carrying elements of both— mixed-methods. Accordingly, the collection of publications and preprints in this dissertation uses these designs where appropriate.

3.1.1 *Quantitative Research Design*

Quantitative research design involves hypothesis testing and collection of numerical data [46]. Analyzing collected data usually requires

1 https://unity.com/

Publications & Preprints	Design	Methods	Analyses
VR Games for Physical Activity			
[CorePub1]	Qualitative	literature review	Quantitative & Qualitative
[CorePub2]	Mixed	controlled experiments, interviews	Quantitative & Qualitative
[CorePub3]	Quantitative	controlled experiments	Quantitative
[CorePre1]	Mixed	interviews, surveys	Quantitative & Qualitative
VR Games for Active Aging			
[CorePub4]	Qualitative	prototype testing, interviews	Qualitative
[CorePub5]	Qualitative	prototype testing, interviews, focus group	Qualitative
[CorePre2]	Mixed	prototype testing, controlled experiments, interviews	Quantitative & Qualitative
[CorePre3]	Mixed	prototype testing, controlled experiments, interviews	Quantitative & Qualitative
VR Games for Social Interaction			
[CorePub6]	Mixed	controlled experiments, interviews	Quantitative & Qualitative
[CorePub7]	Qualitative	literature review	Quantitative & Qualitative
[CorePub8]	Mixed	controlled experiments, interviews	Quantitative & Qualitative
[CorePub9]	Mixed	controlled experiments	Quantitative & Qualitative

Table 3.1: This table provides an overview of the research designs, methods, and analyses used in the contributions of this dissertation. The methods were illustrated using icons: literature review, prototype testing, interviews, focus group, surveys, and controlled experiments.

statistical tests [46]. Typically, quantitative results are reported in text, tables are used to provide descriptive values, and visualization of results is supported by figures. In our publications, we consider only [CorePub3] as an example of using this research design, since it primarily contains numerical data, while also providing brief insights into the responses of participants to an optional open-ended question.

3.1.2 *Qualitative Research Design*

The data collected in this type of research includes non-numerical data and can be in various forms, such as video and text [96]. The collected data are analyzed through qualitative analysis methods such as thematic [37] and content analysis [222]. In this dissertation, four publications [CorePub1, CorePub4, CorePub5, CorePub7] use this research design. These papers primarily include non-numerical data as the unit of data analysis, such as textual data from previously published articles or participants' audio-recorded answers to questions.

3.1.3 *Mixed-methods Research Design*

In mixed-methods research design, researchers combine both quantitative and qualitative methods. They benefit from both the generalizability of quantitative findings and the in-depth understanding provided by qualitative approaches [46, 47]. Therefore, the most of the publications [CorePub2, CorePub6, CorePub8, CorePub9] and preprints [CorePre1, CorePre2, CorePre3] (seven research papers out of 12) of this dissertation uses mixed-methods research design.

3.2 RESEARCH METHODS

This section presents the employed HCI research methods in the publications and preprints of this dissertation. We note that an article may contain multiple methods, as methodological triangulation contributes to a comprehensive understanding and increases validity [102].

3.2.1 *Literature Reviews*

Literature reviews () are sources of cumulative information. They synthesize and summarize existing knowledge and provide evidence about the current state of the chosen topic [43]. Beyond the summary, such research can also offer new insights [210], such as guidelines (e.g., [CorePub1]), frameworks (e.g., [CorePub7]), and taxonomies (e.g., [CorePub1], [OtherPub3]). With the increase in research, literature reviews have become essential. Literature reviews provide an overview of a particular topic, showcase conflicting results, and determine the quality of evidence [42, 142]. The increase in research publications also applies to the field of HCI [OtherPub14]. In parallel, we also see an increase in the number of published HCI literature reviews [OtherPub14]. This richness is also reflected in the types of conducted literature reviews [77, 203].

This dissertation includes two types of literature reviews: (i) a systematic [CorePub7] and (ii) a scoping review [CorePub1]. Systematic reviews synthesize the findings of a specific, narrow research question [203], while scoping reviews focus on providing a broad overview of the current state of research on a particular topic [44, 109, 142]. Thus, while systematic reviews typically apply a critical appraisal step to assess the quality of evidence in primary sources, scoping reviews do not require this step [44]. There are established guidelines and recommendations [134, 159, 166, 184, 211] for conducting reviews, which are also used in HCI research [4, 24, 92], [OtherPub14], such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [159] and Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) [211]. These guidelines consist of a set of items that help improve the reporting of reviews at various stages, such as stating objectives and explaining search strategies [159, 211]. Accordingly, in our systematic review contribution [CorePub7], we adhere to the recommendations PRISMA, whereas in our scoping review [CorePub1] we follow the PRISMA-ScR.

3.2.2 *Prototype Testing*

This dissertation presents nine immersive VR games. All the designed and developed games (i.e., interactive systems [99]) were tested in pilot tests to ensure playability (e.g., clarity of tutorials and game reactions to an event [CorePub9]) and parameter selection (e.g., to smooth movement data [CorePub3]). However, some of these immersive VR games have undergone a more extensive design process and prototype testing (⚙) than others.

To ensure a thorough design process and prototype testing, we used one of the common design approaches in HCI (e.g., [138]): Human-Centered Design (HCD) [99]. Norman [151] was the first researcher to coin the term HCD and pointed out that "*Good design requires good communication, especially from machine to person, indicating what actions are possible, what is happening, and what is about to happen*". While in the early days this approach was known as user-centered design, today it is called HCD to emphasize the involvement of not only end-users but also other stakeholders [99]. This design approach puts users at the center and aims to develop interactive systems that meet their needs and requirements [15, 99, 151]. The International Organization for Standardization (ISO) [99] outlines four iterative steps for HCD: (i) understanding and specifying the context of use, (ii) identifying user requirements, (iii) creating design solutions to meet those requirements, and (iv) evaluating the design. While this approach is costly in terms of time and resources [15], according to the ISO [99], it "*enhances effectiveness and efficiency, improves human well-being, user satisfaction, accessibility and sustainability; and counteracts possible adverse effects of use on human health, safety and performance*".

In our publications [CorePub4, CorePub5] and preprints [CorePre2, CorePre3] in which we target active aging, our immersive VR games were designed and developed using HCD to ensure the well-being of older adults, offer age-appropriate activities, and mitigate the potential adverse effects of VR. Therefore, we involved both end-users and other stakeholders (e.g., physiotherapists) in the design process to gather feedback and observe their interaction with our immersive VR games, and iteratively improved these prototypes.

3.2.3 *Interviews*

Interviews ( ) allow researchers to elicit in-depth responses from participants [119]. In general, interviews can be divided into three groups: (i) structured, (ii) semi-structured, and (iii) unstructured interviews [96, 119]. While structured interviews follow a strict structure and predefined questions, unstructured interviews do not contain pre-prepared questions. Accordingly, semi-structured interviews include predefined questions, but researchers can still ask follow-up questions that may be necessary to gain deeper understanding or clarification from participants' responses [119]. Although interviews allow researchers to interact with their participants, they are costly in terms of time, difficult to conduct, and challenging to analyze [119]. Therefore, many interview studies include only a limited number of participants [119]. HCI research widely uses this method of data collection to capture participants' views, attitudes, and experiences (e.g., [173, 183]). In our research, when we used an interview, we conducted semi-structured interviews, giving us the flexibility to ask follow-up questions (e.g., [CorePre3]) and direct questions specific to the expertise of the participants (e.g., [CorePub2]). Further, in addition to semi-structured interviews, we also used a special interview method called *contextual inquiry* [16] (in [CorePub5]) to obtain user requirements. Although this method can be seen as a form of interview, it also includes distinct features, such as the observation of users in a natural setting, which allows researchers to observe and understand users' needs considering context.

3.2.4 *Focus Group*

Focus group ( ) is a widespread research method used by HCI researchers (e.g., [17, 225, 238]). It involves a moderator who interviews a group of users simultaneously [119, 136]. This method facilitates group discussion and allows researchers to see similarities and differences between participants' views on a particular system [136]. Therefore, it provides an overview of shared opinions and perspectives, but not about every specific individual [136]. They are considered to be more time efficient data collection method compared to conducting individual interviews [119]. However, as with any data collection method, focus groups have some drawbacks. For example, they usually contain fewer questions than interviews to account for the time constraints of interviewing multiple people [119]. The group composition also plays an important role in this method. For example, researchers can see different characters in the group, who are uncomfortable expressing their opinions or who express their opinions excessively [119]. Typically, focus groups are a method of choice for researchers, especially when researchers want to "*understand the diversity*" and "*complex be-*

havior and motivations" [136]. Therefore, in [CorePub5], we used this method to study the interaction of older adults with dementia and their caregivers with our immersive VR game.

3.2.5 *Surveys*

Surveys (📊) are one of the data collection methods commonly used by HCI researchers (e.g., [38, 178, 244]). In surveys, respondents answer a series of questions [158]. They are a very flexible method of data collection where researchers can collect information about "*users' satisfaction, opinions, ideas, and evaluations regarding a system*" [158]. Using this method, researchers can easily reach a large number of participants [119, 158]. According to Lazar et al. [119], surveys provide "*statistically accurate estimates for a population, when structured using random sampling*". Nowadays, many HCI research publications use online recruitment platforms for their surveys. Among numerous online platforms, Prolific[2] has been found to provide high-quality data [58, 64] and was used by many HCI researchers (e.g., [91, 113, 165]). Nevertheless, surveys also have disadvantages. For example, respondents may misinterpret the questions, the received responses may not be detailed, or respondents may not put much attention to answering the questions [119, 158]. Therefore, surveys require clearly articulated questions, a good structure (e.g., including instructions), and well-designed controls [119], such as attention checks [153].

3.2.6 *Controlled Experiments*

Controlled experiments (🔬) allow researchers to establish causal relationships. Researchers using this method manipulate / control variables (i.e., independent variables) and evaluate their effects on outcomes (i.e., dependent variables). This method is widely used in HCI research. In general, for an experiment, after deciding on the research question and hypothesis, three elements are necessary: (i) conditions, (ii) units and (iii) assignment [119]. In HCI, conditions typically refer to the variety of techniques, software applications, and hardware solutions, while units refer to the participants that researchers want to include in the study considering the inclusion criteria [119]. Finally, assignment explains how participants are assigned to specific conditions (e.g., randomization, balanced Latin square [25]) [119]. Many HCI studies conduct controlled experiments that include users completing a task (e.g., playing a game) accompanied by automatically recorded data about their performance (e.g., movement data) and filling out questionnaires about their experience.

2 https://www.prolific.com/

3.3 RESEARCH ANALYSES

This dissertation used both quantitative and qualitative data analysis methods [96]. Here, we only give a brief overview; for more details, we refer to the individual publications and preprints.

3.3.1 *Quantitative Strategies*

To analyze the quantitative data, we used two types of statistical analysis methods: (i) descriptive and (ii) inferential statistics. We would like to point out that even though the data collected is qualitative, researchers can still use some quantification methods [96]. For example, when researchers want to assess users' prior technology experience, participants may answer a yes / no question asking about their prior technology use. However, the researcher may want to quantify this qualitative yes / no answer (i.e., categorical response) and report its frequencies. This can be particularly useful in some situations where, for example, researchers want to show a trend. Thus, the use of qualitative research data collection methods does not necessarily mean that the researcher will follow a qualitative analysis strategy.

DESCRIPTIVE STATISTICS Descriptive statistics provide a summary and overview of the characteristics of the data [96]. Such statistics can be reported based on two components: (i) measures of central tendency and (ii) spread [119]. The former provides information about the central point characteristics of data using values such as mean, median, and mode, while the latter provides an overview of the distribution of the data using values such as interquartile range and standard deviation. As most of the publications and preprints of this dissertation have a mixed-methods research design, they include quantitative data. Even the qualitative research design papers of this dissertation, which involve users [CorePub4, CorePub5], contain quantitative demographic information about participants, such as average age. Additionally, although the literature review publications [CorePub1, CorePub7], have qualitative data, at several points (e.g., frequency reporting of immersive exergames' goals [CorePub1]) we choose to quantify this data. Therefore, all papers of this dissertation are supported by descriptive statistics reports.

INFERENTIAL STATISTICS Inferential statistics allow performing hypothesis testing and to draw generalizable conclusions: "*they help us to confirm or reject our predictions*" [66]. In broad terms, inferential statistics can be divided into two categories: (i) comparison-based and (ii) relationship-based analysis [66, 80, 119]. Comparison-based analysis focuses on differences; it relies on comparing means (if parametric[3]) or medians (if non-parametric) across conditions to detect statistically significant differences [66, 80]. In HCI, for parametric cases, the most commonly used examples of these tests are t-tests and analysis of variances (ANOVAs), while for non-parametric cases, non-parametric versions of t-tests (e.g., Wilcoxon-signed rank test) and ANOVAs (e.g., Friedman's ANOVA), and alternative methods to transform the data (e.g., aligned rank transform (ART) [228]) are preferred [119]. T-tests and their non-parametric versions are used to compare two conditions, while ANOVAs and their alternatives are typically used to compare three or more conditions. Accordingly, using these methods of analysis, we can determine whether the differences between the conditions indicate a statistically significant difference[4] [80]. Relationship-based analysis aims to determine the relationship / association between variables [66, 80, 119]. Common methods used in HCI for this purpose include correlation and regression analysis [119]. While correlation analysis shows the relationship between variables (e.g., if one increases, so does the other), regression analysis goes one step beyond and allows researchers to predict one variable based on another [80, 119]. In this dissertation, we use these inferential tests where appropriate.

3.3.2 *Qualitative Strategies*

Qualitative analysis methods aim to summarize and find patterns in the data. Typically, analyzing this type of data (e.g., videos, interviews, observations) is complex because there may be no clear answers or each researcher may understand the data differently [119]. However, at the same time, qualitative analyses can provide rich data and in-depth understanding of the research question under investigation.

THEMATIC ANALYSIS Thematic analysis is one of the methods commonly used, both in HCI [24] and in qualitative research in general [26]. It enables researchers to summarize patterns in data and "*involves systematic processes of data coding to develop themes*" [37]. Coding refers to descriptive units for labeling data, while themes are created by categorizing and grouping these codes that have a common meaning [37].

In general, thematic analysis has three main orientations: (i) coding reliability, (ii) codebook, and (iii) reflexive approach [27]. Coding relia-

3 When determining whether to use parametric or non-parametric tests, it is necessary to check certain assumptions, including the normality of the data (e.g., by using the Shapiro-Wilk tests) [66]. For more in-depth information on this subject, we refer to Field et al. [66] and statistical analysis resources of Wobbrock [226].

4 To determine a statistically significant difference, significance testing procedures employ a pre-determined probability threshold approach (i.e., significance level) [119]. The probability value obtained after conducting a significance test (i.e., p value) that is lower than the significance level (0.05 in HCI [119]) is considered an indicator of a statistically significant difference.

bility approaches have more quantitative roots, structured coding set, prior theme development, require multiple coders, and strive for agreement between researchers' coding [27]. Codebook approaches involve prior theme development and multiple coders, but having multiple coders is for practical reasons (e.g., reducing data analysis time), not for agreement to demonstrate the validity of the applied codes [27]. However, reflexive thematic analysis follows the qualitative research paradigm [37]. It clearly acknowledges researchers' subjective role in analysis, does not require multiple coders, and recommends against reporting frequencies for themes. According to Clarke and Braun [37], researchers' perspectives, research background, and experiences can affect analysis of data. Therefore, the authors strongly emphasize the importance of reflecting on personal, functional, and disciplinary reflexivity to inform research. Overall, the process of reflexive thematic analysis consists of six main stages: (i) familiarization with data, (ii) coding data, (iii) generating initial themes, (iv) developing and revising themes, (v) refining themes, and (vi) reporting [37]. It is a flexible method that requires a lot of analytical work for researchers.

In the publications and preprints of this dissertation, the choice of thematic analysis orientation was based on the specific scope and RQs of each study. Accordingly, we either employed elements of reflexive thematic analysis, such as using a single coder and maintaining a flexible approach (e.g., in [CorePub5]), or used a hybrid approach that incorporated elements of both reflexive and codebook orientations, which involved collaboration among multiple coders (e.g., in [CorePub6]). Following the recommendations of Clarke and Braun [37], each publication included a reflexivity statement to provide information about how the researchers' perspectives, research backgrounds, and experiences might shape the data analysis.

BEST-FIT FRAMEWORK SYNTHESIS The method of framework analysis (or framework synthesis when used for a literature review [57]) has its roots in health science and is often used for policymaking [31, 57, 167]. It allows researchers to approach their predefined goals and objectives in a highly structured way [167]. In general, framework synthesis consists of five steps: (i) becoming familiar with the raw data, (ii) developing a priori framework, (iii) applying the framework to the data, (iv) summarizing and abstracting the data within the framework overview, (v) mapping and interpreting the data within the framework. Best-fit framework synthesis is a specialized version of framework synthesis [31, 57]. As a theory-driven approach, this method starts with the development of a priori framework based on an existing framework or a combination of frameworks, rather than deriving your own preliminary framework [30, 31]. Researchers then code the data using this priori framework. When the priori framework is not sufficient enough to cover all aspects of the data, researchers use inductive methods to extend the framework, resulting in a post-hoc framework. While the approach is pragmatic, it is also only applicable when there is existing theoretical work related to the research question under investigation. In our systematic literature review paper [CorePub7], where we created a post-hoc framework for asymmetric multiplayer immersive VR games, we used this method.

TAXONOMY DEVELOPMENT METHOD Taxonomies, "*systems of groupings that are derived conceptually or empirically*" [148], are useful tools for classifying and organizing knowledge. A taxonomy should be useful [148]. Accordingly, it should be (i) concise, (ii) robust, (iii) comprehensive, (iv) extendible, and (v) explanatory [148].

To develop a taxonomy, Nickerson et al. [148] proposes an iterative method. It starts with the identification of meta-characteristics: "*the most comprehensive characteristic that will serve as the basis for the choice of characteristics in the taxonomy*" [148]. The choice of meta-characteristics depends on the purpose of the taxonomy and is the most important decision. For example, if the goal of a taxonomy is to identify how users interact with mobile applications at a high level, the meta-characteristic is "*the high-level interaction between the user and the application*" [148]. This leads to the identification of various dimensions along with specific characteristics within each dimension (e.g., temporal dimension: synchronous and asynchronous characteristics) [148].

Once researchers have decided on their meta-characteristics, they need to consider when to finalize the taxonomy development. To this end, Nickerson et al. [148] proposes multiple ending conditions (e.g., "*all objects or a representative sample of objects have been examined*" [148]). Following the decision on meta-characteristics and ending conditions, researchers can decide on their approach: (i) empirical-to-conceptual and (ii) conceptual-to-empirical. The empirical-to-conceptual approach involves using the data set to group objects, while the conceptual-to-empirical approach requires using pre-existing knowledge to group objects. By applying these approaches, researchers create dimensions and characteristics. Every dimension should be mutually exclusive (i.e., "*no object can have two different characteristics in a dimension*" [148]) and collectively exhaustive (i.e., "*each object must have one of the characteristics in a dimension*" [148]). This process continues until the decided ending conditions are met.

In [CorePub1], we followed Nickerson et al. [148]'s approach. For example, we decided on meta-characteristics (goals, people, exercises, design, technologies (GPEDT), inspired by people, activities, contexts, technologies (PACT) framework [15]) and used both conceptual-to-empirical and empirical-to-conceptual approaches in the different steps of our taxonomy creation. As our ending condition, we used the condition "*all objects or a representative sample of objects have been examined*" [148]. However, we also adapted some aspects of the taxonomy development method; we did not adhere to the aspects of mutual exclusivity and collective exhaustivity (similar to [OtherPub3]). We did this because, on the one hand, following these aspects might lead to losing some details in immersive VR exergame design, and on the other hand, not every publication or game description fully describes its games. For example, immersive VR exergames may be designed for more than one target group (i.e., not mutually exclusive).

4 IMMERSIVE VR GAMES FOR PHYSICAL ACTIVITY

Physical activity is essential for our health. Among others, it decreases the risk of severe diseases (e.g., cardiovascular conditions and cancer) and mortality [233]. In addition, it also supports mental well-being [33, 233]. Despite such significant benefits, 1.8 billion adults are not sufficiently physically active [200, 233]—we are facing a sedentary pandemic. This trend has various reasons [137, 163]. For example, digitalization might contribute to physical inactivity, since many of us work at a computer with limited movement during the day. Other reasons include lacking motivation or limited access to exercise opportunities due to financial or structural limitations. Immersive VR exergames could be a solution to overcome these challenges by "*provid[ing] affordable, enjoyable and inclusive opportunities for all people to be active*" [232]. To explore this potential, this chapter focuses on RQ1:

> **? RQ1:** *How can immersive VR games support physical activity?*

4.1 IMMERSIVE VR EXERGAMES

Immersive VR exergames are playful exercise opportunities. In academic research, immersive VR exergames are a relatively new domain. However, it is rapidly growing, with more and more immersive VR exergames being produced and empirical research being conducted. Despite this, a comprehensive overview of the field, including the investigated research questions and produced artifacts, is lacking. This makes it challenging for researchers and practitioners alike to see current research trends and identify promising future research directions. To address this gap, we created an up-to-date overview of the immersive VR exergame research landscape, published in [CorePub1]. Therefore, we conducted a scoping review [142] by following the recommendations of PRISMA-ScR [166, 211] (see supplementary materials in Section 9.1 for the review protocol and additional resources (e.g., extracted data)). To gain a holistic understanding, we also included other types of immersive technology, which fall under the umbrella term of XR; however, almost all identified XR exergames (94.87%) relied only on VR technology. To analyze this field through an HCI lens, we used a foundational framework of HCI: PACT [15]. This framework considers people, activities, contexts, and technologies when designing interactive systems from an HCD perspective. We extended this framework to align with the unique characteristics of exergames,

Technologies					
Display					
VR				AR	
Mobile HMD	PC-tethered HMD	Console-tethered HMD	Projector-based	Mobile HMD	PC-tethered HMD

Figure 4.1: Our derived taxonomy's Technologies subdimension: Display – with a total of two characteristics and six sub-characteristics [CorePub1].

which resulted in the five principles: goals, people, exercises, design, technologies (GPEDT). Thus, our research asked the following RQs[1]:

For more information on [CorePub1], please refer to Section 9.1 or scan the QR code.

?

RQ1.1: **Goals**: What are the goals of the XR exergames?
RQ1.2: **People**: Which different user groups are usually targeted by XR exergames?
RQ1.3: **Exercises**: What kinds of exercises are being designed in the XR exergames?
RQ1.4: **Design**: What kinds of game design aspects are considered for XR exergames?
RQ1.5: **Technologies**: What kinds of technologies are being used in the XR exergames?

We started our scoping review by gathering 1318 potentially-relevant papers from various digital libraries: Association for Computing Machinery (ACM)[2], Scopus (indexing IEEE Xplore and many other HCI-related sources)[3], and PubMed[4]. After completing the screening and eligibility steps, we ended up with 186 articles from which we extracted and analyzed the full-text data focusing on both study specifics and exergame designs. As a first contribution, we provided a descriptive analysis of various metrics describing the current state of XR exergame research. For example, the majority of user studies (72.31%) were conducted in Europe and North America, which illustrates the socio-cultural bias in research. Also, 79.11% of publications featured only single-time testing of the developed exergames, limiting our knowledge of prolonged usage.

As a second contribution, we derived a taxonomy (using Nickerson et al. [148]'s method, see Chapter 3) of XR exergames using GPEDT aspects to guide the design, implementation, and research of future XR exergames. As our dataset, we identified all 195 unique XR exergames

1 The research questions of this chapter have been revised to ensure stylistic consistency in this dissertation while preserving their meaning.
2 https://dl.acm.org/
3 see the Scopus indexed venues: https://www.scopus.com/sources
4 https://pubmed.ncbi.nlm.nih.gov/

used in the analyzed papers (20 custom and 175 self-implemented games). We derived dimensions corresponding to each aspect of GPEDT (see example Figure 4.1, but to see the full taxonomy, please refer to the paper's supplementary materials):

- Goals (i.e., objective of exergame).
- People (i.e., users) includes two *subdimensions:* target age range and condition.
- Exercises (i.e., performed movements) features two *subdimensions:* body part and position.
- Design (i.e., information on game design) has four *subdimensions:* themes, player mode, task type, and adaptation.
- Technologies (i.e., hardware and devices) consist of three *subdimensions:* display, setup, and support hardware.

Similarly to the reporting of the study details, we then also provided frequency and qualitative summaries for each of our taxonomy dimensions. For example, VR is the dominant technology in the immersive XR realm. Out of 195 games, only 10 use AR display technologies. The dominant goal of exergames ($n = 57$) to promote physical activity. Only 51 XR exergames enclose their target user group, mostly targeting older adults ($n = 31$). Our research further provided nine research directions based on our results. Among others, we emphasized the importance of age-targeted game designs and evaluations, referring to a surprising but critical finding: While most games considered older adults as a target group, our descriptive statistics revealed that only 24.10% of studies recruited participants with an average age of above 40. This indicates a mismatch between design and evaluation of XR exergames. Similarly, we only found very limited explorations of the long-term effects of XR exergames, leaving gaps about what and how XR exergames can afford adherence. As immersive exergames are a relatively young field, we also found several inconsistent and inadequate reporting styles in the scoping review articles. For instance, the majority (73.85%) did not report their games' target audience and there was a lack of clarity in how the movements were performed (e.g., what do dodging movements precisely include? [157]), how these games were played (e.g., sitting, standing), or even which display[5] was used. As a first step towards a systematic reporting, we concluded our paper with 15 guiding questions for researchers to follow when reporting the details of their XR exergame, such as: "*How do players precisely perform the featured movements in the XR exergame?*"

4.2 TRAINING IN IMMERSIVE VR EXERGAMES

While our review paper [CorePub1] provided a theoretical overview of XR exergame research and helped uncovering promising research gaps,

5 To collect information about the playing position and display information, we had to consult figures, gameplay videos, and other supplementary materials of the articles.

our remaining papers in this chapter focused on conducting empirical investigations in three areas: (i) jump training, (ii) continuous cueing, and (iii) game elements. In these contributions, we specifically used VR as the immersive technology, as our review revealed its popularity and research value. Among others, our review findings pointed out that many exergames included lower body movements, but these were often limited to walking and squatting. More dynamic, explosive movements, which require rapid acceleration and maximal force generation in a short period of time (e.g., jumping) [164, 217], were rarely part of these games, which is not surprising. Since VR-HMDs block the users' view, performing such exercises while wearing this equipment can bring up some safety concerns. Even more, explosive movements can be technically difficult to detect. However, they are part of many sports (e.g., basketball, volleyball), and training routines (e.g., high-intensity interval training (HIIT)).

While some works included jumping as a feature in their immersive VR [100] and non-immersive exergames [120], none focused on providing vertical jumping training. Although jumping improves personal fitness and supports daily activities, it is a challenging movement with a high risk of injury. Therefore, in [CorePub2], we explored the design process towards a safe and effective VR-exergame-based jump training. Hence, we consulted sports and medical field experts ($N = 9$) from different domains, such as sports research, physiotherapy, and training. Our thematic analysis of the semi-structured interviews revealed four main points. First, jumps are typically part of larger movement patterns (e.g., blocks in volleyball involve stepping with a jump) and are beneficial for general health (e.g., support coordination and injury prevention). Secondly, while jump training includes elements of both reactive and strength exercises, incorrect execution of the jump (e.g., medial collapse of the knees on landing) can cause injuries. Thirdly, experts emphasized that jump training should provide an easy start and be adaptable to serve differences in individuals' fitness and improvement. Finally, the experts highlighted the potential and challenges of providing jump training with VR from their own specialized perspectives. For example, while the possibility of providing real-time feedback or replaying one's performance using an avatar was praised, concerns about safety were raised (e.g., cybersickness caused by the mismatch between virtual and physical jumping).

Based on these insights from experts, we implemented an immersive VR exergame for vertical jump training: *JumpExTra VR*. In the game, players were equipped with three Vive trackers attached to players' shins and waist, a Vive Pro headset, and two controllers. In total, we used six tracking points to animate the virtual avatar of the player using inverse kinematics. *JumpExTra VR* included a total of four levels, three of which trained the required foundations for jumping (i.e., mind-body coordination, balance, endurance) with rhythmic levels

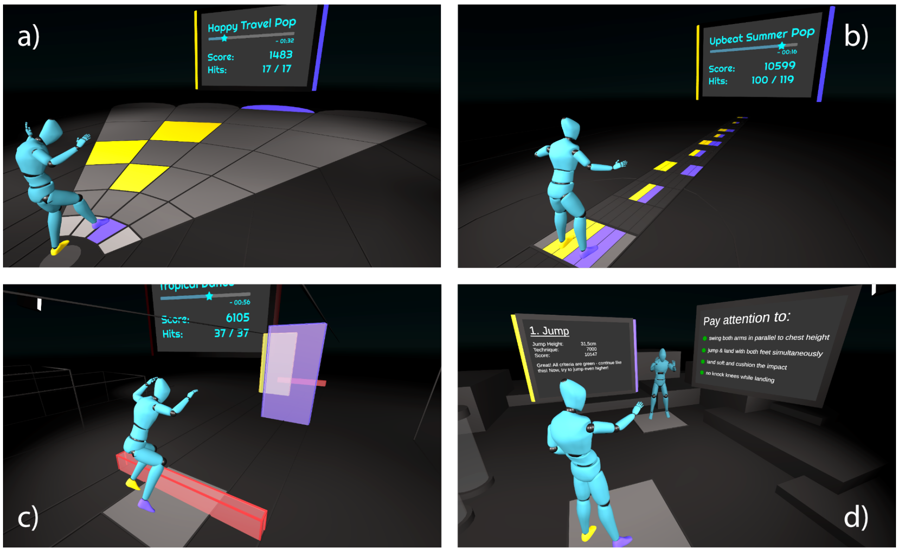

Figure 4.2: The impressions from *JumpExTra VR*: a) A player performs tapping movements in the first level. b) In the second level, the player performs small hops. c) The third level features obstacles which the player needs to move sideways and perform jumps to avoid. d) In the last level, the player performs maximal jumps and receives personalized feedback [CorePub2].

(Tap to the Beat, Hopscotch, Obstacle Course, see Figure 4.2). All these rhythmic levels had dynamic difficulty adaptation in line with the experts' recommendations and prior work [124]. The final level then build on these results and focused on improving the technique when performing vertical jumps. Our game gradually progressed from small tapping movements to performing maximal jumps (see Figure 4.2). In the first level (Tap to the Beat), players performed tapping movements on yellow and purple tiles that approached them. Similar to the first level, the second level (Hopscotch) featured yellow and purple tiles, but this time players performed small hops on them. In the third level (Obstacle Course), players avoided obstacles by moving sideways and jumping over them. Finally, after training the requirements for jumping, players performed five maximal jumps in the last level. They saw their jump replay with an avatar, received personalized feedback on their jump technique after each jump based on four points: (i) jump and land with both feet simultaneously, (ii) land softly (forefeet touch the ground first) and absorb the impact with the entire body, (iii) especially while landing, keep the knees in one line between feet and hips and do not cave them inward, and (iv) swing arms synchronously in a forward-upward arc until about chest height.

After we completed the game implementation, we conducted a study with 25 participants to evaluate players' experience of this novel training exergame. Accordingly, we asked the following RQs:

For more information on [CorePub2], please refer to Section 9.2 or scan the QR code.

?

RQ1.6: How does our immersive VR exergame affect cybersickness symptoms of players?
RQ1.7: How does our immersive VR exergame affect the players' mood?
RQ1.8: How do players evaluate the player experience of our immersive VR exergame and its usability as a training tool?
RQ1.9: How do players perceive the physical activity and perform in our immersive VR exergame?

Our findings revealed that players enjoyed the *JumpExTra VR* game. Player experience constructs and enjoyment of the physical activity were rated generally high. However, they experienced higher cybersickness symptoms (nausea subcategory) after playing the game compared to baseline measures. Nevertheless, our additional analysis show that this significance could be due to the sweating item of the used questionnaire [93, 108], which could be seen as a consequence of exercising. While participants did not improve their jump height, they slightly improved their jumping technique over the five jumps (from 71.45% (*SD*=17.18) to 75.45% (*SD*=15.30)). Our qualitative data also supported the quantitative findings (e.g., enjoyment) while also providing nuanced insights. For example, players reported the need for improving the playzone (i.e., to prevent players from leaving it unintentionally) and the fear of obtaining injuries while playing (e.g., by loosing balance). Based on our findings, we concluded with a set of guidelines (e.g., avoid the unintended forward drift) to aid the design, development, and research of VR training games.

4.3 CONTINUOUS CUES IN IMMERSIVE VR EXERGAMES

In the last level of our jump training game [CorePub2], players received feedback about their technique after each jump, matching our experts' input regarding the benefits of VR for real-time feedback. However, the question remains: How can we best provide real-time performance indicators and guide players to perform the movements correctly in such scenarios? In their movement-based game guidelines, Mueller and Isbister [139] recommends providing moment-to-moment feedback on the movements being performed. At the same time, they point out the importance of avoiding excessive feedback that could overwhelm players and prevent them from reflecting on their own performance. Although a few studies explored the role of feedback in immersive VR exergames [112, 185], research into continuous cues, i.e., real-time indications of how the movements are being performed, is limited. In particular, to our knowledge, no research has examined how to design these cues, what modality to use to represent them. Ultimately, we also do not know how users perceive them (e.g., helpful

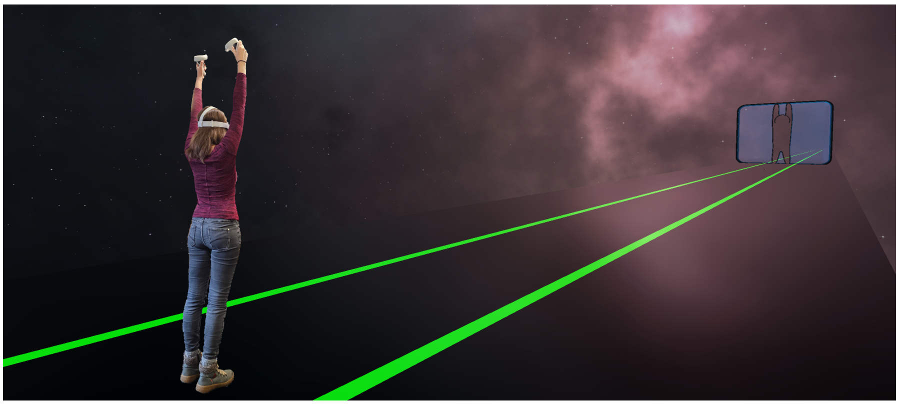

Figure 4.3: Impressions from our immersive VR game that features static poses: A player performs the featured pose on the approaching in-game wall and receives a visual cue (green lines) indicating that the execution is successful [CorePub3].

vs. distracting). Hence, in [CorePub3], we focused on addressing this research gap. We examined the effects of providing continuous cues on player experience and performance, answering the following RQs:

?

RQ1.10: How does having continuous cues affect player experience and performance in immersive VR exergames?
RQ1.11: How do different continuous cue modalities affect player experience and performance in immersive VR exergames?

For more information on [CorePub3], please refer to Section 9.3 or scan the QR code.

To address RQ$_{1.10}$ and RQ$_{1.11}$, we designed and implemented an immersive VR exergame. In our game, players see approaching 3D cut-out walls and perform repeatedly static poses to fit through these walls (see Figure 4.3). The movements displayed on these walls were inspired by Tai Chi and yoga poses [123, 147, 152, 218], such as squat and mountain pose. Our work focused on visual and audio cues (designed based on Ariza et al. [9]), since the majority of games could integrate these modalities of cues into their experience. The auditory cue was a repeatedly playing sound whose frequency gradually changes according to the players' match to the given pose (0% matching: 0.87 Hz—every 1.15 seconds vs. 100% matching: 6.66 Hz—every 0.15 seconds). The visual cue was designed as two lines on the platform path (see Figure 4.3). If the players' pose matched the approaching wall's pose perfectly, they saw green color on these lines (RGB: 0,255,0). Conversely, 0% success corresponded to a red color (RGB: 255,0,0). Depending on the pose match, the color of the lines was linearly interpolated between red and green. To determine the movement and manipulation parameters, we conducted self-play tests and a pilot study[6].

6 To smooth the movement data, we used the easeOutSine function (`https://easings.net/`), which was decided based on a pilot study ($N = 5$).

Figure 4.4: Impressions from *Beat Saber*. The images are retrieved from Steam Store; credit: Beat Games [13] [CorePre1].

With our game ready, we carried out a 2 (auditory cue: on vs. off) x 2 (visual cue: on vs. off) within-participants study ($N = 32$). We mainly collected questionnaire and performance data. Our main findings show that regardless of modality type, providing continuous cues improves player experience. For example, players had more enjoyment, felt like the game was more meaningful and provided better feedback. Overall, the cognitive workload of participants remained unaffected by continuous cues, but audio cueing led to a higher physical workload (i.e., was perceived physically more demanding). We attribute this to the urgency and pressure that repeatedly playing sound might have created. In terms of performance, only bimodal continuous cues significantly improved players' pose accuracy. We also provided design implications that can guide future work in this area, such as the granularity of cues, the exploration of different continuous cue designs, and the relevance of continuous cues to other immersive experiences.

4.4 GAME ELEMENTS IN IMMERSIVE VR EXERGAMES

Our findings show that players can benefit from the real-time guidance of VR to boost their performance and improve the player experience [CorePub3]. While these results highlight the momentary advantages of immersive VR exergames, lasting health benefits can only unfold if players keep using these games over a long period [3], similar to every physical activity. However, our extensive review of exergames [CorePub1] also revealed limited research on long-term engagement in these games. While we know that immersive VR exergames can be a motivating options for exercise, we have only limited knowledge about what is so motivating about these games and how these motivating factors relate to the long-term engagement.

Immersive VR exergames are complex systems, having numerous game design elements. Beyond immersive VR exergames, some researchers even use game elements to motivate users in non-game contexts (i.e., gamification [55, 83] or motivational affordances [111]). Nevertheless, we only have very limited knowledge about how and which game elements motivate users in immersive VR exergames and consequently contribute to their long-term use. Therefore, in our newest work (preprint) [CorePre1], we focused on successful commercial immersive VR exergames to holistically understand which aspects of immersive VR exergames motivate users and consequently contribute to their long-term use. To investigate motivation, we used the inspiration theory [208, 209]. This theory sees motivation as evoked by external stimuli rather than as an inherently existing construct. It involves two essential components: (i) people can be *inspired by* an external stimuli (i.e., immersive VR exergames) and (ii) *inspired to* take an action (i.e., to do physical exercise). Based on this theory [208, 209], we derived the following RQs:

?

RQ1.12: Which game elements influence inspiration in immersive VR exergames?

RQ1.13: How does inspiration in immersive VR exergames influence long-term use?

For more information on [CorePre1] (preprint), please refer to Section 9.4 or scan the QR code.

To answer these questions, we conducted two sequential studies. The first study aimed to identify commercial immersive VR exergames and game elements, while the second study investigated the quantitative impact of game elements on inspiration and long-term use. In Study 1, we identified a total of 110 commercial immersive VR exergames (based on [220] and [CorePub1], see the list in the paper's supplementary materials in Section 9.4) and asked five experts (people with VR experience and who had played immersive VR exergames) to select the immersive VR exergame they played the most; we only asked experts to select one game since each game contains its own unique game elements [55]. We also asked the experts which game elements they found present and important in the immersive VR exergame they played the most. For game elements, we used a comprehensive list [111], which contains 46 game elements such as "*points, score, XP*". As a result, we proceeded with the immersive VR game that each expert voted for: *Beat Saber* [13]. It is a very popular music rhythm game released in 2019 [13, 196]. In this exergame, players use lightsabers to slash cubes while avoiding obstacles (see Figure 4.4). Based on experts' ratings and an expert interview session involving two VR exergame researchers (one being part of the project team, conducting research in Germany, while the other was an external researcher, located in North America), we finalized our game elements list and provided an example for each element.

After identifying the relevant game elements in the top voted (by our experts) and most academically researched immersive VR exergame [CorePub1], *Beat Saber* [13], in the second study, we conducted an online survey with *Beat Saber* players ($N = 256$) in Prolific to see how these game elements of *Beat Saber* inspire and, in turn, influence long-term use. Our survey included empirically validated scales and items [28, 68, 105, 111, 113, 208] targeting four aspects: game elements, *inspired by Beat Saber*, *inspired to physical exercise*, and *long-term intention* *to use Beat Saber* (all $\alpha \geq 0.87$). We provide the survey items with the exact phrasings in the paper's supplementary materials in Section 9.4.

Our stepwise backward multiple regression analysis [66] revealed that while seven game elements (*(i) Points, score, XP, (ii) Badges, achievements, medals, trophies, (iii) Levels, (iv) Competition, (v) Customization, personalization, (vi) Virtual world, 3D world, game world, (vii) Assistance, virtual helpers* [111]) positively influence *inspired by Beat Saber*, two game elements (i.e., *(i) Performance stats, performance feedback and (ii) Progress, status bars, skill trees* [111]) negatively influence it. Our further mediation analysis (using the PROCESS macro mediation model [88]) considering *inspired by playing Beat Saber*, *inspired to do physical exercise*, and *intention of long-term use* showed that there are positive relationship between those variables. In simple terms, being inspired by playing *Beat Saber* inspires players to engage in physical exercise and also supports long-term engagement with the game. We concluded our paper by presenting a number of design and theory implications for future work. For example, we suggested using various game elements to foster inspiration and long-term use, whereas being careful to use some game elements (i.e., *Performance stats, performance feedback and Progress, status bars, skill trees* [111]) to avoid hindering inspiration and long-term engagement.

4.5 CORE TAKEAWAYS

This chapter focused on supporting physical activity through immersive VR exergames. Apart from providing a holistic overview through a scoping review on XR exergames [CorePub1], we presented three empirical studies. These works show that tailored and carefully-designed immersive VR exergames can be used to train even explosive movements like jumps [CorePub2], that in particular bimodal continuous cues are beneficial for providing real-time feedback and enriching the exergame experience [CorePub3], and that game elements like badges and competition are vital in making *Beat Saber* a motivating experience which inspires players to be physically active and use the game in long-term [CorePre1]. Our core takeaways are:

[CorePub1]

- Our holistic review on XR exergames can be considered as a first step to systematically explore the XR exergame space through the HCI lens. With our taxonomy, researchers and practitioners can easily identify non-targeted areas or explored features in XR exergames. Our reporting guidelines contribute to efforts for comprehensibility, reproducibility, and transferability in exergame research.

[CorePub2]

- Explosive movements (i.e., jumping) that involve the entire body can be trained through immersive VR exergames by using enjoyable, layered routines that build up and adapt automatically to players' performance. With *JumpExTraVR*, players experience fun while receiving personalized training.

[CorePub3]

- Immersive VR exergames not only include motion-based tasks, but also help players accomplish these tasks through continuous cues. Especially, bimodal ones can enhance both player experience and performance. We contribute to the creation of purpose-driven immersive VR exergames without compromising the positive player experience.

[CorePre1]

- We disassembled *Beat Saber*, an immersive VR exergame, into its building blocks. Accordingly, we show that while seven game elements (like points and achievements) can positively inspire physical activity and long-term use, two game elements (like performance stats and progress) negatively influence these constructs. We offer insights into purposeful and targeted design of immersive VR exergames.

5 IMMERSIVE VR GAMES FOR ACTIVE AGING

According to a United Nations report [216], the world is experiencing demographic change. The number of the older adults (65+ years)[1] is expected to reach 1.5 billion by 2050. While this indicates improved health, economic, and social standards, the aging population also raises several challenges relating to, for example, preserving the independence of individuals, or limiting the strain on the economy and healthcare systems [216, 231]. The WHO defines active aging as "*the process of optimizing opportunities for health, participation and security in order to enhance quality of life as people age*" and emphasizes that it should be promoted to ensure positive aging experiences and low burden on society [231]. Among other factors like mental health, the WHO highlights the significance of physical activity for active aging. While exercise is essential for people of all ages, it can be particularly beneficial for older adults to live independent lives [231, 232]. Accordingly, the WHO's policy framework for active aging [231] recommends that "*policies and programmes should be based on the rights, needs, preferences and capacities of older people*". To provide older adults with such opportunities for active aging, we, therefore, explore the potential of novel, engaging solutions: immersive VR exergames. Correspondingly, this chapter answers the RQ2 with the aim of providing effective, enjoyable active aging opportunities for older adults:

> **?** **RQ2:** *How can immersive VR games support active aging?*

5.1 COGNITIVE AND PHYSICAL EXERCISES FOR OLDER ADULTS

Immersive VR exergames have great potential to improve the exercise experience for older adults by resolving three major challenges: The first challenge is motivation. These games can keep people exercising due to their playful nature. Secondly, providing cognitive and physical exercise can have a positive impact on older adults' well-being [122, 162], in addition to mitigating the decline in their cognitive abilities and brain health [40, 121]. In particular, simultaneous training featuring both cognitive and physical tasks can even have greater benefits for older adults [206]. Immersive VR exergames games can seamlessly incorporate both cognitive and physical tasks. Another

1 There is no consensus on what age (60 or 65) counts as an older adult. In the scope of this dissertation, we follow the United Nations' definition for older adults, and consider people aged 65 and older as older adults [216].

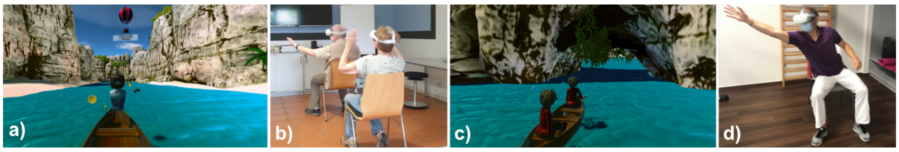

Figure 5.1: Impressions from *Canoe VR*: a) An artificial agent playing the game together with the user. b) Two older adults in the multiplayer version. c) The representation of the two players in the multiplayer version. d) A physiotherapist testing the game [CorePub4].

problem with exercising is access to opportunities [137]. This has been particularly challenging during COVID-19, which has imposed restrictions on access to exercise opportunities, indicating declining physical activity for older adults [239]. Immersive VR exergames can address these challenges simultaneously: These games can motivate older adults to perform physical activity, combine cognitive and physical exercises into a coherent game experience, and provide easy access to exercise opportunities. Immersive VR exergames can serve as a valuable alternative for home-based exercises. They allow social distancing without sacrificing physical activity, especially in the cases that necessitate distancing, such as COVID-19. Despite this potential, immersive VR technology imposes unique challenges related to usability [2] and ergonomics [91], while exergame design requires attention in terms of age-related changes [76] and movements. Therefore, in [CorePub4], we address the following RQ[2]:

For more information on [CorePub4], please refer to Section 10.1 or scan the QR code.

RQ2.1: How can we design immersive VR exergames featuring appropriate cognitive and physical exercises for older adults?

To design and implement an immersive VR exergame that features both cognitive and physical exercises for older adults, we followed an HCD process. We conducted eight prototype testing sessions (e.g., including at the gym of older adults) and interviews throughout our design process. We involved both older adults (with and without dementia), fitness experts, and a healthcare professional in the process. As the result, we iteratively designed and implemented an immersive VR exergame for older adults: *Canoe VR* (see Figure 5.1). This game was a seated experience given the abilities and safety of older adults and required players to perform upper body movements (e.g., torso, reaching, and leaning). Overall, players were tasked with collecting objects that appear while traveling down the river in a canoe in a naturalistic environment. There were a total of 22 levels in our game. For example, we implemented levels that encouraged simultaneous

2 The research questions of this chapter have been revised to ensure stylistic consistency in this dissertation while preserving their meaning.

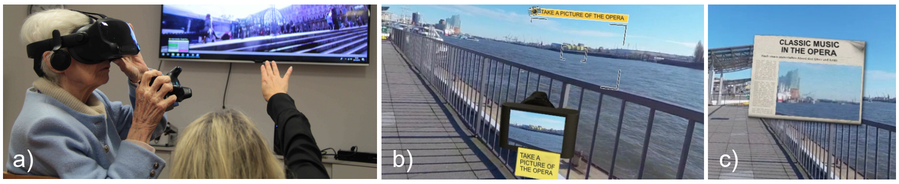

Figure 5.2: Impressions from *Memory Journalist VR*: a) A player explores the game environment with a 3D printed camera, b) communicates with a health professional to find a landmark in the scene, and c) takes a correct picture of the landmark which is shown in a newspaper [CorePub5].

cognitive and physical exercises, such as tasking players to collect red bubbles with their left hand and yellow bubbles with their right hand. To account for social interaction, we added multiplayer levels that could be played with another person or agent. Each prototype test targeted another challenge, ranging from understanding the game's movement value to seeing its use as an additional exercise tool in the gym for older adults.

Our thematic analysis (e.g., observation and field notes, interviews) and questionnaires revealed that *Canoe VR* was received very well by the older adults, the fitness experts, and the health care professional. We also show that an immersive VR exergame can be used in older adults' regular gym sessions. While our results indicate that *Canoe VR* offers suitable exercises for older adults with dementia by adjusting the difficulty of the tasks to their abilities, we also identified required improvements (e.g., creating a more accessible tutorial and calibration sequence). Hence, we recommend further evaluation of this game to design age-appropriate immersive VR exergames, especially targeting older adults with different level of abilities. As one follow-up work, we, therefore, investigated the effect of difficulty adjustments for older adults using *Canoe VR* [OtherPub6][3].

5.2 INCLUSIVITY FOR OLDER ADULTS

One of the major learnings of our work with older adults was that people with specific conditions, like dementia, have different requirements which pose new design challenges when targeting these user groups. Understanding these requirements is essential towards a fully-inclusive game design. Currently, there are more than 55 million people with dementia in the world population [154]. Dementia is a neurological condition that affects one's cognitive, physical, and social abilities, with profound effects on society at large as well [154]. Among others, aging and lack of physical activity increase the risk of dementia [6, 49, 154]. Accordingly, similar to academic research [90], the WHO recommends keeping older adults with dementia physically active

3 The publication by Kruse et al. [OtherPub6] is not a part of this dissertation.

both as a method of dementia prevention and during dementia [154]. Although immersive VR exergames are promising to motivate users, the previously-discussed study [CorePub4] shows that game design requires more attention to deliver accessible immersive VR exergame experiences to this user group. Additionally, using VR technology with this user group may have unknown effects. For example, VR can create a strong sense of presence (as described in Chapter 2) and older adults may experience emotional stress depending on the situation encountered. Therefore, there is a need to carefully investigate the needs and abilities of this user group to create appropriate active aging opportunities. In [CorePub5], we targeted this challenge with the following RQs:

For more information on [CorePub5], please refer to Section 10.2 or scan the QR code.

?

RQ2.2: Does the designed immersive VR exergame provide positive and good player experiences, usability, and accessibility for older adults with dementia?

RQ2.3: Which aspects of the HCD approach should be adapted in the context of immersive VR exergames for older adults with dementia?

To design and implement an immersive VR exergame that fulfills the requirements of older adults with dementia, we followed a four-step HCD approach. First, we focused on understanding these needs and requirements in depth. Therefore, we conducted semi-structured interviews with 11 stakeholders (e.g., clinical experts, physiotherapist, technicians) and contextual inquiry sessions ($N = 8$) with the target group, and additionally included people without dementia to gain additional information about the older adults in general since they share, for example, similar aesthetics values. Through our thematic analysis of interviews and inquiries, we determined the user requirements (e.g., importance of social gaming activities, challenges of learning new interactions, and safety). In the second step, we designed our exergame concept: *Memory Journalist VR*[4] (see Figure 5.2). For example, we benefited from the reminiscence that VEs can create [205] and aimed to offer intuitive ways to interact with the VE. Accordingly, in the third step, we implemented *Memory Journalist VR*. In this seated exergame, players experienced 360° 3D recordings of famous landmarks in two major German cities. They took the role of a reporter who has to take pictures for a newspaper. As game interaction, the players pointed a 3D-printed, positionally-tracked camera towards the intend object, resulting in upper-body movements (e.g., torso, stretching) that could assist daily activities. The game tasks were controlled by caregivers through a browser-based application, and the game was projected onto a TV screen for monitoring and sharing the experience. In the last step, we conducted five focus group sessions (involving 11 older adults

4 The very early stage of this prototype was presented in [175, 176].

with different degrees of dementia, 6 older adults without dementia, 3 caregivers, 1 health professional) and iteratively improved our game.

Our thematic analysis of the focus groups show that older adults with dementia had positive experiences. While our tailored game promoted enjoyment and social interaction for older adults with dementia, it was boring for older adults without dementia. The older adults with dementia immersed themselves in the game. They tried to talk to people or touch soap bubbles in the VE. While this perceived presence can be seen as a positive aspect of the game, it may also raise ethical concerns. Therefore, we recommend careful design of VEs for this user group. Based on user requirements and observations from focus group sessions (e.g., a player getting irritated by a caregiver's touch), we advise against using unrealistic scenarios, as they may lead to negative outcomes (e.g., anxiety). We also found that interaction with controllers is too complex for this user group, and that we should use real-life interaction paradigms (e.g., 3D camera). In terms of the HCD approach, we emphasize the importance of involving stakeholders in the user requirements step through semi-structured interviews. The stakeholders can enrich the design process by providing information about dementia and the daily routines of older adults with dementia. For example, learning about dementia led us to not use some typical mechanics for game design, such as unpredictability. Similarly, we highlight the key role of contextual inquiries that provide information about the routines of people with dementia, their caregivers who play these games with older adults, and the technicians who maintain the technologies. Overall, our results demonstrate the feasibility of designing immersive VR exergames for people with dementia and integrating these games into their daily routines. As a further work to this paper, we, therefore, investigated the long-term use effects of *Memory Journalist VR* on cognitive, physical, and psychological well-being of older adults with dementia, indicating promising results [OtherPub5][5].

5.3 AUGMENTATION FOR OLDER ADULTS

After seeing the general potential of immersive VR exergames for older adults, we focused on utilizing one of the unique capabilities of VR. Similar to other studies [23, 100, 230], we refer to users' virtually enhanced abilities as augmented interaction. Since augmented interaction depicts unrealistic experiences, based on our learning [CorePub5], we only considered older adults without dementia in our research. Although, unrealistic interactions have great potential to boost motivation, accessibility, and efficiency, we know little about such interactions on the experience of older adults. Especially, investigation of augmented interaction in immersive VR and non-immersive exergames focuses mostly younger population [23, 100] and targets exaggerated

5 The publication by Kruse et al. [OtherPub5] is not a part of this dissertation.

jumping abilities [100, 103, 120]. However, augmented interaction may also improve older adults' competence and motivation to perform physical exercises. Yet, previous work emphasized using realistic scenarios and mechanics for older adults [11]. Accordingly, in [CorePre2] (preprint), we focused on investigating augmented interaction by involving older adults. Thus, we created two augmented interaction components considering external validity: (i) movement (augmented locomotion) and (ii) interaction with virtual objects (augmented object range). Both of these elements are at the core of immersive VR exergames, as every exergame involves movement and interaction with objects. Considering that the swimming theme was requested in [CorePub5] and used in a prior study [245], we specifically designed and implemented a seated exergame with this theme: *ExerSwimVR*.

The development of *ExerSwimVR* (see Figure 5.3) involved two pilot studies including both young ($N = 5$) and older adults ($N = 4$), and an interview with a physiotherapist. In this underwater seated exergame, players performed breaststroke arm movements (i.e., stretching arms and hands out in front with hands at a $45°$ outward angle, pulling them slightly down and outward until shoulder level [242]) to move in the VE. To collect the polluted objects, they performed reaching movements. To implement movements in the VE, we calculated the players' virtual velocity based on two factors: thrust (acceleration) and drag (deceleration). Acceleration was based on the controllers' velocity, movement pattern (breaststroke), and hand orientation. The drag constantly slowed players' speed based on their current virtual velocity. Accordingly, in the not augmented locomotion, players' breaststroke arm movements translated into the VE without any adjustment, whereas in the augmented locomotion, they experienced movement with an exaggerated velocity by a factor of six (determined in the pilot studies), leading players to travel farther in the same time frame. When players experienced the no augmented version of the object range, they could collect the polluted cans by performing reaching movement towards the objects and pressing and holding the trigger button of the controller. In the augmented version of object range, pressing a trigger created a laser pointer, which allowed collecting an object from a distance up to three meters. With *ExerSwimVR*, we investigated the RQs below in [CorePre2]:

For more information on [CorePre2] (preprint), please refer to Section 10.3 or scan the QR code.

RQ2.4: How does augmented interaction (i.e., augmented locomotion and augmented range of reachable objects) in immersive VR exergames affect player experience and performance?

RQ2.5: How does the same augmented interaction differ in effect on those same player experience components and performance when experienced by older adults?

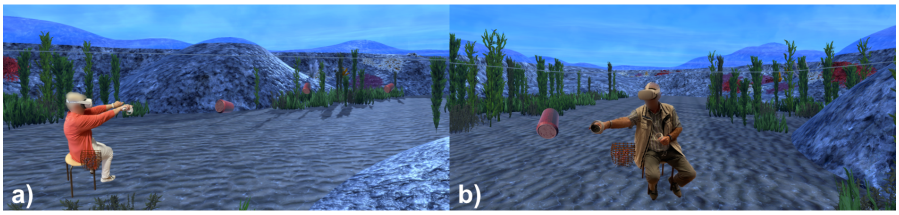

Figure 5.3: Impressions from *ExerSwimVR*: a) A player performs breaststroke arm movements to swim. b) A player collects pollution objects to clean the ocean [CorePre2].

To examine RQ2.4 and RQ2.5, we conducted two 2 (augmented locomotion: on and off) x 2 (augmented range of reachable objects: on and off) within-participants studies: one remotely with young adults ($N = 29$, due to COVID-19) and one in-person with older adults ($N = 24$). Although our focus in this research was on older adults, we specifically conducted a study with young adults to reduce the likelihood that our results were due to our artifact, to identify similarities and differences between user groups by focusing on augmented interaction, and to replicate the findings of previous studies that focused on young adults (e.g., [23, 100]). In our studies, we collected questionnaire, performance, and qualitative data. We found that primarily augmented locomotion positively impacted younger players' experiences (e.g., motivation) and led them to collect more objects without causing increased cybersickness. Younger players did not perform significantly less physical activity despite the advantages of augmented interaction techniques. However, for older adults, augmented interaction did not lead to significantly higher player experience (e.g., motivation); they rated all conditions positively. Older adults performed significantly less physical activity when they experienced the combined version of the augmented interaction, indicating that augmented interaction might decrease the effectiveness of immersive VR exergames for older adults. Overall, our findings contribute to a comprehensive understanding of how age group differences can affect the design of immersive VR games for active aging.

5.4 EXERGAMING TECHNOLOGIES FOR OLDER ADULTS

After exploring the unique interaction possibility of immersive VR exergames for active aging, we questioned what are the strengths and disadvantages of exergaming technologies in providing active aging opportunities for older adults. Previous research has shown that immersive VR exergaming technologies for younger adults lead to higher motivation and performance [22, 235], but we do not know how immersion affects the player experience and performance of older adults in this context. Yet, we know that older adults have different experience and performance outcomes with digital technologies compared to

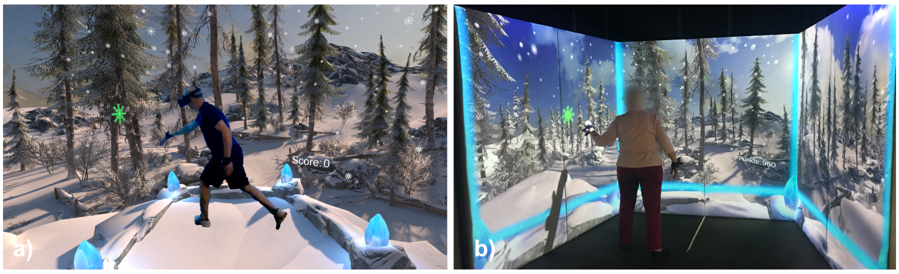

Figure 5.4: Impressions from *Winter Wonderland*: a) A participant approaches a snowflake to catch it in the VR-HMD version of *Winter Wonderland*. b) A player catches a snowflake in the ExerCube version of the game [CorePre3].

younger generations [53, 56, 133]. Given the increase in development of exergames for older adults [CorePub1], it is important to empirically examine the role of display technologies to facilitate the creation of age-appropriate active aging opportunities for older adults. Looking at exergaming research, commercial market, and digital technologies used for exercising in senior living facilities [CorePub1, CorePub4], [174][6], we specifically choose to investigate two technologies varying in their technical immersion: VR-HMD and *ExerCube* [126][7].

ExerCube is a training setup consisting of three projector walls and a motion tracking system. While ExerCube resembles immersive CAVE systems (e.g., displaying a VE on the projector walls), unlike CAVE systems, it does not offer view-dependent stereoscopic images. Therefore, this setup has less technical immersion capabilities. Following the definition we introduced in Chapter 2 for immersiveness and for simplicity, we refer to this setup as not being technically immersive. Although this technology does not fit the definition of being technically immersive, research showed promising results of its use for training purposes (e.g., motivating) [126, 127]. In contrast to ExerCube, we consider a VR-HMD to be a more immersive display since it provides stereoscopic view-dependent experiences. Hence, we label this setup immersive. Therefore, in [CorePre3] (preprint), our RQ was:

For more information on [CorePre3] (preprint), please refer to Section 10.4 or scan the QR code.

> ?
>
> **RQ2.6:** How does the technical immersion of an exergame system (i.e., a VR-HMD vs. ExerCube) affect older adults' player experience and performance?

To investigate RQ2.6, we designed and implemented an exergame that older adults can play using a VR-HMD and the ExerCube system, entitled *Winter Wonderland*. We configured the alignment of the physical and VEs to ensure an exact virtual and spatial mapping between the two technologies. This allowed us to have identical positions

6 https://sphery.ch/en/standorte/
7 The research product has been commercialized, see https://sphery.ch/en/.

of the game objects in the real world and in the VE, and required players to perform exactly the same movements in both versions of the game. The game design was informed by the guidelines and recommendations of previous studies conducted with older adults (e.g., [CorePub4, CorePub5], [OtherPub6], [2]) as well as an HCD approach [99]. Accordingly, we iteratively tested the game with older adults ($N = 4$) and movement experts (a physiotherapist and a trainer) and integrated their feedback into our game design. The game had a winter theme and included different cognitive and physical tasks related to collecting snowflakes. For example, in one level, players had to catch green snowflakes with both hands, while in another level they had to catch snowflakes of certain colors with specific hands.

Using *Winter Wonderland*, we run a within-participants study with 34 older adults. We collected questionnaire, in-game performance, heart rate, and interview data. Our results show that both exergaming technologies resulted in comparably high player experience and physical exertion measured by heart rate. The VR-HMD condition led to a higher level of presence, whereas the ExerCube condition led to higher in-game performance and physical activity. Our qualitative findings provided further nuances in the understanding of older adults' experiences. For example, we observed an emphasis on familiarity with the game (by testing during the study only) and setup, suggesting that novelty may not contribute to older adults' experience. Furthermore, the results indicate that VR-HMDs may be perceptually more difficult for older adults; they felt more challenged and had spatial difficulties with this setup, suggesting that field of view and depth perception difficulties should be considered in VR-HMDs for older adults. Finally, we proposed considering the benefits and drawbacks of both exergaming setups. The ExerCube encourages greater physical activity, but has associated cost and space requirements in general, whereas the VR-HMD offers higher presence at an affordable cost, but requires more attention in terms of safety.

5.5 CORE TAKEAWAYS

This chapter presented four contributions targeting active aging opportunities for older adults. In [CorePub4], we focused on designing appropriate cognitive and physical exercises for older adults. [CorePub5] explored the feasibility of designing immersive VR exergames for older adults with dementia. In [CorePre2], we gave older adults enhanced virtual abilities and explored how these impact their experience in engaging active aging possibilities. Finally, in [CorePre3], we examined the role of technical immersion of display technologies on older adults' experience and performance. Accordingly, we offer the following takeaways:

[CorePub4]

- Immersive VR exergames can provide age-appropriate cognitive and physical exercises, even combined exercises. They can be used in regular gym sessions of older adults as an alternative fitness tool.

[CorePub5]

- To contribute to active aging of older adults with dementia, we can use immersive VR exergames. Considering the interests and abilities of this user group, immersive VR exergames can provide a positive and good player experience, usability, and accessibility. We recommend including stakeholders as well as contextual inquiry sessions in the design process of immersive VR exergames targeting older adults with dementia.

[CorePre2]

- Older adults experience and perform differently from younger adults when they experience enhanced virtual abilities in immersive VR exergames. Older adults' motivation stays significantly not affected, but in one condition, augmentation led to reduced physical activity.

[CorePre3]

- We provide insights into the benefits and tradeoffs between the two technologies for active aging. The VR-HMD provides higher presence, while the ExerCube results in higher physical activity and in-game performance.

6 IMMERSIVE VR GAMES FOR SOCIAL INTERACTION

We use social platforms and technologies everywhere to maintain and build social relationships. Although the use of technology for social interaction may have negative consequences, such as a decrease in face-to-face communications [114], technologies can also bridge the gaps between people [7] and even support well-being (e.g., [52, 72, 78, 221]). The WHO defines well-being as "*a positive state experienced by individuals and societies*" [155] and notes that well-being "*is determined by social, economic and environmental conditions*" [155]. Considering the technological advances, the WHO recommends "*strengthen[ing] the benefits*" [155] and "*addressing the impacts of digital transformation*" [61] to promote well-being. During the COVID-19 pandemic [234], these benefits became particularly apparent, as many people maintained their social connections and fulfilled their social needs using digital communication technologies [72]. Games played an important role during these times [38, 79] since these interactive applications particularly good at creating fun and shared times. Especially, multiplayer immersive VR games can foster higher social presence and player experience compared to their non-immersive-VR versions [35]. Players can perceive the illusion of sharing the same virtual space [190] and being together with each other [20]. For example, in a hypothetical scenario where two friends play a cooperative game in VR, this could allow them to first shake hands to greet each other, and later celebrate their win with a high-five. In this chapter, we, therefore, aim to explore the potential of social immersive VR games, focusing on RQ3:

> ? **RQ3:** *How can immersive VR games support social interaction?*

6.1 SOCIAL MULTIPLAYER IMMERSIVE VR GAMES

Multiplayer immersive VR games have great potential to deliver socially rich and positive player experiences. However, can we use multiplayer immersive VR gaming environments as an alternative to real-life multiplayer gaming scenarios? What does immersive VR still need to improve, or where does VR outperform real-life scenarios? Given that COVID-19 affected social interactions during lockdown, we saw multiplayer immersive VR game scenarios as an opportunity and approached these questions with an extensive user study. As a socially rich real-life gaming scenario, we chose a traditional couch-coop gaming setup.

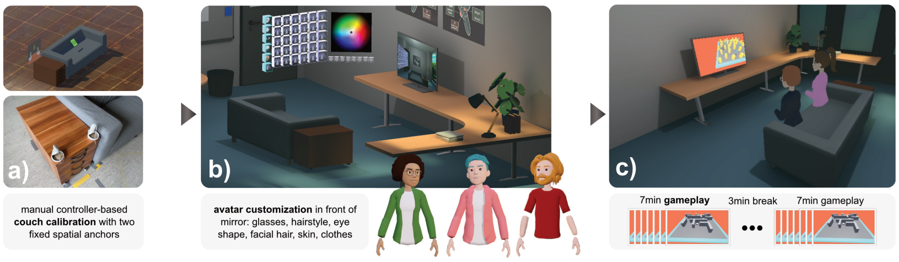

Figure 6.1: Illustrations of the simulated co-located multiplayer immersive VR experience: a) The calibration, b) avatar customization, and c) virtual room scene where players meet to play a multiplayer game on a virtual TV [CorePub6].

Couch-coop games can be differentiated from other gaming settings, since in these games people often gather with friends and family members in their immediate surroundings, get excited together, and have distinct (e.g., physically blocking each other's view while playing) social experiences. However, this possibility was limited during COVID-19. As a potential solution, remote immersive VR games could evoke a similar sensation, mitigating the need for physical co-presence. Hence, in [CorePub6], our research was driven by the following RQs[1]:

For more information on [CorePub6], please refer to Section 11.1 or scan the QR code.

?

RQ3.1: How do the player and social experience in VR compare to the experience in the co-located setting?
RQ3.2: What features enhance or inhibit the player and social experience in VR?

To explore the effect of an immersive VR gaming scenario compared to a real-life couch-coop gaming scenario, we simulated the real couch-effect in VR. We first created a multiplayer simulator. Here, players sat side-by-side on a virtual couch, like in the real world, to play a multiplayer game on a virtual TV (see Figure 6.1). To have a controlled environment, we also included an initial calibration (i.e., to have identical positioning of physical and virtual couches) and customization step (i.e., avatar customization). Our game required players to communicate and strategize to complete given tasks (see Figure 6.1), and was designed based on a prior work [204].

We then conducted a within-participants study ($N = 50$) (including questionnaires, videos, and interviews) with two conditions: (i) co-located vs. (ii) VR. In the co-located condition, two players were physically in the same room, sitting next to each other on a couch and playing the game on a physical TV. In the VR condition, two players were placed in two physically-separated rooms and came together in a VR room to play the same multiplayer game on a virtual TV.

1 The research questions of this chapter have been revised to ensure stylistic consistency in this dissertation while preserving their meaning.

Regardless of how players experienced the scenario, we found that social connectedness between players increased compared to their pre-measurement. However, the VR condition received significantly lower social experience scores, i.e., social richness, co-presence, attention allocation, and message and affective understanding. Both conditions had comparably moderate to high player experience, with one exception: The VR condition led to significantly higher immersive player experience compared to the co-located condition. Despite the significant differences, both conditions resulted in high player experience, social presence, and social richness. Our qualitative data provided insights indicating that the lack of facial expressions, lower-body parts, and body language in the avatars and the limited field of view of the VR hardware hindered the VR experience. In contrast, the novelty effect of VR and the avatar customization feature positively contributed to players' experience. Overall, our findings suggest that the immersive VR gaming scenario is an alternative that players imagined using and preferred over other available digital communication options.

6.2 ASYMMETRIC MULTIPLAYER IMMERSIVE VR GAMES

While participants imagined using our VR setup [CorePub6] and there is an expected increase in sales of VR hardware [8], VR-HMDs are still not the most widely owned hardware. In a typical single-owned VR-HMD scenario, this excludes other players (i.e., non VR-HMD players) and spectators from the experience. Even more, using VR might be challenging in social settings due to its immersive nature. It can lead to technologically and socially isolating experiences [21, 144]. However, to enable multiplayer engagement and rich social experience, we need to design and propose solutions that enable non-VR-HMD players to be part of immersive VR games. The hardware difference (i.e., interface asymmetry) is just one of the differences between players that asymmetric games take into account [87]. Other examples include differences in abilities or skills, although they are not limited to these [87]. These games foster socially more rich experiences compared to symmetric games where both players have the same abilities and interface [86]. Unsurprisingly, asymmetric games, in particular games targeting the asymmetry of interface between users, have gained attention in VR research (e.g., [81, 82]). For example, Gugenheimer et al. [81] have developed a top-down floor projection combined with a hand-held controller system called *ShareVR* that aims to integrate non-VR-HMD users into VR experiences. The system allowed non-VR-HMD players to actively interact with VE and VR-HMD players in the same shared space, for example, using props connected to the controller. Nevertheless, despite the growing research and potential in this field, asymmetric multiplayer immersive VR games are still in their infancy, with different focal points and little work dealing with the theoretical

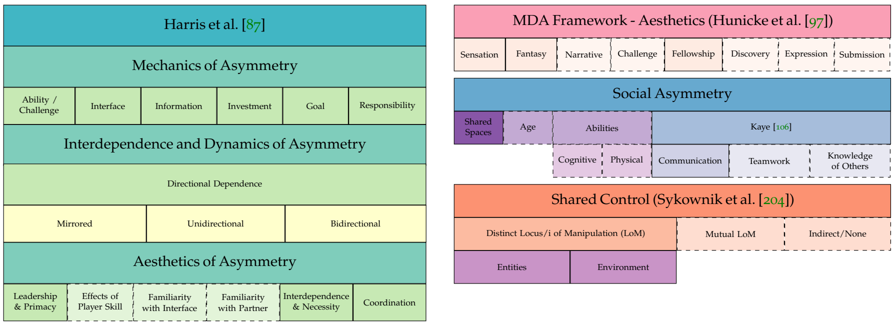

Figure 6.2: Our post-hoc "best-fit" framework for asymmetric multiplayer immersive VR games: Suggested post-hoc "best-fit" framework, with dashed lines and more transparent colour for categories that are under-represented in existing asymmetric immersive multiplayer VR games research so far [CorePub7].

For more information on [CorePub7], please refer to Section 11.2 or scan the QR code.

foundations of these games. To address this challenge, we conducted a systematic literature review [CorePub7] to answer the following RQs:

?

RQ3.3: What kinds of asymmetry are being designed in multiplayer immersive VR games?

RQ3.4: How does VR asymmetry in multiplayer immersive VR games affect player experience?

Starting with 481 articles from various sources (e.g., ACM[2]), we analyzed the final set of 25 articles (after applying screening and eligibility steps) using best-fit framework synthesis [31] and thematic analysis [37] and published our findings in [CorePub7]. Our best-fit framework analysis involved identifying multiple conceptual frameworks, combining them (i.e., creating a priori framework), and mapping the literature based on this priori framework. Our thematic analysis summarized the literature based on this framework, using the framework characteristics as a codebook. With our work, we provide a summary of asymmetric multiplayer immersive VR games, develop a post-hoc framework for the design of these games (see Figure 6.2), and identify research gaps. For example, at a theoretical level, we suggest that goals and responsibilities in asymmetric games should be treated as separate game mechanics, as players may share the same goal (e.g., solving a puzzle), but have different responsibilities (e.g., (i) remembering where the missing puzzle pieces are and (ii) placing them in the puzzle field). Additionally, our post-hoc framework expands the conceptual framework of Harris et al. [87] by adding other asymmetrical considerations in game design, such as social asymmetry (e.g., age and shared space) and shared control (e.g, distinct control of entities and environment). Similarly, we point out that various aspects

2 https://dl.acm.org/

of asymmetric games are underexplored, such as alternative interfaces, familiarity with the co-player, or remote play. Overall, our work contributes to the theoretical understanding of asymmetric multiplayer immersive VR games as well as paves the way for future research possibilities of these games.

6.3 INTERDEPENDENCE IN ASYMMETRIC MULTIPLAYER IMMERSIVE VR GAMES

Having established a theoretical understanding and a foundational framework of asymmetric multiplayer immersive VR games, our remaining publications empirically explored some of the research gaps. We specifically targeted two important aspects of multiplayer immersive VR games: (i) an interface asymmetry and (ii) a social dynamic that can influence the experience of players in asymmetric games. Accordingly, our next paper [CorePub8][3] focused on the interface asymmetry between players. Gugenheimer et al. [81] have shown that integrating non-VR-HMD users into VR experiences leads to higher enjoyment and presence for both players. However, while *ShareVR* [81] improved both VR-HMD and non-VR-HMD players' experiences compared to the baseline setup (i.e., TV and gamepad), most experience constructs remained higher for VR-HMD players compared to non-VR-HMD players. Accordingly, it remains an open research gap whether and how an asymmetric game design would lead to similar experiences for both players, regardless of their setup. While previous work has often focused on innovation in hardware design (e.g., [81, 82]), we targeted asymmetric multiplayer immersive VR game design. Previous studies in game research have shown that interdependence (i.e., reliance of players on each other [87]) between players can positively contribute to player and social experience [51, 86]. However, they did not assess the impact of types of interdependence. Therefore, we focused on creating and evaluating different types of interdependence that could facilitate multiplayer engagement.

First, inspired by several commercial games (e.g., Keep Talking and Nobody Explodes [197]), we implemented a strategic interdependence that requires one player to rely on the other in terms of ability and information. For example, the non-VR-HMD player had additional information about the game world that they needed to communicate with the VR-HMD player to successfully complete the game (e.g., approaching lasers, see Figure 6.3) while the VR-HMD player was the only one who can navigate the VE. Next, we incorporated an alternative interface (i.e., physiological interface) into our asymmetric game design to feature the second interdependence between players: a biometric interdependence. While biometric feedback has shown promise in

3 The data collection and implementation steps of this article were part of my master's thesis [104].

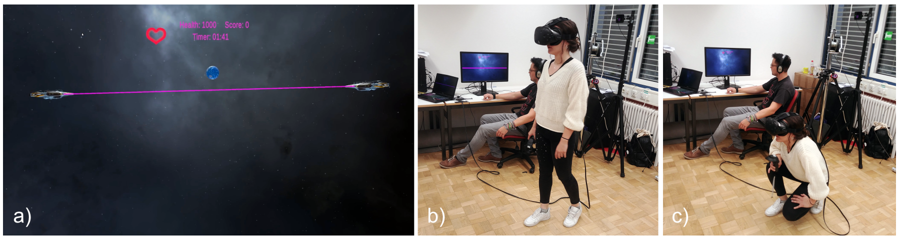

Figure 6.3: Impressions from our asymmetric multiplayer immersive VR game that features strategic and biometric interdependences between the VR-HMD and the non-VR-HMD players: a) The non-VR-HMD player can see and hear approaching lasers, b) while the VR-HMD player cannot. c) The VR-HMD player relies on the non-VR-HMD player's instructions for avoiding lasers. In some variants of the game, the time between spawn of lasers was affected by the heart rate metrics of the non-VR-HMD player [CorePub8].

academic game research [94, 95, 150, 177] and used in commercial games (e.g., Nevermind [67]), it has not received much attention in multiplayer immersive VR games, especially in asymmetric ones. In our game, the biometric interdependence required the VR-HMD player to rely on the physiological data of the non-VR-HMD player. The heart rate of the non-VR-HMD player continuously affected the difficulty of the game, time between laser spawns (see Figure 6.3). Therefore, our resulting RQs in [CorePub8] were:

For more information on [CorePub8], please refer to Section 11.3 or scan the QR code.

?

RQ3.5: Can imbalanced asymmetric information between players lead to one player feeling less in control (i.e., like they are following the other player's instructions without their own agency)?

RQ3.6: Does biometric influence over game difficulty dynamically affect player experience in VR (compared to static difficulty)? Does it make a difference a) whether players are provided an in-game indicator of whether their own biometric state is currently increasing game difficulty, or b) in what modality this indicator is represented in-game?

To investigate RQ3.5 and RQ3.6, we conducted a mixed-design controlled experiment ($N = 30$) using this asymmetric game, which incorporates both interdependences and allows two players to play, with one wearing a VR-HMD whereas the other does not. Participants played the game either as a VR-HMD player or as a non-VR-HMD player (playing the game on a PC monitor) (between-participants factor). We also further tested the effect of biometric influence and its multimodal cues to understand how to provide the best emotion regulation support (within-participants factor). In the study, we collected questionnaire, game interaction, heart rate, and interview data. Our results show that all measured quantitative constructs were comparably high for both players, suggesting that asymmetric game design and the resulting interdependences can be used to create multiplayer engagement across interfaces. We conclude that the strategic interdependence is a type of interdependence that can be leveraged to achieve comparably high quantitative and qualitative player experience between VR-HMD and non-VR-HMD players. However, we did not find any significant effect of biometric interdependence (and their cues) on quantitative outcomes; we attributed this to the subtle implementation of the biometric effect and the need for familiarization, especially for studies involving biometric measurements. Nevertheless, our qualitative data also revealed that when players perceived biometric interdependence, both players felt more immersed and stressed. Overall, our findings demonstrate the feasibility of achieving comparably high player experience using asymmetric game design for both players, regardless of their displays.

6.4 FAMILIARITY IN ASYMMETRIC MULTIPLAYER IMMERSIVE VR GAMES

After showing the potential of asymmetric multiplayer immersive VR games for players in different mediums, we focused on exploring another crucial aspect of multiplayer immersive VR games: a social dynamic, i.e., familiarity between players. While the effects of familiarity between players have been investigated in video games in both competitive [125, 168] and cooperative settings [219], there is a research gap in investigating this factor in asymmetric multiplayer immersive VR games [CorePub7]. Particularly, in an asymmetric multiplayer immersive VR game study [34], a sample consisted mostly of strangers. The authors discussed that recruiting players with prior relationships may improve the player experience, but did not evaluate this factor. Given that asymmetric multiplayer immersive VR games involve mostly verbal communication [CorePub7]—sometimes even gestures [192]—and played in VEs, we do not know whether these games would be affected by existing social bonds, or whether the verbal communication required by these games would be an advantage for players who know each other. Therefore, in the final publication of this part, [CorePub9], we address the following RQ:

For more information on [CorePub9], please refer to Section 11.4 or scan the QR code.

?

RQ3.7: How does familiarity between players impact player experience, social experience, and game performance in an asymmetric multiplayer immersive VR game?

To answer RQ3.7, we designed and developed an asymmetric multiplayer immersive VR game where two players can play while wearing VR-HMDs. Since the study was conducted during COVID-19, and given the increase in the sales of VR-HMDs [195] and the lack of re-

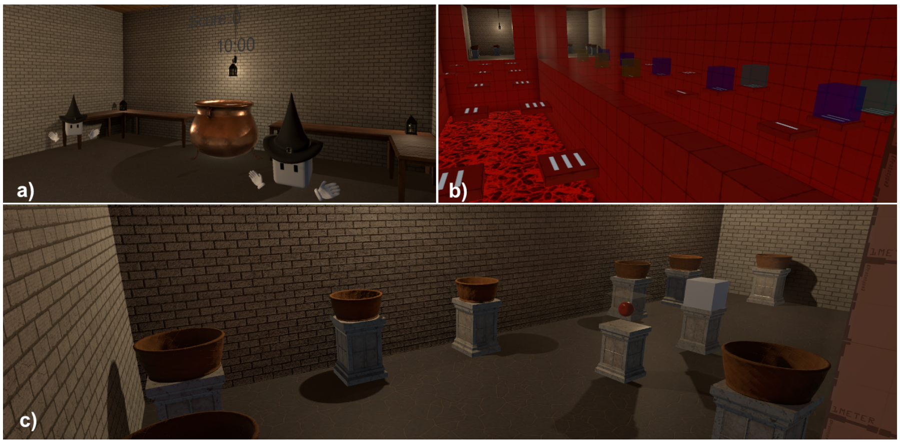

Figure 6.4: Impressions from our asymmetric multiplayer immersive VR game that features a strategic interdependence between two VR players: a) Players embody VR avatars, b) provide empty tile information to their co-player to progress through the maze, and c) share ingredient information with their co-player to complete a recipe. Note that both players can only see the other player's information [CorePub9].

mote exploration setup investigation in immersive VR asymmetric games [CorePub7], we specifically chose to design a remote game (i.e., both players were placed in different rooms) and equipped each player with a VR-HMD. Our asymmetric game only features strategic interdependence between players, based on the positive results from our previous work [CorePub8]. To create a strategic interdependence between players, we use information asymmetry, which has also recently received attention in game research in general [171], in game design [87]. Accordingly, both players shared the same type of knowledge about the game world (e.g., the location of tiles)—their reliance on each other was mutual and identical [87]. However, each player only saw information about the task that the other player can complete in the game. Therefore, they had to communicate these asymmetric information (e.g., required ingredient) with each other to progress in the game. Using this game, we conducted a between-participants ($N = 14$) controlled experiment to examine the effect of familiarity between players (friends vs. strangers) on player and social experience. We collected questionnaire, game performance, and in-game verbal communication data. Our findings revealed that the familiarity factor had no significant effect on player and social experience. For both groups, asymmetric play elicited high social presence, suggesting that these games are inherently social and can provide high social experience even when played by pairs of strangers. For the strangers group, the asymmetric multiplayer immersive VR gameplay even enhanced the social closeness between participants, even if they were only playing remotely in VR. With our work, we show that asymmetric multiplayer immersive VR games can elicit high player and social experience, regardless of familiarity between players, even in VR where we still have limited facial cues and technical constraints (e.g., limited field of view) [CorePub6].

6.5 CORE TAKEAWAYS

In conclusion, in this chapter, we explored the potential of immersive VR games to support social interaction through four publications. First, we questioned whether even multiplayer immersive VR game scenarios can approximate their real-life counterparts [CorePub6]. In the remaining publications, we targeted multiplayer immersive VR games that consider differences in players' abilities, skills, and hardware: asymmetric multiplayer immersive VR games. Thus, in [CorePub7], we presented a summary of asymmetric multiplayer immersive VR games research, proposed a framework, and identified research gaps by conducting a systematic literature review. Following this publication, we used asymmetric game design and created interdependences between players to tackle the single-owned HMD challenge [CorePub8]. Finally, in [CorePub9], we considered an existing social dynamic between players to understand how such a factor affects the players' experience in asymmetric multiplayer immersive VR games. Based on our publications, we derive the following takeaways:

[CorePub6]
- Multiplayer immersive VR gaming scenarios can elicit high player and social experiences, although their real-world counterparts still offer significantly higher social experiences. The inhibiting factors of VR are mostly related to the technical limitations of VR hardware, such as limited field of view or lack of facial expressions, which will eventually be solved with the advancements in technology.

[CorePub7]
- Asymmetric multiplayer immersive VR games offer enormous potential to address and leverage on differences between players, but received only limited attention in research. Our paper contributes to the theoretical understanding of these games by providing a best-fit framework and presents research gaps for future work.

[CorePub8]

- Using asymmetric game design and interdependences between players, it is possible to resolve the "headset-divide" and provide comparably high player experiences for both players. We build a bridge between interfaces using asymmetric game design and create shared social environments across mediums.

[CorePub9]

- Asymmetric multiplayer immersive VR games can be inherently social. They can provide high player and social experience regardless of the familiarity between players. With our work, we demonstrate the possibility of creating social environments using asymmetric games in remote settings, regardless of players' existing social ties.

7 DISCUSSION

We have been able to transfer several tasks into the online world, but this transformation has also contributed to a lack of physical activity [135, 229]. Advances in technology have also increased the life expectancy [229]. However, the resulting demographic change has brought new challenges to overcome, such as keeping older adults active [231]. Similarly, our social interactions have moved into the digital world, enabling us to communicate in real-time with people around the world [7] and supporting our well-being [52, 72, 78, 221]. However, we have yet to learn how to best use technologies, such as VR to offer engaging social interaction opportunities that connect but not divide people.

Among the many application areas where immersive VR technology could unlock new possibilities, one domain has received major attention, both among end-users and researchers: immersive VR games. VR headsets offer capabilities that traditional gaming technologies lack. They create immersive, multisensory experiences that enable players to feel present in VEs [187]. Consequently, they hold great potential not only for entertainment purposes but also for addressing societal challenges. Therefore, the overarching RQ of this dissertation asked:

> ? **RQ:** *How can immersive VR games support addressing contemporary societal challenges?*

We answered this question by investigating the potential of VR games regarding three pressing issues at the societal level: (i) *physical activity*, (ii) *active aging*, and (iii) *social interaction*. To answer our RQ, we presented a total of nine peer-reviewed publications and three preprints, featuring a broad range of investigations [227]: two surveys, nine artifacts, ten empirical studies, and twelve theoretical and opinion contributions.

7.1 IMMERSIVE VR GAMES FOR PHYSICAL ACTIVITY

Immersive VR games have enormous potential to engage players and create enjoyable experiences. In RQ_1, we focused on the insufficient physical activity crisis and explored how immersive VR games, played by physical movements, i.e., exergames, can help support or supply physical activity:

> ❓ **RQ1:** *How can immersive VR games support physical activity?*

First, with our scoping review [CorePub1], we confirmed the importance and timeliness of this topic. Our results revealed the rapid growth of this research field and the design of immersive exergames. We also presented a taxonomy to ease the systematic exploration in this area and discussed pressing research directions. In the remaining three contributions, we addressed three open questions empirically, advancing our knowledge of good immersive VR exergame design significantly. In [CorePub2], we extended the range of use-cases of exergames by demonstrating that they are a viable training method even for more explosive, lower-body-targeted movements (i.e., vertical jumping). We showcased how an expert-driven design process can result in a successful immersive VR exergame training tool, which delivers high player experience and physical activity enjoyment. Even during our study period (i.e., a single-time playing), we observed an improvement in players' jumping technique, suggesting that the game's personalized feedback helped players improve. Next, we explored how to guide players during movement execution via continuous real-time cues [CorePub3]. Our findings show that particularly bimodal, i.e., audiovisual, cues enriched both player experience and pose accuracy. Finally, we explored the game elements' role in the inspiration to do physical activity and long-term use of immersive VR exergames [CorePre1] by focusing on *Beat Saber* [13] and its game elements—one of the most popular and researched commercial immersive VR exergames [CorePub1]. Our results indicate that we should support player competence (e.g., through points and achievements), while being careful when providing performance-related feedback (which affected inspiration negatively) to avoid player frustration. Overall, our contributions are significant steps in demonstrating how immersive VR exergames support physical activity by incorporating explosive lower-body targeted training, guiding players to improve their movement execution, and inspiring players to be physically active through various game elements.

Nevertheless, our contributions only cover a fraction of the open research questions regarding the potential of immersive VR exergames, and there is still more to be discovered. In particular, our scoping review [CorePub1] revealed research gaps as well as areas for improvement. For instance, we found a limited understanding of long-term retention, a mismatch between the game design and evaluation with regard to target groups, a limited exploration of XR exergame experiences beyond VR (e.g., AR), and a lacking investigation of the social aspect. Additionally, we have discussed promising follow-up questions from our other publications as well. For example, our cueing research [CorePub3] did not investigate all modalities (e.g., haptics).

Similarly, we did not explore different design representations of cues. Our contribution [CorePub2] was only tested with non-athletes. However, we see a potential to explore such type of jump training in real sports environments, e.g., with professional basketball players. Finally, we assessed users' perspectives on motivating game elements by asking about their perception (e.g., long-term usage intention) [CorePre1]. However, conducting a controlled experiment may also offer additional behavioral insights for a more holistic understanding. Despite these open research questions, we believe that the future of immersive VR exergames as extremely positive. First of all, VR headsets are becoming cheaper, more capable, lighter and user-friendly, thereby reaching more users [195]. Meanwhile, society desperately needs new and provocative methods of engaging users in physical activity, as also highlighted by the WHO's report [232]. As we build the scientific foundation for this, our results are of major importance for exergame developers, who play a key role in turning our findings into consumer-ready products.

7.2 IMMERSIVE VR GAMES FOR ACTIVE AGING

Despite the potential of immersive VR games for keeping older adults active, their design and implementation requires attention and care. Technological unfamiliarity [133] and the effects of aging—such as changes in mobility and cognitive function [76]—introduce additional challenges for designers and researchers alike. Accordingly, we explored how to develop active aging opportunities for older adults through four contributions to RQ2:

> ? **RQ2:** *How can immersive VR games support active aging?*

We specifically used immersive VR exergames because these games can seamlessly combine both physical and cognitive challenges. While designing *Canoe VR* [CorePub4], we focused on creating age-appropriate cognitive and physical exercises through the involvement of older adults and movement experts. We show that immersive VR exergames have the potential to provide enjoyable experiences that match the expectations of older adults and as movement professionals. Next, we targeted a specific demographic within the older generation: older adults with dementia [CorePub5]. Considering their unique abilities and skills, we show how immersive VR technology can particularly benefit this user group. Older adults with dementia can easily interact with immersive VR exergames due to the natural 3D interaction that resembles real life. While realistic experiences and interaction techniques hold significant importance for VR exergames designed for older adults with dementia, VR also offers opportunities for unrealistic interactions that could be beneficial for motivation. However, we show that while augmented interactions—enhanced virtual abilities—in immersive VR exergames are motivating for the younger generation (matching with [100, 120]), these motivational effects do not transfer to older adults and may even reduce physical activity [CorePre2]. Finally, we focused on the challenges and possibilities associated with exergaming technologies for older adults. We show that while the VR-HMD induces higher presence, the ExerCube leads to higher physical activity for this generation [CorePre3]. In conclusion, we see that immersive VR games can engage older adults into beneficial exercises, even combining physical and cognitive challenges, that help them to retain an active, independent life. However, it is vital, more than for any other user group, to consider the unique requirements: Older adults might not require high levels of immersion or augmented interactions. Instead, exergame designers should focus on evoking a positive atmosphere (e.g., by basing the design on past memories) and aligning the difficulty to the older adults' abilities.

While our research paved the way for more inclusive design in the field of immersive VR exergames, it also left some open questions to be explored in the future. For example, when we designed cognitive and physical exercises for older adults, we did not evaluate the benefits of these exercises on their well-being [CorePub4]. Similarly, while the lessons we learned when designing for older adults with dementia point to the social value of exergames—not only for safety, but also for increasing social interactions— [CorePub5], we are not aware of any specific research targeting this aspect in the immersive VR exergames field, such as through asymmetric multiplayer immersive VR games. Furthermore, we still see the potential to explore enhanced virtual abilities in immersive VR exergames for older adults [CorePre2], for example targeting different interactions. Finally, our conclusions regarding the comparison of exergaming technologies only consider short-term testing of these setups with older adults [CorePre3], and long-term follow-ups may provide additional insights. Taken together, our contributions have demonstrated the feasibility and importance of human-targeted design in immersive VR exergames. In particular, our findings will have lasting importance, given that today's children will be the older adults of tomorrow and are growing up with technology.

7.3 IMMERSIVE VR GAMES FOR SOCIAL INTERACTION

Using VR for social interaction seems contradictory at first glance. After all, users wear a VR-HMD that blocks their entire view, ultimately preventing any interaction with the surrounding world [21]. However, compared to other types of digital interactions, immersive VR experiences are inherently spatial, being able to evoke strong presence and embodiment [188]. These capabilities allow us to shape the users' perceptions, which opens interesting possibilities for social interactions and virtual togetherness. With our final RQ and four publications, we investigated this potential of immersive VR games:

> ❓ **RQ3:** *How can immersive VR games support social interaction?*

We began by examining the role of immersive VR in supporting social interaction in the context of multiplayer gaming compared to its real-world counterpart [CorePub6]. Our findings show that although co-located couch-coop gaming results in significantly higher social experiences, immersive VR gaming scenarios can also deliver high player and social experiences, but need technical advances, especially in supporting facial expressions and extending the field of view. We then focused on the differences arising from both game design and external factors in multiplayer immersive VR games—asymmetries [CorePub7]. Our findings demonstrate the growing interest in HCI into these games, and we provide a theoretical framework to facilitate their investigation. Continuing with asymmetric multiplayer immersive VR games, we create a bridge between VR-HMD and non VR-HMD players' experiences using asymmetric game design and interdependences [CorePub8]. Finally, we looked at how asymmetric immersive VR games are affected by existing social ties. We show that these games can be inherently social. They offer high player and social experiences regardless of the familiarity between players, even if players are only together and interact in virtual worlds [CorePub9]. In summary, our research has broadened the understanding of how immersive VR games can support social interaction by showing the potential of social immersive VR gaming scenarios, presenting differences in game design and external factors in multi-user contexts, bridging the gap between VR-HMD and non-VR-HMD players, and showcasing the inherently social nature of asymmetric multiplayer immersive VR games.

For future work in the area of social immersive VR games, our first publication [CorePub6] provides insight into the technical limitations that VR hardware needs to improve upon, such as the field of view or face and full-body tracking capabilities. However, seeing the progress in technology, we believe that these features will become widely available soon. Already today, first VR-HMDs support tracking of the face and even the upper body. Furthermore, our systematic literature review on asymmetric multiplayer immersive VR games provides concrete answers as to what we can investigate further [CorePub7]. For example, we see intergenerational asymmetric multiplayer immersive VR games as a promising avenue to explore, as the two groups of users differ in their abilities and preferences. While our research on types of interdependence shows promising results, it is limited to two types: (i) strategic and (ii) biometric [CorePub8]. We think that physical interdependence (i.e., reliance of one player on another player in terms of performed movements) has great potential to be explored, especially to elicit social connectedness. However, it also requires care in design to avoid overwhelming or disturbing players. Finally, when assigning players to pairs of friends or strangers, we asked them a binary yes / no question [CorePub9]. We acknowledge the more fine-grained nature of social ties—such as best and close friends—and we recommend further research that considers more nuanced approaches. Despite the potential future challenges and opportunities, our research on immersive VR games shows how these games create social interaction opportunities. With the growing sales of VR hardware [195], our findings offer valuable insights into designing and developing social immersive VR games that can engage multiple users, regardless of their hardware access or pre-existing social connections.

8 CONCLUSION

HCI research can provide engaging, novel technology-based solutions to support addressing today's societal challenges. The work presented in this dissertation contributes to the United Nations' third sustainable development goal "*good health and well-being*" by targeting three societal challenges: (i) *physical activity*, (ii) *active aging*, and (iii) *social interaction*. In each of these challenges, we focused on different aspects that are important for researchers, designers, and developers. By doing so, we presented a total of 12 research papers (nine peer-reviewed publications and three preprints) that encompass a diverse array of research contributions. To this end, this dissertation offers valuable scientific insights by exploring the design and development of enjoyable immersive VR games that address current societal challenges.

In **Immersive VR Games for Physical Activity**, we presented four contributions that focused on using immersive VR games to support physical activity (see Chapter 4). Apart from providing a comprehensive overview and taxonomy of the prior research on immersive exergames, we created an explosive lower-body VR exergame-based training featuring personalized feedback and adaptability. Furthermore, we compared different modalities of continuous cues to guide players in executing movements correctly and provided empirical insights into the benefits of using these cues. Lastly, we evaluated the role of *Beat Saber*'s game elements in inspiring physical activity and long-term use, and we advised careful consideration when using game elements related to performance feedback.

Immersive VR Games for Active Aging featured four contributions involving older adults (see Chapter 5). We started with designing age-appropriate cognitive as well as physical exercises in immersive VR games for older adults, matching with older adults and movement experts' requirements. We then targeted older adults with dementia and showed the feasibility of offering immersive VR exergames for this demographic. Afterwards, we enhanced the virtual abilities of older adults and found that, unlike the younger generation, the positive effects of augmented interactions in immersive VR games were not transferred to older adults. Lastly, we focused on the role of immersion in exergaming technologies for older adults and revealed the unique advantages and disadvantages of both technologies.

The last core chapter of this dissertation contained four publications for **Immersive VR Games for Social Interaction** (see Chapter 6). We questioned if and to which extent immersive VR can approximate socially rich experiences offered in real social gaming setups, with findings suggesting promising outcomes. We further focused on asymmetries between players and in game design that offer social multiplayer immersive VR gaming environments and presented a framework to guide the design and exploration in the field. Next, we showed the value of using interdependences between players to provide comparably high player experiences across players in different mediums. In the last exploration, we questioned whether the experiences of asymmetric multiplayer immersive VR games are influenced by existing social connections between players, and found that these games offer high player and social experiences regardless of existing social ties even in remote settings.

Ultimately, **at the heart of this dissertation and my entire research is to create a positive *societal impact* through immersive VR games**. To achieve this, we presented a comprehensive investigation into how immersive VR games can address three significant contemporary societal challenges. We demonstrated the motivational power of these games to engage diverse user groups, including younger adults, older adults, and older adults with different degrees of dementia. Through these scientific contributions, we lay a foundation for further research in the field of HCI. Accordingly, I also advocate for the use of immersive VR games to create a positive societal impact for people of all ages and conditions. In the future, I want to extend this impact to further application areas, while preserving my inherent motivation: to create engaging, enjoyable, and (most importantly) beneficial experiences.

Part II

CORE PUBLICATIONS AND PREPRINTS

Immersive VR Games for Physical Activity

Immersive VR Exergames	Born to Run, Programmed to Play: Mapping the Extended Reality Exergames Landscape (CHI'24) [CorePub1]
Training in Immersive VR Exergames	Never Skip Leg Day Again: Training the Lower Body with Vertical Jumps in a Virtual Reality Exergame (CHI'23) [CorePub2]
Continuous Cues in Immersive VR Exergames	Move, React, Repeat! The Role of Continuous Cues in Immersive Exergames (CHI PLAY'24) [CorePub3]
Game Elements in Immersive VR Exergames	Unlocking the Potential: The Role of Game Elements on Player Motivation and Intention of Long-Term Use in Virtual Reality Exergames [PREPRINT] [CorePre1]

Immersive VR Games for Active Aging

Cognitive and Physical Exercises for Older Adults	Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults (CHI EA'22) [CorePub4]
Inclusivity for Older Adults	Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia (CHI PLAY'21) [CorePub5]
Augmentation for Older Adults	Evaluating Augmented Locomotion and Range of Reachable Objects for Older Adults in a Virtual Reality Exergame [PREPRINT] [CorePre2]
Exergaming Technologies for Older Adults	ExerCube vs. Virtual Reality: A Comparative Study of Exergame Technologies for Older Adults [PREPRINT] [CorePre3]

Immersive VR Games for Social Interaction

Social Multiplayer Immersive VR Games	VR Almost There: Simulating Co-located Multiplayer Experiences in Social Virtual Reality (CHI'23) [CorePub6]
Asymmetric Multiplayer Immersive VR Games	A Best-Fit Framework and Systematic Review of Asymmetric Gameplay in Multiplayer Virtual Reality Games (Frontiers in VR'21) [CorePub7]
Interdependence in Asymmetric Multiplayer Immersive VR Games	Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games (CHI'21) [CorePub8]
Familiarity in Asymmetric Multiplayer Immersive VR Games	Playing with Friends or Strangers? The Effects of Familiarity between Players in an Asymmetric Multiplayer Virtual Reality Game (CHI PLAY Companion'23) [CorePub9]

Overview of the structure of this dissertation. The icons for research contributions: survey, artifact, empirical, theoretical, and opinion contribution [227]. The icon indicates honorable mentions.

9 IMMERSIVE VR GAMES FOR PHYSICAL ACTIVITY

9.1 BORN TO RUN, PROGRAMMED TO PLAY: MAPPING THE EXTENDED REALITY EXERGAMES LANDSCAPE

[CorePub1] **Sukran Karaosmanoglu**, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. Born to run, programmed to play: Mapping the extended reality exergames landscape. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, (CHI '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 979-8-4007-0330-0/24/05. DOI 10.1145/3613904.3642124. URL https://doi.org/10.1145/3613904.3642124.

Supplementary materials for this article are available here: https://doi.org/10.1145/3613904.3642124.

Born to Run, Programmed to Play: Mapping the Extended Reality Exergames Landscape

Sukran Karaosmanoglu
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
sukran.karaosmanoglu@uni-hamburg.de

Sebastian Cmentowski
HCI Games Group, Stratford School of Interaction Design and Business, University of Waterloo
Waterloo, Canada
sebastian.cmentowski@uwaterloo.ca

Lennart E. Nacke
HCI Games Group, Stratford School of Interaction Design and Business, University of Waterloo
Waterloo, Canada
lennart.nacke@acm.org

Frank Steinicke
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
frank.steinicke@uni-hamburg.de

ABSTRACT

Many people struggle to exercise regularly, raising the risk of serious health-related issues. Extended reality (XR) exergames address these hurdles by combining physical exercises with enjoyable, immersive gameplay. While a growing body of research explores XR exergames, no previous review has structured this rapidly expanding research landscape. We conducted a scoping review of the current state of XR exergame research to (i) provide a structured overview, (ii) highlight trends, and (iii) uncover knowledge gaps. After identifying 1318 papers in human-computer interaction and medical databases, we ultimately included 186 papers in our analysis. We provide a quantitative and qualitative summary of XR exergame research, showing current trends and potential future considerations. Finally, we provide a taxonomy of XR exergames to help future design and methodological investigation and reporting.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; **Virtual reality**; **HCI theory, concepts and models**; • **Software and its engineering → Interactive games**.

KEYWORDS

extended reality, mixed reality, augmented reality, virtual reality, exercise, exergames, movement games, motion games, active video games, active games, sports games, games, review, taxonomy

ACM Reference Format:
Sukran Karaosmanoglu, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. 2024. Born to Run, Programmed to Play: Mapping the Extended Reality Exergames Landscape. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 28 pages. https://doi.org/10.1145/3613904.3642124

CHI '24, May 11–16, 2024, Honolulu, HI, USA

1 INTRODUCTION

Technological advances have changed our lifestyle in an unprecedented way. Today, many people spend most of their day in front of a computer, working in cognitively demanding but physically underwhelming jobs. According to the World Health Organization (WHO), 27.5% of adults and even 81% of adolescents do not meet the recommended physical activity levels [261]. This sedentary lifestyle affects our well-being: insufficient physical activity is associated with severe health issues, such as cardiovascular diseases, dementia, or depression [261]. Consequently, the WHO launched the Global Action Plan on Physical Activity (GAPPA) to raise awareness of the importance of regular exercise [260]. Unfortunately, establishing and maintaining a healthy, physically active lifestyle is challenging. It requires time and motivation to become a habit [194]. However, modern technology is not only the problem's cause but can also contribute to its resolution. A promising way of supporting people's efforts towards a healthier lifestyle is to use digital games: exergames. These games combine enjoyable gameplay with physical activities to achieve engaging experiences. Although this genre is not in its infancy anymore and includes several milestone games, such as Dance Dance Revolution [10] or Wii Fit [181], recent technical innovations have inspired a new wave of research on exergames.

The key to the next generation of exergames is extended reality (XR). XR is an umbrella term for virtual reality (VR), augmented reality (AR), augmented virtuality (AV), and mixed reality (MR) [91, 198, 203]. These technologies offer invaluable benefits to designing engaging and responsive exergames for full-body exertion. Above all, the unique advantage of XR is its spatial nature. Physical interactions in a three-dimensional (3D) environment are the foundation of most XR applications. Exergame developers can extend these interaction patterns to blend exercises naturally into the gameplay. Unlike previous technologies, modern XR systems feature built-in six degrees of freedom (6-DOF) tracking, which is often essential for exergames. Furthermore, view-dependent, stereoscopic images improve the perceived realism and immersion in the virtual environment, thereby increasing motivation in VR exergames [17]. XR technologies also grant full control over the players' surroundings, which can be used to create highly engaging scenarios that are impossible in the real world (e.g., superhuman powers [104]). Lastly, the full control over virtual worlds is ideal for customization and personalization: XR exergames can be tailored to meet users' exercise needs (e.g., to design exercise scenarios for people with dementia [112]). These advantages make XR an ideal fit to take fitness applications to the next level. At the same time, the apparent differences compared to more traditional exergames (e.g., free 360°view, limited proprioception, focus on motion-based interactions) create significant challenges and open questions that need addressing to ensure the games' safety, efficiency, and enjoyability.

The release of more broadly adopted, consumer-ready hardware, such as Meta Quest 2 [157] or Microsoft HoloLens 2 [160], has shifted focus to this promising domain. As a result, the field is producing plenty of promising new contributions. Despite the breadth of research, a comprehensive, systematic organization of the field is currently lacking. This creates a complex landscape; it is difficult to get a complete picture of the areas already covered and the open questions. This lack of structure hinders the information flow between different domains in this interdisciplinary field (e.g., movement and computer science). Further, it makes it difficult to identify research trends and promising directions and prevents researchers from following a systematic and efficient approach to advance our knowledge. Providing an organization of such a rapidly growing area is crucial and timely to help the field progress. By conducting a scoping review, we provide practitioners with invaluable resources for designing XR exergames. Most importantly, our results help researchers, especially those early in their careers, to think critically and identify promising research directions.

While previous research covered specific aspects of exergames, no prior review or taxonomy has been established for organizing the growing field of XR exergames. Existing frameworks of non-immersive exergames are not easily applicable to the XR domain due to the significant differences in design. While XR exergame designers can profit from increased immersion and enjoyment [17], they must consider unique challenges, such as safety concerns or tracking problems [44]. Prior reviews also do not include the latest advances in this rapidly evolving field (over 50% of our identified corpus was published since 2021). Accordingly, our goal is to create a taxonomy and analysis of XR exergames research using an adapted version of the foundational People-Activities-Contexts-Technologies (PACT) principles [12] of the human-computer interaction (HCI) field. With our work, we approach these concerns and contribute an up-to-date, comprehensive analysis of the existing body of XR exergame research. To answer our research questions, we conducted a scoping review and identified 1318 relevant papers. Following an eligibility step, we quantitatively (i.e., by frequency reporting) and qualitatively analyzed the final corpus of 186 papers. Focusing on Goals, People, Exercises, Design, and Technologies (GPEDT) principles (inspired by and adapted from PACT [12]), we created a hierarchical taxonomy of XR exergames following the taxonomy development method of Nickerson et al. [180].

With our taxonomy, we disassemble the current landscape of XR exergames into a complete set of dimensions. This taxonomy illustrates the primary goals and design elements explored by domain researchers and reveals trends and underresearched areas. We use these insights to derive nine central research directions that guide the design, implementation, and future study of XR exergames. Together, these contributions help researchers and practitioners gain an overview of the field, identify promising research questions, and establish interdisciplinary collaborations. Further, our analysis of the XR exergame corpus also highlighted another key challenge for future research: clear reporting of exergame design and study findings. We identified significant differences in terminology and reported information impeding comprehensibility, reproducibility, and transferability. To address this issue, we provide guiding questions that help in the systematic reporting of research on XR exergames. We believe that establishing reporting standards will contribute to clear communication in this young and emerging field. We hope to facilitate the compilation of existing knowledge and spark new research initiatives that will advance the capabilities of XR systems as a platform for gamified exercise. To summarize, our contributions are as follows:

- a comprehensive overview of the XR exergame research,
- a taxonomy of XR exergames based on ● Goals, ● People, ● Exercises, ● Design, and ● Technologies meta-characteristics,
- nine future research directions for the field, and
- a set of guiding questions to formalize the reporting of future XR exergame research.

2 BACKGROUND

This section provides an overview of exergames and XR, and explains how authors in previous literature combined those domains to create playful exercise applications.

2.1 Exergames

Research uses synonymous terms to describe applications that involve movement-related gameplay, such as exergames [112], exertion games [170], or motion-based games [78]. In this paper, we use the term "*exergames*" to refer to "*digital game[s] where the outcome [...] is predominantly determined by physical effort*" [170]. In our review, we consider any game that matches this specification and is defined as an exergame (or variant words), regardless of its physical activity level (e.g., breathing, high-intensity interval training).

Since the 1990s, exergames have been among the top games. In Dance Dance Revolution [10], players perform dance-inspired rhythmic stepping movements. Another notable example was the Wii console bringing various exergames, called WiiFit games [181], to people's living rooms. Later, Microsoft Kinect [159] introduced full-body motion-tracking as a novel input for commercial exergames. This fascinating coupling of digital games with exercising is moving to a new level with novel, immersive XR technologies.

2.2 Immersive Extended Reality

The reality-virtuality continuum [162, 163] is a popular taxonomy for XR-related term definitions. This continuum encompasses the physical real environment on one end and the digital virtual environment on the other. The real environment consists of only real-world objects, whereas the virtual environment (i.e., VR) includes only artificial objects. According to Milgram and Kishino [162], MR describes the blending of these two realities. However, according to Speicher et al. [231], experts differ in how they understand the term MR, with some using it following the explanation of [162, 163], while others using different meanings (e.g., strong AR). We follow the definition of [162, 163] and use the term MR as everything between reality and VR, including AR and AV. AR overlays virtual objects on real surroundings, while AV overlays real objects on virtual environments. To encompass all these concepts, we use the umbrella term XR. We note that other sources [204] do not understand XR as extended reality but as xReality, with X serving as a placeholder to denote different realities. However, the overall use as an umbrella term remains the same [91, 198, 203].

Unlike its colloquial use, XR nomenclature solely indicates the relationship between virtual and real content; it does not make assumptions about perceptual effects. To describe how these systems (partially) replace users' sensations with an artificial environment, researchers use the terms *immersion* and *presence*. Slater and Sanchez-Vives [229] define immersion as the technical quality of a setup. It is a continuum that depends on various characteristics such as stereoscopic vision, field-of-view, resolution, latency, or sensor substitution. These factors allow us to compare XR systems with one another. Slater and Sanchez-Vives [229] suggest that head-mounted displays (HMDs) are more immersive than CAVE systems because they can simulate the latter, but not vice versa. In general, an immersive system should at least provide a head-view-point-dependent, stereoscopic image of the virtual environment [72, 229]. For our review, we follow these definitions and consider any system with stereoscopic vision and an effective display area that exceeds the players' field of view without requiring input other than head rotation. These requirements can be achieved with spatially-tracked HMDs (i.e., VR HMD, AR HMD) or by surrounding the player with projection walls or displays (i.e., CAVE). However, single monitors, floor projections, or hand-held AR do not meet these criteria; thus we do not consider these to be immersive XR.

2.3 Immersive Extended Reality Exergames

A growing body of research is exploring the convergence of XR and exergames, highlighting the advantages of XR exergames. Born et al. [17] showed that VR exergaming provides higher motivation, embodiment, and performance than non-immersive exergaming. Similarly, Xu et al. [268] found that VR exergames improve player performance compared to playing the identical game on a large display. HMD-VR and CAVE-based VR exergaming increase flow and presence compared to non-VR exercise [218].

Immersive XR motivates, engages, and enables players to do activities. Ioannou et al. [104] showed that virtual augmentation of running and jumping contributes to intrinsic motivation. Similarly, Born et al. [20] found that players perform longer voluntary strenuous activities if they experience augmentation. Further, VR's realism can also be used to build cognitively and physically stimulating exergames for people with dementia [112].

With new technological developments—HoloLens 2 released in 2016 [160]—immersive AR has begun to be used for exergaming [69, 207, 273]. According to a recent study, AR exergaming can lead to a significantly lower level of collision anxiety (i.e., being aware of surroundings) compared to its VR counterpart. This unique AR feature—blending virtual and real environments—has also increased exergame advancements in rehabilitation [69, 89, 248].

2.4 Existing Literature Reviews & Taxonomies

Increasing interest in XR-based exercising has led to more recent review articles. While some of these works focus on specific subtypes or use cases of XR exergames, no work has yet provided an up-to-date overview of this rapidly evolving domain. Some reviews focused on general physical activity, not particularly on exergames (e.g., [79, 184]). Other reviews covered non-immersive "VR" exergames (e.g., [36, 45, 49]) and examined their health-related outcomes (e.g., [49, 141, 165]). For instance, Mo et al. [165] concluded that exergames are overall safe but not significantly effective for musculoskeletal pain in older adults. Kappen et al. [110] identified various focuses of non-immersive exergames (e.g., cognitive training) for older adults, which we incorporated in the goal dimension of our taxonomy. Lastly, only a few papers considered the intersection of XR technology and physical activity (e.g., [79, 184]). Odenigbo et al. [184] reviewed 39 VR, AR, and MR physical activity interventions and showed that most of them included exergaming. Only one paper [247] provided a narrative review on the overlapping space of XR and exergames, featuring 29 HMD-VR "*health games*". The authors found that most games used obstacle-based gameplay and extrinsic rewards. However, this review focused on health-centered exergames and HMD-based VR, without covering the broader XR field or providing a taxonomy of XR exergames.

Structuring research on sports systems is an ongoing effort in HCI research. Reilly et al. [206] classified computer-augmented sports systems based on two high-level dimensions: form and function. While the function dimension contains the system's purposes and abilities (e.g., sports entertainment, refereeing), the form dimension concerns its implementation (e.g., hardware, software). Similarly, Frevel et al. [71] present a *SportsTech Matrix*, considering sports and technology from two angles: The user angle captures user groups interacting with sports technology (e.g., athletes, consumers), whereas the tech angle comprises technology used with/for sports. Inspired by Reilly et al. [206], Postma et al. [200] provide a taxonomy of sports interaction technology to bridge sports science and HCI. For instance, they introduce new forms of sports interaction technology relating to space, time, game nature, feedback, and integration of interaction. Although these taxonomies hold importance for advancing the field, they do not particularly focus on exergames. Additionally, not all exergames match the characteristics of sports according to Jenny et al. [107]'s definition; for example, an exergame does not need to "*include competition (outcome of a winner and loser)*". Similarly, a sport can be supported by technology without requiring additional gameplay.

3 RESEARCH FOCUS: WHAT WE ADD TO THE LITERATURE

Likely most related to our research is the work by Mueller et al. [171], who approach the field through a social lens and present a taxonomy with four dimensions: non-exertion vs. exertion, non-competitive vs. competitive, non-parallel vs. parallel, and combat vs. object. For example, the exertion dimension provides an understanding of what an exergame is (in line with the definition we followed in [170]), whereas the competitive unit covers exergames featuring opponents. While this taxonomy helps to define social aspects in exergames, it does not provide classifications on other aspects and does not target XR technology. However, XR exergames have unique advantages (e.g., immersion, real-time feedback) and challenges (e.g., safety, technical constraints) [44]. The unique potential of such applications—offering engaging and motivating exercises—and continued advances in XR hardware motivate additional concentrated research. We believe a structured review identifying trends and knowledge gaps could strengthen community efforts.

We comprehensively review exergames that use XR technology (e.g., CAVE-VR, AR) using a scoping review, and also include non-health exergames to provide an extensive picture of the domain through an HCI lens. With our review, we first provide a summary of XR exergame research (e.g., studies). Based on our review, we then give a quantitative and qualitative analysis of the XR exergames featured in research and create a hierarchical taxonomy. We base this step on the established PACT framework [12]. This framework serves as a guide for creating interactive systems using a human-centered design (HCD) perspective [68]. Since XR exergames are highly interactive and humans are at the core of these systems, we used the PACT framework elements as an inspiration. However, we adapted them to fit the specific characteristics of exergames. We split the Activities dimension into Goals and Exercises since exergames typically have an overarching interventional goal (e.g., rehabilitation) and specific exercises that contribute to this purpose[1]. Also, we renamed Contexts into Design to better reflect which aspect of the exergame we aimed to capture. In total, our GPEDT consists of the five principles: ● Goals, ● People, ● Exercises, ● Design, and ● Technologies. Overall, we summarize our research questions motivating the taxonomy as follows:

(RQ1) ● **Goals**: What are the goals of the XR exergames?
(RQ2) ● **People**: Which different user groups are usually targeted by XR exergames?
(RQ3) ● **Exercises**: What kinds of exercises are being designed in the XR exergames?
(RQ4) ● **Design**: What kinds of game design aspects are considered for XR exergames?
(RQ5) ● **Technologies**: What kinds of technologies are being used in the XR exergames?

4 SCOPING REVIEW OF EXTENDED REALITY EXERGAMES

To assess current XR exergame contributions, we conducted a scoping review [173, 243]. Scoping reviews provide an overview of a specific problem and identify potential research directions [173]. These reviews are precursors to systematic reviews and do not involve a critical appraisal stage (typical for systematic reviews) to evaluate the methodological quality of articles [173]. For our review, we follow the recommendations of the PRISMA extension for scoping reviews (PRISMA-ScR) [249] and best practices of Peters et al. [196, 197]. Our main steps for the scoping review are (i) identification of the corpus, (ii) screening, (iii) eligibility, (iv) data extraction, (v) data synthesis, and (vi) reporting (see Figure 1).

[1]We note that some may see exercises as part of the design. However, since exercises are the critical components of exergames and correspond to activities within the PACT framework, we keep this aspect separate.

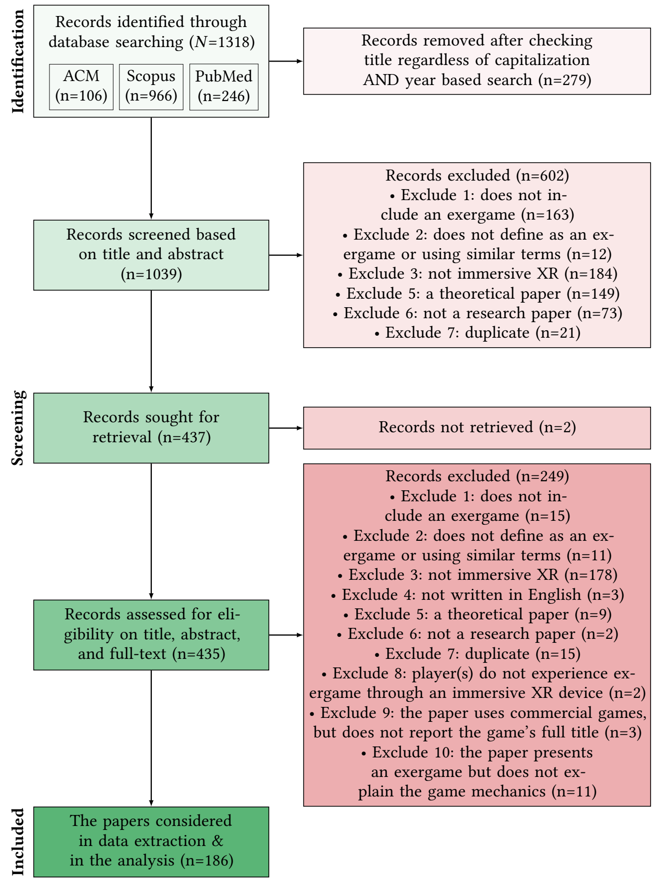

Figure 1: We illustrate our scoping review process in a PRISMA flow diagram to detail our steps [186].

4.1 Protocol, Databases, & Search

Following the best practices of conducting a scoping review [196, 249], we prepared a protocol that reports our steps (see supplementary materials). We decided on Association for Computing Machinery (ACM)[2], Scopus (indexing IEEE Xplore and many other HCI-related sources)[3], and PubMed[4] as our databases because XR exergames is an interdisciplinary research area and mostly HCI and medical science scholars publish in our selected databases.

We performed informal searches on these databases to uncover exergame definition phrases and include less prevalent synonyms like motion-based games or movement-based games. We combined all identified phrases in our search query using `OR` operators. XR-related keywords were decided based on Milgram et al. [163]'s reality-virtuality continuum. We also used related phrases like immersive or virtual environments. Similar to the exergames-related keywords, we concatenated all keywords using `OR` operators. Consequently, our search query contained XR-related keywords `AND` exergames-related keywords (see Table 1). We repeated our search

[2]https://dl.acm.org/
[3]see the Scopus indexed venues: https://www.scopus.com/sources
[4]https://pubmed.ncbi.nlm.nih.gov/

Table 1: This table shows the XR- and Exergames-related keywords. Within the same group of keywords, we used the `OR` operator, while for between the keywords groups, we used `AND` operator. The final search was conducted on 11 August 2023 in ACM (The ACM Guide to Computing Literature), Scopus, and PubMed digital libraries.

XR-related keywords		Exergames-related keywords
("immersive" OR "VR" OR "AR" OR "AV" OR "MR" OR "XR" OR "virtual realit*" OR "augmented realit*" OR "extended realit*" OR "mixed realit*" OR "augmented virtualit*" OR "virtual environment*")	AND	("exergame*" OR "exercise game*" OR "physical game*" OR "movement game*" OR "motion game*" OR "motion-based game*" OR "movement-based game*" OR "training game*" OR "exertion game*" OR "sport game*" OR "sports game*")

query in the selected databases on different days to double-check for inconsistencies within the databases [149, 212]. We provide our exact search queries for each database in the Appendix (see Table 3); the final search was conducted on August 11, 2023.

During data collection, we considered reviewing commercial games as well. However, we deliberately decided against this step because of XR's spatial interaction nature: Many VR games require at least a little physical effort to complete tasks. Still, this does not mean that they are designed as exergames [276]. Thus, we let XR exergame researchers decide: if a commercial game is considered an exergame in any research paper, we included it in our corpus.

4.2 Identified Corpus & Duplicate Removal

Our search query had no constraints (e.g., publication date) because we intended to cover the current state of XR exergames research [173]. Our search yielded 1318 papers (The ACM Guide to Computing Literature = 106, Scopus = 966, PubMed = 246). After the identification of articles, we removed duplicates testing for matching year and title (regardless of capitalization) using Python scripts (n=279), leaving a total of 1039 articles for the screening phase. Please refer to our supplementary materials for all papers.

4.3 Screening

For the screening step, we first decided on inclusion criteria (IC):

(1) IC1: The paper includes an "exergame", "exercise game", "exertion game", "physical game", "movement game", "motion game", "motion-based game", "movement-based game", "training game", or "sport(s) game" application.
(2) IC2: The paper defines the application as an "exergame", "exercise game", "exertion game", "physical game", "movement game", "motion game", "motion-based game", "movement-based game", "training game", "sport(s) game" or using similar terms (e.g., rowing game, table tennis game)[5].
(3) IC3: The exergame is an immersive XR exergame[6].
(4) IC4: The paper is written in English.
(5) IC5: The paper is not a theoretical paper (e.g., a literature review, a paper that does not include any self or commercial exergame implementation or its testing)[7].
(6) IC6: The paper is a research work (e.g., not proceedings preface).
(7) IC7: The paper does not have duplicates in the corpus.

[5]Without limiting the corpus to papers using the term exergame or variant words, we would have had to include almost every XR experience, given that even very simple interactions in spatial XR environments require some physical effort.
[6]Following the definition of immersive XR (see subsection 2.2)
[7]The paper does not have to include an evaluation (e.g., implementation-only papers are also included in our review).

Using these inclusion criteria, two authors screened the corpus based on the titles and abstracts using the software *Dovetail*[8]. Some papers, such as [47], had no abstracts, so we screened them based on their title and introduction. To screen the papers, two authors first reviewed the same ∼ 15% (N=156) papers individually. Then, we conducted Cohen's kappa test [73] to assess the inter-rater reliability between the two coders' binary decisions, which indicated almost perfect agreement (%94 of papers had agreed on, κ=0.872, z=10.9, p<0.001); for the disagreements, one additional author acted as a tie-breaker in case the screening authors did not reach consensus after discussion, but this was never the case. Since we had almost perfect agreement [133, 154], we split the remaining corpus between two coders, who screened the articles independently (one N=441, N=442). Such practices are common (e.g., [212]), and agreement strategies have been used in HCI literature to divide the data set between multiple coders [153] to facilitate the coding process.

We also note that we excluded some articles during our screening phase that met our inclusion criteria because they did not use the required terms in the intended sense (e.g., training game for beekeepers [119] and physical game to indicate that the game has played in a physically shared environment [121]). If cases were not clear, we included those papers to avoid missing any relevant work, such as work using the Kinect for interactions in VR (e.g., [46, 92, 178]) since this technology might be used as a tracking system with immersive XR. After completing the screening, we kept 437 papers to be checked in the eligibility step.

4.4 Eligibility

For eligibility, we tried accessing all papers resulting from the screening (n=437). However, we could not retrieve two papers. Therefore, in this step, we reviewed a total of 435 papers based on their title, abstract, and full-text. In addition to the screening inclusion criteria, we also applied the following exclusion criteria (EC) for the eligibility check:

(1) EC8: Player(s) do not experience the exergame(s) through an immersive XR device.
(2) EC9: The paper uses commercial games without reporting the games' full titles.
(3) EC10: The paper presents an exergame but does not explain the game mechanics, i.e., which movements are performed.

Similar to the screening step, two authors coded ∼15% of papers (n=66) independently. We used this step to check the inter-rater reliability between the two coders for the binary eligibility decision (agreement for 97% of papers, κ= 0.939, z=7.64, p<0.001). For any

[8]https://dovetail.com/

disagreements, we followed the same strategy from the screening step. Ultimately, we divided the remaining papers (n=371) between the same two authors (one n=185, n=186) because the inter-rater reliability showed almost perfect agreement [133, 154]. Based on our ICs and ECs, we excluded 249 articles. We note that we observed some edge cases that resulted in exclusion: papers that include exercises in XR worlds but do not define the application as an exergame (e.g., [83, 136]), papers that include a CAVE-like interface but do not offer a stereoscopic view (e.g., [13, 151]), and papers that define their game as exergame but do not describe the included movements (e.g., [259]). Overall, this step resulted in 186 papers (see Appendix Table 5 and 6) being included for data extraction.

4.5 Data Extraction

We created a data extraction form using Airtable[9] to extract the data from the included full-text papers. Initially, the lead author created a first version of the form, which was reviewed by the other co-authors. We then extracted the data for the first ten random papers and used the process to refine our form iteratively. Afterward, we applied the finalized data extraction form to all papers by dividing the remaining corpus between two coders (n=88, n=88). The final form had three parts (we supply the form and the extracted full data in the supplementary materials):

(a) *General information.* First, we extracted general details of each paper (i.e., title, authors, publication year, and main objective).
(b) *Study details.* Next, we assessed information on the conducted studies, focusing on the study design (i.e., type, independent and dependent variables) and recruited participants (i.e., number, sample details). If publications had multiple studies, we completed the form once for each. We only considered studies where participants actively played at least one version of the exergame and provided feedback. For example, we would complete the form separately for a multi-session HCD process and a subsequent independent user study of the final prototype (e.g., [221]). In contrast, we excluded purely exploratory gameplay sessions of arbitrary VR games (i.e., not exergames) or domain expert interviews preceding the actual implementation phase (e.g., contextual inquiry [112]). We also did not extract information about experts' feedback, where those did not actively play the games but only reacted to video material (e.g., [98]).
(c) *Information about the used/developed exergame.* Lastly, we extracted details on the featured exergames, considering only commercial or custom exergames using immersive XR. Apart from general information (i.e., number, names, and duration of games), we mainly focused on GPEDT-related information.

Overall, this process led to a total of 200 filled forms (i.e., studies) for 186 papers, which can be seen in the supplementary materials. During the data extraction, we encountered papers using commercial exergames. For these games, we performed an additional data extraction step and gathered information from the following resources: Steam[10], Meta Store[11], PlayStation Store[12], and the game publisher's website. We only used this official information because (i) authors typically include only relevant information for their study, and (ii) the publisher is the primary source for their game.

[9]https://airtable.com/
[10]https://store.steampowered.com/
[11]https://www.meta.com/experiences/
[12]https://store.playstation.com/

4.6 Synthesis

For our corpus, we used two types of analysis method: frequency-based reporting and taxonomy development methodology [180].

4.6.1 Frequency-based Analysis. Our goal with this analysis was to provide an overview of existing XR exergames research, covering published papers, conducted studies, and featured games. This information provides an understanding of the current state, a reference point for future studies, and can help identify research gaps.

4.6.2 Taxonomy Development Methodology. To create a hierarchical taxonomy of XR exergames applications, we followed the principles of the widely used taxonomy development methodology of Nickerson et al. [180], who define taxonomies as "*systems of groupings that are derived conceptually or empirically*". A taxonomy should be *useful*, which means it should be *concise*, *robust*, *comprehensive*, *extendible*, and *explanatory*. Nickerson et al. [180] recommends a set of objective ending conditions to create a taxonomy iteratively. From those, we used the first ending condition for our taxonomy ("*All objects or a representative sample of objects have been examined*") and we analyzed all identified papers in our scoping review iteratively (see below for more details).

To start the taxonomy development, researchers should decide on *meta-characteristics*, which are "*the most comprehensive characteristic[s] that will serve as the basis for the choice of characteristics in the taxonomy*" [180]. Based on these meta-characteristics, researchers develop dimensions and characteristics within each dimension. Each dimension is *mutually exclusive* (i.e., "*no object can have two different characteristics in a dimension*" [180]) and *collectively exhaustive* (i.e., "*each object must have one of the characteristics in a dimension*" [180]). We note that we chose not to follow the mutual exclusivity and collective exhaustivity aspects (similar to other works [90]), as they may result in the loss of details (e.g., XR exergames might be designed for multiple user groups), and not every game description sufficiently reports the necessary details.

Meta-characteristics & GPEDT. According to Nickerson et al. [180], the choice of meta-characteristics "*should be based on the purpose of the taxonomy*". The objective of our taxonomy is to systematically analyze, organize, and map the current XR exergame landscape through the lens of HCI. Accordingly, we based our choice on the PACT framework and our prior domain knowledge. We decided on our final meta-characteristics based on a discussion between four researchers, resulting in ● Goals, ● People, ● Exercises, ● Design, and ● Technologies (GPEDT). ● Goals refers to the purpose of the designed/used XR exergame(s). ● People considers any information about the target audience of the XR exergame(s). ● Exercises focus on information about the exercises included in the XR exergame(s). ● Design is concerned with any game design information relating to the XR exergame(s). Finally, ● Technologies captures information about the technology (hardware, devices) used in XR exergame(s).

Approach. For the taxonomy development, the lead author first created deductive codes based on our meta-characteristics: For

Goals, we consulted various resources to derive our deductive codes [110, 175] (e.g., cognitive training). For People, we created deductive categories that capture the users' age groups (e.g., older adults) and, if applicable, their clinical conditions (e.g., dementia). Exercises covered three categories (e.g., body parts) containing different codes (e.g., full body). Design focused on four overarching categories like task adaptation, each with their own codes (e.g., physiological measure-based adaptation). Finally, Technologies included three deductive categories (e.g., display) and corresponding codes (e.g., VR HMD).

For coding, we followed a hybrid approach. Our strategy involved the aforementioned deductive codes and data-driven inductive codes. First, two authors randomly selected ten exergames to code iteratively using the deductive categories and codes. This step attempted to produce uniformly understood codes. After each paper, we reviewed our codebook understanding in a meeting. If necessary, we refined and created new inductive codes. For example, although we used deductive codes from Kappen et al. [110] for the goals dimension (e.g., rehabilitation), we crafted new codes (e.g., preservation) for a more nuanced understanding of the aim and design of exergames. After creating a shared understanding, we assigned the remaining dataset to be coded by the two coders independently. Both authors met multiple times to discuss challenging topics and newly arisen codes. They refined their codebook during the process and established coder consistency if needed. The full-text list of our codes is in our supplementary materials.

After coding, we created digital notes from our final codes and performed affinity mapping (see the supplementary materials) with the two researchers using a Miro Board[13]. After establishing an initial taxonomy, we discussed it with a third author with games research expertise to finalize the dimensions and characteristics.

Overall, we followed both *conceptual-to-empirical*—employing preexisting knowledge to a set of data– and *empirical-to-conceptual*—creating a taxonomy based on a set of data— approaches to create the taxonomy [180]. We consider our priori focus of GPEDT as conceptual-to-empirical since the GPEDT concept was created based on the PACT framework and the research team's prior domain knowledge to ensure a clear focus on the essential elements from an HCI lens. The remaining steps of taxonomy creation (e.g., characteristics) followed an empirical-to-conceptual approach; the final taxonomy, including subdimensions and characteristics, arose through the scoping review and iterative data-driven approach.

We provide a qualitative summary to explain our taxonomy and characteristics. Here, we used the codes from our codebook to produce summaries. Further, we provide a positionality statement on the authors' background [22, 179]. Even after inter-rater-reliability steps and unifying the codes between the coders, the involved researchers might still affect the interpretation of results. The two authors who conducted the analyses have worked on the intersection of XR and exergames technologies for several years; they published papers in the XR exergame field, studied different XR technologies, and played and implemented XR exergames. The other two involved authors, working on games and XR respectively, added their own perspectives to the discussion and analysis. We also list the potential implications of this bias in our limitations.

[13]https://miro.com/

5 RESULTS OF SCOPING REVIEW ON EXTENDED REALITY EXERGAMES

Here, we report the quantitative findings of our scoping review before explaining our hierarchical taxonomy in section 6. We included a total of 186 articles ranging from 1998 to 2023 in our corpus. The steep increase visible in Figure 2—over 50% of the papers were published in the last three years—shows that XR exergame research is booming and producing new publications rapidly.

5.1 XR Spectrum

The majority of research focused on PC-tethered VR headsets (n=136). This group is dominated by the HTC Vive (including Pro, n=80), Oculus Rift (n=39), and Valve Index (n=5). Other headsets, such as the HP Reverb, Lenovo Explorer, or PlayStation VR, were only used in one paper each. Although much younger than PC-VR, mobile VR headsets take second place (n=42). Besides the popular Meta Quest platform (comprising Meta Quest 1 and 2, n=33), eight papers also used older 3DOF headsets like Google Cardboard, Oculus Go, or Samsung Gear VR. In contrast, projection-based VR systems, like CAVE, were only featured in seven publications. Lastly, nine papers explored AR exergames, utilizing Microsoft's HoloLens (n=6), the pass-through functionality of the HTC Vive Pro (n=1), or a custom-built solution (n=1).

5.2 Study & Sample Characteristics

Thirty-three papers (17.74%) presented only an exergame implementation. Of the remaining 153 papers, 9 conducted an HCD/iterative design approach with, on average, 15.5 participants (SD=10.27, one not reported). Four papers complemented the HCD with a final user study to test their product. Except for one publication, all HCD studies targeted older adults or people with dementia. Unfortunately, only four of these HCD papers reported sufficient demographic data: the 80 participants in these publications were mostly older female adults with a mean age of 80.17 (SD=9.05, woman=63, man=17).

Apart from the HCD approaches, we extracted 158 evaluatory user studies. Although most papers only covered one user study, eight papers featured two or even three studies. The 158 studies included, on average, 24.63 participants (SD=26.50, three not reported) and covered a broad range from single-participant case studies to large evaluations reaching 250 players.

Unfortunately, many papers miss crucial demographic information. Even after recovering the missing data to the best of our capabilities (i.e., performing age merge of different groups tested in the papers), only 83 studies (54.25%) report both the mean and standard deviation of the sample population's age. Across these studies, the mean age was 31.67 (SD=19.76, N=2166, see Figure 3).

Similarly, gender distribution was only fully reported in 106 cases (67.09%). The sample was slightly skewed, with 53.70% men compared to 46.30% women (N=2421, no non-binary or other).

For 65 studies, papers reported the sample culture: most studies were conducted in Europe (UK: 12, Germany: 7, Norway: 4, Spain: 4, Finland: 2, Greece: 2) and North America (USA: 15, Canada: 1). Only nine studies were conducted in Asia (China: 3, Japan: 2, South Korea: 2, Taiwan: 1, Malaysia: 1) and Oceania (Australia: 7, New Zealand: 2). Africa is only represented by one study [275] that was run simultaneously in the USA and Nigeria.

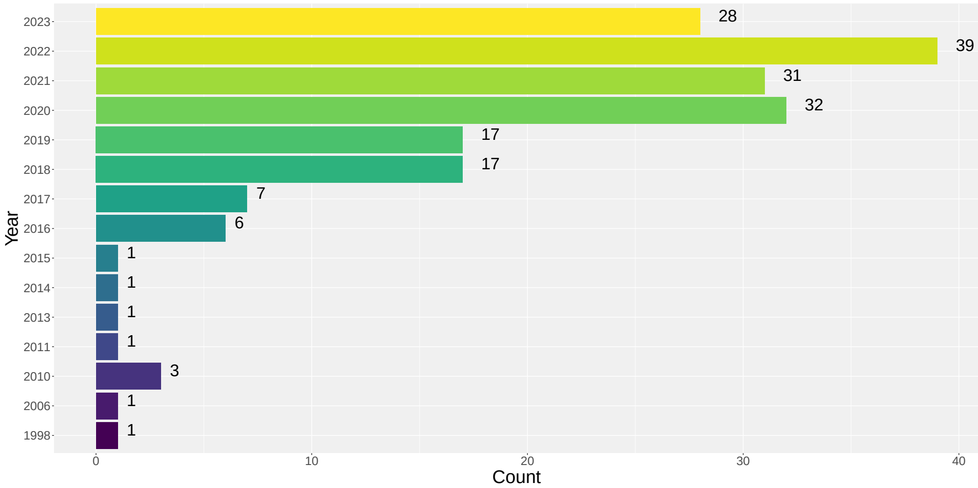

Figure 2: The distribution of our corpus (N=186) papers by year. We see a steep increase in published XR exergame papers.

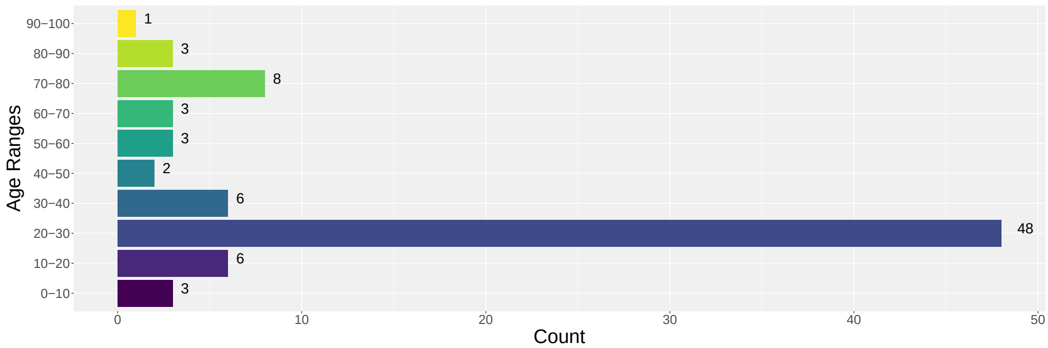

Figure 3: The average age distribution of our corpus' studies (n=83) by age range group. We see that most studies included participants with an average age range of 20-30 (n=48).

5.3 Study Design & Data Collection

Our corpus contains 158 user studies. In 79.11% of the cases, players experienced the exergame only once (n=125). Only 33 studies featured repeated play sessions, ranging from two runs on consecutive days to long-term exercise programs. To understand what the included papers explored in their evaluations, we analyzed the independent variables. Fifty studies evaluated factors relating to *time*, such as differences between pre- and post-test scores (n=29), two consecutive play sessions (n=5), longer intervention periods (n=9), or exercise durations (n=1). Similarly, *game design*-related studies (n=47) received much attention. A notable example is incorporating gameplay elements into physical exercise (n=18). Others explored the influence of avatars, music, haptic elements, or narratives. Other common topics include comparisons between different *platforms* (n=24, e.g., VR vs. non-VR), *exercise-related factors* (n=18, e.g., changes in difficulty), or the use of *non-player characters* (n=11, e.g., for competition). Surprisingly, only a few studies compared different *age and user groups* (n=8). The remaining categories are similarly underrepresented: *augmenting movements* in XR (n=8), personalizing the exergame experience (n=6), creating *multiplayer* experiences (n=5, i.e., multiple users play together), or providing visualized *feedback* (n=3).

To explore the outcomes of independent variables, the studies employed a variety of subjective and objective measures. Above all, 79.11% of the studies included quantitative subjective metrics (e.g., questionnaires) to assess players' experience (n=125). The most commonly covered aspects are cybersickness (Simulator Sickness Questionnaire (SSQ) [120], n=31), intrinsic motivation (Intrinsic Motivation Inventory (IMI) [214], n=23), perceived physical exertion (Borg Rating of Perceived Exertion Scale [16], n=21), game experience (Game Experience Questionnaire (GEQ) [103, 108], n=19), and usability (System Usability Scale (SUS) [23], n=19). To quantify the effects of exergame interventions, authors recorded gameplay and movement data (n=53) and assessed physiological measures (n=60, e.g., heart rate). Few studies also relied on standardized physiological/cognitive tests (n=16), such as the Stroop test (n=5). Lastly, qualitative data was collected for 58 studies, typically through interview sessions (n=31) or open-ended questions (n=24).

Goals						
Solely Movement	Promoting Physical Activity	Cognitive Training	Physical Training	Medical-Focused Training	Social Interaction	Psychological Factors

Figure 4: Our derived taxonomy's Goals dimension – with a total of seven characteristics.

5.4 Commercial vs. Custom XR exergames

We faced three challenges in calculating the total number of featured XR exergames in extracted papers: (i) Some articles featured similar games, sometimes using the same name and similar visuals (i.e., we checked the figures). We considered these games to be unique because there were some adjustments to their design and implementation. (ii) Some articles included commercial XR exergames, which we considered only once. (iii) Some papers included systems/games that include multiple mini-games (e.g., Nvidia VR Fun house [242], FitXR [67], [80, 88, 166]); in these cases, we also treated these systems as one. Based on these criteria, we found 195 games: 20 commercial games (see Table 4) and 175 custom games.

6 TAXONOMY OF EXTENDED REALITY EXERGAMES

We aimed to provide a taxonomy that guides the design, implementation, and research of future XR exergames. As our data unit, we consider all custom-built and commercial exergames as long as they were used in one of the papers. By analyzing the identified games according to the GPEDT aspects, we derived a taxonomy that is *concise*, *robust*, *comprehensive*, *extendible*, and *explanatory* [180]. Since some dimension characteristics are not mutually exclusive, the reported numbers of games do not necessarily sum up. For example, the safety equipment characteristic comprises both stabilization harnesses (*n*=3) and monitoring devices (*n* = 7). Since one specific game [166] requires both a harness and a monitoring screen, the broader safety category only contains nine distinct games. In the following, we describe our taxonomy's dimensions and characteristics, and provide a qualitative summary of the corpus (see supplementary materials for the complete taxonomy illustration).

6.1 Dimension I: Goals

The first dimension of our taxonomy is Goals, explaining the inherent purpose of the XR exergames. For this dimension, we identified a set of seven distinct characteristics. Since many applications target multiple goals, this dimension is not mutually exclusive, i.e., XR exergames can have more than one goal. Our taxonomy's goal dimension is illustrated in Figure 4.

Overall, we identified 129 games that reported their purpose. Unsurprisingly, many XR exergames drive for physical training (*n*=39); some games aim for endurance training (*n*=14) and involve high-intensity interval training protocols, such as [227, 235]. Others offer strength training for their users to improve muscular power (*n*=7). The remaining types of physical training targeted by XR exergames are balance (*n*=8, [21, 168]), skill (*n*=8, [32, 282]), flexibility (*n*=4, [31, 238]), and coordination/reaction training (*n*=7, [7, 139]).

Medical-focused training is another essential characteristic of our goal dimension (*n*=52). A notable number of XR exergames (*n*=31) aim to provide rehabilitation opportunities for their users (e.g., [89, 166, 245]). Other exergames (*n*=14) are designed as prevention/preservation tools; these games aim to prevent a deterioration of people's conditions and preserve their current state of abilities. Some games (*n*=8) are used to monitor/test the current abilities of users (e.g., range of motion [190] or reactive control [222]). Finally, XR exergames are also used to train for everyday activities (*n*=5).

Despite being less common than the first two characteristics, social interaction is still an important goal of many analyzed games (*n*=15), which aim to connect multiple players in the virtual world. The psychological factors (*n*=10) characteristic comprises games that target the psyche by improving psychological wellbeing (*n*=3, e.g., "*to face depression effect*" [177]) or by offering relaxation (*n*=7, "*control of breathing: relaxation/mindfullness training*" [3]). Furthermore, cognitive training is targeted in 14 games, e.g., [80, 111].

Another prevalent goal was to promote physical activity (*n*=57). This characteristic represents exergames designed to motivate users to perform more physical activity, for example: "*The main focus of the exergame is to motivate full body movements to promote exercise [...]*" [238]. Interestingly, in one game, the aim was "*having the purpose of no purpose*" [148]; the authors argued that moving is fun on its own without needing "another" purpose. We feature this intrinsic motivation in the characteristic solely movement.

6.2 Dimension II: People

People are the core elements of interactive systems. In the design of XR exergames, we identified two dimensions that concern people.

6.2.1 Subdimension: Target Age Range. This subdimension represents the target age group. Only 51 exergames provide information about their target age group range. Again, this subdimension is not mutually exclusive, i.e., a game can target different age groups.

This subdimension consists of a total of six characteristics. Nine games are designed for children (e.g., [40, 187]). At the same time, adolescents [59] and young adults [3] are only targeted by one XR game each. Although we have seen the term "younger generations" or "young people" in the descriptions of a few games [25, 109], we did not code them for this category. It was unclear what exactly the younger generation refers to; is it children or younger adults? While two games [109, 268] are particularly designed for middle-aged adults, the bulk of age-targeting exergames focuses on older adults (*n*=31). Lastly, eight games aim to be inclusive to all ages. Figure 5 shows our target age range subdimension of our taxonomy.

6.2.2 Subdimension: Condition. This subdimension maps people's medical conditions (i.e., disabilities and impairments) into characteristics (see Figure 5 (right)). As one game can be designed for

People					
Target Age Range					
Children	Adolescents	Younger Adults	Middle-aged Adults	Older Adults	All Ages

People		
Condition		
Developmental Disorders/Impairments	Neurological Disorders/Impairments	Physical Disorders/Impairments

Figure 5: (left) Our derived taxonomy's People subdimension: Target Age Range – with a total of six characteristics. (right) Our derived taxonomy's People subdimension: Condition – with a total of three characteristics.

Exercises			
Body Part			
Lungs	Upper-body	Lower-body	Full-body

Exercises			
Position			
Seated	Standing	Mixed seated/standing	Supports both seated/standing

Figure 6: (left) Our derived taxonomy's Exercises subdimension: Body Part – with a total of four characteristics. (right) Our derived taxonomy's Exercises subdimension: Body Part – with a total of four characteristics.

people with different health conditions, one game could have more than one condition characteristic.

In our corpus, 138 games do not report their user groups' condition, which leads to only 57 remaining XR exergames falling into one of three categories. First, the developmental disorders/impairments characteristic (*n*=8) describes people who typically have difficulties with attention, learning, or using certain skills, such as language. In our corpus, XR exergames primarily target people with attention deficit hyperactivity disorder (*n*=4, [40, 187]), autism (*n*=3, [60, 61, 63]), and intellectual disabilities (*n*=1, [38]).

Neurological disorders/impairments (*n*=26) covers all neurological conditions, such as people with dementia (*n*=11, [65, 174, 208]), Parkinson's disease (*n*=6, [248, 279, 280]), and ataxia (*n*=2, [69, 70]).

The final characteristic focuses on physical disorders/impairments (*n*=24), for example, upper limb conditions (*n*=4, [56, 89, 281]), hypertension (*n*=3, [24, 235], and neck/back pain (*n*=2, [161, 233]).

6.3 Dimension III: Exercises

Exercises dimension provides details on the movements performed in the XR exergames. We derived two subdimensions focusing on the targeted body part and supported play position to provide a comprehensive look at the movement aspect of exergames.

6.3.1 Subdimension: Body Part. This subdimension categorizes the XR exergames according to the body parts involved in the exercises (see Figure 6 (left)). This subdimension is mutually exclusive, i.e., every game has one of the characteristics of the subdimension.

Most games (*n*=92) feature full-body movements involving both the upper and lower body, e.g., by combining punches and squats. Secondly, 72 exergames focus only on upper-body training, which we define as movements of the upper limbs, head, or torso: "*After discussions with exercise therapists and considering safety, we decided to focus on upper body motions, including hand, shoulder, and head motions [...]*" [52]. In contrast, fewer exergames (*n*=26) integrate solely lower-body movements, i.e., using the lower extremities, into their gameplay: "*A player moves through the scene by physically walking in the game area, a virtual narrow, winding platform over a virtual river of lava overlaid onto an empty hall or room*" [117]. Finally, some exergames (*n*=5) also provide exercises involving the lungs. For example, in *Focus Tree*, "*On inhaling, clouds cover the island to block the view of players; and on exhaling, clouds get blown away allowing players to view the island*" [195].

6.3.2 Subdimension: Position. With this subdimension, we captured in which configuration the XR exergames are played: seated or standing. Determining this characteristic for every game was often challenging, since many games did not clearly state this information. Although many XR experiences could be playable in both configurations, we cannot simply guess without sufficient information on the implementation. Therefore, we decided to use the available sources in the following order: (i) review description, (ii) figures, (iii) supplementary materials, and (iv) available videos of the games. Based on this strategy, every game was coded under one characteristic (see Figure 6 (right)).

It was possible to retrieve the necessary information only for 187 games. Of these cases, 91 games (e.g., [44]) feature exercises to be performed in a standing position. A similar number of exergames (*n*=82) supports playing in a seated position: e.g., [34, 116, 158]. Only two games ([140, 266]) included mixed positions, i.e., players changed between those options to complete the entire game. Finally, 12 games supported both seated and standing position gameplay: these games were mainly commercial games, e.g., [74, 145].

6.4 Dimension IV: Design

This dimension, featuring four subdimensions, focuses specifically on the game design.

6.4.1 Subdimension: Theme. The theme aims to describe the environment and inspiration of the visuals and gameplay of the exergame design. Our analyzed XR exergames fall into one of seven distinct characteristics to represent their themes (see Figure 7 (left)).

Design						
Themes						
Exercise	Sport	Daily	Historical	Sci-fi	Fantasy	Abstract

Design		
Player Mode		
Single-player	Multiplayer	Supports both single-player/multiplayer

Figure 7: (left) Our derived taxonomy's Design subdimension: Themes – with a total of seven characteristics. (right) Our derived taxonomy's Design subdimension: Player Mode Setting – with a total of three characteristics.

Design							
Task Type							
Exercise-based	Path-based	Target-based	Collecting/ catching-based	Obstacle-based	Rhythmic-based	Cognitive-based	Meditation-based

Figure 8: Our derived taxonomy's Design subdimension: Task type – with a total of eight characteristics.

The exercise theme (*n*=19) features gym and exercise environments (e.g., gym hall) or solely focuses on exercise purposes; in Xu et al. [266]'s game, players perform gestures in a mostly empty virtual world. Closely connected is the sports theme (*n*=35), which imitates real-world sports, for example, rowing on a lake [127] or skiing [202]. Daily life activities (*n*=58) are also featured in many XR exergames, e.g., collecting apples from a tree [248] or blowing candles [88]. In contrast, the historical theme is rarely used (*n*=2); Born et al. [18]'s game presents a "*small medieval village that Vikings have just plundered*". Exergames with elements of space flight and future technologies (*n*=11) were categorized as Sci-fi. In *Astrojumper* [63], the players avoid obstacles in a space environment. As one of the most common themes (*n*=51), the fantasy theme incorporates fantastic scenarios, non-real objects, and superpowers (e.g., [106, 267, 272]). For instance, in the *GhostStand* game, players beat ghosts [106]. Finally, the abstract theme consists of games with a minimal design (*n*=19), such as [74, 271]; Beat Saber [74]'s game elements include stylized cubes and rectangles.

6.4.2 Subdimension: Player Mode. This subdimension considers the player mode of XR exergames (see Figure 7 (right)). This subdimension has three characteristics; each game is assigned to one. Single-player covers XR exergames that are designed to be played only by one player at a time (*n*=171), e.g., [33, 55]. Multiplayer games (*n*=12) enable more than one player to play the game together (e.g., [34, 98]). However, some XR exergames, particularly commercial ones, support both single-player and multiplayer modes, such as FitXR [67], HoloFit [93], and Karaosmanoglu et al. [111].

6.4.3 Subdimension: Task Type. This subdimension presents the featured tasks in XR exergames. In total, we identified eight overarching characteristics. Since games are complex systems and typically include many tasks, one game can have multiple tasks.

Exercise-based tasks (*n*=40) use standard workout exercises in a gamified environment. However, these movements are typically only partially embedded into the game's narrative. For example, players could perform squats that trigger magic attacks [268, 272, 278]. With path-based tasks (*n*=66), players travel along or follow a specific path in the game: In [9], "*the player cycles along a straight path with a speed proportional to cycling revolutions per minute*". In target-based tasks (*n*=79), players must shoot, throw, or hit targets with their hands, weapons, or projectiles (e.g., *VRabl* [25]).

Conversely, collecting/catching-based tasks (*n*=45) feature approaching items that players catch or collect with their hands, weapons, or tools. While the previous two task types require players to interact with objects, obstacle-based tasks (*n*=45) focus on avoiding objects. In [238], players must fit through holes in moving obstacles. Next, rhythmic-based tasks (*n*=10) include exercises and movements that players need to follow in rhythmic patterns (e.g., [241, 274]). The XR exergames rarely featured cognitive-based tasks (*n*=8). A notable example is *Beat Saber* [74], where players cut objects according to the direction shown on the objects. Lastly, meditation-based tasks feature tasks with a relaxation focus (*n*=5), for example, in the form of"*halos [that] expanded and contracted when the player inhaled or exhaled, respectively, in real-time*" [3].

6.4.4 Subdimension: Adaptation. The adaptation dimension includes game task adaptations. A game may employ several adaption strategies. Figure 9 illustrates this dimension.

The existing XR exergames mainly apply no adaptation to their game tasks (*n*=125). We note that we do not consider calibration of the players' body proportions (e.g., height) if it is not framed as critical to the game's task since calibration should be done for every VR game to ensure that players can interact and that the avatar matches the players. Additionally, some games had adaptations that were controlled by clinicians or therapists (e.g., [21, 112]).

Another strategy was to apply game adaptation before gameplay (*n*=45). Sixteen games adapt their task to the players' individual range of motion/body proportion. For example, [17] developed a

Design								
Adaptation								
No Adaptation	Clinican-induced	Before-Game				In-Game		
		Subjective-rating-based	Range of Motion/ Body Proportions	Performance-based	Physiological Measure-based	Subjective-rating-based	Performance-based	Physiological Measure-based

Figure 9: Our derived taxonomy's Design subdimension: Adaptation – with a total of four characteristics and seven sub-characteristics.

Technologies					
Display					
VR				AR	
Mobile HMD	PC-tethered HMD	Console-tethered HMD	Projector-based	Mobile HMD	PC-tethered HMD

Technologies		
Setup		
Mobile	Stationary	Supports both mobile/stationary

Figure 10: (left) Our derived taxonomy's Technologies subdimension: Display – with a total of two characteristics and six sub-characteristics. (right) Our derived taxonomy's Technologies subdimension: Setup – with a total of three characteristics.

game in which players pass through the holes in the walls and calibrated the cutouts to the players' body proportions to ensure comparable difficulties. When subjective-rating-based adaptation (*n*=18) is used, players can typically choose their preferred difficulty level (e.g., [74, 104]). Some games also employed performance-based adaptation (*n*=7) before the gameplay, e.g., by adjusting the game based on the players' performance in a preceding task, such as [116, 158]. Lastly, few games used physiological-based adaptation (*n*=5) before the gameplay; for example, [130, 158] adapted the resistance of their sports hardware based on players' body mass.

Besides adapting the task in an initial calibration phase, some games used in-game adaptation strategies (*n*=25) during the gameplay. In-game subjective-rating-based adaptation is only featured in one game (*n*=1) in the form of perceived exertion scores that contribute to adjusting the difficulty [138]. More commonly, in-game performance-based adaptation (*n*=17) changes the game difficulty based on the players' success and performance (e.g., [44, 255]. Finally, in-game physiological-based adaptation uses the players' physiological data to individualize the gameplay (*n*=9): for example, "*if the heart rate is too low the procedurally generated rings are placed higher requiring more stroke power to pass through them*" [227].

6.5 Dimension V: Technologies

● Technology is the last dimension of our taxonomy and provides information about the used technologies and their specifications.

6.5.1 Subdimension: Display. The display subdimension categorizes which display technology is used by XR exergames (see Figure 10 (left)). Since one game can support more than one type of display, this subdimension is not mutually exclusive. For example, commercial games are often available for multiple platforms.

Most XR exergames use VR displays (*n*=186). Many exergames rely on PC-tethered VR HMDs, (e.g., HTC Vive or Oculus Rift), to display the games (*n*=141). Despite their recent popularity, mobile VR HMDs, (i.e., standalone devices), only take the second place [11, 150, 254] (*n*=48). Rarely, we also encountered the use of console-tethered VR HMDs (*n*=6), (e.g., [66, 74, 75]). Apart from HMD-based VR, some games used projector-based VR (*n*=7), (e.g., [55, 63, 218]), which typically consists of projectors, walls, and 3D glasses.

The remaining XR exergames display their game world through AR technologies, *n*=10. Seven of those games (e.g., [25, 70, 89]) used mobile AR HMDs, such as Microsoft HoloLens [160]. The other three games relied on a PC and an attached PC-tethered AR HMD to render the game, e.g., [185, 193, 273].

6.5.2 Subdimension: Setup. In this dimension, we give an overview of the supported setup conditions (see Figure 10 (right)). A single characteristic was allocated to each game.

The term mobile (*n*=36) comprises games that work with portable hardware and thus can be played almost everywhere: *Seas the Day* [174] only requires a mobile VR HMD, so players can run the game without being bound to one location. Stationary setups (*n*=149) cannot be easily moved to other places since they use hardware that is typically fixed to one location. Games in this category use headsets that are connected to a desktop PC [131], bulky sports machines [128], or permanently installed and calibrated motion-tracking hardware [166]. Finally, ten commercial XR exergames support both mobile/stationary setups, e.g., [87, 145, 240].

6.5.3 Subdimension: Support hardware. The support hardware dimension comprises all additional hardware needed in addition to the regular display setup (e.g., VR HMD with controllers and base

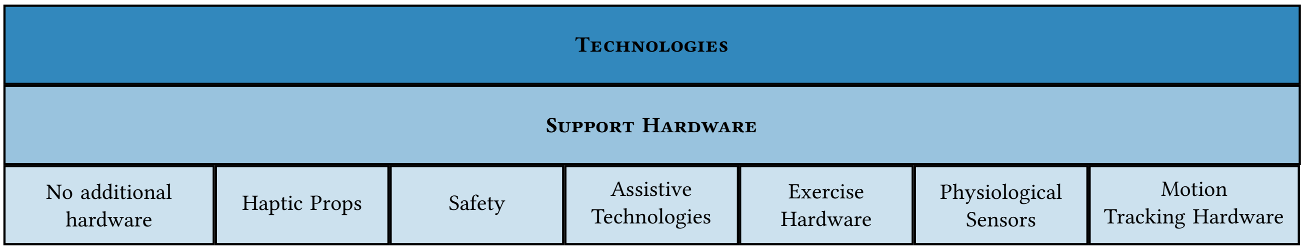

Figure 11: Our derived taxonomy's Technologies subdimension: Support Hardware – with a total of seven characteristics.

stations). Since games might require more than one type of supporting hardware, this dimension is not mutually exclusive.

Seventy-six games did not necessarily require any additional hardware to be played (e.g., [161, 221, 276]). In contrast, haptic props (*n*=20) were used by some exergames. For example, [94] use a physical wakeboard for their exergame. Regarding safety equipment (*n*=9), we encountered the use of harnesses (*n*=3) and monitoring devices (*n*=7, e.g., displays [3, 53]). In rare cases, XR exergames also included assistive technologies (*n*=4), such as an exoskeleton (*n*=3, [89, 201]) or pneumatic gel muscles (*n*=1, [202]).

Exercise hardware (*n*=48) includes all balance, strength, and sports machines. Only two games use balance-exercising hardware: a robotic balance platform [2] and a balance board [140]. Only four games feature strength exercising hardware, which targets the strength development of muscles: weighted straps [56], suspension bands [94], flywheel ergometer [118], and cable resistance equipment [255]. Finally, sports hardware refers to equipment we use for training regardless of games, such as bicycle (*n*=30) [100], elliptical (*n*=2) [245], rowing (*n*=9) [218], and treadmill machines (*n*=4) [137].

Physiological sensors (*n*=18) capture the hardware used to integrate physiological measures, such as heart rate (*n*=11) [224] and breath (*n*=8) [127], in gameplay. In [224], heart rate was used to adjust the speed of a virtual competitor. In Kojić et al. [127]'s game, players were provided feedback based on their breathing rhythm.

Motion tracking hardware is typically used in conjunction with the XR technology to track the players' body parts. In our corpus, 60 games used such hardware; the most often used were Vive Trackers (*n*=26, [44, 143, 279]), followed by Kinect (*n*=20, [40, 102, 239]).

7 DISCUSSION

This section reflects on previous XR exergame research, highlights future research directions, and provides recommendations for XR designers and developers. Second, we explain how to use our taxonomy: GPEDT. Finally, we provide guiding questions for systematic reporting of research on XR exergames.

7.1 Reflecting on XR Exergame Research and Uncovering Research Directions

Continuum of Goals: XR Exergames are Enforcers? The current XR exergame landscape has seven distinct goals. Almost all of them represent serious purposes; for example, promoting physical activity, providing medical-focused training, or physical training. But do XR exergames need to have serious goals? Should exergames force players to perform movements, or instead rely on players' inherent motivation to exercise [148]? Answering these questions is complex and requires serious discussion within the exergame research community. However, our taxonomy shows that the current decision of XR exergame design is on enforcing physical activity. Nevertheless, rare examples also show the promise of relying on players' inherent motivation. In [20], players were provided a custom grip controller to perform an optional strenuous activity that triggered virtual performance augmentation; as a result, players voluntarily performed strenuous activity for a longer period. Yet, we also note that our results are specific to the XR exergame research. Other movement-based applications (e.g., sports) may offer new means to interact with sports without enforcing players.

Different User Groups Require Different Tasks and Perhaps a Different Exergame. Comparing the people dimension to our study data, only very few studies examined differences between user groups (*n*=8). This limited research attention is problematic since users with different abilities or physiologies play differently and prefer different tasks; Karaosmanoglu et al. [112] reported that dementia-specific tasks bored older participants without dementia. Thus, further user group comparison studies are needed to inform the feature design of XR exergames to cater to user preferences and needs. Moreover, every XR application features at least some degree of movement because of its spatial interactions (e.g., grabbing). This raises the question of what to consider as exergames. Our taxonomy answers this question with "it depends". In our review, we followed the definition of Mueller et al. [170]: "*digital game[s] where the outcome [...] is predominantly determined by physical effort*". In our analysis, we saw that Nvidia Fun House [242] and Carnival Games VR [75] were considered as exergames [6, 30] despite little physical activity. We believe that the physical effort required for a digital game depends highly on the user group, and researchers and developers should pay attention to this aspect.

XR Exergames Knowledge is Limited When it Comes to Age Groups. Our evaluation showed significant design and evaluation inconsistencies for the targeted age range. Most games with an age target focused on older adults. Also, the identified nine HCD approaches all targeted either older adults or people with dementia. On the other hand, most studies recruited a young audience: Only 24.10% of studies had a sample with an average age above 40. This discrepancy between the target user group and evaluation is problematic: Simply, the effectiveness of the exergame intervention and validity of measured player experience remains unclear. Furthermore, our results reveal that middle-aged adults [283] are rarely targeted by design or evaluation. Lastly, since abilities change drastically throughout life, we cannot assume that all exergames work for any age group. In contrast, we could even potentially harm people while driving for the good. In particular, the different physiology of children and adolescents requires special care. Hence, these age groups need more targeted exergame concepts.

Generalizability of XR Exergame Research to Non-Western Countries. When designing XR exergames, designers should consider people's cultural background because sociocultural factors influence how we use and accept technology [269]. Hence, the unbalanced culture sample identified in our corpus is worrisome. European and North American studies accounted for 72.31%. The Global South, Asia, and Oceania are severely underrepresented. This raises doubts regarding the generalizability and relevance of XR exergames studies to non-Western cultures [142]. Accordingly, additional research in non-Western nations is needed to understand better how sociocultural influences affect XR exergame design.

Targeting Social Interaction and Multiplayer Engagement. Social interaction and multiplayer support received little attention. However, social settings are a critical factor in providing motivation to exercise. We assume this effect can apply to XR exergames, too. We suspect two primary explanations for the lack of coverage. First, most games feature single-owned HMD experiences because of limited hardware access. Here, we believe that asymmetric XR exergame experiences could be a promising solution [212]. Second, real-life activities (e.g., video conversations) can satisfy social needs. Compared to these, XR technologies may not create the same social experience. Hence, people choose established options. Nonetheless, a recent work found that "*VR (is) almost there*" [244] to mimic its real-world gaming counterpart. Therefore, we believe that future technology may improve XR exergame participants' connections.

Further, we see potential in social interactions beyond traditional multiplayer setups (i.e., multiple users play together), such as the involvement of people who are not active players. For example, having non-player audiences that provide feedback [278], or caregivers that guide the exergame experience [112] can contribute to the experience and performance of players. Although these types of social interaction have been covered in few publications [112, 278], or commercial games [67], we see the importance of further research as such interactions play a vital role in traditional training.

Designers and Developers Should Decide the Gameplay Position of Their Application at an Early Stage. The games in our corpus are split almost equally between seated and standing gameplay. This spread is beneficial because both roles have crucial use cases. The standing setting can enable full-body exercises to combat our current lifestyle of sitting too much. However, many people may be unable to do this (e.g., because of age or space restrictions). Interestingly, research [266] showed that seated exergames increase physiological and perceived exertion. Contrary to common belief, seated exergames do not inevitably reduce physical activity. To be inclusive, it is best to support both seated and standing positions. However, only 6.42% of our games supported this option. Of course, supporting both options might be difficult since most games must be designed from the ground up with this goal in mind. Hence, we advise developers to consider these early in their design process.

Dominance of VR-HMD-based Exergames. In the past, XR exergames required stationary hardware, limiting their usability in the general public. New and powerful mobile headsets, such as Meta Quest or HoloLens, are a prominent step towards mobile and accessible exergames. However, the drawback of most consumer hardware is limited motion tracking. Motion tracking of all four extremities is needed to create exergames that use the entire body. This is impossible without non-portable tracking devices like Vive Trackers or Kinect. But, Sony's Mocopi, HTCs inside-out trackers, and other future systems make us optimistic that this problem will be overcome quickly. Accordingly, many of the current stationary games might become mobile applications in the near future. However, other support hardware (e.g., rowing machines) still limits more advanced exergames to dedicated places.

Our review focused on XR exergames, but VR dominated. Only 5.10% of the analyzed games support AR HMDs. We presume the unavailability of AR HMDs is the primary cause of this lack of AR-focused research. The primary AR system used by games in our corpus is the Microsoft HoloLens. Although it advances mobile and lightweight AR glasses, the HoloLens has a restricted field of view, gesture-only interactions, and a substantial retail price. These shortcomings hinder AR development and acceptance, even though, for specific user groups, AR might be preferable over HMD-based VR systems. For example, people with neurological conditions like dementia likely profit from preserving a connection to the real world [112]. Similar problems also exist for CAVE-based systems. While they boost proprioception since users can see their bodies, the major financial and structural restrictions precluded wider use.

Long-Term Effects and Adherence of XR Exergames Remain Unclear. Exergames may promote physical activity and help prevent sedentary behavior globally. However, it is uncertain if games provide enough incentives for adherence needed for behavior change. Also, the long-term effects of XR exergames are poorly studied—only nine studies evaluated effects over more than two play sessions. Thus, the practicality of such games beyond gameplay novelty remains unknown. Hence, further research in two complementary directions is crucial: (i) conducting more long-term studies and (ii) exploring the raised knowledge gaps to create a good foundation for designing games with strong adherence.

Transferability of Our Results to General XR & HCI Research. Our results represent the subsample of XR and HCI research, and align with previous papers examining a broader area in HCI. For example, Linxen et al. [142] found that non-Western participants are underrepresented at CHI. Similarly, the novelty effects of XR reported by many researchers [113, 199, 213] emphasize the need for long-term exploration of XR technologies in general.

We see the potential for our results to be applied to broader HCI research beyond XR exergames. We believe that many of the guiding questions and taxonomy dimensions can be easily applied to other XR games or applications. Every XR research application has a purpose, there are people who will use the system, every XR application typically involves spatial interaction (even if limited), every application has a design that matches its research intention, and is used with certain technologies. However, we also emphasize that the more exergame-specific dimensions (e.g., exercise dimension) are not easily transferable. Lastly, given that convenience sampling is a common approach to recruit participants (e.g., among students), we speculate that bias in the representation of certain age groups may be applicable to broaden HCI research, but further research is needed to inform this.

7.2 How to Use the Taxonomy of XR Exergames

With our taxonomy, our goal is to create a more standardized and systematic approach to XR exergame design, research, and communication in the field. For example, exergames aimed at specific outcomes could be developed through a taxonomical lens presented within the GPEDT. With a building block approach, researchers and practitioners can explore how exergame features can effectively incentivize individuals, for example, to promote social interaction. Starting with the predetermined factors, they would explore the taxonomy iteratively to determine the best-fitting characteristic for every dimension. A populated example of such an exergame could ultimately look like this:

`To improve [social interaction] among [adolescents], use [full-body] movements in a [standing] position to build a [sci-fi]-themed exergame featuring a [rhythmic-based] [multiplayer] task using [mobile AR-HMDs].`

In addition, our taxonomy can act as a gap analysis tool. Researchers can readily find underrepresented areas or untapped opportunities by mapping XR exergames based on the frequencies reported in each characteristic (see supplementary materials for each coded game); if the taxonomy reveals that there are few exergames designed for older adults, that is a clear area for future research. Similarly, if most existing games focus on physical training but not social interaction, that is another avenue for innovation.

7.3 Reporting Standards in XR Exergames Research

To promote comprehensibility, reproducibility, and transferability, academia relies on transparent communication of methodology and results. Good reporting standards are crucial, especially in highly interdisciplinary areas or when conducting reviews to structure a domain. XR exergame research is a young and emerging field that no reporting consensus exists for every aspect yet. Unsurprisingly, individual explanation styles caused some challenges during our data collection and coding process. Next, we provide some examples of incomplete, inconclusive, or problematic reporting types:

(1) Many papers do not report their XR exergame's goals (n=66). Similarly, 144 of 195 games do not mention their target audience. But, we cannot use the same games for the same goal or expect every exergame to be playable by or be harmless to everyone.
(2) We created our taxonomy by finding common patterns. However, while coding the exergames' movement data, we could not distinguish a pattern. The rationale was the reporting level of the XR exergame movements (please refer to the affinity mapping activity in the supplementary material to see the movement codes). Some XR exergames descriptions explain the movements from a higher level; for example, players have to perform leg and arm movements, dancing movements, or dodging movements, but what these movements feature is unclear.
(3) Similar to previous concerns, it was unclear if exergames were intended to be played seated or standing. We had to consult additional resources (e.g., figures, supplementary materials, videos) to track down the missing information.
(4) We also used additional resources when the retrieved data did not report the display used for gameplay (e.g., figures, supplementary materials, gameplay videos).

The only way to fix diverging reporting is to develop a shared understanding and common reporting patterns. As a first step towards a more systematic approach, we close this paper with guiding questions for researchers, designers, and developers to answer and communicate (see Table 2). Similar to other efforts [77], we emphasize the value of providing audiovisual supporting materials—videos showing gameplay sequences to illustrate how the movements translate into the game or executable of developed games to provide clear communication about the featured gameplay.

7.4 Limitations

In this section, we address the limitations of our methodology as well as the limitations of the corpus.

7.4.1 Limitations of the Methodology. We consulted prior XR and exergames publications, and conducted multiple informal searches to decide on our search query, but like every review, we cannot claim completeness. Despite our best efforts, our query might not have covered all terms. For example, different fields (e.g., movement science vs. HCI) may use different terms to refer to exergames, such as *"active video games"* and *"active games"* (e.g., [152]). Additionally, our query might have missed articles that used only the name of a sport (e.g., virtual rowing) instead of an exergame-related term. Since it is impossible to account for all sports in a query, we urge researchers to use at least one of the exergame terms (e.g., movement games) to ensure that their research is included in relevant literature reviews. Nevertheless, we believe that with our included search terms (23 keywords) (see Table 3), we give a comprehensive overview of the research and design of XR exergames.

Aligning with Nickerson et al. [180]'s methodology, our taxonomy can be easily expanded given the rapidly growing field. For example, while this article was under review, Kontio et al. [129]'s paper on non-standing locomotion techniques (potentially useful for VR exergames) has been published, providing opportunities for further extension of the position subdimension (e.g., lying). Accordingly, we emphasize that taxonomies are rarely static constructs but will grow over time to include new research directions.

Lastly, we elaborate on the reflexivity further [22, 179]. The authors of this paper work on the intersection of the XR and exergames and have published on those topics for several years. However, none of them has in-depth knowledge of movement science or physiology. Although we believe their expertise contributes to understanding the XR exergame landscape, we acknowledge that an in-depth focus on movement science is not in this review.

7.4.2 Limitations of the Corpus. HCI and medical areas use different terminologies. For example, in HCI literature, VR refers to

Table 2: The list of guiding questions to follow when reporting details of implemented/used XR exergames.

No	Guiding Questions
Goals	
#1	What is the purpose of the XR exergame?
People	
#1	Which age group does the XR exergame target?
#2	For which clinical group was the XR exergame designed or is it safe for everyone to play it? Why?
Exercises	
#1	Which movements do players perform in the XR exergame?
#2	How do players precisely perform the featured movements in the XR exergame?
#3	Which body parts are targeted by the XR exergame?
#4	In which position can/should the XR exergame be played? Why?
Design	
#1	Which game tasks does the XR exergame feature?
#2	Is there any adaptation applied in the game to match the game tasks to the players' game skill, physiological state, or condition If so, in which form is this adaptation applied?
#3	How are the game tasks completed (e.g., performing a throwing action using the arms)?
#4	What is the design theme of the XR exergame? Why?
#5	How many players can play the game at once?
Technologies	
#1	Which technology is used to display immersive XR environment? Why (e.g., advantages)?
#2	Is any additonal hardware necesssary to play the XR exergame?
#3	Which type of setup is used for the XR exergame? Mobile or stationary?

immersive digital worlds, whereas several medical papers use VR to refer to non-immersive digital worlds (e.g., games that are played on a TV). Another example was the description of hardware: HCI typically uses the term Vive trackers to refer to HTC Vive motion tracking hardware. However, a medical paper referred to this equipment as *"pucks"* [205]. While every research field has its own established terms, we believe in the importance of having shared terminologies. The absence of mutual understanding limits communication and information flow between these fields. We hope that our work will be a first step towards finding shared terminologies, thereby supporting the growth of this interdisciplinary field.

Finally, we saw that medical literature describes performed movements in detail (e.g., shoulder flexion, abduction), while for HCI literature, the explanation is typically limited to the general action (e.g., arm movements, throwing). Moreover, for some articles, it was not clear which specific movement (e.g., which arm movement [114]) was used, how certain exercises were accomplished (e.g., squatting while sitting [205]), or which movements were performed for the specific action (e.g., dodging in the *Fishing Master* [187]). When we did not see how these movements translated to the gameplay, we omitted them from our frequency reporting. Hence, we do not claim that we provide the complete list of movements performed in XR exergames but provide an approximation in our best capabilities.

8 CONCLUSION

Motivation is a central requirement for preserving one's engagement in physical activity. XR exergames can support this by offering immersive virtual worlds with enjoyable gameplay. The increased public attention on this timely topic, paired with recent advances in XR technology, has led to a surge of research focusing on XR exergames. The rapidly expanding field calls for a comprehensive organization to steer the community's efforts toward unexplored but promising topics. Therefore, we conducted a scoping review of the current state of XR exergame research.Based on analysis of 186 papers, we give a quantitative and qualitative summary of XR exergame research. Further, we provide a taxonomy with five central dimensions that map the design space of XR exergames. Finally, we conclude with nine research directions and guiding questions to guide future research and reporting in the XR exergame field.

ACKNOWLEDGMENTS

This work was supported by the European Union, the German Federal Ministry of Education and Research (BMBF), the German Research Foundation (DFG), and partly funded under the Excellence Strategy of the Federal Government and the Länder. Additionally, this project was supported by the SSHRC INSIGHT Grant (grant number: 435-2022-0476), NSERC Discovery Grant (grant number: RGPIN-2023-03705), CIHR AAL EXERGETIC operating grant (number: 02237-000), CFI John R. Evans Leaders Fund or CFI JELF (grant number: 41844), and the Provost's Program for Interdisciplinary Postdoctoral Scholars at the University of Waterloo.

REFERENCES

[1] Hojun Aan, Sangsun Han, Hyeonkyu Kim, and Kibum Kim. 2020. Remote Virtual Showdown. In *SIGGRAPH Asia 2020 XR* (Virtual Event, Republic of Korea) *(SA*

'20). Association for Computing Machinery, New York, NY, USA, Article 10, 3 pages. https://doi.org/10.1145/3415256.3421488 pages #28
[2] Juan D. Abril, Oswaldo Rivera, Paola Niño-Suárez, Eduardo Castillo-Castañeda, and Oscar F. Aviles. 2023. Virtual Reality Serious Game Session for Rehabilitation Through Calculation of the Center of Pressure in a Stabilometric Platform. *Smart Innovation, Systems and Technologies* 318 (2023), 35 – 45. https://doi.org/10.1007/978-981-19-6347-6_4 pages #13, #27
[3] Dalya Al-Mfarej, Jennifer M. Vojtech, Serge H. Roy, Elise Townsend, Julie J. Keysor, Nancy Kuntz, Vamshi Rao, Joshua C. Kline, and Bhawna Shiwani. 2023. A Virtual Reality Exergame: Clinician-Guided Breathing and Relaxation for Children with Muscular Dystrophy. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 270–276. https://doi.org/10.1109/VRW58643.2023.00065 pages #9, #11, #13, #27
[4] Iannis Albert, Nicole Burkard, Dirk Queck, and Marc Herrlich. 2022. The Effect of Auditory-Motor Synchronization in Exergames on the Example of the VR Rhythm Game BeatSaber. *Proc. ACM Hum.-Comput. Interact.* 6, CHI PLAY, Article 253 (oct 2022), 26 pages. https://doi.org/10.1145/3549516 pages #26, #28
[5] LLC Alzan Studios. 2020. *Holofit by Holodia*. Game [VR]. Alzan Studios, LLC. Longwood, Florida, USA. http://www.alzanstudios.com/. Meta Reviews: https://www.meta.com/experiences/2202354219893697/. pages #26
[6] Evaggelia Amprasi, Nikolaos Vernadakis, Eleni Zetou, and Panagiotis Antoniou. 2022. Effect of a Full Immersive Virtual Reality Intervention on Selective Attention in Children. *International Journal of Instruction* 15, 1 (2022), 565 – 582. https://doi.org/10.29333/iji.2022.15132a pages #13, #26, #28
[7] Akash Ashok, Mazen Darwish, and Burkhard Claus Wünsche. 2023. Investigating Changes in Hand-Eye Coordination and Reaction Time Using a VR Squash Simulation. In *Proceedings of the 2023 Australasian Computer Science Week* (Melbourne, VIC, Australia) *(ACSW '23)*. Association for Computing Machinery, New York, NY, USA, 257–260. https://doi.org/10.1145/3579375.3579422 pages #9, #27
[8] Soumya C. Barathi, Daniel J. Finnegan, Matthew Farrow, Alexander Whaley, Pippa Heath, Jude Buckley, Peter W. Dowrick, Burkhard C. Wuensche, James L. J. Bilzon, Eamonn O'Neill, and Christof Lutteroth. 2018. Interactive Feedforward for Improving Performance and Maintaining Intrinsic Motivation in VR Exergaming. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173982 pages #27
[9] Soumya C. Barathi, Michael Proulx, Eamonn O'Neill, and Christof Lutteroth. 2020. Affect Recognition Using Psychophysiological Correlates in High Intensity VR Exergaming. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376596 pages #11, #27
[10] Bemani. 1998. *Dance Dance Revolution*. Game [Arcade]. https://www.konami.com/games/asia/en/products/ddr_a/. Accessed: 13.08.2023. pages #1, #2
[11] Semih Bertuğ Benim, Mehmet İlker Berkman, and Güven Çatak. 2021. Design Process of a VR Sports Games Trilogy for Paraplegic Players: VR4Inclusion Case Study. *Springer Series in Design and Innovation* 13 (2021), 181 – 192. https://doi.org/10.1007/978-3-030-65060-5_15 pages #12, #27
[12] David Benyon, Phil Turner, and Susan Turner. 2005. *Designing interactive systems: People, activities, contexts, technologies*. Pearson Education. pages #2, #4
[13] Emilia Biffi, Cristina Maghini, Alessia Marelli, Eleonora Diella, Daniele Panzeri, Ambra Cesareo, Chiara Gagliardi, Gianluigi Reni, and Anna Carla Turconi. 2016. Immersive Virtual Reality Platform for Cerebral Palsy Rehabilitation. In *Proceedings of the 4th Workshop on ICTs for Improving Patients Rehabilitation Research Techniques* (Lisbon, Portugal) *(REHAB '16)*. Association for Computing Machinery, New York, NY, USA, 85–88. https://doi.org/10.1145/3051488.3051497 pages #6
[14] Blueteak. 2018. *QuiVr*. Game [VR]. Blueteak. Islamorada, FL, USA. http://blueteak.io/. Steam Reviews: https://store.steampowered.com/app/489380/QuiVr/. pages #26
[15] John Bolton, Mike Lambert, Denis Lirette, and Ben Unsworth. 2014. PaperDude: A Virtual Reality Cycling Exergame. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 475–478. https://doi.org/10.1145/2559206.2574827 pages #28
[16] Gunnar Borg. 1998. *Borg's perceived exertion and pain scales*. Human kinetics. pages #8
[17] Felix Born, Sophie Abramowski, and Maic Masuch. 2019. Exergaming in VR: The Impact of Immersive Embodiment on Motivation, Performance, and Perceived Exertion. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, 1–8. https://doi.org/10.1109/VS-Games.2019.8864579 pages #1, #2, #3, #11, #27
[18] Felix Born, Linda Graf, and Maic Masuch. 2021. Exergaming: The Impact of Virtual Reality on Cognitive Performance and Player Experience. In *2021 IEEE Conference on Games (CoG)*. IEEE, 1–8. https://doi.org/10.1109/CoG52621.2021.9619105 pages #11, #27
[19] Felix Born, Maic Masuch, and Antonia Hahn. 2020. Ghost Sweeper: Using a Heavy Passive Haptic Controller to Enhance a Room-Scale VR Exergame. In *2020 IEEE Conference on Games (CoG)*. 221–228. https://doi.org/10.1109/CoG47356.2020.9231867 pages #27
[20] Felix Born, Adrian Rygula, and Maic Masuch. 2021. Motivating Players to Perform an Optional Strenuous Activity in a Virtual Reality Exergame Using Virtual Performance Augmentation. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 225 (oct 2021), 21 pages. https://doi.org/10.1145/3474652 pages #3, #13, #27
[21] Lars Peder Vatshelle Bovim, Lauritz Valved, Bendik Bleikli, Atle Birger Geitung, Harald Soleim, and Bård Bogen. 2021. Theoretical Rationale for Design of Tasks in a Virtual Reality-Based Exergame for Rehabilitation Purposes. *Frontiers in Aging Neuroscience* 13 (2021). https://doi.org/10.3389/fnagi.2021.734223 pages #9, #11, #27
[22] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47. https://doi.org/10.1002/capr.12360 pages #7, #15
[23] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (11 1995). pages #8
[24] Ilona Buchem, Susan Vorwerg, Oskar Stamm, Kristian Hildebrand, and Yvonne Bialek. 2021. Gamification in mixed-reality exergames for older adult patients in a mobile immersive diagnostic center: A pilot study in the BewARe Project. *Proceedings of 2021 7th International Conference of the Immersive Learning Research Network, iLRN 2021* (2021). https://doi.org/10.23919/iLRN52045.2021.9459337 pages #10, #28
[25] Tim Buckers, Boning Gong, Elmar Eisemann, and Stephan Lukosch. 2018. VRabl: Stimulating Physical Activities through a Multiplayer Augmented Reality Sports Game. In *Proceedings of the First Superhuman Sports Design Challenge: First International Symposium on Amplifying Capabilities and Competing in Mixed Realities* (Delft, Netherlands) *(SHS '18)*. Association for Computing Machinery, New York, NY, USA, Article 1, 5 pages. https://doi.org/10.1145/3210299.3210300 pages #9, #11, #12, #28
[26] Joey Campbell and Mike Fraser. 2019. Switching It Up: Designing Adaptive Interfaces for Virtual Reality Exergames. In *Proceedings of the 31st European Conference on Cognitive Ergonomics* (BELFAST, United Kingdom) *(ECCE '19)*. Association for Computing Machinery, New York, NY, USA, 177–184. https://doi.org/10.1145/3335082.3335087 pages #28
[27] Pablo Campo-Prieto, José Mª Cancela-Carral, and Gustavo Rodríguez-Fuentes. 2022. Feasibility and Effects of an Immersive Virtual Reality Exergame Program on Physical Functions in Institutionalized Older Adults: A Randomized Clinical Trial. *Sensors* 22, 18 (2022). https://doi.org/10.3390/s22186742 pages #26, #28
[28] Pablo Campo-Prieto, José Mª Cancela-Carral, and Gustavo Rodríguez-Fuentes. 2022. Wearable Immersive Virtual Reality Device for Promoting Physical Activity in Parkinson's Disease Patients. *Sensors* 22, 9 (2022). https://doi.org/10.3390/s22093302 pages #26, #28
[29] Pablo Campo-Prieto, José Mª Cancela-Carral, Borja Alsina-Rey, and Gustavo Rodríguez-Fuentes. 2022. Immersive Virtual Reality as a Novel Physical Therapy Approach for Nonagenarians: Usability and Effects on Balance Outcomes of a Game-Based Exercise Program. *Journal of Clinical Medicine* 11, 13 (2022). https://doi.org/10.3390/jcm11133911 pages #26, #28
[30] Pablo Campo-Prieto, Gustavo Rodríguez-Fuentes, and José Ma Cancela-Carral. 2021. Immersive virtual reality exergame promotes the practice of physical activity in older people: An opportunity during covid-19. *Multimodal Technologies and Interaction* 5, 9 (2021). https://doi.org/10.3390/mti5090052 pages #13, #26, #28
[31] Lizhou Cao, Chao Peng, and Yangzi Dong. 2021. Ellic's exercise class: promoting physical activities during exergaming with immersive virtual reality. *Virtual reality* 25 (2021), 597–612. pages #9, #28
[32] Polona Caserman, Shule Liu, and Stefan Göbel. 2022. Full-Body Motion Recognition in Immersive- Virtual-Reality-Based Exergame. *IEEE Transactions on Games* 14, 2 (2022), 243–252. https://doi.org/10.1109/TG.2021.3064749 pages #9, #27
[33] Arnaldo Cesco, Francesco Ballardin, and Gustavo Marfia. 2021. CeVRicale: A VR app for cervical rehabilitation. *Proceedings - 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VRW 2021* (2021), 641 – 642. https://doi.org/10.1109/VRW52623.2021.00203 pages #11, #28
[34] Bo-Han Chen, Sai-Keung Wong, Wei-Che Chang, and Roy Ping-Hao Fan. 2022. LAGH: Towards Asymmetrical Collaborative Bodily Play between 1st and 2nd Person Perspectives. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 447 (nov 2022), 26 pages. https://doi.org/10.1145/3555548 pages #10, #11, #27
[35] Bo-Han Chen, Sai-Keung Wong, Wei-Che Chang, and Roy Ping-Hao Fan. 2021. Towards Social Interaction between 1st and 2nd Person Perspectives on Bodily Play. In *Adjunct Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 1–3. https://doi.org/10.1145/3474349.3480211 pages #28

[36] Po-Jung Chen, Hui-Fen Hsu, Kuei-Min Chen, and Frank Belcastro. 2023. VR exergame interventions among older adults living in long-term care facilities: A systematic review with Meta-analysis. *Annals of Physical and Rehabilitation Medicine* 66, 3 (2023), 101702. https://doi.org/10.1016/j.rehab.2022.101702 pages #3

[37] Weiqin Chen, Martin Bang, Daria Krivonos, Hanna Schimek, and Arnau Naval. 2020. An Immersive Virtual Reality Exergame for People with Parkinson's Disease. In *Computers Helping People with Special Needs*, Klaus Miesenberger, Roberto Manduchi, Mario Covarrubias Rodriguez, and Petr Peňáz (Eds.). Springer International Publishing, Cham, 138–145. pages #27

[38] Manuela Chessa, Gabriele Balocchi, Michela Busi, Antonio Novellino, and Fabio Solari. 2016. An oculus rift based exergame to improve awareness in disabled people. *VISIGRAPP 2016 - Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Doctoral Consortium* 4 (2016), 770 – 777. https://doi.org/10.5220/0005852607700777 Cited by: 0; All Open Access, Green Open Access, Hybrid Gold Open Access. pages #10, #27

[39] Chain Yi Chu, Ho Yin Ng, Chia Hui Lin, and Ping Hsuan Han. 2022. PressyCube: An Embeddable Pressure Sensor with Softy Prop for Limb Rehabilitation in Immersive Virtual Reality. In *2022 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. 1–1. https://doi.org/10.1109/ICMEW56448.2022.9859433 pages #28

[40] Chia-Ru Chung, Shih-Ching Yeh, Eric Hsiao-Kuang Wu, and Sheng-Yang Lin. 2023. Qigong Master: A Qigong-based attention training game using action recognition and balance analysis. *IEEE Transactions on Games* (2023), 1–8. https://doi.org/10.1109/TG.2023.3274148 pages #9, #10, #13, #27

[41] Yu-Yen Chung, Thiru M. Annaswamy, and Balakrishnan Prabhakaran. 2023. Performance and User Experience Studies of HILLES: Home-Based Immersive Lower Limb Exergame System. In *Proceedings of the 14th Conference on ACM Multimedia Systems* (Vancouver, BC, Canada) *(MMSys '23)*. Association for Computing Machinery, New York, NY, USA, 62–73. https://doi.org/10.1145/3587819.3590985 pages #28

[42] Julia Ciążyńska, Michał Janowski, and Janusz Maciaszek. 2022. Effects of a Modern Virtual Reality 3D Head-Mounted Display Exergame on Simulator Sickness and Immersion Under Specific Conditions in Young Women and Men: Experimental Study. *JMIR Serious Games* 10, 4 (2022). https://doi.org/10.2196/41234 pages #26, #28

[43] Julia Ciążyńska and Janusz Maciaszek. 2023. Effects of Low-Immersive vs. High-Immersive Exercise Environment on Postural Stability and Reaction and Motor Time of Healthy Young Adults. *Journal of Clinical Medicine* 12, 1 (2023). https://doi.org/10.3390/jcm12010389 pages #26, #28

[44] Sebastian Cmentowski, Sukran Karaosmanoglu, Lennart E. Nacke, Frank Steinicke, and Jens Harald Krüger. 2023. Never Skip Leg Day Again: Training the Lower Body with Vertical Jumps in a Virtual Reality Exergame. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 786, 18 pages. https://doi.org/10.1145/3544548.3580973 pages #2, #4, #10, #12, #13, #27

[45] Ana Isabel Corregidor-Sánchez, Antonio Segura-Fragoso, Marta Rodriguez-Hernandez, Juan José Criado-Alvarez, Jaime Gonzalez-Gonzalez, and Begona Polonio-Lopez. 2020. Can exergames contribute to improving walking capacity in older adults? A systematic review and meta-analysis. *Maturitas* 132 (2020), 40–48. https://doi.org/10.1016/j.maturitas.2019.12.006 pages #3

[46] Joyce Cristina Cândido Soares, Bárbara Letícia Costa de Moraes, Clarissa Cardoso dos Santos Couto Paz, and Lívia de Castro Magalhães. 2019. Influence of the Microsoft Kinect® games on the motor and functional performance of a child with developmental coordination disorder; [Influência do uso de jogos do Microsoft Kinect® sobre o desempenho motor e funcional de criança com transtorno do desenvolvimento de coordenação]. *Brazilian Journal of Occupational Therapy* 27, 4 (2019), 710 – 717. https://doi.org/10.4322/2526-8910.CTOAO1630 pages #5

[47] Silas de Oliveira Damasceno, Eduardo Pizzo Junior, Leonardo Kesrouani Lemos, Taíse Mendes Biral, Allysiê Priscilla de Souza Cavina, and Franciele Marques Vanderlei. 2022. Virtual reality-based therapy associated with blood flow restriction in older adults: A proposal for integration of techniques. *Frontiers in Physiology* 13 (2022). https://doi.org/10.3389/fphys.2022.958823 pages #5

[48] Lucas Martins de Souza, Irem Gokce Yildirim, Anya Kolesnichenko, and Taiwoo Park. 2016. World Of Riders: Exercising is Fun. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Austin, Texas, USA) *(CHI PLAY Companion '16)*. Association for Computing Machinery, New York, NY, USA, 55–60. https://doi.org/10.1145/2968120.2971807 pages #28

[49] Brittany F Drazich, Sarah LaFave, Breanna M Crane, Sarah L Szanton, Michelle C Carlson, Chakra Budhathoki, and Janiece L Taylor. 2020. Exergames and depressive symptoms in older adults: A systematic review. *Games for Health Journal* 9, 5 (2020), 339–345. https://doi.org/10.1089/g4h.2019.0165 pages #3

[50] Eva Dulau, Chrisna R Botha-Ravyse, and Mika Luimula. 2019. Virtual reality for physical rehabilitation: A Pilot study How will virtual reality change physical therapy?. In *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. 277–282. https://doi.org/10.1109/CogInfoCom47531.2019.9089980 pages #27

[51] Martina Eckert, José Zarco, Juan Meneses, and José-Fernán Martínez. 2017. Usage of VR Headsets for Rehabilitation Exergames. In *Bioinformatics and Biomedical Engineering*, Ignacio Rojas and Francisco Ortuño (Eds.). Springer International Publishing, Cham, 434–442. pages #27

[52] Mazhar Eisapour, Shi Cao, and Jennifer Boger. 2020. Participatory design and evaluation of virtual reality games to promote engagement in physical activity for people living with dementia. *Journal of rehabilitation and assistive technologies engineering* 7 (2020), 2055668320913770. pages #10, #27

[53] Mahzar Eisapour, Shi Cao, Laura Domenicucci, and Jennifer Boger. 2018. Participatory Design of a Virtual Reality Exercise for People with Mild Cognitive Impairment. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3170427.3174362 pages #13, #27

[54] Mahzar Eisapour, Shi Cao, Laura Domenicucci, and Jennifer Boger. 2018. Virtual Reality Exergames for People Living with Dementia Based on Exercise Therapy Best Practices. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 62, 1 (2018), 528–532. https://doi.org/10.1177/1541931218621120 pages #27

[55] Aviv Elor, Michael Powell, Evanjelin Mahmoodi, Nico Hawthorne, Mircea Teodorescu, and Sri Kurniawan. 2020. On Shooting Stars: Comparing CAVE and HMD Immersive Virtual Reality Exergaming for Adults with Mixed Ability. *ACM Trans. Comput. Healthcare* 1, 4, Article 22 (sep 2020), 22 pages. https://doi.org/10.1145/3396249 pages #11, #12, #27

[56] Aviv Elor, Michael Powell, Evanjelin Mahmoodi, Mircea Teodorescu, and Sri Kurniawan. 2022. Gaming beyond the Novelty Effect of Immersive Virtual Reality for Physical Rehabilitation. *IEEE Transactions on Games* 14, 1 (2022), 107 – 115. https://doi.org/10.1109/TG.2021.3069445 pages #10, #13, #27

[57] Cassondra M. Eng, Dominic M. Calkosz, Sebastian Y. Yang, Nathan C. Williams, Erik D. Thiessen, and Anna V. Fisher. 2020. Doctoral Colloquium-Enhancing Brain Plasticity and Cognition Utilizing Immersive Technology and Virtual Reality Contexts for Gameplay. *Proceedings of 6th International Conference of the Immersive Learning Research Network, iLRN 2020* (2020), 395 – 398. https://doi.org/10.23919/iLRN47897.2020.9155120 pages #26, #28

[58] Mingming Fan, Vinita Tibdewal, Qiwen Zhao, Lizhou Cao, Chao Peng, Runxuan Shu, and Yujia Shan. 2023. Older Adults' Concurrent and Retrospective Think-Aloud Verbalizations for Identifying User Experience Problems of VR Games. *Interacting with Computers* 34, 4 (01 2023), 99–115. https://doi.org/10.1093/iwc/iwac039 pages #27

[59] Nuša Farič, Lee Smith, Adrian Hon, Henry W W Potts, Katie Newby, Andrew Steptoe, and Abi Fisher. 2021. A Virtual Reality Exergame to Engage Adolescents in Physical Activity: Mixed Methods Study Describing the Formative Intervention Development Process. *J Med Internet Res* 23, 2 (4 Feb 2021), e18161. https://doi.org/10.2196/18161 pages #9, #27

[60] Samantha Finkelstein, Tiffany Barnes, Zachary Wartell, and Evan A. Suma. 2013. Evaluation of the exertion and motivation factors of a virtual reality exercise game for children with autism. In *2013 1st Workshop on Virtual and Augmented Assistive Technology (VAAT)*. IEEE, 11–16. https://doi.org/10.1109/VAAT.2013.6786186 pages #10, #27

[61] Samantha Finkelstein, Andrea Nickel, Tiffany Barnes, and Evan A. Suma. 2010. Astrojumper: Motivating Children with Autism to Exercise Using a VR Game. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI EA '10)*. Association for Computing Machinery, New York, NY, USA, 4189–4194. https://doi.org/10.1145/1753846.1754124 pages #10, #28

[62] Samantha Finkelstein, Andrea Nickel, Zachary Lipps, Tiffany Barnes, Zachary Wartell, and Evan A. Suma. 2011. Astrojumper: Motivating Exercise with an Immersive Virtual Reality Exergame. *Presence: Teleoperators and Virtual Environments* 20, 1 (02 2011), 78–92. https://doi.org/10.1162/pres_a_00036 arXiv:https://direct.mit.edu/pvar/article-pdf/20/1/78/1625108/pres_a_00036.pdf pages #28

[63] Samantha L. Finkelstein, Andrea Nickel, Tiffany Barnes, and Evan A. Suma. 2010. Astrojumper: Designing a virtual reality exergame to motivate children with autism to exercise. In *2010 IEEE Virtual Reality Conference (VR)*. IEEE, 267–268. https://doi.org/10.1109/VR.2010.5444770 pages #10, #11, #12, #28

[64] Jenna Finlayson, Jamie Peterson, Joshua Free, Michael Lo, Lindsay A. Shaw, Christof Lutteroth, and Burkhard C. Wünsche. 2017. Tour de Tune - Auditory-game-motor Synchronisation in Exergames. In *ICAT-EGVE 2017 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, Robert W. Lindeman, Gerd Bruder, and Daisuke Iwai (Eds.). The Eurographics Association. https://doi.org/10.2312/egve.20171335 pages #27

[65] James M. Finley, Marientina Gotsis, Vangelis Lympouridis, Shreya Jain, Aram Kim, and Beth E. Fisher. 2021. Design and Development of a Virtual Reality-Based Mobility Training Game for People With Parkinson's Disease. *Frontiers in Neurology* 11 (2021). https://doi.org/10.3389/fneur.2020.577713 pages #10,

#28
[66] FitXR. 2019. *BoxVR.* Game [VR]. FitXR. London, England. https://fitxr.com/boxvr. Steam Reviews: https://store.steampowered.com/app/641960/BOXVR/. pages #12, #26
[67] FitXR. 2019. *FitXR - Box. HIIT. Dance. Sculpt. Combat.* Game [VR]. FitXR. London, England. Meta Reviews: https://www.meta.com/experiences/2327205800645550/. pages #9, #11, #14, #26
[68] International Organization for Standardization. 2019. ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. https://www.iso.org/standard/77520.html International Organization for Standardization. Geneva, Switzerland. pages #4
[69] Michela Franzo', Simona Pascucci, Mariano Serrao, Franco Marinozzi, and Fabiano Bini. 2022. Exergaming in mixed reality for the rehabilitation of ataxic patients. *2022 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2022 - Conference Proceedings* (2022). https://doi.org/10.1109/MeMeA54994.2022.9856552 pages #3, #10, #28
[70] Michela Franzo, Simona Pascucci, Mariano Serrao, Franco Marinozzi, and Fabiano Bini. 2023. Breakthrough in Occupational Therapy with Mixed-Reality exergaming for cerebellar ataxia patients. *2023 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2023 - Conference Proceedings* (2023). https://doi.org/10.1109/MeMeA57477.2023.10171955 pages #10, #12, #27
[71] Nicolas Frevel, Sascha L. Schmidt, Daniel Beiderbeck, Benjamin Penkert, and Brian Subirana. 2020. *Taxonomy of Sportstech.* Springer International Publishing, Cham, 15–37. https://doi.org/10.1007/978-3-030-50801-2_2 pages #3
[72] Borko Furht (Ed.). 2008. *Immersive Virtual Reality.* Springer US, Boston, MA, 345–346. https://doi.org/10.1007/978-0-387-78414-4_85 pages #3
[73] Matthias Gamer, Jim Lemon, Maintainer Matthias Gamer, A Robinson, and W Kendall's. 2012. Package 'irr'. *Various coefficients of interrater reliability and agreement* (2012). https://cran.r-project.org/web/packages/irr/irr.pdf. pages #5
[74] Beat Games. 2019. *Beat Saber.* Game [VR]. Beat Games. Los Angeles, CA, USA. https://beatsaber.com/. Steam Reviews: https://store.steampowered.com/app/620980/Beat_Saber/. pages #10, #11, #12, #26
[75] Cat Daddy Games. 2016. *Carnival Games® VR.* Game [VR]. Cat Daddy Games. London, United Kingdom. https://2k.com/en-US/game/carnival-games-vr/. pages #12, #13, #26
[76] Cristina García-Muñoz, María-Dolores Cortés-Vega, Juan-Carlos Hernández-Rodríguez, Lourdes M. Fernández-Seguín, Isabel Escobio-Prieto, and María Jesús Casuso-Holgado. 2022. Immersive Virtual Reality and Vestibular Rehabilitation in Multiple Sclerosis: Case Report. *JMIR Serious Games* 10, 1 (2022). https://doi.org/10.2196/31020 pages #26, #28
[77] Kathrin Gerling and Max V. Birk. 2022. Reflections on Rigor and Reproducibility: Moving Toward a Community Standard for the Description of Artifacts in Experimental Games Research. In *Extended Abstracts of the 2022 Annual Symposium on Computer-Human Interaction in Play* (Bremen, Germany) *(CHI PLAY '22)*. Association for Computing Machinery, New York, NY, USA, 266–267. https://doi.org/10.1145/3505270.3558360 pages #15
[78] Kathrin Gerling, Ian Livingston, Lennart Nacke, and Regan Mandryk. 2012. Full-Body Motion-Based Game Interaction for Older Adults. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1873–1882. https://doi.org/10.1145/2207676.2208324 pages #2
[79] Frano Giakoni-Ramírez, Andrés Godoy-Cumillaf, Sebastián Espoz-Lazo, Daniel Duclos-Bastias, and Pablo del Val Martín. 2023. Physical Activity in Immersive Virtual Reality: A Scoping Review. *Healthcare* 11, 11 (2023). https://doi.org/10.3390/healthcare11111553 pages #3
[80] Christos Goumopoulos, Emmanouil Drakakis, and Dimitris Gklavakis. 2022. Augmented and Virtual Reality Based Exergames in GAME2AWE for Elderly Fall Prevention. In *2022 18th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 100–105. https://doi.org/10.1109/WiMob55322.2022.9941651 pages #9, #27
[81] Christos Goumopoulos, Emmanouil Drakakis, and Dimitris Gklavakis. 2023. Feasibility and Acceptance of Augmented and Virtual Reality Exergames to Train Motor and Cognitive Skills of Elderly. *Computers* 12, 3 (2023). https://doi.org/10.3390/computers12030052 pages #27
[82] Yusuke Goutsu and Tetsunari Inamura. 2023. Instant Difficulty Adjustment: Predicting Success Rate of VR Kendama When Changing the Difficulty Level. In *Proceedings of the Augmented Humans International Conference 2023* (Glasgow, United Kingdom) *(AHs '23)*. Association for Computing Machinery, New York, NY, USA, 346–348. https://doi.org/10.1145/3582700.3583954 pages #28
[83] Antti Granqvist, Tapio Takala, Jari Takatalo, and Perttu Hämäläinen. 2018. Exaggeration of Avatar Flexibility in Virtual Reality. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (Melbourne, VIC, Australia) *(CHI PLAY '18)*. Association for Computing Machinery, New York, NY, USA, 201–209. https://doi.org/10.1145/3242671.3242694 pages #6
[84] Sidney Grosprêtre, Philémon Marcel-Millet, Pauline Eon, and Bettina Wollesen. 2023. How Exergaming with Virtual Reality Enhances Specific Cognitive and Visuo-Motor Abilities: An Explorative Study. *Cognitive Science* 47, 4 (2023). https://doi.org/10.1111/cogs.13278 pages #26, #28
[85] Joshua C. Haller, Young H. Jang, Jack Haller, Lindsay Shaw, and Burkhard C. Wünsche. 2019. HIIT The Road: Using Virtual Spectator Feedback in HIIT-Based Exergaming. In *Proceedings of the Australasian Computer Science Week Multi-conference* (Sydney, NSW, Australia) *(ACSW '19)*. Association for Computing Machinery, New York, NY, USA, Article 47, 9 pages. https://doi.org/10.1145/3290688.3290752 pages #28
[86] Hana Hanifah, Yuko Ito, Daryl Patrick Gamboa Yao, Natsuka Suyama, and Kaoru Inoue. 2022. Promoting Sports Engagement during the COVID-19 Pandemic via Virtual Reality Games. *Occupational Therapy International* 2022 (2022). https://doi.org/10.1155/2022/4824152 pages #26, #28
[87] Inc. Harmonix Music Systems. 2019. *Dance Central.* Game [VR]. Harmonix Music Systems, Inc. Boston, MA, USA. https://www.dancecentral.com/. Meta Reviews: https://www.meta.com/experiences/2453152771391571/. pages #12, #26
[88] Oo Jin Heng and Quek Albert. 2020. Bubble Tower: Breathing Based Virtual Reality Action Game. In *Proceedings of the 2020 4th International Conference on Big Data and Internet of Things* (Singapore, Singapore) *(BDIOT '20)*. Association for Computing Machinery, New York, NY, USA, 33–37. https://doi.org/10.1145/3421537.3421552 pages #9, #11, #27
[89] Nadia Vanessa Garcia Hernandez, Stefano Buccelli, Matteo Laffranchi, and Lorenzo De Michieli. 2023. Mixed Reality-based Exergames for Upper Limb Robotic Rehabilitation. *ACM/IEEE International Conference on Human-Robot Interaction* (2023), 447 – 451. https://doi.org/10.1145/3568294.3580124 pages #3, #9, #10, #12, #13, #28
[90] Julia Hertel, Sukran Karaosmanoglu, Susanne Schmidt, Julia Bräker, Martin Semmann, and Frank Steinicke. 2021. A Taxonomy of Interaction Techniques for Immersive Augmented Reality based on an Iterative Literature Review. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 431–440. https://doi.org/10.1109/ISMAR52148.2021.00060 pages #6
[91] Teresa Hirzle, Florian Müller, Fiona Draxler, Martin Schmitz, Pascal Knierim, and Kasper Hornbæk. 2023. When XR and AI Meet - A Scoping Review on Extended Reality and Artificial Intelligence. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 730, 45 pages. https://doi.org/10.1145/3544548.3581072 pages #1, #3
[92] Mohamad Hoda, Haiwei Dong, Dewan Ahmed, and Abdulmotaleb El Saddik. 2014. Cloud-based rehabilitation exergames system. *2014 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2014* (2014). https://doi.org/10.1109/ICMEW.2014.6890660 pages #5
[93] Holodia. 2021. *Holofit by Holodia.* Game [VR]. Holodia. Schiltigheim, France. https://www.holodia.com/. Meta Reviews: https://www.meta.com/experiences/2853664488067174/. pages #11, #26
[94] Yu-Jun Hong, Chen-Yuan Hsieh, Keng-Ta Yang, and Liwei Chan. 2017. Wakeboarding: An Exertion Game in Virtual Reality. In *ACM SIGGRAPH 2017 VR Village* (Los Angeles, California) *(SIGGRAPH '17)*. Association for Computing Machinery, New York, NY, USA, Article 15, 2 pages. https://doi.org/10.1145/3089269.3089271 pages #13, #27
[95] Kuo-Ting Huang. 2020. Exergaming Executive Functions: An Immersive Virtual Reality-Based Cognitive Training for Adults Aged 50 and Older. *Cyberpsychology, Behavior, and Social Networking* 23, 3 (2020), 143 – 149. https://doi.org/10.1089/cyber.2019.0269 pages #26, #28
[96] Tobias Huber, Silvan Mertes, Stanislava Rangelova, Simon Flutura, and Elisabeth André. 2021. Dynamic Difficulty Adjustment in Virtual Reality Exergames through Experience-driven Procedural Content Generation. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. 1–8. https://doi.org/10.1109/SSCI50451.2021.9660086 pages #27
[97] Emil R. Høeg, Nikolaj B. Andersen, Nicolai Malmkjær, Adrian H. Vaaben, and Jacob Uth. 2023. Hospitalized older adults' experiences of virtual reality-based group exercise therapy with cycle ergometers: An early feasibility study. *Computers in Human Behavior Reports* 11 (2023). https://doi.org/10.1016/j.chbr.2023.100301 pages #26, #28
[98] Emil Rosenlund Høeg, Jon Ram Bruun-Pedersen, Shannon Cheary, Lars Koreska Andersen, Razvan Paisa, Stefania Serafin, and Belinda Lange. 2023. Buddy biking: a user study on social collaboration in a virtual reality exergame for rehabilitation. *Virtual Reality* (2023). https://doi.org/10.1007/s10055-021-00544-z pages #6, #11, #27
[99] Kiran Ijaz, Yifan Wang, Naseem Ahmadpour, and Rafael A. Calvo. 2017. Physical activity enjoyment on an immersive VR exergaming platform. In *2017 IEEE Life Sciences Conference (LSC)*. 59–62. https://doi.org/10.1109/LSC.2017.8268143 pages #28
[100] Kiran Ijaz, Yifan Wang, Naseem Ahmadpour, and Rafael A. Calvo. 2019. Immersive VR Exergames for Health and Wellbeing. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10.1145/3290607.3313281 pages #13, #27
[101] Kiran Ijaz, Yifan Wang, David Milne, and Rafael A. Calvo. 2016. Competitive vs Affiliative Design of Immersive VR Exergames. In *Serious Games*, Tim Marsh, Minhua Ma, Manuel Fradinho Oliveira, Jannicke Baalsrud Hauge, and Stefan

Göbel (Eds.). Springer International Publishing, Cham, 140–150. pages #27

[102] Kiran Ijaz, Yifan Wang, David Milne, and Rafael A. Calvo. 2016. VR-Rides: Interactive VR Games for Health. In *Serious Games*, Tim Marsh, Minhua Ma, Manuel Fradinho Oliveira, Jannicke Baalsrud Hauge, and Stefan Göbel (Eds.). Springer International Publishing, Cham, 289–292. pages #13, #28

[103] Wijnand A IJsselsteijn, Yvonne AW De Kort, and Karolien Poels. 2013. The game experience questionnaire. (2013). pages #8

[104] Christos Ioannou, Patrick Archard, Eamonn O'Neill, and Christof Lutteroth. 2019. Virtual Performance Augmentation in an Immersive Jump & Run Exergame. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300388 pages #1, #3, #12, #27

[105] Shun Ishii, Mika Luimula, Anna Yokokubo, and Guillaume Lopez. 2020. VR Dodge-ball: Application of Real-time Gesture Detection from Wearables to ExerGaming. In *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. 000081–000082. https://doi.org/10.1109/CogInfoCom50765.2020.9237824 pages #27

[106] João Jacob, Ana Lopes, Rui Nóbrega, Rui Rodrigues, and António Coelho. 2018. Towards Player Adaptivity in Mobile Exergames. In *Advances in Computer Entertainment Technology*, Adrian David Cheok, Masahiko Inami, and Teresa Romão (Eds.). Springer International Publishing, Cham, 278–292. pages #11, #27

[107] Seth E Jenny, R Douglas Manning, Margaret C Keiper, and Tracy W Olrich. 2017. Virtual(ly) Athletes: Where eSports Fit Within the Definition of "Sport". *Quest* 69, 1 (2017), 1–18. https://doi.org/10.1080/00336297.2016.1144517 pages #3

[108] Daniel Johnson, M John Gardner, and Ryan Perry. 2018. Validation of two game experience scales: the player experience of need satisfaction (PENS) and game experience questionnaire (GEQ). *International Journal of Human-Computer Studies* 118 (2018), 38–46. pages #8

[109] Alexandra Kalaitzidou, Nathalie Senechal, Paschalis Dimitriou, Krishnan Chandran, and Matthew Mcginity. 2022. "E-WAFE" - A Full Body Embodied Social Exergame. In *Extended Abstracts of the 2022 Annual Symposium on Computer-Human Interaction in Play* (Bremen, Germany) *(CHI PLAY '22)*. Association for Computing Machinery, New York, NY, USA, 286–290. https://doi.org/10.1145/3505270.3558375 pages #9, #27

[110] Dennis L. Kappen, Pejman Mirza-Babaei, and Lennart E. Nacke. 2019. Older Adults' Physical Activity and Exergames: A Systematic Review. *International Journal of Human–Computer Interaction* 35, 2 (2019), 140–167. https://doi.org/10.1080/10447318.2018.1441253 pages #3, #7

[111] Sukran Karaosmanoglu, Lucie Kruse, Sebastian Rings, and Frank Steinicke. 2022. Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 342, 7 pages. https://doi.org/10.1145/3491101.3519736 pages #9, #11, #28

[112] Sukran Karaosmanoglu, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. 2021. Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 252 (oct 2021), 27 pages. https://doi.org/10.1145/3474679 pages #2, #3, #6, #11, #13, #14, #27

[113] Sukran Karaosmanoglu, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. 2021. Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 443, 15 pages. https://doi.org/10.1145/3411764.3445492 pages #14

[114] Abdel Ghani Karkar, Somaya AlMaadeed, Rehab Salem, Mariam AbdelHady, Sara Abou-Aggour, and Hafsa Samea. 2018. KinFit: A factual aerobic sport game with stimulation support. *International Journal of Emerging Technologies in Learning* 13, 12 (2018), 50 – 66. https://doi.org/10.3991/ijet.v13i12.8626 pages #16, #28

[115] Stamos Katsigiannis, Rhys Willis, and Naeem Ramzan. 2019. A QoE and Simulator Sickness Evaluation of a Smart-Exercise-Bike Virtual Reality System via User Feedback and Physiological Signals. *IEEE Transactions on Consumer Electronics* 65, 1 (2019), 119–127. https://doi.org/10.1109/TCE.2018.2879065 pages #27

[116] Aaron Keesing, Matthew Ooi, Ocean Wu, Xinghao Ye, Lindsay Shaw, and Burkhard C. Wünsche. 2019. HIIT With Hits: Using Music and Gameplay to Induce HIIT in Exergames. In *Proceedings of the Australasian Computer Science Week Multiconference* (Sydney, NSW, Australia) *(ACSW '19)*. Association for Computing Machinery, New York, NY, USA, Article 36, 10 pages. https://doi.org/10.1145/3290688.3290740 pages #10, #12, #27

[117] Marie Kegeleers, Shivam Miglani, Gijs M. W. Reichert, Nestor Z. Salamon, J. Timothy Balint, Stephan G. Lukosch, and Rafael Bidarra. 2018. STAR: Superhuman Training in Augmented Reality. In *Proceedings of the First Superhuman Sports Design Challenge: First International Symposium on Amplifying Capabilities and Competing in Mixed Realities* (Delft, Netherlands) *(SHS '18)*. Association for Computing Machinery, New York, NY, USA, Article 7, 6 pages. https://doi.org/10.1145/3210299.3210306 pages #10, #28

[118] Nathan Keller, Richard S. Whittle, Neil McHenry, Adam Johnston, Colton Duncan, Lori Ploutz-Snyder, Gabriel G. De La Torre, Melinda Sheffield-Moore, Gregory Chamitoff, and Ana Diaz-Artiles. 2022. Virtual Reality "exergames": A promising countermeasure to improve motivation and restorative effects during long duration spaceflight missions. *Frontiers in Physiology* 13 (2022). https://doi.org/10.3389/fphys.2022.932425 pages #13, #28

[119] Ursina Kellerhals, Naomi Burgess, and Richard Wetzel. 2021. Let it Bee: A Case Study of Applying Triadic Game Design for Designing Virtual Reality Training Games for Beekeepers. *ACM International Conference Proceeding Series* (2021). https://doi.org/10.1145/3472538.3472558 pages #5

[120] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220. https://doi.org/10.1207/s15327108ijap0303_3 pages #8

[121] Eng Tat Khoo, Shang Ping Lee, Adrian David Cheok, Sameera Kodagoda, Yu Zhou, and Gin Siong Toh. 2006. Age Invaders: Social and Physical Inter-Generational Family Entertainment. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI EA '06)*. Association for Computing Machinery, New York, NY, USA, 243–246. https://doi.org/10.1145/1125451.1125503 pages #5

[122] Chaowanan Khundam and Frédéric Nöel. 2021. A study of physical fitness and enjoyment on virtual running for exergames. *International Journal of Computer Games Technology* 2021 (2021), 1–16. https://doi.org/10.1155/2021/6668280 pages #28

[123] Kolja Kirsch, Christian Schatzschneider, Carina Garber, Aila Rosenberger, Konya Kirsten, Oscar Ariza, Frank Steinicke, and Gerd Bruder. 2019. KiVR Sports: Influencing the Users Physical Activity in VR by Using Audiovisual Stimuli in Exergames. In *Proceedings of Mensch Und Computer 2019* (Hamburg, Germany) *(MuC '19)*. Association for Computing Machinery, New York, NY, USA, 777–781. https://doi.org/10.1145/3340764.3344907 pages #27

[124] Oona Kivela, Paula Alavesa, Aku Visuri, and Timo Ojala. 2019. Study on the motivational and physical effects of two VR exergames. *2019 11th International Conference on Virtual Worlds and Games for Serious Applications, VS-Games 2019 - Proceedings* (2019), 1DUUMY. https://doi.org/10.1109/VS-Games.2019.8864544 pages #26, #28

[125] Tanja Kojic, Steven Schmidt, Sebastian Möller, and Jan-Niklas Voigt-Antons. 2019. Influence of Network Delay in Virtual Reality Multiplayer Exergames: Who is actually delayed?. In *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. 1–3. https://doi.org/10.1109/QoMEX.2019.8743342 pages #27

[126] Tanja Kojic, Uliana Sirotina, Sebastian Möller, and Jan-Niklas Voigt-Antons. 2019. Influence of UI Complexity and Positioning on User Experience During VR Exergames. In *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. 1–6. https://doi.org/10.1109/QoMEX.2019.8743273 pages #27

[127] Tanja Kojić, Lan Thao Nugyen, and Jan-Niklas Voigt-Antons. 2019. Impact of Constant Visual Biofeedback on User Experience in Virtual Reality Exergames. In *2019 IEEE International Symposium on Multimedia (ISM)*. IEEE, 307–3073. https://doi.org/10.1109/ISM46123.2019.00068 pages #11, #13, #27

[128] Tanja Kojić, Jan-Niklas Voigt-Antons, Steven Schmidt, Lukas Tetzlaff, Bruno Kortowski, Uliana Sirotina, and Sebastian Möller. 2018. Influence of Virtual Environments and Conversations on User Engagement During Multiplayer Exergames. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 1–3. https://doi.org/10.1109/QoMEX.2018.8463370 pages #12, #27

[129] Reetu Kontio, Markus Laattala, Robin Welsch, and Perttu Hämäläinen. 2023. "I Feel My Abs": Exploring Non-Standing VR Locomotion. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY, Article 423 (oct 2023), 26 pages. https://doi.org/10.1145/3611069 pages #15

[130] Jordan Koulouris, Zoe Jeffery, James Best, Eamonn O'Neill, and Christof Lutteroth. 2020. Me vs. Super(Wo)Man: Effects of Customization and Identification in a VR Exergame. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3313831.3376661 pages #12, #27

[131] Lucie Kruse, Sukran Karaosmanoglu, Sebastian Rings, Benedikt Ellinger, and Frank Steinicke. 2021. Enabling immersive exercise activities for older adults: A comparison of virtual reality exergames and traditional video exercises. *Societies* 11, 4 (2021). https://doi.org/10.3390/soc11040134 pages #12, #28

[132] Lucie Kruse, Sukran Karaosmanoglu, Sebastian Rings, and Frank Steinicke. 2022. Evaluating Difficulty Adjustments in a VR Exergame for Younger and Older Adults: Transferabilities and Differences. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction* (Online, CA, USA) *(SUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 7, 11 pages. https://doi.org/10.1145/3565970.3567684 pages #27

[133] J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (1977), 159–174. http://www.

jstor.org/stable/2529310 pages #5, #6
[134] Hoo Kyung Lee and Jooyeon Jin. 2023. The effect of a virtual reality exergame on motor skills and physical activity levels of children with a developmental disability. *Research in Developmental Disabilities* 132 (2023). https://doi.org/10.1016/j.ridd.2022.104386 pages #26, #28
[135] Wanwan Li. 2022. Procedural Marine Landscape Synthesis for Swimming Exergame in Virtual Reality. In *2022 IEEE Games, Entertainment, Media Conference (GEM)*. 1–8. https://doi.org/10.1109/GEM56474.2022.10017505 pages #27
[136] Wanwan Li. 2022. Simulating Ice Skating Experience in Virtual Reality. *2022 7th International Conference on Image, Vision and Computing, ICIVC 2022* (2022), 706 – 712. https://doi.org/10.1109/ICIVC55077.2022.9887260 pages #6
[137] Wanwan Li. 2023. Terrain Synthesis for Treadmill Exergaming in Virtual Reality. *Proceedings - 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VRW 2023* (2023), 263 – 269. https://doi.org/10.1109/VRW58643.2023.00064 pages #13, #27
[138] Xiang Li and Yuzheng Chen. 2020. Auto-Hierarchical Data Algorithm: Focus on Increasing Users' Motivation and Duration in Virtual Reality. *2020 5th IEEE International Conference on Big Data Analytics, ICBDA 2020* (2020), 150 – 153. https://doi.org/10.1109/ICBDA49040.2020.9101254 pages #12, #28
[139] Hui Liang, Fanyu Bao, Yusheng Sun, Qian Zhang, Mingge Pan, and Jian Chang. 2021. Children's Early Educational Game under the Background of Chinese Three Kingdoms Culture — To Borrow Arrows with Thatched Boats. In *2021 IEEE 7th International Conference on Virtual Reality (ICVR)*. IEEE, 180–188. https://doi.org/10.1109/ICVR51878.2021.9483814 pages #9, #28
[140] Agris Liepa, Jason Tang, Ingemara Jaundaldere, Evita Dubinina, Viesturs Larins, et al. 2022. Feasibility randomized controlled trial of a virtual reality exergame to improve physical and cognitive functioning in older people. *Acta Gymnica* (2022). pages #10, #13, #27
[141] Chenli Lin, Yuanyuan Ren, and Aming Lu. 2023. The effectiveness of virtual reality games in improving cognition, mobility, and emotion in elderly post-stroke patients: a systematic review and meta-analysis. *Neurosurgical Review* 46, 1 (2023), 167. https://doi.org/10.1007/s10143-023-02061-w pages #3
[142] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How WEIRD is CHI?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 143, 14 pages. https://doi.org/10.1145/3411764.3445488 pages #14
[143] Huimin Liu, Zhiquan Wang, Christos Mousas, and Dominic Kao. 2020. Virtual Reality Racket Sports: Virtual Drills for Exercise and Training. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 566–576. https://doi.org/10.1109/ISMAR50242.2020.00084 pages #13, #28
[144] Sealost Interactive LLC. 2019. *The Thrill of the Fight.* Game [VR]. Sealost Interactive LLC. Knoxville, TN, USA. http://sealostinteractive.com/. Meta Reviews: https://www.meta.com/experiences/3008315795852749/. pages #26
[145] Halfbrick Studios Pty Ltd. 2016. *Fruit Ninja VR.* Game [VR]. Halfbrick Studios Pty Ltd. Red Hill, Australia. https://www.halfbrick.com/games/fruit-ninja-vr. Steam Reviews: https://store.steampowered.com/app/486780/Fruit_Ninja_VR/. pages #10, #12, #26
[146] Kenziko Ltd. 2017. *Snow Games VR.* Game [VR]. Kenziko Ltd. Steam Reviews: https://store.steampowered.com/app/579080/Snow_Games_VR/. pages #26
[147] Amy Shirong Lu, Victoria Pelarski, Dar Alon, Aleksandra Baran, Emma McGarrity, Neha Swaminathan, and Caio Victor Sousa. 2023. The effect of narrative element incorporation on physical activity and game experience in active and sedentary virtual reality games. *Virtual Reality* 27, 3 (2023), 1607 – 1622. https://doi.org/10.1007/s10055-023-00754-7 pages #26, #28
[148] Alexander Hvidbjerg Kjær Lund, Amalie Finnemannn Sørensen, Lars Elbæk, and Maximus D. Kaos. 2021. Insights from design processes used in developing exergames. *Proceedings of the European Conference on Games-based Learning* 2021-September (2021), 490 – 498. https://doi.org/10.34190/GBL.21.102 pages #9, #13, #27
[149] Cayley MacArthur, Arielle Grinberg, Daniel Harley, and Mark Hancock. 2021. You're Making Me Sick: A Systematic Review of How Virtual Reality Research Considers Gender & Cybersickness. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 401, 15 pages. https://doi.org/10.1145/3411764.3445701 pages #5
[150] Joe Y. Mambu, Rismayani, Jay Idoan Sihotang, and Vivi Peggy Rantung. 2022. VR REAL RUN: An immersive Oculus Quest 2-Based Virtual Reality Exergaming. In *2022 4th International Conference on Cybernetics and Intelligent System (ICORIS)*. IEEE, 1–6. https://doi.org/10.1109/ICORIS56080.2022.10031600 pages #12, #28
[151] Anna Lisa Martin-Niedecken, Andrea Mahrer, Katja Rogers, Eling D. de Bruin, and Alexandra Schättin. 2020. "HIIT" the ExerCube: Comparing the Effectiveness of Functional High-Intensity Interval Training in Conventional vs. Exergame-Based Training. *Frontiers in Computer Science* 2 (2020). https://doi.org/10.3389/fcomp.2020.00033 pages #6
[152] Pamela N. Martinez. 2022. Design recommendations for active games. *Frontiers in Digital Health* 4 (2022). https://doi.org/10.3389/fdgth.2022.814226 pages #15
[153] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-Rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (nov 2019), 23 pages. https://doi.org/10.1145/3359174 pages #5
[154] Mary L McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia medica* 22, 3 (2012), 276–282. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/ pages #5, #6
[155] Don D. McMahon, Brenda Barrio, Amanda K. McMahon, Kristen Tutt, and Jonah Firestone. 2020. Virtual Reality Exercise Games for High School Students With Intellectual and Developmental Disabilities. *Journal of Special Education Technology* 35, 2 (2020), 87 – 96. https://doi.org/10.1177/0162643419836416 pages #26, #28
[156] Luis Andres Mendez, Ho Yin Ng, and Ping-Hsuan Han. 2023. MovableBag: Exploring Asymmetric Interaction for Multi-User Exergame in Extended Reality. In *Adjunct Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2022 ACM International Symposium on Wearable Computers* (Cambridge, United Kingdom) *(UbiComp/ISWC '22 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 515–519. https://doi.org/10.1145/3544793.3561316 pages #27
[157] Meta Platforms, Inc. 2020. Meta Quest 2. https://www.meta.com/ca/quest/products/quest-2/ Released: 13 October 2020. pages #2
[158] Alexander Michael and Christof Lutteroth. 2020. Race Yourselves: A Longitudinal Exploration of Self-Competition Between Past, Present, and Future Performances in a VR Exergame. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3313831.3376256 pages #10, #12, #27
[159] Microsoft. 2010. Kinect. Released: 4 November 2010. pages #2
[160] Microsoft. 2019. Microsoft HoloLens 2. https://www.microsoft.com/en-us/hololens pages #2, #3, #12
[161] Zeljka Mihajlovic, Sinisa Popovic, Karla Brkic, and Kresimir Cosic. 2018. A system for head-neck rehabilitation exercises based on serious gaming and virtual reality. *Multimedia Tools and Applications* 77 (2018), 19113–19137. pages #10, #13, #27
[162] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329. https://search.ieice.org/bin/summary.php?id=e77-d_12_1321 pages #2
[163] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1995. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, Vol. 2351. Spie, 282–292. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/2351/1/Augmented-reality--a-class-of-displays-on-the-reality/10.1117/12.197321.full pages #2, #4
[164] Nikesh Mishra and Eelke Folmer. 2018. Measuring Physical Exertion in Virtual Reality Exercise Games. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 106, 2 pages. https://doi.org/10.1145/3281505.3281611 pages #27
[165] Nan Mo, Jin yu Feng, Hai xia Liu, Xiao yu Chen, Hui Zhang, Hui Zeng, et al. 2023. Effects of Exergaming on Musculoskeletal Pain in Older Adults: Systematic Review and Meta-analysis. *JMIR Serious Games* 11, 1 (2023), e42944. pages #3
[166] Merete Endresen Moan, Elise Klæbo Vonstad, Xiaomeng Su, Beatrix Vereijken, Marit Solbjør, and Nina Skjæret-Maroni. 2021. Experiences of Stroke Survivors and Clinicians With a Fully Immersive Virtual Reality Treadmill Exergame for Stroke Rehabilitation: A Qualitative Pilot Study. *Frontiers in Aging Neuroscience* 13 (2021). https://doi.org/10.3389/fnagi.2021.735251 pages #9, #12, #27
[167] Mitchell S. Mologne, Jonathan Hu, Erik Carrillo, David Gomez, Trent Yamamoto, Stevin Lu, Jonathan D. Browne, and Brett A. Dolezal. 2023. The Efficacy of an Immersive Virtual Reality Exergame Incorporating an Adaptive Cable Resistance System on Fitness and Cardiometabolic Measures: A 12-Week Randomized Controlled Trial. *International Journal of Environmental Research and Public Health* 20, 1 (2023). https://doi.org/10.3390/ijerph20010210 pages #26, #28
[168] Marta Mondellini, Marco Sacco, and Luca Greci. 2020. Virtual Fitness Trail: A Complete Program for Elderlies to Perform Physical Activity at Home. In *Augmented Reality, Virtual Reality, and Computer Graphics*, Lucio Tommaso De Paolis and Patrick Bourdot (Eds.). Springer International Publishing, Cham, 168–176. pages #9, #27
[169] Maria F. Montoya, Julian F. Villada, John Muñoz, and Oscar A. Henao. 2021. Exploring coordination patterns in vr-based rehabilitation for stroke using the kinect sensor. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 12790 LNCS (2021), 372 – 387. https://doi.org/10.1007/978-3-030-77414-1_27 pages #26, #28
[170] Florian 'Floyd' Mueller, Darren Edge, Frank Vetere, Martin R. Gibbs, Stefan Agamanolis, Bert Bongers, and Jennifer G. Sheridan. 2011. Designing Sports: A Framework for Exertion Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2651–2660. https://doi.org/10.1145/1978942.1979330 pages #2, #3, #13

[171] Florian 'Floyd' Mueller, Martin R. Gibbs, and Frank Vetere. 2008. Taxonomy of Exertion Games. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat* (Cairns, Australia) *(OZCHI '08)*. Association for Computing Machinery, New York, NY, USA, 263–266. https://doi.org/10.1145/1517744.1517772 pages #3

[172] Stanley Mugisha, Vamsi Krisha Guda, Christine Chevallereau, Matteo Zoppi, Rezia Molfino, and Damien Chablat. 2022. Improving Haptic Response for Contextual Human Robot Interaction. *Sensors* 22, 5 (2022). https://doi.org/10.3390/s22052040 pages #27

[173] Zachary Munn, Micah DJ Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur, and Edoardo Aromataris. 2018. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC medical research methodology* 18 (2018), 1–7. https://doi.org/10.1186/s12874-018-0611-x pages #4, #5

[174] John Muñoz, Samira Mehrabi, Yirou Li, Aysha Basharat, Laura E Middleton, Shi Cao, Michael Barnett-Cowan, and Jennifer Boger. 2022. Immersive Virtual Reality Exergames for Persons Living With Dementia: User-Centered Design Study as a Multistakeholder Team During the COVID-19 Pandemic. *JMIR Serious Games* 10, 1 (19 Jan 2022), e29987. https://doi.org/10.2196/29987 pages #10, #12, #28

[175] National Institute of Aging. 2021. Four Types of Exercise Can Improve Your Health and Physical Ability. Website. Retrieved September 4, 2023 from https://www.nia.nih.gov/health/four-types-exercise-can-improve-your-health-and-physical-ability. pages #7

[176] Aravind Nehrujee, Lenny Vasanthan, Anjali Lepcha, and Sivakumar Balasubramanian. 2019. A Smartphone-based gaming system for vestibular rehabilitation: A usability study. *Journal of Vestibular Research* 29, 2-3 (2019), 147–160. https://doi.org/10.3233/VES-190660 pages #28

[177] Leticia Neira-Tovar and Marco Eduardo Elizondo Elizondo. 2018. A Virtual Reality Serious Game to Prevent Type 2 Diabetes Mellitus. In *Smart Technology*, Francisco Torres Guerrero, Jorge Lozoya-Santos, Eduardo Gonzalez Mendivil, Leticia Neira-Tovar, Pablo G. Ramírez Flores, and Jorge Martin-Gutierrez (Eds.). Springer International Publishing, Cham, 178–188. pages #9, #27

[178] Luceli Eunice Da Silva Neves, Mariza Paver Da Silva Cerávolo, Elisangela Silva, Wagner Zeferino De Freitas, Fabiano Fernandes Da Silva, Wonder Passoni Higino, Wellington Roberto Gomes Carvalho, and Renato Aparecido De Souza. 2015. Cardiovascular effects of Zumba® performed in a virtual environment using XBOX Kinect. *Journal of physical therapy science* 27, 9 (2015), 2863–2865. https://doi.org/10.1589/jpts.27.2863 pages #5

[179] Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand, and Jon Howard Raphael. 2012. No room for reflexivity? Critical reflections following a systematic review of qualitative research. *Journal of Health Psychology* 17, 6 (2012), 866–885. https://doi.org/10.1177/1359105311427615 pages #7, #15

[180] Robert C Nickerson, Upkar Varshney, and Jan Muntermann. 2013. A method for taxonomy development and its application in information systems. *European Journal of Information Systems* 22, 3 (2013), 336–359. https://doi.org/10.1057/ejis.2012.26 pages #2, #6, #7, #9, #15

[181] Nintendo EAD. 2007. *Wii Fit.* Game [Wii]. pages #1, #2

[182] Genta Ochi, Ryuta Kuwamizu, Tomomi Fujimoto, Koyuki Ikarashi, Koya Yamashiro, and Daisuke Sato. 2022. The Effects of Acute Virtual Reality Exergaming on Mood and Executive Function: Exploratory Crossover Trial. *JMIR Serious Games* 10, 3 (28 Sep 2022), e38200. https://doi.org/10.2196/38200 pages #26, #28

[183] Oculus. 2020. *First Steps for Quest 2.* Game [VR]. Oculus. Menlo Park, CA, USA. Meta Reviews: https://www.meta.com/experiences/3675568169182204. pages #26

[184] Ifeanyi Paul Odenigbo, Jaisheen Kour Reen, Chimamaka Eneze, Aniefiok Friday, and Rita Orji. 2022. Virtual, Augmented, and Mixed Reality Interventions for Physical Activity: A Systematic Review. In *2022 IEEE 10th International Conference on Serious Games and Applications for Health(SeGAH)*. IEEE, 1–9. https://doi.org/10.1109/SEGAH54908.2022.9978564 pages #3

[185] Toshikazu Ohshima, Kiyohide Satoh, Hiroyuki Yamamoto, and Hideyuki Tamura. 1998. AR2Hockey: A case study of collaborative augmented reality. *Proceedings - Virtual Reality Annual International Symposium*, 268 – 275. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0031680416&partnerID=40&md5=7c9910e9cafb590bd77dad8ac4ba0291 pages #12, #28

[186] Ottawa Hospital Research Institute (OHRI): PRISMA Group. 2020. PRISMA 2020 flowchart. http://prisma-statement.org/prismastatement/flowdiagram.aspx. Last accessed: December 23, 2021.. pages #4

[187] Yang-Kun Ou, Yu-Lin Wang, Hua-Cheng Chang, Shih-Yin Yen, Yu-Hua Zheng, and Bih-O. Lee. 2020. Development of virtual reality rehabilitation games for children with attention-deficit hyperactivity disorder. *Journal of Ambient Intelligence and Humanized Computing* 11, 11 (2020), 5713 – 5720. https://doi.org/10.1007/s12652-020-01945-9 pages #9, #10, #16, #27

[188] Daniel Palacios-Alonso, Agustín López-Arribas, Guillermo Meléndez-Morales, Esther Núñez-Vidal, Andrés Gómez-Rodellar, José Manuel Ferrández-Vicente, and Pedro Gómez-Vilda. 2022. A Pilot and Feasibility Study of Virtual Reality as Gamified Monitoring Tool for Neurorehabilitation. In *Artificial Intelligence in Neuroscience: Affective Analysis and Health Applications*, José Manuel Ferrández Vicente, José Ramón Álvarez-Sánchez, Félix de la Paz López, and Hojjat Adeli (Eds.). Springer International Publishing, Cham, 239–248. pages #28

[189] Daniel Palacios-Alonso, Agustín López-Arribas, Guillermo Meléndez-Morales, Esther Núñez-Vidal, Andrés Gómez-Rodellar, José Manuel Ferrández-Vicente, and Pedro Gómez-Vilda. 2022. A Pilot and Feasibility Study of Virtual Reality as Gamified Monitoring Tool for Neurorehabilitation. In *Artificial Intelligence in Neuroscience: Affective Analysis and Health Applications*, José Manuel Ferrández Vicente, José Ramón Álvarez-Sánchez, Félix de la Paz López, and Hojjat Adeli (Eds.). Springer International Publishing, Cham, 239–248. pages #28

[190] Shanmugam Muruga Palaniappan and Bradley S. Duerstock. 2018. Developing Rehabilitation Practices Using Virtual Reality Exergaming. In *2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. IEEE, 090–094. https://doi.org/10.1109/ISSPIT.2018.8642784 pages #9, #27

[191] Shanmugam Muruga Palaniappan, Shruthi Suresh, Jeffrey M. Haddad, and Bradley S. Duerstock. 2020. Adaptive Virtual Reality Exergame for Individualized Rehabilitation for Persons with Spinal Cord Injury. In *Computer Vision – ECCV 2020 Workshops*, Adrien Bartoli and Andrea Fusiello (Eds.). Springer International Publishing, Cham, 518–535. pages #28

[192] Shanmugam Muruga Palaniappan, Ting Zhang, and Bradley S. Duerstock. 2019. Identifying Comfort Areas in 3D Space for Persons with Upper Extremity Mobility Impairments Using Virtual Reality. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 495–499. https://doi.org/10.1145/3308561.3353810 pages #28

[193] Jong-Seung Park, TaeYong Kim, and Jong-Hyun Yoon. 2006. AR Table Tennis: A Video-Based Augmented Reality Sports Game. In *Advances in Artificial Reality and Tele-Existence*, Zhigeng Pan, Adrian Cheok, Michael Haller, Rynson W. H. Lau, Hideo Saito, and Ronghua Liang (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 197–206. pages #12, #28

[194] Russell R Pate, Ruth P Saunders, Jennifer R O'Neill, and Marsha Dowda. 2011. Overcoming barriers to physical activity: helping youth be more active. *ACSM's Health & Fitness Journal* 15, 1 (2011), 7–12. https://doi.org/10.1249/FIT.0b013e318201c8ee pages #1

[195] Rakesh Patibanda, Florian 'Floyd' Mueller, Matevz Leskovsek, and Jonathan Duckworth. 2017. Life Tree: Understanding the Design of Breathing Exercise Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17)*. Association for Computing Machinery, New York, NY, USA, 19–31. https://doi.org/10.1145/3116595.3116621 pages #10, #27

[196] Micah DJ Peters, Christina Godfrey, Patricia McInerney, Hanan Khalil, Palle Larsen, Casey Marnie, Danielle Pollock, Andrea C Tricco, and Zachary Munn. 2022. Best practice guidance and reporting items for the development of scoping review protocols. *JBI evidence synthesis* 20, 4 (2022), 953–968. https://doi.org/10.11124/JBIES-21-00242 pages #4

[197] Micah DJ Peters, Christina M Godfrey, Hanan Khalil, Patricia McInerney, Deborah Parker, and Cassia Baldini Soares. 2015. Guidance for conducting systematic scoping reviews. *JBI Evidence Implementation* 13, 3 (2015), 141–146. https://doi.org/10.1097/XEB.0000000000000050 pages #4

[198] Alexander Plopski, Teresa Hirzle, Nahal Norouzi, Long Qian, Gerd Bruder, and Tobias Langlotz. 2022. The Eye in Extended Reality: A Survey on Gaze Interaction and Eye Tracking in Head-Worn Extended Reality. *ACM Comput. Surv.* 55, 3, Article 53 (mar 2022), 39 pages. https://doi.org/10.1145/3491207 pages #1, #3

[199] John Porter and Andrew Robb. 2019. An Analysis of Longitudinal Trends in Consumer Thoughts on Presence and Simulator Sickness in VR Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 277–285. https://doi.org/10.1145/3311350.3347159 pages #14

[200] Dees B.W. Postma, Robby W. van Delden, Jeroen H. Koekoek, Wytse W. Walinga, Ivo M. van Hilvoorde, Bert Jan F. van Beijnum, Fahim A. Salim, and Dennis Reidsma. 2022. A Design Space of Sports Interaction Technology. *Foundations and Trends in Human-Computer Interaction* 15, 2-3 (25 Aug. 2022), 132–316. https://doi.org/10.1561/1100000087 pages #3

[201] Vera Z. Pérez, Juan C. Yepes, John F. Vargas, Juan C. Franco, Natalia I. Escobar, Leonardo Betancur, Juanita Sánchez, and Manuel J. Betancur. 2022. Virtual Reality Game for Physical and Emotional Rehabilitation of Landmine Victims. *Sensors* 22, 15 (2022). https://doi.org/10.3390/s22155602 pages #13, #28

[202] Priyanka Ramasamy, Swagata Das, and Yuichi Kurita. 2021. Ski for Squat: A Squat Exergame with Pneumatic Gel Muscle-Based Dynamic Difficulty Adjustment. In *Universal Access in Human-Computer Interaction. Access to Media, Learning and Assistive Environments*, Margherita Antona and Constantine Stephanidis (Eds.). Springer International Publishing, Cham, 449–467. pages #11, #13, #28

[203] Jack Ratcliffe, Francesco Soave, Nick Bryan-Kinns, Laurissa Tokarchuk, and Ildar Farkhatdinov. 2021. Extended Reality (XR) Remote Research: A Survey of Drawbacks and Opportunities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 527, 13 pages. https:

//doi.org/10.1145/3411764.3445170 pages #1, #3

[204] Philipp A. Rauschnabel, Reto Felix, Chris Hinsch, Hamza Shahab, and Florian Alt. 2022. What is XR? Towards a Framework for Augmented and Virtual Reality. *Computers in Human Behavior* 133 (2022), 107289. https://doi.org/10.1016/j.chb.2022.107289 pages #3

[205] Clifford A Reilly, Aimee Burnett Greeley, David S Jevsevar, and Ida Leah Gitajn. 2021. Virtual reality-based physical therapy for patients with lower extremity injuries: Feasibility and acceptability. *OTA International* 4, 2 (2021). https://doi.org/10.1097/OI9.0000000000000132 pages #16, #28

[206] Sean Reilly, Peter Barron, Vinny Cahill, Kieran Moran, and Mads Haahr. 2009. A General-Purpose Taxonomy of Computer-Augmented Sports Systems. *Digital Sport for Performance Enhancement and Competitive Evolution: Intelligent Gaming Technologies: Intelligent Gaming Technologies* (2009), 19. pages #3

[207] Patrizia Ring and Maic Masuch. 2023. Measuring Collision Anxiety in XR Exergames. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 627–628. https://doi.org/10.1109/VRW58643.2023.00156 pages #3, #27

[208] Sebastian Rings, Sukran Karaosmanoglu, Lucie Kruse, Daniel Apken, Tobias Picker, and Frank Steinicke. 2020. Using Exergames to Train Patients with Dementia to Accomplish Daily Routines. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Canada) *(CHI PLAY '20)*. Association for Computing Machinery, New York, NY, USA, 345–349. https://doi.org/10.1145/3383668.3419883 pages #10, #28

[209] Sebastian Rings, Frank Steinicke, Tobias Picker, and Caspar Prasuhn. 2020. Enabling Patients with Neurological Diseases to perform Motor-Cognitive Exergames under Clinical Supervision for Everyday Usage. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 360–364. https://doi.org/10.1109/VRW50115.2020.00078 pages #28

[210] Sebastian Rings, Frank Steinicke, Tobias Picker, and Caspar Prasuhn. 2020. Memory Journalist: Creating Virtual Reality Exergames for the Treatment of Older Adults with Dementia. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 686–687. https://doi.org/10.1109/VRW50115.2020.00194 pages #28

[211] Sebastian Rings, Frank Steinicke, Tobias Picker, and Caspar Prasuhn. 2020. Seated Immersive Exergaming for Fall Prevention of Older Adults. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 289–290. https://doi.org/10.1109/VRW50115.2020.00063 pages #28

[212] Katja Rogers, Sukran Karaosmanoglu, Dennis Wolf, Frank Steinicke, and Lennart E. Nacke. 2021. A Best-Fit Framework and Systematic Review of Asymmetric Gameplay in Multiplayer Virtual Reality Games. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.694660 pages #5, #14

[213] Katja Rogers, Giovanni Ribeiro, Rina R. Wehbe, Michael Weber, and Lennart E. Nacke. 2018. Vanishing Importance: Studying Immersive Effects of Game Audio Perception on Player Experiences in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173902 pages #14

[214] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion* 30, 4 (2006), 344–360. https://doi.org/10.1007/s11031-006-9051-8 The IMI items are available online under this link: http://selfdeterminationtheory.org/intrinsic-motivation-inventory/. pages #8

[215] Nouran Sadek, Passant Elagroudy, Ali Khalil, and Slim Abdennadher. 2022. The Superhero Pose: Enhancing Physical Performance in Exergames by Embodying Celebrity Avatars in Virtual Reality. In *Nordic Human-Computer Interaction Conference* (Aarhus, Denmark) *(NordiCHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 1, 11 pages. https://doi.org/10.1145/3546155.3546707 pages #27

[216] Inc Sanzaru Games. 2016. *VR Sports Challenge.* Game [VR]. Sanzaru Games, Inc. Foster City, CA, USA. https://www.sanzaru.com/. Meta Reviews: https://www.meta.com/experiences/pcvr/1132870140059346/. pages #26

[217] Sarah Sauchelli and Jeffrey M. Brunstrom. 2022. Virtual reality exergaming improves affect during physical activity and reduces subsequent food consumption in inactive adults. *Appetite* 175 (2022). https://doi.org/10.1016/j.appet.2022.106058 pages #26, #28

[218] Steven Schmidt, Patrick Ehrenbrink, Benjamin Weiss, Jan-Niklas Voigt-Antons, Tanja Kojic, Andrew Johnston, and Sebastian Möller. 2018. Impact of Virtual Environments on Motivation and Engagement During Exergames. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 1–6. https://doi.org/10.1109/QoMEX.2018.8463389 pages #3, #12, #13, #27

[219] Matthias C. Schrempf, Johannes Zanker, Tim Tobias Arndt, Dmytro Vlasenko, Matthias Anthuber, Gernot Müller, Florian Sommer, and Sebastian Wolf. 2023. Immersive Virtual Reality Fitness Games to Improve Recovery After Colorectal Surgery: A Randomized Single Blind Controlled Pilot Trial. *Games for Health Journal* (2023). https://doi.org/10.1089/g4h.2023.0004 pages #26, #28

[220] Syed Hammad Hussain Shah, Ibrahim A. Hameed, Anniken Susanne T. Karlsen, and Mads Solberg. 2022. Towards a Social VR-based Exergame for Elderly Users: An Exploratory Study of Acceptance, Experiences and Design Principles. In *Virtual, Augmented and Mixed Reality: Design and Development*, Jessie Y. C. Chen and Gino Fragomeni (Eds.). Springer International Publishing, Cham, 495–504. pages #28

[221] Syed Hammad Hussain Shah, Anniken Susanne T. Karlsen, Mads Solberg, and Ibrahim A. Hameed. 2022. A social VR-based collaborative exergame for rehabilitation: codesign, development and user study. *Virtual Reality* (2022). https://doi.org/10.1007/s10055-022-00721-8 pages #6, #13, #28

[222] Alex Shaw, Jude Buckley, Paul Corballis, Christof Lutteroth, and Burkhard Wuensche. 2020. Exploring the Possibility of Virtual Reality Exergaming as a Cognitive Screening System. (2020). pages #9, #28

[223] Alexa J. Shaw and Anat V. Lubetzky. 2021. A Short Bout of Exercise With and Without an Immersive Virtual Reality Game Can Reduce Stress and Anxiety in Adolescents: A Pilot Randomized Controlled Trial. *Frontiers in Virtual Reality* 1 (2021). https://doi.org/10.3389/frvir.2020.598506 pages #28

[224] Lindsay A. Shaw, Jude Buckley, Paul M. Corballis, Christof Lutteroth, and Burkhard C. Wuensche. 2016. Competition and cooperation with virtual players in an exergame. *PeerJ Computer Science* 2016, 10 (2016). https://doi.org/10.7717/peerj-cs.92 pages #13, #27

[225] Lindsay Alexander Shaw, Burkhard Claus Wuensche, Christof Lutteroth, Jude Buckley, and Paul Corballis. 2017. Evaluating Sensory Feedback for Immersion in Exergames. In *Proceedings of the Australasian Computer Science Week Multiconference* (Geelong, Australia) *(ACSW '17)*. Association for Computing Machinery, New York, NY, USA, Article 11, 6 pages. https://doi.org/10.1145/3014812.3014823 pages #28

[226] Lindsay Alexander Shaw, Burkhard Claus Wünsche, Christof Lutteroth, Stefan Marks, and Rodolphe Callies. 2015. Challenges in virtual reality exergame design. (2015). pages #27

[227] Harrison She, Jared Young, Ofek Wittenberg, Patrick Poole, Alexander Mercer, Alex Shaw, and Burkhard C. Wünsche. 2021. HIITCopter: Analysis of the Suitability of VR Exergaming for High-Intensity Interval Training. In *Proceedings of the 32nd Australian Conference on Human-Computer Interaction* (Sydney, NSW, Australia) *(OzCHI '20)*. Association for Computing Machinery, New York, NY, USA, 293–302. https://doi.org/10.1145/3441000.3441009 pages #9, #12, #28

[228] Nur Aqilah Shoib, Mohd Shahrizal Sunar, Nurshamine Nazira Mohd Nor, Azizul Azman, Mohd Najeb Jamaludin, and Hadafi Fitri Mohd Latip. 2020. Rowing Simulation using Rower Machine in Virtual Reality. In *2020 6th International Conference on Interactive Digital Media (ICIDM)*. 1–6. https://doi.org/10.1109/ICIDM51048.2020.9339603 pages #27

[229] Mel Slater and Maria V. Sanchez-Vives. 2016. Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI* 3 (2016). https://doi.org/10.3389/frobt.2016.00074 pages #3

[230] Peng Song, Shuhong Xu, Wee Teck Fong, Ching Ling Chin, Gim Guan Chua, and Zhiyong Huang. 2010. An immersive VR system for sports education. *IEICE TRANSACTIONS on Information and Systems* 95, 5 (2010), 1324–1331. pages #28

[231] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. 2019. What is Mixed Reality?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300767 pages #2

[232] Oskar Stamm and Rebecca Dahms. 2023. Habituation to Simulator Sickness in Older Adults During a Chronic Back Pain Therapy Program with Repeated Virtual Reality Exposure. In *Human Aspects of IT for the Aged Population*, Qin Gao and Jia Zhou (Eds.). Springer Nature Switzerland, Cham, 627–641. pages #27

[233] Oskar Stamm, Rebecca Dahms, Norbert Reithinger, Aaron Ruß, and Ursula Müller-Werdan. 2022. Virtual reality exergame for supplementing multimodal pain therapy in older adults with chronic back pain: a randomized controlled pilot study. *Virtual Reality* 26, 4 (2022), 1291 – 1305. https://doi.org/10.1007/s10055-022-00629-3 pages #10, #27

[234] Oskar Stamm and Susan Vorwerg. 2021. Virtual Reality Sickness Evaluation in Exergames for Older Hypertensive Patients: A Comparative Study of Training Methods in a Virtual Environment. In *Virtual, Augmented and Mixed Reality*, Jessie Y. C. Chen and Gino Fragomeni (Eds.). Springer International Publishing, Cham, 134–146. pages #28

[235] Oskar Stamm, Susan Vorwerg, Michele Haink, Kristian Hildebrand, and Ilona Buchem. 2022. Usability and Acceptance of Exergames Using Different Types of Training among Older Hypertensive Patients in a Simulated Mixed Reality. *Applied Sciences* 12, 22 (2022). https://doi.org/10.3390/app122211424 pages #9, #10, #27

[236] Trenton H. Stewart, Kirsten Villaneuva, Amanda Hahn, Julissa Ortiz-Delatorre, Chandler Wolf, Randy Nguyen, Nicole D. Bolter, Marialice Kern, and James R. Bagley. 2022. Actual vs. perceived exertion during active virtual reality game exercise. *Frontiers in Rehabilitation Sciences* 3 (2022). https://doi.org/10.3389/fresc.2022.887740 pages #26, #28

[237] Thomas Stranick and Christian Lopez. 2021. Adaptive Virtual Reality Exergame: Promoting Physical Activity Among Workers. *Journal of Computing and Information Science in Engineering* 22, 3 (12 2021), 031002. https://doi.org/10.1115/1.4053002 pages #27

[238] Thomas Stranick and Christian Lopez. 2021. Virtual Reality Exergames: Promoting Physical Health Among Industry Workers *(International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Vol. Volume 2: 41st Computers and Information in Engineering Conference (CIE))*. V002T02A076. https://doi.org/10.1115/DETC2021-67608 pages #9, #11, #27

[239] Thomas Stranick and Christian E. Lopez. 2021. Leveraging Virtual Reality and Exergames to Promote Physical Activity. In *HCI International 2021 - Posters*, Constantine Stephanidis, Margherita Antona, and Stavroula Ntoa (Eds.). Springer International Publishing, Cham, 392–399. pages #13, #28

[240] Armature Studio. 2019. *Sports Scramble*. Game [VR]. Armature Studio. Austin, Texas, USA. https://www.armature.com/games/sports-scramble/. Meta Reviews: https://www.meta.com/experiences/705576999566582/. pages #12, #26

[241] Kinemotik Studios. 2022. *Audio Trip*. Game [VR]. Kinemotik Studios. San Francisco, CA, USA. http://www.kinemotik.com/audiotrip/. Steam Reviews: https://store.steampowered.com/app/826540/Audio_Trip/. pages #11, #26

[242] Lightspeed Studios™. 2016. *NVIDIA® VR Funhouse*. Game [VR]. Lightspeed Studios™. https://www.lightspeed-studios.com/. Steam Reviews: https://store.steampowered.com/app/468700/NVIDIA_VR_Funhouse/. pages #9, #13, #26

[243] Anthea Sutton, Mark Clowes, Louise Preston, and Andrew Booth. 2019. Meeting the review family: exploring review types and associated information retrieval requirements. *Health Information & Libraries Journal* 36, 3 (Sept. 2019), 202–222. https://doi.org/10.1111/hir.12276 pages #4

[244] Philipp Sykownik, Sukran Karaosmanoglu, Katharina Emmerich, Frank Steinicke, and Maic Masuch. 2023. VR Almost There: Simulating Co-Located Multiplayer Experiences in Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 789, 19 pages. https://doi.org/10.1145/3544548.3581230 pages #14

[245] Syncsense. 2019. *Syncsense*. Game [VR]. Syncsense. Copenhagen, Denmark. https://www.syncsense.io/. pages #9, #13, #26

[246] Ancret Szpak, Stefan Carlo Michalski, and Tobias Loetscher. 2020. Exergaming with Beat Saber: An Investigation of Virtual Reality Aftereffects. *Journal of Medical Internet Research* 22, 10 (2020). https://doi.org/10.2196/19840 pages #26, #28

[247] Gordon Tao, Bernie Garrett, Tarnia Taverner, Elliott Cordingley, and Crystal Sun. 2021. Immersive virtual reality health games: a narrative review of game design. *Journal of NeuroEngineering and Rehabilitation* 18 (2021), 1–21. https://doi.org/10.1186/s12984-020-00801-3 pages #3

[248] Marios Thomos, Vasileios Charisis, Stelios Hadjidimitriou, Dimitrios Iakovakis, Sevasti Bostantzopoulou, and Leontios Hadjileontiadis. 2020. HoloPrognosis - An AR-Based Serious Exercise Game for Early Stage Parkinson's Disease Patients. In *Universal Access in Human-Computer Interaction. Design Approaches and Supporting Technologies*, Margherita Antona and Constantine Stephanidis (Eds.). Springer International Publishing, Cham, 417–430. https://doi.org/10.1007/978-3-030-49282-3_30 pages #3, #10, #11, #27

[249] Andrea C Tricco, Erin Lillie, Wasifa Zarin, Kelly K O'Brien, Heather Colquhoun, Danielle Levac, David Moher, Micah DJ Peters, Tanya Horsley, Laura Weeks, et al. 2018. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine* 169, 7 (2018), 467–473. https://doi.org/10.7326/M18-0850 pages #4

[250] Mateus Trombetta, Patrícia Paula Bazzanello Henrique, Manoela Rogofski Brum, Eliane Lucia Colussi, Ana Carolina Bertoletti De Marchi, and Rafael Rieder. 2017. Motion Rehab AVE 3D: A VR-based exergame for post-stroke rehabilitation. *Computer Methods and Programs in Biomedicine* 151 (2017), 15–20. https://doi.org/10.1016/j.cmpb.2017.08.008 pages #27

[251] José Varela-Aldás, Esteban M. Fuentes, Guillermo Palacios-Navarro, and Iván García-Magariño. 2020. A Comparison of Heart Rate in Normal Physical Activity vs. Immersive Virtual Reality Exergames. In *Human Systems Engineering and Design II*, Tareq Ahram, Waldemar Karwowski, Stefan Pickl, and Redha Taiar (Eds.). Springer International Publishing, Cham, 684–689. pages #28

[252] VirZOOM. 2016. *VirZOOM Arcade*. Game [VR]. VirZOOM. Cambridge, Massachusetts. https://virzoom.com/. Steam Reviews: https://store.steampowered.com/app/448710/VirZOOM_Arcade/. pages #26

[253] Inc VirZOOM. 2021. *VZfit*. Game [VR]. VirZOOM, Inc. Cambridge, Massachusetts. https://vzfit.com/. Meta Reviews: https://www.meta.com/experiences/2088366894520136/. pages #26

[254] Daniel Vogel, Susanne Schmidt, and Frank Steinicke. 2022. Interactive Chinese Character Learning in Virtual Reality. In *SIGGRAPH Asia 2022 XR* (Daegu, Republic of Korea) *(SA '22)*. Association for Computing Machinery, New York, NY, USA, Article 9, 2 pages. https://doi.org/10.1145/3550472.3558415 pages #12, #27

[255] Black Box VR. 2022. *Black Box VR*. Game [VR]. Black Box VR. Boise, USA. https://www.blackbox-vr.com/. pages #12, #13, #26

[256] Shun-Yu Wang, Chia-Yu Cheng, Shu-Meng Huang, Weng Io Chan, Yu-Hsuan Huang, and Jhih-Wei Lin. 2023. Breathero: Not Another Slow Breathing Game — Exploring Faster-Paced VR Breathing Exercise Games. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 592, 7 pages. https://doi.org/10.1145/3544549.3583829 pages #27

[257] Yifan Wang, Kiran Ijaz, and Rafael A. Calvo. 2017. A software application framework for developing immersive virtual reality experiences in health domain. In *2017 IEEE Life Sciences Conference (LSC)*. 37–30. https://doi.org/10.1109/LSC.2017.8268135 pages #27

[258] Yifan Wang, Kiran Ijaz, Dong Yuan, and Rafael A. Calvo. 2020. VR-Rides: An object-oriented application framework for immersive virtual reality exergames. *Software: Practice and Experience* 50, 7 (2020), 1305–1324. https://doi.org/10.1002/spe.2814 pages #28

[259] Krergkiat Wongutai, Patison Palee, and Noppon Choosri. 2021. The Effect of Sound in VR Exergame to Adult Player: A Primary Investigation. *2021 Joint 6th International Conference on Digital Arts, Media and Technology with 4th ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering, ECTI DAMT and NCON 2021* (2021), 1 – 4. https://doi.org/10.1109/ECTIDAMTNCON51128.2021.9425698 Cited by: 1. pages #6

[260] World Health Organization. 2018. The global action plan on physical activity 2018 - 2030. Website. Retrieved September 4, 2023 from https://www.who.int/news-room/initiatives/gappa/action-plan. pages #1

[261] World Health Organization. 2022. *Global status report on physical activity 2022*. World Health Organization. https://www.who.int/publications/i/item/9789240059153 (visited September 4, 2023). pages #1

[262] Matthijs Ferdinand Wouda, Jon-Arve Gaupseth, Espen Ingvald. Bengtson, Truls Johansen, Espen Andreas Brembo, and Eivind Lundgaard. 2023. Exercise intensity during exergaming in wheelchair-dependent persons with SCI. *Spinal Cord* 61, 6 (2023), 338 – 344. https://doi.org/10.1038/s41393-023-00893-3 pages #26, #28

[263] Burkhard C. Wünsche, Michael Abernethy, Sylwia Hyniewska, Zhang Luo, Roshan Prakash Pawar, Sheng Cao, Christopher Brock, Alex Shaw, and Christof Lutteroth. 2021. Rift Racers - Effect of Balancing and Competition on Exertion, Enjoyment, and Motivation in an Immersive Exergame. In *2021 36th International Conference on Image and Vision Computing New Zealand (IVCNZ)*. 1–6. https://doi.org/10.1109/IVCNZ54163.2021.9653159 pages #28

[264] Biao Xie, Yongqi Zhang, Haikun Huang, Elisa Ogawa, Tongjian You, and Lap-Fai Yu. 2018. Exercise Intensity-Driven Level Design. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (2018), 1661–1670. https://doi.org/10.1109/TVCG.2018.2793618 pages #27

[265] Wenge Xu, Hai-Ning Liang, Nilufar Baghaei, Xiaoyue Ma, Kangyou Yu, Xuanru Meng, and Shaoyue Wen. 2021. Effects of an Immersive Virtual Reality Exergame on University Students' Anxiety, Depression, and Perceived Stress: Pilot Feasibility and Usability Study. *JMIR Serious Games* 9, 4 (22 Nov 2021), e29330. https://doi.org/10.2196/29330 pages #26, #28

[266] Wenge Xu, Hai-Ning Liang, Qiuyu He, Xiang Li, Kangyou Yu, and Yuzheng Chen. 2020. Results and Guidelines From a Repeated-Measures Design Experiment Comparing Standing and Seated Full-Body Gesture-Based Immersive Virtual Reality Exergames: Within-Subjects Evaluation. *JMIR Serious Games* 8, 3 (27 Jul 2020), e17972. https://doi.org/10.2196/17972 pages #10, #11, #14, #27

[267] Wenge Xu, Hai-Ning Liang, Xiaoyue Ma, and Xiang Li. 2020. VirusBoxing: A HIIT-Based VR Boxing Game. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Canada) *(CHI PLAY '20)*. Association for Computing Machinery, New York, NY, USA, 98–102. https://doi.org/10.1145/3383668.3419958 pages #11, #27

[268] Wenge Xu, Hai-Ning Liang, Kangyou Yu, and Nilufar Baghaei. 2021. Effect of Gameplay Uncertainty, Display Type, and Age on Virtual Reality Exergames. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 439, 14 pages. https://doi.org/10.1145/3411764.3445801 pages #3, #9, #11, #27

[269] Wenge Xu, Hai-Ning Liang, Kangyou Yu, Shaoyue Wen, Nilufar Baghaei, and Huawei Tu. 2023. Acceptance of Virtual Reality Exergames Among Chinese Older Adults. *International Journal of Human–Computer Interaction* 39, 5 (2023), 1134–1148. https://doi.org/10.1080/10447318.2022.2098559 pages #14, #26, #28

[270] Wenge Xu, Hai-Ning Liang, Yifan Yu, Diego Monteiro, Khalad Hasan, and Charles Fleming. 2019. Assessing the Effects of a Full-Body Motion-Based Exergame in Virtual Reality. In *Proceedings of the Seventh International Symposium of Chinese CHI* (Xiamen, China) *(Chinese CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3332169.3333574 pages #28

[271] Wenge Xu, Hai-Ning Liang, Zeying Zhang, and Nilufar Baghaei. 2020. Studying the Effect of Display Type and Viewing Perspective on User Experience in Virtual Reality Exergames. *Games for Health Journal* 9, 6 (2020), 405 – 414. https://doi.org/10.1089/g4h.2019.0102 pages #11, #28

[272] Wenge Xu, Kangyou Yu, Xuanru Meng, Diego Monteiro, Dominic Kao, and Hai-Ning Liang. 2023. Exploring the effect of the Group Size and Feedback of non-player character spectators in virtual reality exergames. *Frontiers in Psychology* 14 (2023). https://doi.org/10.3389/fpsyg.2023.1079132 pages #11, #27

[273] Hui Yang, Junyan Li, Jingnan Liu, Yulong Bian, and Juan Liu. 2022. Development of an Indoor Exergame Based on Moving-Target Hitting Task for COVID-19

Epidemic: A Comparison Between AR and VR Mode. *IEEE Transactions on Games* 14, 3 (2022), 511 – 521. https://doi.org/10.1109/TG.2021.3118035 pages #3, #12, #27

[274] Betty Yin, Samuel Bailey, Emma Hu, Milinda Jayarekera, Alex Shaw, and Burkhard C. Wünsche. 2021. Tour de Tune 2 - Auditory-Game-Motor Synchronisation with Music Tempo in an Immersive Virtual Reality Exergame. In *Proceedings of the 2021 Australasian Computer Science Week Multiconference* (Dunedin, New Zealand) *(ACSW '21)*. Association for Computing Machinery, New York, NY, USA, Article 11, 10 pages. https://doi.org/10.1145/3437378.3437379 pages #11, #27

[275] Jinhee Yoo, Seungbum Lee, and Eugene Agboifo Ohu. 2022. A cross-cultural analysis of VR gaming psychological needs and motivations: a self-determination theory approach. *Managing Sport and Leisure* (2022). https://doi.org/10.1080/23750472.2022.2046490 pages #7, #26, #28

[276] Soojeong Yoo, Marcus Carter, and Judy Kay. 2018. VRmove: Design Framework for Balancing Enjoyment, Movement and Exertion in VR Games. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Melbourne, VIC, Australia) *(CHI PLAY '18 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 295–307. https://doi.org/10.1145/3270316.3272054 pages #5, #13, #28

[277] Soojeong Yoo and Judy Kay. 2016. VRun: Running-in-Place Virtual Reality Exergame. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (Launceston, Tasmania, Australia) *(OzCHI '16)*. Association for Computing Machinery, New York, NY, USA, 562–566. https://doi.org/10.1145/3010915.3010987 pages #27

[278] Kangyou Yu, Shaoyue Wen, Wenge Xu, Maurizio Caon, Nilufar Baghaei, and Hai-Ning Liang. 2023. Cheer for me: effect of non-player character audience feedback on older adult users of virtual reality exergames. *Virtual Reality* (2023), 1–17. pages #11, #14, #27

[279] Seo Jung Yun, Sung Eun Hyun, Byung-Mo Oh, and Han Gil Seo. 2023. Fully immersive virtual reality exergames with dual-task components for patients with Parkinson's disease: a feasibility study. *Journal of NeuroEngineering and Rehabilitation* 20, 1 (2023). https://doi.org/10.1186/s12984-023-01215-7 pages #10, #13, #27

[280] Naghmeh Zamani, Adim Abass, Manjunath Shetkar, Saumya Dureja, Menghan Li, Heather Culbertson, and James Finley. 2021. Integrating Haptic Feedback into a Virtual Reality Mobility Training Game for People with Parkinson's Disease. *2021 IEEE World Haptics Conference, WHC 2021* (2021), 875. https://doi.org/10.1109/WHC49131.2021.9517136 pages #10, #27

[281] Xu-Hui Zhang, Xin-Yuan Lv, Miao-Yun Wang, Qian-Nan Wei, Heng Wang, Yu-Meng Zhang, and Yu-Yang Du. 2021. Upper limb rehabilitation training system based on virtual reality. *Proceedings of SPIE - The International Society for Optical Engineering* 11720 (2021). https://doi.org/10.1117/12.2589431 pages #10, #27

[282] Xiu-Wei Zhang, Fong-Ming Shyu, and Geeng-Neng You. 2018. The exergame for Tae-Bo learning with virtual reality. In *2018 IEEE International Conference on Applied System Invention (ICASI)*. IEEE, 1018–1021. https://doi.org/10.1109/ICASI.2018.8394448 pages #9, #28

[283] Jochen P. Ziegelmann, Sonia Lippke, and Ralf Schwarzer. 2006. Adoption and maintenance of physical activity: Planning interventions in young, middle-aged, and older adults. *Psychology & Health* 21, 2 (2006), 145–163. https://doi.org/10.1080/1476832050018891 pages #13

A SUPPLEMENTARY TABLES

Table 3: The final search queries in the syntax of each database. The final search was conducted on 11 August 2023.

Database	Query
The ACM Guide to Computing Literature	`Title:(("immersive" OR "VR" OR "AR" OR "AV" OR "MR" OR "XR" OR "virtual realit*" OR "augmented realit*" OR "extended realit*" OR "mixed realit*" OR "augmented virtualit*" OR "virtual environment*") AND ("exergame*" OR "exercise game*" OR "physical game*" OR "movement game*" OR "motion game*" OR "motion-based game*" OR "movement-based game*" OR "training game*" OR "exertion game*" OR "sport game*" OR "sports game*")) OR Abstract:(("immersive" OR "VR" OR "AR" OR "AV" OR "MR" OR "XR" OR "virtual realit*" OR "augmented realit*" OR "extended realit*" OR "mixed realit*" OR "augmented virtualit*" OR "virtual environment*") AND ("exergame*" OR "exercise game*" OR "physical game*" OR "movement game*" OR "motion game*" OR "motion-based game*" OR "movement-based game*" OR "training game*" OR "exertion game*" OR "sport game*" OR "sports game*")) OR Keyword:(("immersive" OR "VR" OR "AR" OR "AV" OR "MR" OR "XR" OR "virtual realit*" OR "augmented realit*" OR "extended realit*" OR "mixed realit*" OR "augmented virtualit*" OR "virtual environment*") AND ("exergame*" OR "exercise game*" OR "physical game*" OR "movement game*" OR "motion game*" OR "motion-based game*" OR "movement-based game*" OR "training game*" OR "exertion game*" OR "sport game*" OR "sports game*"))`
Scopus	`TITLE-ABS-KEY (("immersive" OR "VR" OR "AR" OR "AV" OR "MR" OR "XR" OR "virtual realit*" OR "augmented realit*" OR "extended realit*" OR "mixed realit*" OR "augmented virtualit*" OR "virtual environment*") AND ("exergame*" OR "exercise game*" OR "physical game*" OR "movement game*" OR "motion game*" OR "motion-based game*" OR "movement-based game*" OR "training game*" OR "exertion game*" OR "sport game*" OR "sports game*"))`
PubMed	`(("immersive"[Title/Abstract] OR "VR"[Title/Abstract] OR "AR"[Title/Abstract] OR "AV"[Title/Abstract] OR "MR"[Title/Abstract] OR "XR"[Title/Abstract] OR "virtual realit*"[Title/Abstract] OR "augmented realit*"[Title/Abstract] OR "extended realit*"[Title/Abstract] OR "mixed realit*"[Title/Abstract] OR "augmented virtualit*"[Title/Abstract] OR "virtual environment*"[Title/Abstract]) AND ("exergame*"[Title/Abstract] OR "exercise game*"[Title/Abstract] OR "physical game*"[Title/Abstract] OR "movement game*"[Title/Abstract] OR "motion game*"[Title/Abstract] OR "motion-based game*"[Title/Abstract] OR "movement-based game*"[Title/Abstract] OR "training game*"[Title/Abstract] OR "sport game*"[Title/Abstract] OR "sports game*"[Title/Abstract] OR "exertion game*"[Title/Abstract]))`

Table 4: The list of used commercial games in XR exergames research.

No	Name	Papers	No	Name	Papers
#1	Audio Trip [241]	[42, 43]	#11	HoloPoint [5]	[236]
#2	Beat Saber [74]	[4, 57, 76, 84, 124, 147, 236, 246, 269]	#12	NVIDIA VR Fun House [242]	[30]
#3	Black Box VR [255]	[167]	#13	QuiVr [14]	[124]
#4	BoxVR [66]	[27, 29, 30]	#14	Snow Games VR [146]	[169]
#5	Carnival Games VR [75]	[6]	#15	Sports Scramble [240]	[76, 86]
#6	Dance Central [87]	[269]	#16	SyncSense [245]	[97]
#7	First Steps [183]	[76]	#17	Thrill of the Fight [144]	[262]
#8	FitXR [67]	[28, 182, 265, 269]	#18	VirZoom [252]	[155, 217]
#9	Fruit Ninja VR [145]	[95, 236]	#19	VRSports Challenge [216]	[275]
#10	HoloFit [93]	[219]	#20	VZFit [253]	[134]

Table 5: The list of papers included in the taxonomy development (N=186).

No	Authors & Papers	No	Authors & Papers
#1	Hong et al. [94]	#2	Moan et al. [166]
#3	Li [135]	#4	Lund et al. [148]
#5	Stranick and Lopez [238]	#6	Bovim et al. [21]
#7	Cmentowski et al. [44]	#8	Zamani et al. [280]
#9	Katsigiannis et al. [115]	#10	Farič et al. [59]
#11	Xie et al. [264]	#12	Xu et al. [268]
#13	Shaw et al. [226]	#14	Huber et al. [96]
#15	Eisapour et al. [52]	#16	Ishii et al. [105]
#17	Thomos et al. [248]	#18	Xu et al. [266]
#19	Elor et al. [55]	#20	Born et al. [19]
#21	Benim et al. [11]	#22	Born et al. [20]
#23	Chen et al. [34]	#24	Shoib et al. [228]
#25	Franzo et al. [70]	#26	Ou et al. [187]
#27	Michael and Lutteroth [158]	#28	Heng and Albert [88]
#29	Yun et al. [279]	#30	Zhang et al. [281]
#31	Karaosmanoglu et al. [112]	#32	Dulau et al. [50]
#33	Al-Mfarej et al. [3]	#34	Kojic et al. [126]
#35	Stamm et al. [233]	#36	Born et al. [17]
#37	Abril et al. [2]	#38	Chen et al. [37]
#39	Sadek et al. [215]	#40	Yoo and Kay [277]
#41	Mishra and Folmer [164]	#42	Palaniappan and Duerstock [190]
#43	Xu et al. [272]	#44	Barathi et al. [8]
#45	Li [137]	#46	Kojic et al. [125]
#47	Yu et al. [278]	#48	Kojić et al. [127]
#49	Koulouris et al. [130]	#50	Mondellini et al. [168]
#51	Goumopoulos et al. [81]	#52	Keesing et al. [116]
#53	Høeg et al. [98]	#54	Kojić et al. [128]
#55	Stamm et al. [235]	#56	Trombetta et al. [250]
#57	Ring and Masuch [207]	#58	Wang et al. [257]
#59	Ashok et al. [7]	#60	Jacob et al. [106]
#61	Kalaitzidou et al. [109]	#62	Barathi et al. [9]
#63	Chung et al. [40]	#64	Eisapour et al. [54]
#65	Wang et al. [256]	#66	Schmidt et al. [218]
#67	Fan et al. [58]	#68	Stamm and Dahms [232]
#69	Kruse et al. [132]	#70	Neira-Tovar and Elizondo Elizondo [177]
#71	Stranick and Lopez [237]	#72	Eisapour et al. [53]
#73	Mendez et al. [156]	#74	Finlayson et al. [64]
#75	Kirsch et al. [123]	#76	Xu et al. [267]
#77	Elor et al. [56]	#78	Shaw et al. [224]
#79	Yang et al. [273]	#80	Born et al. [18]
#81	Patibanda et al. [195]	#82	Eckert et al. [51]
#83	Liepa et al. [140]	#84	Chessa et al. [38]
#85	Caserman et al. [32]	#86	Ijaz et al. [101]
#87	Ijaz et al. [100]	#88	Vogel et al. [254]
#89	Goumopoulos et al. [80]	#90	Ioannou et al. [104]
#91	Mugisha et al. [172]	#92	Mihajlovic et al. [161]
#93	Yin et al. [274]	#94	Finkelstein et al. [60]

Table 6: Table 5 continued: The list of papers included in the taxonomy development (N=186).

No	Authors & Papers	No	Authors & Papers
#95	Mambu et al. [150]	#96	Shaw et al. [225]
#97	Shah et al. [221]	#98	Finkelstein et al. [61]
#99	Shah et al. [220]	#100	Finkelstein et al. [63]
#101	Chu et al. [39]	#102	Ohshima et al. [185]
#103	Pérez et al. [201]	#104	Bolton et al. [15]
#105	Franzo' et al. [69]	#106	Reilly et al. [205]
#107	Muñoz et al. [174]	#108	Yoo et al. [276]
#109	Stranick and Lopez [239]	#110	Ciążyńska and Maciaszek [43]
#111	Shaw and Lubetzky [223]	#112	Keller et al. [118]
#113	Liang et al. [139]	#114	de Souza et al. [48]
#115	Chung et al. [41]	#116	Park et al. [193]
#117	Finley et al. [65]	#118	Ramasamy et al. [202]
#119	Buchem et al. [24]	#120	Goutsu and Inamura [82]
#121	Kruse et al. [131]	#122	Aan et al. [1]
#123	Rings et al. [209]	#124	Buckers et al. [25]
#125	Wang et al. [258]	#126	Chen et al. [35]
#127	Rings et al. [210]	#128	Kegeleers et al. [117]
#129	Liu et al. [143]	#130	Ciążyńska et al. [42]
#131	Rings et al. [211]	#132	Lu et al. [147]
#133	Xu et al. [271]	#134	Albert et al. [4]
#135	Li and Chen [138]	#136	Grosprêtre et al. [84]
#137	Rings et al. [208]	#138	Xu et al. [269]
#139	She et al. [227]	#140	García-Muñoz et al. [76]
#141	Haller et al. [85]	#142	Szpak et al. [246]
#143	Varela-Aldás et al. [251]	#144	Eng et al. [57]
#145	Palaniappan et al. [191]	#146	Kivela et al. [124]
#147	Palaniappan et al. [192]	#148	Stewart et al. [236]
#149	Stamm and Vorwerg [234]	#150	Mologne et al. [167]
#151	Campbell and Fraser [26]	#152	Campo-Prieto et al. [29]
#153	Shaw et al. [222]	#154	Campo-Prieto et al. [27]
#155	Khundam and Nöel [122]	#156	Campo-Prieto et al. [30]
#157	Nehrujee et al. [176]	#158	Amprasi et al. [6]
#159	Xu et al. [270]	#160	Xu et al. [269]
#161	Ijaz et al. [99]	#162	Ochi et al. [182]
#163	Karaosmanoglu et al. [111]	#164	Campo-Prieto et al. [28]
#165	Zhang et al. [282]	#166	Xu et al. [265]
#167	Ijaz et al. [102]	#168	Huang [95]
#169	Palacios-Alonso et al. [189]	#170	Montoya et al. [169]
#171	Hernandez et al. [89]	#172	Hanifah et al. [86]
#173	Palacios-Alonso et al. [188]	#174	Høeg et al. [97]
#175	Song et al. [230]	#176	Wouda et al. [262]
#177	Karkar et al. [114]	#178	Sauchelli and Brunstrom [217]
#179	Finkelstein et al. [62]	#180	McMahon et al. [155]
#181	Wünsche et al. [263]	#182	Yoo et al. [275]
#183	Cao et al. [31]	#184	Lee and Jin [134]
#185	Cesco et al. [33]	#186	Schrempf et al. [219]

9.2 Never Skip Leg Day Again: Training the Lower Body with Vertical Jumps in a Virtual Reality Exergame

[CorePub2] ✱🎗 Sebastian Cmentowski, **Sukran Karaosmanoglu**, Lennart E. Nacke, Frank Steinicke, and Jens Harald Krüger. Never skip leg day again: Training the lower body with vertical jumps in a virtual reality exergame. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. DOI 10.1145/3544548.3580973. URL https://doi.org/10.1145/3544548.3580973.

Supplementary materials for this article are available here: https://doi.org/10.1145/3544548.3580973.

Never Skip Leg Day Again: Training the Lower Body with Vertical Jumps in a Virtual Reality Exergame

Sebastian Cmentowski*
High-Performance Computing
University of Duisburg-Essen
Duisburg, Germany
sebastian.cmentwoski@uni-due.de

Sukran Karaosmanoglu*
Human-Computer Interaction
Universität Hamburg
Hamburg, Germany
sukran.karaosmanoglu@uni-hamburg.de

Lennart E. Nacke
HCI Games Group & Stratford School of Interaction Design and Business
University of Waterloo
Waterloo, Canada
lennart.nacke@acm.org

Frank Steinicke
Human-Computer Interaction
Universität Hamburg
Hamburg, Germany
frank.steinicke@uni-hamburg.de

Jens Krüger
High-Performance Computing
University of Duisburg-Essen
Duisburg, Germany
jens.krueger@uni-due.de

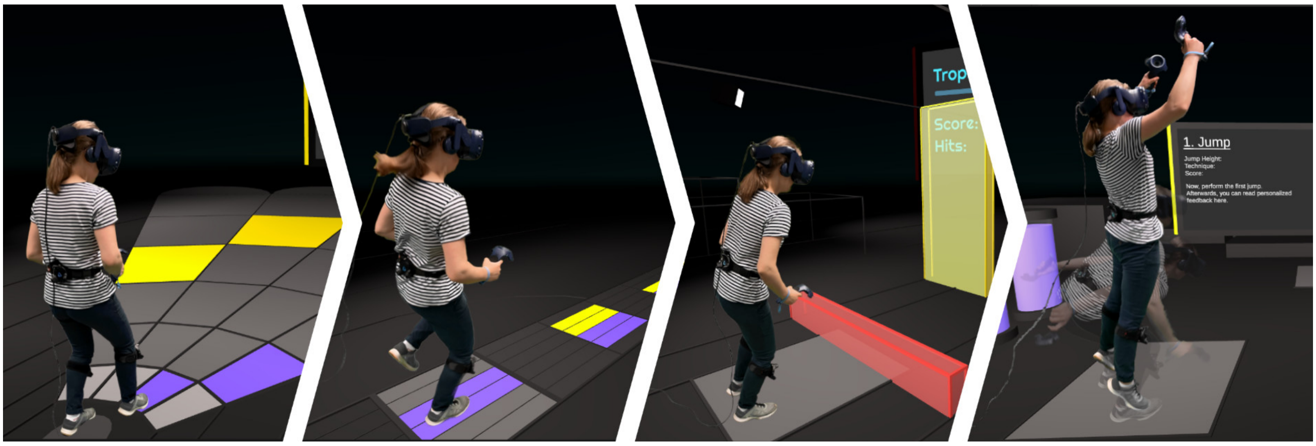

Figure 1: Demonstration of a VR-based full-body training with our exergame *JumpExTra VR*: In the first three levels (from left to right), players perform various movements, including taps, hops, and jumps, that train lower body coordination, stability, and endurance. Finally, the right figure shows a player who trains maximal vertical jumps.

ABSTRACT

Virtual Reality (VR) exergames can increase engagement in and motivation for physical activities. Most VR exergames focus on the upper body because many VR setups only track the users' heads and hands. To become a serious alternative to existing exercise programs, VR exergames must provide a balanced workout and train the lower limbs, too. To address this issue, we built a VR exergame focused on vertical jump training to explore full-body exercise applications. To create a safe and effective training, nine domain experts participated in our prototype design. Our mixed-methods study confirms that the jump-centered exercises provided a worthy challenge and positive player experience, indicating long-term retention. Based on our findings, we present five design implications to guide future work: avoid an unintended forward drift, consider technical constraints, address safety concerns in full-body VR exergames, incorporate rhythmic elements with fluent movement patterns, adapt difficulty to players' fitness progression status.

*Both authors contributed equally to this research.

CHI '23, April 23–28, 2023, Hamburg, Germany

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**; • **Software and its engineering** → **Interactive games**.

KEYWORDS

virtual reality, VR, exergame, vertical jump, training, sport, health, dynamic difficulty, serious games

ACM Reference Format:
Sebastian Cmentowski, Sukran Karaosmanoglu, Lennart E. Nacke, Frank Steinicke, and Jens Krüger. 2023. Never Skip Leg Day Again: Training

the Lower Body with Vertical Jumps in a Virtual Reality Exergame. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3544548.3580973

1 INTRODUCTION

Regular physical exercise is vital for our bodily and mental well-being. Athletic activities not only increase our overall fitness but can also delay the natural aging process [43] and even benefit the brain's cognitive functions [114]. However—because physical activity is strenuous on our bodies—many people hesitate to transition toward an active lifestyle and to create lasting exercise habits if they do not receive incentives [30, 49]. Apart from popular approaches, such as peer-support or fitness trackers, exergames—*"digital game[s] where the outcome [...] is predominantly determined by physical effort"* [108]—promise to motivate users by providing an engaging experience. In the virtual reality (VR) domain, fitness games are among the highest-grossing titles and a crucial reason for headset purchases [51]. The advantages are apparent: Affordable mobile VR headsets with handheld controllers allow players to combine enjoyable gaming activities with healthy physical exercise while staying in the comfort of their homes.

Exergames like Beat Saber[1] [52] or FitXR [47] mainly use head and hand movements for their gameplay while featuring some lower-body movements. Although there are also VR applications that target full-body exercises (e.g., OhShape [83] and VRWorkout [139]), the majority only use lower body movements indirectly, such as when ducking under obstacles. Such exercises are easily possible with the default VR setup, i.e., headset and controller tracking, without requiring additional hardware, such as Vive trackers. Consequently, the training effect of these exergames mainly targets cardiovascular improvements and upper body fitness. As a result, players do not get the full benefits of balanced full-body activities, such as improved coordination, stability, and balance [5, 115, 123]. Also, training only individual muscle groups can ultimately lead to muscular imbalances promoting bad posture [20, 75] and increasing the risk for injuries [36, 124]. Lastly, lower body exercises and physical activities like walking, running, or jumping are vitally important in our society because an average person spends most of their day sitting and moving insufficiently [138]. To make VR exergames more valuable as exercise environments, their traditional hand-focused gameplay must be adapted to incorporate activities for the lower body. While we currently do not know if VR exergames are as effective as gym classes or personal training, we know from short-time studies that exergames provide great motivation for exercising [15, 46, 93, 97]. However, we note that long-term exergame studies present mixed results on users' adherence [121, 130].

Unfortunately, using lower body or full-body movement in VR exergames introduces many challenges. These exercises have a higher risk of swift or unstable movements which might lead to dangerous collisions. Explosive movements, such as running or jumping, can suffer from poor tracking stability [67, 142], which negatively affects player experience. Full-body movements require expert knowledge for safe training because they are more complex than simple arm swings in Beat Saber [52] and bear a greater risk of injury or wrong execution. Limitingly, existing design guidelines and best practices primarily target physical exercises and movements for non-VR exergames [96, 106, 107].

Our research fills this knowledge gap by following the feedback from domain experts to create a full-body VR exergame. We focus on one particular use case: training people's vertical jump performance using a VR exergame. We chose the vertical jump exercise as our full-body movement because of the following reasons: Jumping is a fundamental human movement that is not only required for many sports, such as basketball [37] or volleyball [119], but is also used to assess general fitness [55], body composition [17], and functional performance [100]. Vertical jumps are a perfect, yet challenging, core movement for our research. They also work inside the tracking areas of current VR headsets. At the same time, vertical jumps are a highly explosive movement that challenges tracking stability. Finally, jumps can be improved through many training modalities and combined with other movements to achieve a diverse exercise experience.

We focus primarily on improving general fitness and motivation through training. For this, we conducted a semi-structured interview with experts from different domains (e.g., sports research or physical therapy). We discuss the potential benefits and challenges of VR-based jump training identified in our thematic analysis and provide guidelines for structuring gamified training routines. We developed a VR exergame to train the vertical jump based on these insights. In our design process, we closely follow the recommendations of experts. In particular, our exergame is composed of four levels of increasing difficulty to prevent injuries and foster the learning process (see Figure 1).

In the second phase of our work, we investigate how users perceive our exergame prototype and what implications can be drawn for future designs and research projects. As our first step, we conducted an exploratory study with 25 participants to evaluate how users perceive this new training experience. The results confirm that our jump-centered exercises provided a worthy challenge and led to a positive player experience. Our study also revealed the technical limitations of current VR systems, and the participants provided substantial suggestions for improving the training experience. Subsequently, we condensed these insights into design implications and lessons learned.

In the design implications, we first discuss potential safety issues. In our case, frequent jumps on one spot often led to an unnoticed forward movement—the unintended forward drift—which eventually leads to players leaving the intended play area. Hence, we recommend particular care to avoid dangerous collisions, especially when non-stationary movements, such as forward jumps, are used.

Next, we examine the design implications arising from technical limitations of current VR systems. Slipping hardware trackers and insufficient tracking accuracy challenged the precision of our individualized jump feedback. Therefore, we recommend empowering users towards self-correction (e.g., by visualizing a replay of their movements) and supporting this process with automated feedback.

Finally, we talk about the design implications based our efforts to provide a pleasant game experience and increase replayability. Above all, the participants praised using our exergame *JumpExTra VR* for jump training. We discuss lessons learned for enabling a

[1]We note that Beat Saber is a VR rhythm game. However, some users can use this game for exercise purposes, and this game has been the subject of research as an exergame [4, 125].

natural and fluent movement sequence, such as the suitability of different patterns (e.g., the "walking style") or incorporating frequent resting periods. We emphasize the importance of aligning the difficulty with the players' capabilities and improvements. In particular, beginners profit from adapting the difficulty automatically.

The main contributions of this research are:

(1) Identifying the benefits and requirements of VR jump training through semi-structured interviews with domain experts,
(2) Designing and developing a VR exergame prototype for training the vertical jump: *JumpExTra VR*,
(3) Conducting an exploratory study to evaluate the feasibility of our exergame as a training tool, and
(4) Deriving a set of design guidelines and lessons learned for developing full-body VR exergames.

Our explorative exergame study and the resulting five design guidelines (unintended forward drift fix, technical constraints consideration, safety concern mitigation, rhythmic elements using fluent movement patterns, difficulty adaptation to players' fitness level) constitute a first step to the creation of safer and more engaging exergame VR training environments. Our findings help players train effectively and without the risk of injury in small spaces using immersive technology. We believe this represents one possible future of technology-augmented sports exercises.

2 RELATED WORK

Our exergame design builds on two domain knowledge sources: prior research and expert interviews. Therefore, we introduce VR-based training, the relevant biomechanical foundations of the vertical jump and provide an overview of the related research on jump training, jumping in VR, and exergames in general.

2.1 VR Training

Using immersive experiences for training offers unique benefits and allows users to assess and improve their individual performance effectively [23, 131]. Especially when physical training at the target location is difficult or dangerous, like for pilots [57] or firefighters [122], VR applications can provide an accessible alternative [69]. In recent years, VR-based training has been successfully applied to many domains, including healthcare [95], medicine [3], and navigation [99]. In sports, VR has been used in research projects to analyze athletic performances [11, 34] and understand motor and perceptual skills [29, 35, 44, 143]. Furthermore, VR works well for training movement patterns, like golf swings [73] or dance moves [42], and improving hand-eye coordination (e.g., for table tennis [103], darts [127], or juggling [84]). Similarly, VR training can also support rehabilitation for stroke or cerebral palsy patients [31, 77, 105].

Besides boosting motivation, virtual experiences also have benefits for sports training. VR improves observational learning by displaying correct action patterns immersively compared to traditional computer applications [126]. Especially for sports, stereoscopic information is essential to trigger the correct motor responses [89, 141]. Additionally, mobile headsets also provide an easy way of monitoring the own performance [110]. For availability, VR applications can—to some degree—eliminate the need for specialized sporting equipment, dedicated training environments, or workout partners [103]. Adaptive training routines [41] and personalized feedback increase the individual gain and can complement professional trainers [76]. Despite these benefits, the effectiveness of VR applications for sports training remains an open research field. While studies have shown a positive impact on the execution of a particular exercise, the transferability to actual activities is not guaranteed [7, 85]. VR training can improve real-world skills for some use cases [13, 25, 102], but others cannot profit in the same way [79] and might even suffer from reduced performance [128].

2.2 Jumping

Jumping is a fundamental motor skill humans learn at an early age and improve upon throughout their lives. While most people do not jump regularly in their day-to-day lives, jumping is used in sports, fitness, and rehabilitation. In particular, jumps are a good indicator of a person's general fitness level [55], functional performance [100], and muscle composition [17]. Apart from being required for typical jumping-intensive sports, such as basketball [37],jumps are used in rehabilitation to measure changes in pathology [58, 59] and predict the individual risk for injuries [27]. Aside from the positive effects on athletic performance [22, 118], jump training can also benefit daily activities and occupational tasks [9, 80].

2.2.1 *Composition and Execution of the Vertical Jump.*

Jumping is a "complex polyarticular dynamic movement requiring intermuscular coordination" [116] with typical execution times of less than 4 seconds [132]. The literature differentiates between various types of jumps [14], including squat jumps, drop jumps [66], or *countermovement jumps*. In the scope of this paper, we focus on the last type. It is typically initiated from an upright standing position [21], followed by a brief downward phase before the upward impulse is generated by extending the body explosively and swinging the arms in a forward-upward arc [62].

Jumping can exert high loads on the lower body's joints and tissues, increasing the risk of injury (e.g., ruptures of the anterior cruciate ligaments (ACL) [117] caused by hard or incorrect landings that lead to high vertical ground reaction forces (vGRF) [60]). Proper instructions can help athletes reduce the vGRF immediately [101, 120]. Other risk factors, such as the athletes' joint stiffness [39] and maturity [86], cannot be eliminated as easily with proper form.

2.2.2 *Jump Tests.*

Measuring a person's jump height is not a trivial task. Professional athletes and researchers often rely on motion-capturing systems [8, 10, 87] or dedicated vertical jump tests [78]. One of the oldest techniques is to jump next to a wall and mark the highest reachable points while standing and jumping. Then, one can calculate the effective jump height from the difference between both points [68]. Whereas this jump and reach test follows a simple principle, it also splits the athlete's attention and limits arm movement, which easily reduces performance. Another approach is to jump on force platforms and calculate the jump height from the athlete's airtime [17]. However, athletes can easily distort the result by flexing their legs to delay ground contact [78]. Instead, the most precise results are achieved by measuring the vertical displacement of the athlete's center of mass [132]. Consequently, our VR application can easily and comfortably measure the precise jump height by using a hardware tracker attached to the user's hip.

2.2.3 Jump Training. Critical for improving jump height is increasing an athlete's take-off velocity. Training usually concentrates on improving the extensor muscles' forces and contraction velocity [82] because the maximal dynamic force of the lower extremities and the rate of force development (RFD) directly correspond to the final jump height [12, 135]. A well-timed arm swing [24], good neuromuscular control to initiate joint extension with minimal delay [65], and a countermovement with the optimal squat depth [53] have all been shown to improve jump height significantly.

Various training modalities have been found useful in achieving lasting training results. Firstly, plyometric exercises [2], such as drop jumps or alternate-leg bounding [48], improve the muscles' stretch-shortening cycle (SSC) and are primarily beneficial activities that involve rapid concentric contractions and high-intensity eccentric contractions [94]. The positive effect on jump performance has been extensively researched [14, 91] and varies between 5% and 35% depending on the athletes' proficiency [116]. Alternatively, athletes may conduct weight training [136] with heavy loads to increase the maximal dynamic strength [50] or light loads for explosive movements [92]. For jump training, the best results were reported using light loads and high speeds and range from 2% to 25% [116]. Some muscles, such as the hip joint extensors, are used only maximally in maximal jumps.Performance should be trained using maximal jumps to achieve the best training results [87].

2.3 Jumping in VR and Exergames

Jumping in virtual and mixed reality has also been the subject of other research projects. Prior literature has approached jumping in the context of locomotion [56, 90, 137], exergames [45, 67, 70, 88] or training [32]. Wolf et al. [137] presented an augmented locomotion technique where users performed physical vertical jumps, which translated into hyper-realistic forward jumps in VR. Their findings indicated that hyper-realistic jumps can enhance some factors of user experience (e.g., immersion). A recent work [32] presented a prototype of a VR jump training where players jumped and received feedback on their performance.

VR exergames are one of the most popular types of games in the VR gaming community (e.g., Beat Saber [52]) and have been shown to elicit a higher level of motivation compared to their non-VR headset exergame counterparts [15]. Despite the potential drawbacks associated with the use of VR exergames (e.g., cybersickness [125]), we see these games as an opportunity because they offer fun, physical activity, and accessibility to training regardless of location or health condition [72, 81]. Although prior research also covered exergames featuring full-body training [67, 70, 98] and investigated full-body movement recognition [26], many papers focus specifically on upper-body exercises [16, 71, 72]. Whereas this focus may be even preferable in some cases (e.g., due to safety [72]), full-body training could benefit more muscle structures, and has not been widely explored yet [98].

Many papers explored the design space for exergames and provided implications for designers, developers, and researchers [96, 106, 107]. However, to our knowledge, these design guidelines have not been specifically focused on VR exergames. Márquez Segura et al. [96] considered technological, social, and physical factors in designing *"body games"*. In their approach, they primarily focus on understanding the social and physical factors around a game and accordingly support the users with the technology. Furthermore, Mueller and Isbister [106] provided ten comprehensive guidelines for movement-based games, such as (i) using tracking inaccuracies, (ii) rhythmic elements, and (iii) utilizing risk reasonably as elements in movement-based gameplay. Similarly, Mueller et al. [107] presented five design recommendations for exertion games, or exergames: (i) providing an easy start, (ii) presenting short-term achievable goals for long-term motivation, (iii) considering individual skill levels of players, (iv) giving feedback on the performance, and (v) employing social play to promote motivation. Whereas these guidelines are helpful for movement-based games in general, full-body VR exergames may present additional advantages (e.g., higher motivation in VR exergames [15]) or challenges (e.g., wearing a headset) that require attention.

Several researchers have used jumping in their games. Finkelstein et al. [45] designed an CAVE-based exergame, *AstroJumper*, for children with autism. In the game, the players performed jumping movements to avoid objects, and the initial findings with neurotypical players indicated positive experiences. Kajastila et al. [70] examined the impact of three conditions on players' learning trampoline skills. The authors showed that the players were more engaged in the gaming conditions (a trampoline-based mixed reality game with and without exaggerated jumps) compared to the self-training condition, but their performance improved regardless of the conditions. Similarly, another study [88] designed and tested a multiplayer mixed reality trampoline game in a field study. The results indicated positive player experiences (e.g., autonomy and physical activity enjoyment). Many jump-based exergames have focused on player experience rather than providing a structured physical training for jumping using VR exergames. This paper extends the previous work [32] by designing and testing a detailed jump training VR exergame, *JumpExTra VR*, and involving the domain experts in the process.

3 THEMATIC ANALYSIS OF EXPERT INTERVIEWS

Unlike the prior literature, our research was not only on professional athletes and their sports performance training but we focus on the motivation of players to exercise. We gathered information from experts in the sports and medical field through semi-structured interviews to design and implement *JumpExTra VR* (see Figure 2). After internal discussion, the first authors created the interview questions and selected areas of expertise for interviewees, such as sports physicians and trainers. Based on these decisions, they selected several experts and invited them to this study via email. Nine experts from different domains (i.e., sports research, physiotherapy, and training) were recruited for the interviews. Given the variety of domains, the first authors roughly followed the initial interview guidelines (listed in the supplemental materials), but occasionally deviated from these questions to account for the experts' specialties. The interviews were conducted using a video conference tool in German and took 42.4 minutes on average.

To prepare the interview data for analysis, we used the Dovetail [40] software for transcription. One of the first authors checked these transcriptions to correct and cut unnecessary details. Then,

Figure 2: The design process followed in this research project: We reviewed the literature, conducted semi-structured interviews with multiple experts, analyzed the requirements of our VR jump training application, designed and implemented the VR exergame, and conducted a user study to test *JumpExTra VR*.

these texts were translated into English using DeepL Pro [38] and then checked again for any errors. To analyze the interview data, we used characteristics of both reflexive and codebook approaches of thematic analysis [18, 19]. Before the analysis, the first authors identified deductive categories related to the research questions: *execution*, *importance of jumping*, *safety in jumping*, and *jump training*. Then, these two authors independently and inductively coded the interview in groups, using descriptive codes (e.g., "correct execution of a counter-movement jump with arms", "landing is a major source for injuries", and "few short term improvements") under these categories. The groups consisted of two interviews in each and three interviews in the last one. After each group, the first authors met to discuss discrepancies between their coding and each author's understanding of the data. This process led to the creation, refinement, combination or exclusion of codes. Following the last group of interviews, the first two authors developed initial themes by creating a affinity map from those codes. Finally, they discussed and reshaped the affinity map and these themes in several meetings, which led to the formation of the following four themes.

3.1 Theme 1: Jumps are mainly used in sports and have numerous benefits for general health.

Both in everyday life and sports activities, we *jump*. However, the actual form depends on the particular context, as "*[...] we don't jump that often when we go shopping in the supermarket, we probably walk more. But of course you also have situations in which you might jump down in everyday life*"*(**E3**)*. Instead, jumps are mostly **integrated into compound movements**. For example, they might be combined with forward or side movements to dodge obstacles. Especially in sports, jumps are mostly incorporated into bigger movement patterns, such as block jumps in volleyball.

We perform jumps because they have several benefits for general health. The jump movement activates multiple muscle structures and allows us to exercise all of them at once. Moreover, it can help to improve fundamental human skills like **coordination**. A good execution requires the jumper to "*[...] coordinate the impulse from the arms with the impulse on the legs [...]*"*(**E2**)*. A physiotherapist emphasized the importance of jumps for **injury prevention** and gave an example of why it can help us: "*If I am an untrained person and I trip over a curb I could twist my ankle. But if I have trained [to jump] and [thereby] manage to activate my muscles quickly, I might be able to stabilize my body in time so that this accident doesn't happen*"*(**E6**)*. Despite this potential for injury prevention, jump training rarely finds applications in rehabilitation because health insurance often covers "*only the bare necessities*"*(**E6**)*. By jumping, the risk of some physical health conditions, such as "*osteoporosis*"*(**E2**)*, can be decreased. Interestingly, jumping, in fact, might be a helpful tool to trigger the regrowth of bones for older adults with reduced bone density. However, it might also be dangerous as "*maybe they don't have the stability yet to catch themselves and break away*"*(**E5**)*. Therefore, most experts indicated that jump training is generally not favorable for older adults. In particular, one expert pointed out that alternatives should be considered to reach similar benefits: "*Do jumps then make much sense or don't you achieve that rather, for example, with more strength training equipment?*"*(**E2**)*.

In general, we mainly use jumps for **testing, training, and in sports games**. First, jumps are a good instrument to test people's performance, and can lay down a foundation to understand "*[...] the rate of force development, i.e., how quickly can I generate force, which is well represented by jumps*"*(**E1**)*. However, the primary application of jumps lies in the actual gameplay. Especially in certain sports such as basketball and volleyball, jumping can be a decisive part of the game. For instance, an interviewed basketball trainer noted that a good jump performance could be a game-changing advantage for some players: "*Jumping power itself is particularly relevant in basketball because you can compensate for your physical size a bit, if necessary*"*(**E7**)*. Finally, using jumps in training, e.g., for volleyball, can be beneficial to improve the fatigue capacity: "*Someone who practices a lot and is well trained in jumping can hold the jump height much longer with a lot of jumps before it gets less*"*(**E1**)*.

3.2 Theme 2: Jump training combines reactive and strength exercises. Incorrect executions of maximal jumps leads to injuries.

Domain experts frequently gave the obvious answer to how to best train for better jumps: "*by jumping*"*(**E1**)*. However, apart from this conspicious statement, it is essential to frame individual exercise goals: increasing the jump power requires different training concepts from working on the jump technique.

Jump power is mainly determined by the muscles' rate of force development. Hence, it can be improved through various training modalities. In particular, **explosive movements** and **speed-based exercises** are especially effective. An interviewed basketball trainer described a typical reactive training they regularly performs: "*We put several [boxes] in a row behind each other and did bounce training there with our legs closed. Always both feet on a box and then further, up, down, up, down and so moved through the hall*"*(**E7**)*. Additionally, exercises that target the muscles' lift capabilities, e.g., traditional strength training, also improve the overall jump height. Also, the experts recommended **combining jump training with other physical exercises** and including "*different variations*"*(**E2**)* as jumps are usually not performed in isolation. Unfortunately, short-term improvements are not to be expected as changes in muscle mass and composition typically take many weeks: "*It usually takes 4 to 8 weeks minimum, until you can really prove something muscular or microscopically*"*(E5)*.

Even though small hops are generally unproblematic, most experts agreed that the maximal vertical jump is a highly complex and error-prone movement. Thus, the foremost goal of maximal jump training should be **correct execution**. Given the variety of used jump techniques, we explicitly asked the experts about the correct execution of a countermovement jump with an arm swing. This movement is typically executed by starting in an upright shoulder-wide stance. After a quick downward movement, the athlete extends their body explosively. The arms swing in an upward arc and are stopped roughly at chest height before the feet lose contact with the ground: "*If I want to jump to the maximum, I actually have to slow the arms down at shoulder height and at the right moment*"*(E2)*. Despite this general routine, the individual body shape and the sports context can lead to differences in the execution, e.g., athletes cannot reach lower squat depths if the gluteal muscles are too weak.

Most experts deemed **landing** after a jump the major source of **injuries**. A particular pain point is the knee movement. The knees should always remain in one line between the ankle and the hip. However, high-impact forces can cause the knees to collapse medially. This knock-knee position is perilous and puts extreme pressure on the knee and the surrounding ligaments. Apart from a weakly developed musculature, gender differences contribute to this condition as females are generally more prone to having knock-knees: "*There are also anatomical reasons, it's a little bit due to the hip position of women*"*(E2)*. Lastly, preinjuries can increase the risk of further accidents while jumping. Apart from an incorrect landing, the experts mainly attributed the **exercise environment** as an important injury factor and suggested using a mat as a protective measure. In contrast to the muscular changes, improvements in jump execution are quickly achievable but hard to quantify. One expert proposed measuring the knee deviation during landing as a possible improvement.

3.3 Theme 3: Jump training should increase gradually and account for individual differences, goals, and improvements.

In contrast to general physical activity, training is always **goal-oriented** and consists of exercises that exert a sufficient stimulus to trigger progression, such as muscle growth. However, our interviewed experts underlined the importance of carefully weighing training intensity, repetition count, and recovery time to maximize improvements and avoid injuries. Special attention should be placed on beginners who are not regularly exercising. For instance, "*if they are not used to it, [their knees and ankles] are very susceptible to evasive movements*"*(E5)*.

As the risk of injury depends primarily on the range of motion, training should **start with simple exercises**, such as mini hops, before gradually increasing the difficulty. Doing so also has another advantage: many experts agreed that small jumps are generally safe and do not require a prior warmup. In contrast, intensive or longer training sessions should be preceded with warmup movements, such as small hops, to prevent muscle strains. One expert whose research focuses on people with special needs proposed to even start with simple steps to make the application accessible for users with coordination and balance problems: "*So before it's even about doing a jump. To first step over an obstacle, sometimes with the right, sometimes with the left, in order to promote balance*"*(E9)*.

Apart from **adapting** the **training difficulty** according to the users' abilities and fitness level, experts also emphasized the importance of **providing proper feedback**. Guiding the users and building competence is vital for sustainable training results. One interviewed sports didact reported the benefits of recording the users' movements "*so that [they] can see their own jump to create a movement image of themselves*"*(E3)*. Combining this intrinsic learning with extrinsic feedback is particularly useful for continuous improvements. Additionally, the experts recommended focusing primarily on repetitive situations that permit users to incorporate their insights in the subsequent execution.

3.4 Theme 4: VR jump training can provide real-time feedback and boost motivation. Yet, safety might be an issue with VR.

In general, all experts saw potential benefits and challenges in using VR for jump training. Firstly, a major advantage of VR and AR is the ability to provide directly applicable **real-time feedback**. Also, whereas some experts recommended using mirrors or video recordings to show the users their movements, many people do not like seeing themselves. In this case, seeing a replay of their own jump might be even counterproductive: "*Of course, it's not helpful at all to then replay a video of the own moves that aren't working out*"*(E3)*. VR can help avoid such potential alienizing effects by introducing an additional abstraction layer. For example, users could see a generic avatar performing a replay of their jump.

Furthermore, VR exergames could benefit mental well-being by improving mood, reducing stress, and increasing the users' **motivation to train**. In particular, one expert suggested that gamified exercising might provide similar strong incentives like peer support and outperform wearables, as "*just a fitness tracker [...] can only change something, if so, in the very short term*"*(E4)*. However, as we have seen with popular augmented reality (AR) exergames, such as Pokemon Go [112], the users' interest can decrease over time.

Finally, the experts raised concerns regarding the **safety** of VR-based jump training. For instance, a mismatch between the feedback from the virtual world and the physical movement likely causes cybersickness and could even lead to dangerous situations. Therefore, it is necessary to consider potential issues early in the design pipeline because "*if [...] the risk of falling or somehow feeling unwell is greater than the benefit I generate, then it's immediately a problem*"*(E1)*. Lastly, one expert expressed doubts about whether VR should be used to measure the jump height, as there are likely more affordable and precise approaches.

4 EXERGAME: JUMPEXTRA VR

Based on the insights from our expert interviews, we designed a VR exergame with the Unity game engine [129] to train the vertical jump. Our primary focus was on motivating players to exercise in a gameful way (i.e., using the motivational pull of games). However, we also wanted to avoid any injuries (i.e., injuries that can be seen

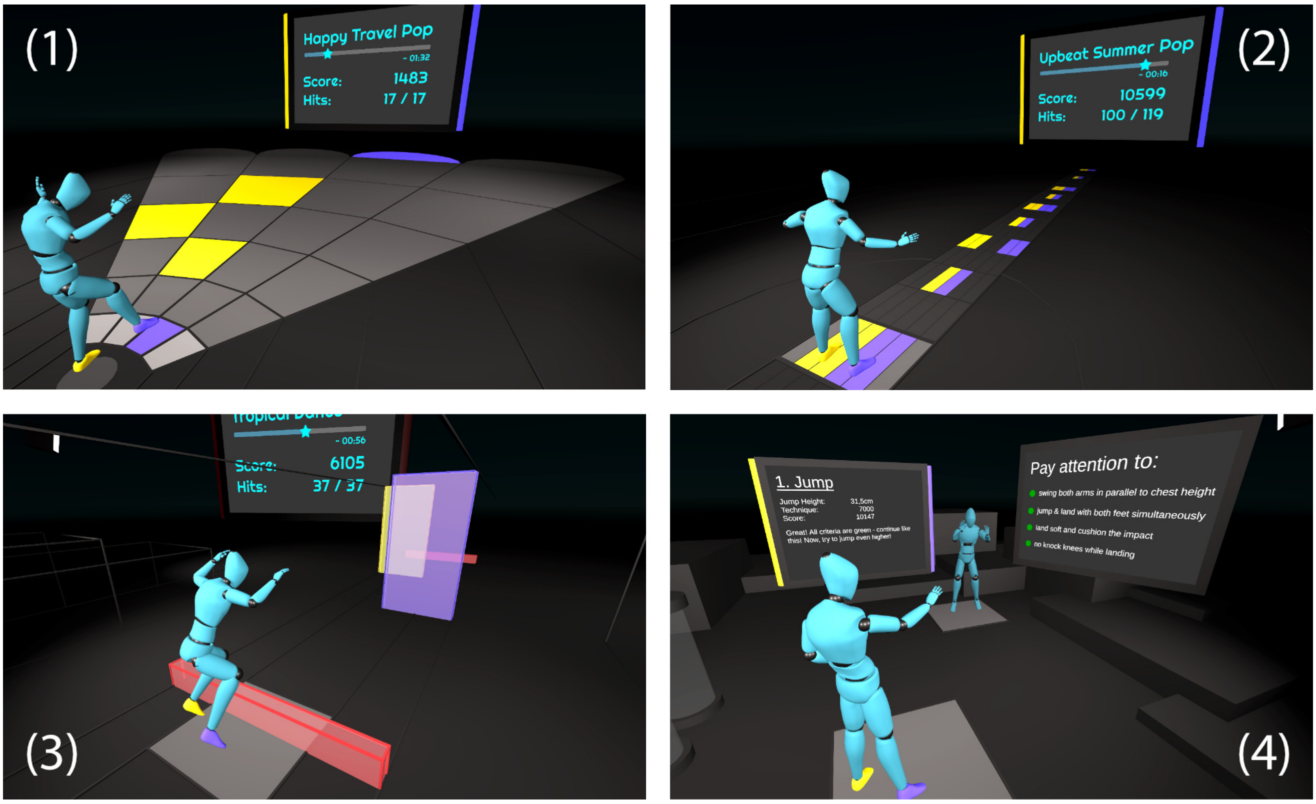

Figure 3: Our exergame *JumpExTra VR* features four sequential levels. In the first level, players tap on colored tiles with their feet. The second level features a hopscotch game. In the third level, players avoid obstacles. In the last level, players train their maximal vertical jump and receive personalized feedback.

during the gameplay, such as hurting a body part due to falling) or frustration by aligning the gameplay with the players' capabilities and giving feedback to assist them in improving their jump.

The first challenge for implementing a jumping-based exergame is tracking the entire body's movements because most VR systems use only an HMD and tracked controllers. We experimented with various approaches throughout our design process, such as using a Kinect 2 for Windows [104]. However, the high latency and inferior tracking quality when jumping made markerless motion capturing an undesirable choice for our use case. Instead, we opted for using a Vive Pro [63] and attaching Vive Trackers [64] to the players' shins and waist. Even though tracking players' feet was our initial first choice, our pilot tests revealed that placing the trackers on the shins drastically improves tracking accuracy and comfortability. Together with the controllers and headset, we used a total of six tracking points to animate a virtual avatar using inverse kinematics.

4.1 Rhythmic Levels

Many experts emphasized the importance of raising the difficulty gradually by starting with small hops before attempting higher jumps. This design not only prepares players for more intensive sections but can also serve as a warmup. In contrast to maximal jumps, small hops have a negligible risk of injury. One expert, who focuses on user groups with special needs, raised the concern that some players might not be ready for hops at all since the frequency someone uses jumps mainly depends on their exercise practice. Instead, the expert proposed starting with steps and balancing exercises before continuing with hops and larger jumps. As our goal was to design an engaging experience for everyone regardless of prior experience, we structured our game into four sequential levels, starting with tapping before increasing the intensity with hops, small jumps, and finally, maximal vertical jumps (see Figure 3).

The first three levels are structured similarly and tie the players' actions to the beats of a song. Such rhythmic movements are suggested by the literature [106] and have been used with great success in some of the most famous VR games, such as Beat Saber [52] or Ragnaröck [133]. Whereas all levels in our application feature a different song, the length is always roughly two minutes with 128 beats per minute. Before starting the action, players receive a short introduction to every level and must perform each relevant movement correctly. During the level, a screen in the background displays the song progression and the players' performance, measured in successful hits and a derived score. Also, the game provides motivational feedback in the resting break between the levels.

To account for individual differences and fitness levels, we implemented three difficulty levels differing in the number and complexity of necessary movements. Starting at the lowest difficulty, the game automatically and unnoticeably switches to a higher level if players achieve a precision of at least 90% or to a lower level if they miss more than 40% of the last notes. With this design decision, we want to ensure that the game challenges all players without causing frustration. In particular, the game should ensure that players experience a feeling of success by performing at least half of the movements successfully. Conversely, it should raise the difficulty if players are not challenged and succeed in most interactions. Ultimately, we aimed for an average hit rate between 60% and 90%. With this adaptive difficulty, we ensured that most players would achieve a high score boosting their motivation and confidence.

Level 1: Tap to the Beat. In the first level (duration 01:54 min), yellow and purple tiles approach the players on four adjacent lanes. As the tiles reach the front line, players must tap on them with the correct foot — yellow tiles require the left foot, whereas purple tiles map to the right foot. After tapping, players must retract their feet as they must not enter the playing field with both feet simultaneously. In the course of the song, the movement patterns become more complex and require players to switch feet quickly or tap crosswise. On the easiest difficulty, the game confronts players with 110 tiles, as opposed to 135 tiles on the hardest difficulty. This level is mainly intended as a warmup and trains the players' lower body coordination, stability, and reaction time.

Level 2: Hopscotch. The second level (duration 01:56 min) advances on the first one by incorporating small hops. Similar to the child's game hopscotch, players have to hop with their left, right, or both feet on the correct tile. However, these hops are performed in place as the approaching tiles reach the players' position. The tiles move on five overlapping lanes, of which either one or two light up in yellow or purple to indicate the correct feet. Players must remain in the last pose between two hops, e.g., standing on one leg until another tile reaches them. Throughout our design phase, we learned that longer phases on one leg and fast switches between two and one leg are highly challenging and often do not fit the music. Instead, we mainly used slower two-legged jumping patterns or faster one-legged "walking-style" patterns. As before, the movement becomes more complex with time and incorporates the outer lanes more to force players to move from side to side. Depending on the difficulty level, players must perform between 108 and 129 hops. This second level builds on the already trained balance and stability. Also, players must time their hop correctly to land on the tile when it lights up. This feature further trains coordination and neuromuscular control.

Level 3: Obstacle Course. The last song-based level (duration: 02:08 min) follows a different principle than the first two. This time, obstacles approach the players at every first beat of a bar. To gain points, players must avoid touching these impediments with their bodies. The most common obstacle is a low wall forcing players to jump at medium height. As our domain experts emphasized the importance of diversified training and dynamic exercises, we interleave the jump-over obstacles with lateral walls to both sides and barriers hanging from the ceiling. These force players to move sideways and duck down before jumping again. In total, this level features 65 obstacles, of which 24 are jump hurdles. As personal size differences could pose an unfair disadvantage due to height differences, we scale the lower and upper walls according to the players' height during calibration. Also, the three difficulty levels affect only the obstacle size, not their frequency. Consequently, players must jump higher with increasing difficulty. This third level mainly focuses on muscle strength and endurance while preparing the players for the last level featuring maximal vertical jumps.

4.2 Maximal Vertical Jumps

After training the prerequisites for a good jump — general fitness, balance, and coordination — our last level focuses on teaching players a proper jump technique. In the beginning, players see an exemplary jump execution as part of a short introduction to the level. Next, players have to perform maximal vertical jumps and receive personalized feedback on their performance based on the four criteria below. The game continuously records the players' movements. After detecting a jump, this data is analyzed with regards to four criteria of a safe and efficient jump that the domain experts mentioned:

(1) jump and land with both feet simultaneously
(2) land softly (forefeet touch the ground first) and absorb the impact with the entire body
(3) especially while landing, keep the knees in one line between feet and hips and do not cave them inward
(4) swing arms synchronously in a forward-upward arc until about chest height

Analyzing most of these criteria, such as arm and leg synchronicity, is easily achieved by comparing positional differences to precalibrated thresholds. However, determining the players' landing style is more challenging since we do not track players' feet directly. Since the landing happens in only a split second, it is too fast for the precise infrared sync of the tracking system. Instead, the Vive trackers have to rely on their less-precise accelerometer. We use this tracking limits of the Vive system to our advantage; if players land hard without catching their impact through feet and legs, the measured position of the shin trackers descends for a brief time well below floor level. This vertical displacement directly correlates with the landing impact - a softer landing leads to less displacement and vice versa. So, our algorithm can use this tracking error as an indicator of the players' landing quality. Furthermore, we implemented offset values until a violation of the above factors is considered insignificant to account for tracking imprecisions. After performing a jump, players see their performance in the four categories on one screen and receive a personalized message on another. This message summarizes their improvement from the previous jump, their worst-performing criterion in this jump, and practical instructions on how to improve on it in the next turn. Additionally, we visualize the last jump as a looping replay in front of the players, which allows them to study their movements. Multiple experts deemed combining extrinsic instructions with the opportunity for intrinsic feedback through jump visualizations highly valuable.

Apart from providing personalized instructions, we also calculate a jump score from the technical criteria and the effective jump height to reflect the performance. Together with the jump height, this value is listed on a highscore board, informing players of their improvements from jump to jump. In total, players perform five consecutive jumps and try to incorporate the feedback from the previous execution. After the last jump, the game ends.

5 EVALUATION

After designing our exergame *JumpExTra VR*, we conducted an exploratory user study to explore how users perceive our novel training experience. Our primary research goal for our study was to confirm that our prototype provides an enjoyable user experience without causing unwellness or endangering players. Since players constantly remain in one spot through the experience, we are confident that the application is not likely to induce cybersickness. Additionally, we are interested in how the game's usability, appeal, and feedback contribute to the players' overall game experience.

Apart from these perceptional factors, we want to explore our exergame's motivational and physical effects on the players. Firstly, following our experts' feedback, we assume it improves the players' mood. Also, we hope that the players find the game's physical exercises challenging without frustrating or overly tiring them. Consequently, we are interested in how players perform in the various levels and where they see the future potential of such an exergame-based jump training. Considering these motivations for our exploratory study, we employed various methods, including pre- and post-questionnaires, game performance data, and qualitative feedback, to answer our research questions:

- RQ1: How does our exergame affect cybersickness symptoms of players?
- RQ2: How does our exergame affect the players' mood?
- RQ3: How do players evaluate the player experience of our exergame and its usability as a training tool?
- RQ4: How do players perceive the physical activity and perform in our exergame?

5.1 Pre-Post Questionnaires

These questionnaires were administered before and after the gameplay. To answer RQ1, we administered the simulator sickness questionnaire (SSQ) [74] in German [61]. It measures three sub-categories, i.e., nausea, oculomotor disturbance, and disorientation, through 16 items on a 4-point scale. For RQ2, we assessed the players' mood with the energetic and valence sub-categories of the German version of the Multidimensional Mood Questionnaire (MMQ) [134], consisting of four bipolar items on a 7-point Likert Scale (ranging from 0 to 6).

5.2 Post-Questionnaires

To measure the general game experience to answer RQ3, we used multiple sub-categories of the German version of the Player Experience Inventory [1, 54]: mastery, immersion, progress feedback, audiovisual appeal, challenge, ease of control, clarity of goals, and enjoyment (7-point Likert scale, ranging from -3 to 3). Additionally, we used the physical activity enjoyment scale (PACES-8) [109] for RQ4, which includes eight bipolar items on a 7-point Likert scale (ranging from 1 to 7). The German item translation of PACES-8 was provided by one of the researchers, who used this scale in another research project.

5.3 Game Performance Measures

As we were interested in how players perform in our exergame to answer RQ4, we logged the necessary performance data for all participants. For each of the three rhythmic levels, we recorded the dynamic difficulty development and the hit ratio, i.e., how many steps, hops, and jumps over obstacles were successfully performed. Additionally, we collected the jump height and the performance in the four jump criteria for each jump in the final level.

5.4 Qualitative Feedback

We gathered qualitative feedback from the players through open-ended questions to understand their experience with *JumpExTra VR*. In these questions, we particularly focused on four topics: risks and benefits, safety and usability, long-term participation, and improvements. We used the following open-ended questions to capture the perspectives of the players on these aspects:

- *"In your opinion, what are the risks and benefits of this VR game for you? Why?"*
- *"Have you encountered any issues that affected your safety and usability during the VR gameplay? Why?"*
- *"If you could continue to use this VR game, how do you think that this would affect your long-term participation in jump training? Why?"*[2]
- *"Considering the VR game you played, what aspects would you like to change, and what aspects would you like to keep as they are? Why?"*

6 RESULTS

Twenty five participants (15 female, 10 male, *M*=24, *SD*=6.01 years) were recruited for our study. Twenty of them had prior VR experience. However, only two reported using VR frequently (1-2 times per month). Of the rest, twelve participants rarely used VR devices (1-2 times per year), and the remaining had only one to two prior sessions. Asked for their exercising habits, only two participants stated not to be exercising. Among the rest, ten participants exercised one or two times per month at most. Eleven reported exercising at least once per week, and two trained daily. Additionally, the participants generally rated physical exercise enjoyment slightly positive (*M*=1.24, *SD*=1.45, range -3 to 3).

After learning about our research objectives and signing informed consent, participants completed the first part of the questionnaire, assessing demographics, pre-SSQ, and mood. Upon completion, we introduced the participants to the Vive Pro VR headset [63] and assisted them in attaching the Vive Trackers [64] to their shins and waist. After starting the game, the participants received an introduction to the controls and calibrated their avatar before playing the four levels. After the playthrough, they completed the post experience questionnaires and answered the open-ended questions. The duration of the study was 35 minutes on average.

[2]This question is aimed at understanding participants' attitudes toward using this exergame in the long term, but we note that this cannot be taken as evidence that people would actually continue playing this game.

Table 1: The table shows the descriptive and the results of statistical tests of pre-post SSQ and MMQ (* *indicates significance*).

Questionnaires	Pre-Que.		Post-Que.		Statistical Test	P Value	Effect Size	Confidence Interval (.95)
	Mean	*SD*	*Mean*	*SD*				
SSQ Total	15.41	13.41	21.09	14.76	$t(24)$=-1.71	0.100	d = -0.34	[-12.533, 1.163]
SSQ-nau.	15.26	11.02	33.58	17.44	V=4.5	<0.001*	r = -0.55	[-28.620, -14.310]
SSQ-ocu. dis.	15.77	16.80	11.82	14.19	V=72.5	0.220	r = -0.17	[-7.580, 22.740]
SSQ-dis.	6.12	12.76	7.80	13.38	V=11	0.170	r = -0.19	[-13.920, 13.920]
MMQ-ene.	3.6	1.35	4.1	1.61	$t(24)$=-1.50	0.146	d = -0.30	[-1.187, 0.187]
MMQ-val.	4.6	1.24	4.2	1.55	V=112	0.092	r = -0.24	[-0.000, 1.250]

Table 2: The table shows the descriptive values of the PXI and PACES-8 questionnaires: The player experience and physical activity enjoyment of players were generally high.

PXI-enj.		PXI-mas.		PXI-imm.		PXI-pro. fee.		PXI-aud. app.		PXI-cha.		PXI-eas.		PXI-cla.		Paces-8	
Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
1.95	0.79	0.93	1.19	2.03	0.75	1.6	0.88	1.33	0.98	1.71	0.80	1.99	0.84	2.2	0.93	5.44	0.91

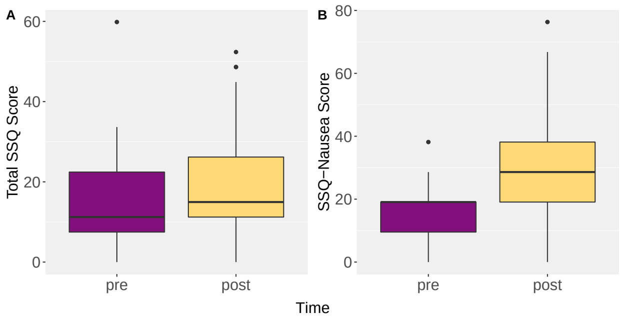

Figure 4: Although the participants' Total SSQ ratings were not significantly different between pre- and post-time points, post-SSQ-Nausea scores were significantly higher compared to the pre-SSQ-Nausea scores.

6.1 Pre-Post Questionnaires

We conducted Shapiro-Wilk tests to check the normality assumption of the data. When the data was normally distributed, we used paired t-tests and reported the effect size with Cohen's d. In the case of non-normally distributed data, we used Wilcoxon-signed rank tests and reported r, Pearson's correlation coefficient, as the effect size measure. We followed Cohen [33]'s recommendations to intepret these effect sizes. Table 1 lists the results of statistical tests as well as descriptives of each questionnaire and their sub-categories.

A wilcoxon-signed rank test indicated that participant's post-nausea scores were significantly higher than pre-scores, V=4.5, p<0.001, Figure 4b. However, Total SSQ scores and the other sub-categories of SSQ did not indicate significant differences between pre- and post-values.

Due to the nature of exercise, people tend to sweat. Whereas sweating can be a symptom of cybersickness, in our case, it is more likely an effect of the physical effort during training [6]. Therefore, we also performed the SSQ analysis while excluding the sweating item. Since this item is considered only in the calculation of Total SSQ and SSQ-Nausea categories, we report only their analysis. Total SSQ scores were not significantly different between pre- (M=12.42, SD=13.30) and post-time points (M=12.42, SD=13.73), $t(24)$=0, p=1, d=0. Similarly, the players' nausea ratings did not significantly differ between pre- (M=7.63, SD=9.54) and post-measurements (M=11.45, SD=13.49), $t(24)$=-1.29, p=0.211, d=-0.26.

Neither the energetic nor the valence sub-scale of the MMQ showed significant differences between before and after scores of playing the game.

6.2 Post-Questionnaires

We report the descriptive values of the questionnaires administered as only post-game measures in Table 2. The findings indicate that all measures were rated with a positive tendency. Whereas most constructs highlight a highly positive experience, the PXI-mastery sub-category indicates that the players did not feel a particularly high mastery in the game.

6.3 Game Performance

For the first three levels, we logged the participants' hit ratio as the main performance measure. By adding the dynamic difficulty adaption in our game design process, we aimed for an overall success rate between 60% and 90%. This goal was achieved for all three levels *Tap to the Beat* (M=78.4%, SD=15.8), *Hopscotch* (M=69.6%, SD=15.4), and *Obstacle Course* (M=85.1%, SD=8.6). On average, participants spent 41.77% of the time in the medium difficulty (SD=23.63), followed by the hardest difficulty (M=35.33%, SD=25.11) and the easiest difficulty (M=22.88%, SD=19.33).

The participants mainly did not improve their jump height in the final level over the five jumps. The average height remained roughly the same from 38.28cm (SD=9.12) in the first round to

37.66cm (*SD*=9.10) in the last execution. However, participants improved their technique score slightly over the five jumps, starting from an average rating of 71.45% (*SD*=17.18) and ending at 75.45% (*SD*=15.30).

Lastly, we also logged how often the game provided feedback to the players. Although this data is purely descriptive, it gives a good impression of how our system responded to the players' movements. Generally, the game advised participants for 62% of the jumps to pay attention to the correct knee movement during the landing. This criterion was followed by a too-hard landing, an issue in 26% of the cases. Lastly, the arm swing was too high for 20% of the jumps and not fully parallel for 8% of the jumps. In contrast, asynchronous feet movements during the lift-off and landing were not an issue.

6.4 Qualitative Feedback

We collected qualitative data in the form of open-ended questions. Before the analysis, the data was translated into English using DeepL Pro [38] and was checked by one of the first authors for inaccurries. After this, we used Dovetail tool to code the data [40].

One of the first authors analyzed this data using a reflexive thematic analysis approach [18, 19]. Before the analysis, the author decided on four deductive categories: *risks and benefits*, *safety and usability*, *participation*, and *improvements*. The author coded the interview data using inductive codes (e.g., "injury risk", "accessibility of exercise opportunities", "standing zone should be improved") under these categories. Following this step, they performed an affinity mapping activity and based on this, they created following themes.

Theme 1: The main advantages of JumpExTra VR are accessibility and enjoyment, however, participants also reported injury concerns. The participants mainly attributed the advantages of this game to two factors. The first factor is the accessibility of physical exercise opportunities: "*[...] you don't need to go to the gym since you can easily exercise at home*"*(**P4**)*. Secondly, they emphasized the enjoyment aspect of *JumpExTra VR*: "*More fun while exercising*"*(**P9**)*. Interestingly, one participant reported both the pros and cons of immersion in this game: "*Forgetting the real world is a disadvantage and being completely immersed in the world is an advantage*"*(**P6**)*. Another player highlighted the positive side of this game by comparing it to another commercial alternative: "*More movement for players than in other VR games like Beat Saber*"*(**P16**)*.

JumpExTra VR was found to be associated with some drawbacks. Some players reported the possibility of losing physical-world awareness while playing this game: "*A disadvantage would be that you might get too "infatuated" with the virtual world and neglect real life*"*(**P4**)*. Additionally, many participants were concerned about the potential injury risks. These were attributed to various causes, but mostly to falling: "*I found it risky to fall down*"*(**P14**)*. The risk of physical collisions were also pointed out: "*You might bump into objects in the real environment*"*(**P8**)*. Referring to the technical feedback in the last level, one participant stressed the need for similar training instructions for the more gameful levels, too, "*because, if [they] do the tasks wrong, injuries could follow*"*(**P3**)*. Keeping the balance was mentioned as an issue for a few: "*It's kinda hard to keep balance sometimes, the risk is that you can fall on the ground*"*(**P7**)*.

Theme 2: Overall JumpExTra VR did not cause serious safety and usability issues. Yet, there were some instances reported by players. Even though participants had no severe issues regarding safety and usability, a few had difficulties understanding where they were physically located: "*I couldn't remember where I was in the room*"*(**P2**)*. Some echoed the safety issues related to having accident, falling down, and losing balance. The difficulty of staying in the standing zone of *JumpExTra VR* was also noted: "*I had trouble staying within the designated play area. I got too far ahead in places, and so I couldn't get the triggers to work properly on the first play*"*(**P5**)*. Notably, the players reported on some usability problems due to the technical VR setup: first, "*Problems with sharp vision, which strained the eyes a lot*"*(**P23**)*. Second, "*Tracking did not work perfectly*"*(**P16**)*. Third, "*The cable on [their] back was disturbing*"*(**P8**)*. Nevertheless, only one participant encountered issues with the leg trackers not staying in position, and another participant had experienced usability problem due to color blindness.

Theme 3: For most players, JumpExTra VR would positively impact participation in jump training. Most participants agreed that the chance of using *JumpExTra VR* would positively affect their long-term participation in jump training. Some commented that "*it would be good for [their] fitness*"*(**P6**)*. According to one player, especially over time, one would get better at performing jumps and at the evaluated aspects, which "*increases the average jumping power and also reduces the risk of injury*"*(**P3**)*. The feedback feature of the game was particularly appreciated by some participants and mentioned as a reason for long-term use: "*My jump will probably improve through feedback*"*(**P10**)*. For a few participants, the motivating or fun nature of *JumpExTra VR* played a role: "*The design and the atmosphere is very pleasant. It would motivate me*"*(**P19**)*. In particular, one player emphasized the advantage of using games for physical exercises: "*[...] during the games you don't really notice how you practice your jumps and therefore it is not so monotonous*"*(**P24**)*.

However, some participants would not continue using *JumpExTra VR*. For a few, this was due to safety and comfort issues that can occur when jumping in VR: "*I didn't feel safe enough to jump with full power. The headset wobbled too much for that and I was afraid of not landing properly or damaging the headset.*"*(**P5**)*. Tracking imprecision was also pointed out as a reason: "*I also found the jump training segment to be too inaccurate from a tracking standpoint*"*(**P25**)*. Additionally, there were other reasons reported by the players, such as a general dislike for VR games, loss of interest after a while, or difficulties in maintaining habits.

Theme 4: Players suggest improvements for JumpExTra VR about the game world, feedback, standing zone, and varition of game levels. Overall, we received a lot of feedback for our game design. A few participants would even "*change nothing*"*(**P13**)*. Others emphasized the parts they liked the most, like "*[...] definitely keep the music and tutorials*"*(**P6**)*. Still, we received many ideas of how to improve *JumpExTra VR*. Some players proposed improvements to the game world, such as "*mak[ing] the game environment more colorful, beautiful, lively*"*(**P8**)*, and "*a more accurate virtual body would be desirable to play better*"*(**P14**)*. A few wished to have more feedback: "*What I would change is the feedback. A sound or buzzer was missing if the jump or obstacle evasion were successful or not.*"*(**P20**)*. A part of the players did not like constantly looking down in the gameplay and improving the standing zone was recommended: "*I wished for an alert once you are outside the designated play zone*"*(**P3**)*. More variety was also suggested for jump training exercises: "*Combine the three levels for more variety (tap, jump, dodge)*"*(**P10**)*.

7 DISCUSSION

The main goal of this research project was to provide playful and safe VR-based jump exergame. Therefore, we considered both prior literature and the findings of expert interviews to design *JumpExTra VR*. As the final step, we evaluated our prototype in a mixed-methods user study (N=25). In the following section, we discuss the findings by focusing on the hypothesis and research questions and provide design implications that can help researchers and practitioners to expand on full-body exertion experiences.

7.1 RQ1: How does our exergame affect cybersickness symptoms of players?

Our results show that the players reported higher SSQ-nausea scores after playing the game. This finding supports previous research [125] showing that VR exergames can cause symptoms of cybersickness. However, we did not observe increased ratings for the remaining categories of SSQ: oculomotor disturbance, disorientation, and the total score. We suspect the increased sickness scores were due to the sweating item of SSQ. The effect of physical training on sweating has been shown in the literature [6]. Our additional SSQ analysis, excluding the sweating item, supports this assumption; the results show no significant effect between pre- and post-measurements. As suggested by [125], we also emphasize the potential overlap between cybersickness and physical activity symptoms. Our results underline that high-intensity full-body exercises can be safely performed in VR without risking discomfort.

7.2 RQ2: How does our exergame affect the players' mood?

Feedback from our domain experts and literature [28] suggests that the physical activity in our exergame is likely to improve the players' mood. To explore this potential effect, we assessed the players' energetic arousal and valence. Our results from the MMQ show that neither of the two subscales revealed any significant difference. We explain this outcome two-fold. In case of the energetic arousal subcategory, most players stated being physically exhausted and sweaty after playing the game. They explicitly attributed their state to two external reasons: summer heat and headset fit. Unfortunately, the study coincided with an extreme summer drought with high temperatures. Also, COVID-19 measures forced us to use easily cleanable headset foams which accumulated extra heat. Both factors added to the expected exhaustion and potentially caused players to feel tired rather than stimulated. Still, as players generally appreciated the physical challenge, we are confident that such full-body exergames can benefit the players' physical and mental well-being in the long run.

In contrast, the slight decrease in valence scores might be connected to the low scores on the PXI's mastery dimension, indicating that participants did not feel particularly good at playing the game. Due to the time constraints of the study, we could let participants try the game only once. However, such fast and timed movements are always hard when attempting them for the first time. Therefore, we speculate that players would feel more relaxed and capable with further repetitions. In turn, a feeling of success that was missing in this first run might also positively impact participants' valence.

7.3 RQ3: How do players evaluate the player experience of our exergame and its usability as a training tool?

JumpExTra VR was implemented based on prior literature and expert interviews. Therefore, we see the high ratings of PXI as a result of the detailed design process of our game. Aligning with previous studies [45, 67, 88], we show that jump-based exergames can lead to positive player experiences. With this, we extend the results of these studies into the training realm of jumping in VR.

Overall, the players highly enjoyed playing this game and found it immersive and appealing. The high PXI progress feedback scores show that the game provided comprehensive feedback regarding players' progress. This finding is further supported with qualitative data as the players appreciated the game's feedback feature: "*feedback will probably improve my jump*"*(**P10**)*. We believe that the instructions given before each level (clarity of goals), interactive tutorials (ease of control), and dynamic difficulty (challenge) are reflected in the high scores in the respective PXI sub-categories. However, players criticized the need to look down at their feet constantly. This pose is uncomfortable and could potentially lead to postural degradation. Instead, visual and audio feedback informing players upon a hit or miss could reduce the urge to check on their own movements.

Even though most players stated they would continue using *JumpExTra VR*, we cannot take this feedback as evidence for long-term retention. Instead, we emphasize the need for longer, focussed studies. Furthermore, players requested visual improvements, namely a more lively world and a better-matching avatar. Also, they proposed combining the levels for more variety in gameplay. Especially this last point is vital for the long-term success of future full-body exergames. To secure retention, developers should focus on varied gameplay with high replayability. Rhythmic elements used in our game or famous titles like Beat Saber can be easily extended by adding new songs and mappings. The repetitive technical jump training in our last level bears a much greater challenge and will likely suffer from decreasing interest. Hence, it remains an open research question on how to enhance retention for advanced training routines as well.

7.4 RQ4: How do players perceive the physical activity and perform in our exergame?

Overall, all participants performed well in our exergame. In particular, our results reveal that the individual success rates for all levels are well within our anticipated window of 60% to 90%. Hence, we can assume that our dynamic difficulty adaptation worked well in aligning the exergame's challenge to the players' individual abilities. Within our sample, the participants were also able to increase their personal technique score in the final level by 5.60 % on average. In contrast, we did not see any improvements in jump height. These observations match the domain experts' feedback that muscular changes take multiple weeks, whereas technical improvements are quickly realizable. However, we note that our results are purely descriptive. Determining the system's accuracy would require an additional evaluation in a professional movement analysis lab. In terms of perceptual effects, the high values in the challenge subcategory of the PXI show that the participants were positively challenged without feeling overly taxed. Additionally, the PACES-8 results highlight that the participants enjoyed the physical activity in *JumpExTra VR*. Lastly, we observed improvements in the players' skills. Initially, many participants had immense difficulties with coordination, used the wrong foot, or could not follow the song's rhythm. As their performance improved throughout the course of the game, we believe that further sessions would positively influence players' body control and neuromuscular coordination.

7.5 Design Implications

Avoid the Unintended Forward Drift (UFD). In our study, the frequent tapping and jumping on the spot led to small, unnoticed forward movements that required the players to monitor their position to prevent leaving the play area. This effect is similar to the unintended positional drift (UPD) [113] that is observed for walking-in-place locomotion techniques. In an analogy, we name our observed effect unintended forward drift (UFD). Unfortunately, we do not see an easy solution except for integrating a warning and pausing the game when players leave the designated area (matches with guidelines of Xu et al. [140]). Forcing players to remain precisely in one spot would increase the necessary mental effort and potentially spoil the game experience. An option would be to integrate omnidirectional gameplay where players jump in all directions and not just forward. However, such game design is challenging because it adds events behind the players' backs.

Consider Technical Constraints. Currently available VR hardware was not designed for tracking explosive full-body movements. Its limitations manifest in the need for additional hardware (e.g., Vive trackers) and the likeliness of tracking errors. Although these technical limitations limit the practical applicability for today's commercial games, we are confident that they are resolvable with future generations of VR headsets. Instead, we share the lessons we have learned by using the current state of technology to achieve a jump-centered exergame. In our pilot tests, we experimented with different approaches to attaching the trackers to the players' legs. In the end, we decided on the shins. Compared to typical shoe adapters, trackers are more stable, more comfortable to wear, and easier to attach in this position. However, during the study, we sometimes noticed that trackers slipped slightly, especially if participants wore slick pants. Even if trackers remain in place, Vive lighthouse tracking is not accurate enough for fast leg movements despite being one of the most precise VR systems. Consequently, we noticed considerable jittering during the game. This problem is not severe in the rhythmic levels and was not noticed. However, when using the tracking data to display a replay of the vertical jump, users see this suboptimal tracking quality. Consequently, we recommend applying a slight low-pass filter to remove artifacts in cases where players see their own movements, like in our replay.

Besides visual artifacts, the tracking-related inaccuracies also influence the quality of the jump-technique feedback. Paired with a variety of body types and postures, these technical constraints can increase the false-positive rate of mistake detections, such as problematic knee movement. To avoid wrong corrections manifesting in the players' jump technique, automated feedback should be a tool only to support their own critical reflection on their movements. In this context, we noticed that players often did not recall what the optimal jump, shown in the tutorial, looked like and could not draw proper conclusions for their jump. Hence, we recommend tying visualizations of the "optimal performance" with the replay and complementing both with the less-precise automatic feedback. This combination further improves the players' critical view of their performance and may lead to lasting technical improvements.

Account for Safety Concerns in Fully-Body VR Exergames. We extend the work of Mueller and Isbister [106] by providing safety implications specifically for full-body VR exergames. As we have seen with the UFD, performing full-body movements in VR has safety concerns that must be accounted for early in the game design pipeline. In particular, one of the main reasons we initially chose to focus on the vertical jump is its compatibility with normal-sized tracking areas. Other movements, such as forward jumps, transport users quickly to the play space's borders or beyond. This consideration is particularly important for game developers who must account for the varying consumer play spaces that often do not have the generous size of a $16m^2$ VR lab.

Another critical concern is the potential risk when using a cable-bound VR headset, such as the Vive Pro. In our case, players mostly did not move enough in the tracking area to encounter cable-related issues. However, one of the authors was always present to monitor the participants' behavior. Of course, this approach is no solution for commercial applications. Lastly, we noticed that participants generally had more problems with balance than in real life. Since accidental tripping in smaller tracking areas could easily lead to collisions, developers should consider this observation.

Incorporate Rhythmic Elements with Fluent Movement Patterns. Similar to other successful exergames like Beat Saber, we aligned the players' movements in the first three levels to the beats of a song. This game design was widely appreciated by the participants who asked us to "*[...] definitely keep the music and tutorials*"***(P6)***. However, it is essential to limit the length of each level, as fast jumping and hopping exercises are exhausting. In our case, we used only two-minute songs. Also, we incorporated frequent "resting passages" featuring fewer and slower interactions to allow players to regain their breath.

In our game design process, we experimented with various movement patterns and learned which worked well or harmed the game experience. First, frequent switches between single-leg and double-leg movements are highly challenging and require excellent coordination and balance. Also, the rhythm of steps and jumps should be mostly regular and seldom change. For instance, patterns on every, every second, or every fourth beat work perfectly fine. However, repeated transitions between these intervals interrupt the players' flow and cause frustration due to missed notes. For single-leg tapping and hopping, we found that a "walking-like" pattern, i.e., switching the feet between every note, works well. A good equivalent for two-legged moves is repeatedly jumping into a narrow and a wide stance.

Adapt the Difficulty According to Players' Fitness Level and Improvements. Similar to prior exergame studies [107, 140], we emphasize the importance of using a dynamic difficulty adaptation—a key feature in ensuring a satisfying game experience for all players. Our game constantly monitors the players' performance and switches to another difficulty level on demand. This functionality allows us to challenge every player regardless of their fitness level and ability without overtaxing beginners. Our analysis of the game's performance measures revealed that the participants spent most of the time in the medium and hard difficulties. In contrast, the easiest difficulty was rarely used. Hence, we see potential in better balancing the difficulty levels and improving the transition parameters. However, estimating the effective difficulty of a particular song mapping is not easy. For instance, we explained in the previous design implication that fewer interactions do not automatically translate into an easy gameplay. In general, our three difficulty levels provided an optimal challenge for most players. However, our logged performance metrics revealed that few players still underperformed with the easiest difficulty, whereas others were not even challenged by the hardest level. Hence, we recommend adding more difficulty levels to fit every fitness level.

7.6 Limitations

We used a Vive Pro headset with Vive Trackers in our study. This decision was mainly motivated by the improved tracking quality compared to alternatives, such as the Kinect. However, the use of a cable-bound system also led to usability and safety concerns. Whereas we are confident that active monitoring during the exergame prevented any dangerous situations, we accept that participants might have felt more insecure or disturbed by the VR system. Additionally, we only conducted an exploratory user study to understand how players experience our exergame. Consequently, we cannot make conclusive statements about the accuracy of the automated feedback suggestions in the last level. Overall, we also note that the use of additional hardware limits access to this exergame, and we hope that future technological advancements will open up new possibilities for tracking full-body movements in VR.

For the qualitative data in our research, we reflect on our background and potential research interests that might have impacted the analyzing process [19, 111]. One of the first authors involved in the expert interview analysis has a computer science background and has published on VR, locomotion, and games research before. The other first author, who conducted both qualitative data analyses, has research experience in VR and exergames for varying user groups and a background in psychology and cognitive systems. Hence, they may have introduced bias into the analysis due to their interests and background. However, we also consider that the combination of these different specialties and perspectives enriches the data processing step.

8 CONCLUSION

VR exergames can provide a great motivation to pursue a more active lifestyle and exercise regularly at home. However, most available games track and focus primarily only hand movements in their routines, which leaves the lower body, already weakened by all-day sitting, severely undertrained and untargeted. In this work, we explored the potential of full-body VR exergames using the example of vertical jump training. Therefore, we interviewed nine domain experts and combined their feedback with insights from prior research into our exergame prototype *JumpExTra VR*. In the first three levels, the game trains lower body coordination, stability, and endurance, before providing technical feedback on the execution of maximal vertical jumps.

Additionally, we conducted an exploratory user study to evaluate how players perceive the training experience with *JumpExTra VR*. Our results reveal that the participants appreciated the physical challenge and enjoyed our jump-centered exergame. Based on the participants' feedback, we provided a set of design implications that can guide future work on full-body VR exergames and help developers design engaging experiences. In future work, we want to extend our research by evaluating the long-term effects of our exergame and compare the training effects of this game with supervised training.

ACKNOWLEDGMENTS

This work was supported by the German Federal Ministry of Education and Research (BMBF).

REFERENCES

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370. https://doi.org/10.1016/j.ijhcs.2019.102370 pages #9

[2] Kent Adams, John P O'Shea, Katie L O'Shea, and Mike Climstein. 1992. The effect of six weeks of squat, plyometric and squat-plyometric training on power production. *Journal of applied sport science research* 6, 1 (1992), 36–41. pages #4

[3] Ali Al-Khalifah, Rachel McCrindle, Paul Sharkey, and Vassil Alexandrov. 2006. Using virtual reality for medical diagnosis, training and education. *International Journal on Disability and Human Development* 5, 2 (2006), 187–194. pages #3

[4] Iannis Albert, Nicole Burkard, Dirk Queck, and Marc Herrlich. 2022. The Effect of Auditory-Motor Synchronization in Exergames on the Example of the VR Rhythm Game BeatSaber. *Proc. ACM Hum.-Comput. Interact.* 6, CHI PLAY, Article 253 (oct 2022), 26 pages. https://doi.org/10.1145/3549516 pages #2

[5] Kenneth Anderson and David G Behm. 2005. The impact of instability resistance training on balance and stability. *Sports medicine* 35, 1 (2005), 43–53. pages #2

[6] LE Armstrong and CM Maresh. 1998. Effects of training, environment, and host factors on the sweating response to exercise. *International journal of sports medicine* 19, S 2 (1998), S103–S105. pages #10, #12

[7] Timothy T Baldwin and J Kevin Ford. 1988. Transfer of training: A review and directions for future research. *Personnel psychology* 41, 1 (1988), 63–105. pages #3

[8] Wendy Balster, ELC Xue, and WK Pui. 2016. Effects of a deeper countermovement on vertical jump biomechanics after three weeks of familiarisation–preliminary findings. *International Journal of Human Movement and Sports Sciences* 4, 4 (2016), 51–60. pages #3

[9] E Joan Bassey, Maria A Fiatarone, Evelyn F O'neill, Margaret Kelly, William J Evans, and Lewis A Lipsitz. 1992. Leg extensor power and functional performance in very old men and women. *Clinical science (London, England: 1979)* 82, 3 (1992), 321–327. pages #3

[10] Nathaniel A Bates, Kevin R Ford, Gregory D Myer, and Timothy E Hewett. 2013. Impact differences in ground reaction force and center of mass between the first and second landing phases of a drop vertical jump and their implications for injury risk assessment. *Journal of biomechanics* 46, 7 (2013), 1237–1241. pages #3

[11] B Bideau, R Kulpa, N Vignais, S Brault, F Multon, and C Craig. 2010. Virtual reality, a serious game for understanding performance and training players in sport. *IEEE Computer Graphic Applications* 30, 2 (2010), 14–21. pages #3

[12] Jonathan R Blackburn and Matthew C Morrissey. 1998. The relationship between open and closed kinetic chain strength of the lower limb and jumping performance. *Journal of Orthopaedic & Sports Physical Therapy* 27, 6 (1998), 430–435. pages #4

[13] James P Bliss, Philip D Tidwell, and Michael A Guest. 1997. The effectiveness of virtual reality for administering spatial navigation training to firefighters. *Presence: Teleoperators & Virtual Environments* 6, 1 (1997), 73–86. pages #3
[14] Maarten F Bobbert. 1990. Drop jumping as a training method for jumping ability. *Sports medicine* 9, 1 (1990), 7–22. pages #3, #4
[15] Felix Born, Sophie Abramowski, and Maic Masuch. 2019. Exergaming in VR: The Impact of Immersive Embodiment on Motivation, Performance, and Perceived Exertion. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, New York City, NY, USA, 1–8. https://doi.org/10.1109/VS-Games.2019.8864579 pages #2, #4
[16] Felix Born, Adrian Rygula, and Maic Masuch. 2021. Motivating Players to Perform an Optional Strenuous Activity in a Virtual Reality Exergame Using Virtual Performance Augmentation. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 225 (oct 2021), 21 pages. https://doi.org/10.1145/3474652 pages #4
[17] C. Bosco, P. V. Komi, J. Tihanyi, G. Fekete, and P. Apor. 1983. Mechanical power test and fiber composition of human leg extensor muscles. *European Journal of Applied Physiology and Occupational Physiology* 51, 1 (1983), 129–135. https://doi.org/10.1007/BF00952545 pages #2, #3
[18] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806 pages #5, #11
[19] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47. https://doi.org/10.1002/capr.12360 pages #5, #11, #14
[20] Eva Buchtelová, M Tichy, and Kateřina Vaniková. 2013. Influence of muscular imbalances on pelvic position and lumbar lordosis: a theoretical basis. *Journal of Nursing, Social Studies, Public Health and Rehabilitation* 1, 2 (2013), 25–36. pages #2
[21] Lee N Burkett, Wayne T Phillips, and Joana Ziuraitis. 2005. The best warm-up for the vertical jump in college-age athletic men. *The Journal of Strength & Conditioning Research* 19, 3 (2005), 673–676. pages #3
[22] Paul K Canavan and Jason D Vescovi. 2004. Evaluation of power prediction equations: peak vertical jumping power in women. *Medicine & Science in Sports & Exercise* 36, 9 (2004), 1589–1593. pages #3
[23] Alberto Cannavò, Filippo Gabriele Pratticò, Giuseppe Ministeri, and Fabrizio Lamberti. 2018. A Movement Analysis System Based on Immersive Virtual Reality and Wearable Technology for Sport Training. In *Proceedings of the 4th International Conference on Virtual Reality* (Hong Kong, Hong Kong) *(ICVR 2018)*. Association for Computing Machinery, New York, NY, USA, 26–31. https://doi.org/10.1145/3198910.3198917 pages #3
[24] Kevin Carlson, Marshall Magnusen, and Peter Walters. 2009. Effect of various training modalities on vertical jump. *Research in Sports Medicine* 17, 2 (2009), 84–94. pages #4
[25] Patrick Carlson, Anicia Peters, Stephen B Gilbert, Judy M Vance, and Andy Luse. 2015. Virtual training: Learning transfer of assembly tasks. *IEEE transactions on visualization and computer graphics* 21, 6 (2015), 770–782. pages #3
[26] Polona Caserman, Shule Liu, and Stefan Göbel. 2022. Full-Body Motion Recognition in Immersive- Virtual-Reality-Based Exergame. *IEEE Transactions on Games* 14, 2 (2022), 243–252. https://doi.org/10.1109/TG.2021.3064749 pages #4
[27] Guilherme M. Cesar, Curtis L. Tomasevicz, and Judith M. Burnfield. 2016. Frontal plane comparison between drop jump and vertical jump: implications for the assessment of ACL risk of injury. *Sports Biomechanics* 15, 4 (Oct. 2016), 440–449. https://doi.org/10.1080/14763141.2016.1174286 pages #3
[28] John S. Y. Chan, Guanmin Liu, Danxia Liang, Kanfeng Deng, Jiamin Wu, and Jin H. Yan. 2019. Special Issue – Therapeutic Benefits of Physical Activity for Mood: A Systematic Review on the Effects of Exercise Intensity, Duration, and Modality. *The Journal of Psychology* 153, 1 (2019), 102–125. https://doi.org/10.1080/00223980.2018.1470487 arXiv:https://doi.org/10.1080/00223980.2018.1470487 PMID: 30321106. pages #12
[29] A Chardenon, Gilles Montagne, MJ Buekers, and Michel Laurent. 2002. The visual control of ball interception during human locomotion. *Neuroscience letters* 334, 1 (2002), 13–16. pages #3
[30] Gary Charness and Uri Gneezy. 2009. Incentives to exercise. *Econometrica* 77, 3 (2009), 909–931. pages #2
[31] Yu-Ping Chen, Lin-Ju Kang, Tien-Yow Chuang, Ji-Liang Doong, Shwn-Jan Lee, Mei-Wun Tsai, Suh-Fang Jeng, and Wen-Hsu Sung. 2007. Use of virtual reality to improve upper-extremity control in children with cerebral palsy: a single-subject design. *Physical therapy* 87, 11 (2007), 1441–1457. pages #3
[32] Sebastian Cmentowski and Jens Krueger. 2021. Exploring the Potential of Vertical Jump Training in Virtual Reality. In *Extended Abstracts of the 2021 Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Austria) *(CHI PLAY '21)*. Association for Computing Machinery, New York, NY, USA, 179–185. https://doi.org/10.1145/3450337.3483503 pages #4
[33] Jacob Cohen. 1992. Quantitative methods in psychology: A power primer. *Psychological Bulletin* 112, 1 (1992), 153–159. https://doi.org/10.1.1.1043.9095 pages #10
[34] Cathy Craig. 2013. Understanding perception and action in sport: how can virtual reality technology help? *Sports Technology* 6, 4 (2013), 161–169. pages #3
[35] Cathy M Craig, Eric Berton, Guillaume Rao, Laure Fernandez, and Reinoud J Bootsma. 2006. Judging where a ball will go: the case of curved free kicks in football. *Naturwissenschaften* 93, 2 (2006), 97–101. pages #3
[36] Jean-Louis Croisier. 2004. Muscular imbalance and acute lower extremity muscle injuries in sport. *International SportMed Journal* 5, 3 (2004), 169–176. pages #2
[37] Don Decker and Mike Vinson. 1996. Using periodization to improve vertical jump performance. *Strength & Conditioning Journal* 18, 1 (1996), 13–17. pages #2, #3
[38] DeepL SE. 2017. DeepL Translator. DeepL Pro version. Cologne, Germany. https://www.deepl.com/pro?cta=header-pro/. pages #5, #11
[39] PAUL Devita and William A Skelly. 1992. Effect of landing stiffness on joint kinetics and energetics in the lower extremity. *Med Sci Sports Exerc* 24, 1 (1992), 108–115. pages #3
[40] Dovetail Research PTY. LTD. 2016. Dovetail. Surry Hills, Australia. https://dovetailapp.com/. pages #4, #11
[41] Peter Düking, Hans-Christer Holmberg, and Billy Sperlich. 2018. The potential usefulness of virtual reality systems for athletes: a short SWOT analysis. *Frontiers in Physiology* 9 (2018), 128. pages #3
[42] Daniel L Eaves, Gavin Breslin, Paul Van Schaik, Emma Robinson, and Iain R Spears. 2011. The short-term effects of real-time virtual reality feedback on motor learning in dance. *Presence: Teleoperators and Virtual Environments* 20, 1 (2011), 62–77. pages #3
[43] Kirk I Erickson, Michelle W Voss, Ruchika Shaurya Prakash, Chandramallika Basak, Amanda Szabo, Laura Chaddock, Jennifer S Kim, Susie Heo, Heloisa Alves, Siobhan M White, et al. 2011. Exercise training increases size of hippocampus and improves memory. *Proceedings of the national academy of sciences* 108, 7 (2011), 3017–3022. pages #2
[44] Philip W Fink, Patrick S Foo, and William H Warren. 2009. Catching fly balls in virtual reality: A critical test of the outfielder problem. *Journal of vision* 9, 13 (2009), 14–14. pages #3
[45] Samantha Finkelstein, Andrea Nickel, Tiffany Barnes, and Evan A. Suma. 2010. Astrojumper: Motivating Children with Autism to Exercise Using a VR Game. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI EA '10)*. Association for Computing Machinery, New York, NY, USA, 4189–4194. https://doi.org/10.1145/1753846.1754124 pages #4, #12
[46] Samantha Finkelstein, Andrea Nickel, Zachary Lipps, Tiffany Barnes, Zachary Wartell, and Evan A Suma. 2011. Astrojumper: Motivating exercise with an immersive virtual reality exergame. *Presence: Teleoperators and Virtual Environments* 20, 1 (2011), 78–92. https://doi.org/10.1162/pres_a_00036 pages #2
[47] FitXR. 2019. *FitXR*. Game [VR]. FitXR, London, United Kingdom.. pages #2
[48] Steven J Fleck and William Kraemer. 2014. *Designing resistance training programs, 4E*. Human Kinetics, Champaign, Illinois. pages #4
[49] Jenny Gabel, Sukran Karaosmanoglu, Celeste Mason, Sebastian Rings, and Frank Steinicke. 2021. Corona Beat - Kicking The Sedentary Habit Induced By Prolonged Social Distancing. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, Lisbon, Portugal, 687–688. https://doi.org/10.1109/VRW52623.2021.00226 pages #2
[50] David A Gabriel, Gary Kamen, and Gail Frost. 2006. Neural adaptations to resistive exercise. *Sports medicine* 36, 2 (2006), 133–149. pages #4
[51] Mat Gallagher. 2022. 3 reasons why you should buy a VR headset. https://www.t3.com/news/3-reasons-why-you-should-buy-a-vr-headset pages #2
[52] Beat Games. 2019. *Beat Saber*. Game [VR]. Beat Games. Los Angeles, CA, USA. https://beatsaber.com/. Steam Reviews: https://store.steampowered.com/app/620980/Beat_Saber/. Accessed: 30.12.2021. pages #2, #4, #7
[53] Rodrigo Ghedini Gheller, Juliano Dal Pupo, Luis Antonio Pereira de Lima, Bruno Monteiro de Moura, and Saray Giovana dos Santos. 2014. Effect of squat depth on performance and biomechanical parameters of countermovement vertical jump. pages #4
[54] Linda Graf, Maximilian Altmeyer, Katharina Emmerich, Marc Herrlich, Andrey Krekhov, and Katta Spiel. 2022. Development and Validation of a German Version of the Player Experience Inventory (PXI). In *Proceedings of Mensch Und Computer 2022* (Darmstadt, Germany) *(MuC '22)*. Association for Computing Machinery, New York, NY, USA, 265–275. https://doi.org/10.1145/3543758.3543763 pages #9
[55] John Graham. 1994. Guidelines for providing valid testing of athletes' fitness levels. *Strength & Conditioning Journal* 16, 6 (1994), 7–14. pages #2, #3
[56] Daigo Hayashi, Kazuyuki Fujita, Kazuki Takashima, Robert W. Lindeman, and Yoshifumi Kitamura. 2019. Redirected Jumping: Imperceptibly Manipulating Jump Motions in Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, New York City, NY, USA, 386–394. https://doi.org/10.1109/VR.2019.8797989 pages #4
[57] Robert T Hays, John W Jacobs, Carolyn Prince, and Eduardo Salas. 1992. Flight simulator training effectiveness: A meta-analysis. *Military psychology* 4, 2 (1992), 63–74. pages #3

[58] Eric J Hegedus, Suzanne McDonough, Chris Bleakley, Chad E Cook, and G David Baxter. 2015. Clinician-friendly lower extremity physical performance measures in athletes: a systematic review of measurement properties and correlation with injury, part 1. The tests for knee function including the hop tests. *British journal of sports medicine* 49, 10 (2015), 642–648. pages #3

[59] Christian Helland, Jens Bojsen-Møller, Truls Raastad, Olivier R Seynnes, Marie M Moltubakk, Vidar Jakobsen, Håvard Visnes, and Roald Bahr. 2013. Mechanical properties of the patellar tendon in elite volleyball players with and without patellar tendinopathy. *British Journal of Sports Medicine* 47, 13 (Sept. 2013), 862–868. https://doi.org/10.1136/bjsports-2013-092275 pages #3

[60] Timothy E Hewett, Gregory D Myer, Kevin R Ford, Robert S Heidt Jr, Angelo J Colosimo, Scott G McLean, Antonie J Van den Bogert, Mark V Paterno, and Paul Succop. 2005. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate ligament injury risk in female athletes: a prospective study. *The American journal of sports medicine* 33, 4 (2005), 492–501. pages #3

[61] Anne Hösch. 2018. *Simulator Sickness in Fahrsimulationsumgebungen-drei Studien zu Human Factors.* Ph. D. Dissertation. Technische Universität Ilmenau. https://www.db-thueringen.de/receive/dbt_mods_00037865 pages #9

[62] Paul A Hough, Emma Z Ross, and Glyn Howatson. 2009. Effects of dynamic and static stretching on vertical jump performance and electromyographic activity. *The Journal of Strength & Conditioning Research* 23, 2 (2009), 507–512. pages #3

[63] HTC Corporation. 2020. HTC Vive Pro. Website. Retrieved October 26, 2020 from https://www.vive.com/eu/product/vive-pro/. pages #7, #9

[64] HTC Corporation. 2022. HTC Vive Tracker. Website. Retrieved July 12, 2022 from https://www.vive.com/us/accessory/tracker3/. pages #7, #9

[65] Jackie L Hudson. 1986. Coordination of segments in the vertical jump. *Medicine & Science in Sports & Exercise* 18, 2 (1986), 241–252. pages #4

[66] Joseph P Hunter and Robert N Marshall. 2002. Effects of power and flexibility training on vertical jump technique. *Medicine and science in sports and exercise* 34, 3 (2002), 478–486. pages #3

[67] Christos Ioannou, Patrick Archard, Eamonn O'Neill, and Christof Lutteroth. 2019. Virtual Performance Augmentation in an Immersive Jump & Run Exergame. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300388 pages #2, #4, #12

[68] Clayne R Jensen and A Garth Fisher. 1972. Scientific Basis of Athletic Conditioning: A Unique Book Dealing Exclusively with the Practical Application of Scientific Information to Athletic Conditioning. *Journal of Health, Physical Education, Recreation* 43, 8 (1972), 6–7. pages #3

[69] Jason Jerald. 2015. *The VR Book: Human-Centered Design for Virtual Reality.* Association for Computing Machinery and Morgan & Claypool, New York, New York, US. pages #3

[70] Raine Kajastila, Leo Holsti, and Perttu Hämäläinen. 2014. Empowering the Exercise: a Body-Controlled Trampoline Training Game. *International Journal of Computer Science in Sport* 1, 13 (2014), 6–23. pages #4

[71] Sukran Karaosmanoglu, Lucie Kruse, Sebastian Rings, and Frank Steinicke. 2022. Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 342, 7 pages. https://doi.org/10.1145/3491101.3519736 pages #4

[72] Sukran Karaosmanoglu, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. 2021. Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 252 (oct 2021), 27 pages. https://doi.org/10.1145/3474679 pages #4

[73] Philip Kelly, Aoife Healy, Kieran Moran, and Noel E. O'Connor. 2010. A Virtual Coaching Environment for Improving Golf Swing Technique. In *Proceedings of the 2010 ACM Workshop on Surreal Media and Virtual Cloning* (Firenze, Italy) *(SMVC '10)*. Association for Computing Machinery, New York, NY, USA, 51–56. https://doi.org/10.1145/1878083.1878098 pages #3

[74] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220. https://doi.org/10.1207/s15327108ijap0303_3 pages #9

[75] Ho-Jun Kim, Seokhee Chung, Sungsoo Kim, Hyundae Shin, Jongsoo Lee, Sehyun Kim, and Mi-Yeon Song. 2006. Influences of trunk muscles on lumbar lordosis and sacral angle. *European Spine Journal* 15, 4 (2006), 409–414. pages #2

[76] Jungjin Kim, Jaebum Son, Nayeon Ko, and BumChul Yoon. 2013. Unsupervised virtual reality-based exercise program improves hip muscle strength and balance control in older adults: a pilot study. *Archives of physical medicine and rehabilitation* 94, 5 (2013), 937–943. pages #3

[77] Joong Hwi Kim, Sung Ho Jang, Chung Sun Kim, Ji Hee Jung, and Joshua H You. 2009. Use of virtual reality to enhance balance and ambulation in chronic stroke: a double-blind, randomized controlled study. *American Journal of physical medicine & rehabilitation* 88, 9 (2009), 693–701. pages #3

[78] P Klavora et al. 2000. Vertical-jump tests: A critical review. *Strength and Conditioning Journal* 22, 5 (2000), 70–75. pages #3

[79] JJ Kozak, Peter A Hancock, EJ Arthur, and Susan T Chrysler. 1993. Transfer of training from virtual reality. *Ergonomics* 36, 7 (1993), 777–784. pages #3

[80] William J Kraemer, Scott A Mazzetti, Bradley C Nindl, Lincoln A Gotshalk, Jeff S Volek, Jill A Bush, JIM O MARX, KEI DOHI, Ana L Gomez, MARY MILES, et al. 2001. Effect of resistance training on women's strength/power and occupational performances. *Medicine & Science in Sports & Exercise* 33, 6 (2001), 1011–1025. pages #3

[81] Lucie Kruse, Sukran Karaosmanoglu, Sebastian Rings, Benedikt Ellinger, Daniel Apken, Thandiwe Feziwe Mangana, and Frank Steinicke. 2021. A Long-Term User Study of an Immersive Exergame for Older Adults with Mild Dementia during the COVID-19 Pandemic. In *ICAT-EGVE 2021 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, Jason Orlosky, Dirk Reiners, and Benjamin Weyers (Eds.). The Eurographics Association, Eindhoven, The Netherlands, 9–18. https://doi.org/10.2312/egve.20211322 pages #4

[82] H Kyröläinen, J Avela, JM McBride, S Koskinen, JL Andersen, S Sipilä, TES Takala, and PV Komi. 2004. Effects of power training on mechanical efficiency in jumping. *European journal of applied physiology* 91, 2 (2004), 155–159. pages #4

[83] Odders Lab. 2019. *OhShape*. Game [VR]. Odders Lab. Seville, Spain. https://ohshapevr.com/. Steam Reviews: https://store.steampowered.com/app/1098100/OhShape/. Accessed: 22.11.2022. pages #2

[84] Rotem Lammfromm and Daniel Gopher. 2011. Transfer of skill from a virtual reality trainer to real juggling. *BIO web of conferences* 1 (2011), 00054. https://doi.org/10.1051/bioconf/20110100054 pages #3

[85] Corinna E Lathan, Michael R Tracey, Marc M Sebrechts, Deborah M Clawson, and Gerald A Higgins. 2002. Using virtual environments as training simulators: Measuring transfer. In *Handbook of Virtual Environments*. CRC Press, Boca Raton, Florida, US, 443–454. pages #3

[86] Savvas Lazaridis, Eleni Bassa, Dimitrios Patikas, Giannis Giakas, Albert Gollhofer, and Christos Kotzamanidis. 2010. Neuromuscular differences between prepubescents boys and adult men during drop jump. *European journal of applied physiology* 110, 1 (2010), 67–74. pages #3

[87] Adrian Lees, Jos Vanrenterghem, and Dirk De Clercq. 2004. The maximal and submaximal vertical jump: implications for strength and conditioning. *The Journal of Strength & Conditioning Research* 18, 4 (2004), 787–791. pages #3, #4

[88] Lauri Lehtonen, Maximus D. Kaos, Raine Kajastila, Leo Holsti, Janne Karsisto, Sami Pekkola, Joni Vähämäki, Lassi Vapaakallio, and Perttu Hämäläinen. 2019. Movement Empowerment in a Multiplayer Mixed-Reality Trampoline Game. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 19–29. https://doi.org/10.1145/3311350.3347181 pages #4, #12

[89] Matthieu Lenoir, Eliane Musch, and Nancy La Grange. 1999. Ecological relevance of stereopsis in one-handed ball-catching. *Perceptual and motor skills* 89, 2 (1999), 495–508. pages #3

[90] Yi-Jun Li, De-Rong Jin, Miao Wang, Jun-Long Chen, Frank Steinicke, Shi-Min Hu, and Qinping Zhao. 2021. Detection Thresholds with Joint Horizontal and Vertical Gains in Redirected Jumping. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, New York City, NY, USA, 95–102. https://doi.org/10.1109/VR50410.2021.00030 pages #4

[91] Phil Lundin and William Berg. 1991. Plyometrics: a review of plyometric training. *Strength & Conditioning Journal* 13, 6 (1991), 22–34. pages #4

[92] Andrew D Lyttle, Greg J Wilson, and Karl J Ostrowski. 1996. Enhancing performance: Maximal power versus combined weights and plyometrics training. *Journal of strength and conditioning research* 10 (1996), 173–179. pages #4

[93] Andrew Macvean and Judy Robertson. 2013. Understanding Exergame Users' Physical Activity, Motivation and Behavior over Time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1251–1260. https://doi.org/10.1145/2470654.2466163 pages #2

[94] Laurent Malisoux, Marc Francaux, Henri Nielens, and Daniel Theisen. 2006. Stretch-shortening cycle exercises: an effective training paradigm to enhance power output of human single muscle fibers. *Journal of applied physiology* 100, 3 (2006), 771–779. pages #4

[95] Fabrizia Mantovani, Gianluca Castelnuovo, Andrea Gaggioli, and Giuseppe Riva. 2003. Virtual reality training for health-care professionals. *CyberPsychology & Behavior* 6, 4 (2003), 389–395. pages #3

[96] Elena Márquez Segura, Annika Waern, Jin Moen, and Carolina Johansson. 2013. The Design Space of Body Games: Technological, Physical, and Social Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 3365–3374. https://doi.org/10.1145/2470654.2466461 pages #2, #4

[97] Anna Lisa Martin-Niedecken, Andrea Mahrer, Katja Rogers, Eling D. de Bruin, and Alexandra Schättin. 2020. "HIIT" the ExerCube: Comparing the Effectiveness of Functional High-Intensity Interval Training in Conventional vs. Exergame-Based Training. *Frontiers in Computer Science* 2 (2020), 15 pages. https://doi.

org/10.3389/fcomp.2020.00033 pages #2

[98] Anna Lisa Martin-Niedecken, Katja Rogers, Laia Turmo Vidal, Elisa D. Mekler, and Elena Márquez Segura. 2019. ExerCube vs. Personal Trainer: Evaluating a Holistic, Immersive, and Adaptive Fitness Game Setup. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300318 pages #4

[99] Arnaud Mas, Idriss Ismaël, and Nicolas Filliard. 2018. Indy: a virtual reality multiplayer game for navigation skills training. In *2018 IEEE Fourth VR International Workshop on Collaborative Virtual Environments (3DCVE)*. IEEE, IEEE, New York City, NY, USA, 1–4. pages #3

[100] Daniel Travis McMaster, Nicholas Gill, John Cronin, and Michael McGuigan. 2014. A Brief Review of Strength and Ballistic Assessment Methodologies in Sport. *Sports Medicine* 44, 5 (May 2014), 603–623. https://doi.org/10.1007/s40279-014-0145-2 pages #2, #3

[101] Peter J McNair, Harry Prapavessis, and Karen Callender. 2000. Decreasing landing forces: effect of instruction. *British journal of sports medicine* 34, 4 (2000), 293–296. pages #3

[102] Stefan C Michalski, Ancret Szpak, and Tobias Loetscher. 2019. Using virtual environments to improve real-world motor skills in sports: A systematic review. *Frontiers in psychology* 10 (2019), 2159. pages #3

[103] Stefan Carlo Michalski, Ancret Szpak, Dimitrios Saredakis, Tyler James Ross, Mark Billinghurst, and Tobias Loetscher. 2019. Getting your game on: Using virtual reality to improve real table tennis skills. *PloS one* 14, 9 (2019), e0222351. pages #3

[104] Microsoft Corporation. 2014. Kinect 2 for Windows. Website. Discontinued Product. pages #7

[105] Anat Mirelman, Benjamin L Patritti, Paolo Bonato, and Judith E Deutsch. 2010. Effects of virtual reality training on gait biomechanics of individuals post-stroke. *Gait & posture* 31, 4 (2010), 433–437. pages #3

[106] Florian Mueller and Katherine Isbister. 2014. Movement-Based Game Guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2191–2200. https://doi.org/10.1145/2556288.2557163 pages #2, #4, #7, #13

[107] Florian Mueller, Rohit Ashok Khot, Kathrin Gerling, and Regan Mandryk. 2016. Exertion Games. *Foundations and Trends® in Human–Computer Interaction* 10, 1 (2016), 1–86. https://doi.org/10.1561/1100000041 pages #2, #4, #14

[108] Florian 'Floyd' Mueller, Darren Edge, Frank Vetere, Martin R. Gibbs, Stefan Agamanolis, Bert Bongers, and Jennifer G. Sheridan. 2011. Designing Sports: A Framework for Exertion Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2651–2660. https://doi.org/10.1145/1978942.1979330 pages #2

[109] Sean P Mullen, Erin A Olson, Siobhan M Phillips, Amanda N Szabo, Thomas R Wójcicki, Emily L Mailey, Neha P Gothe, Jason T Fanning, Arthur F Kramer, and Edward McAuley. 2011. Measuring enjoyment of physical activity in older adults: invariance of the physical activity enjoyment scale (paces) across groups and time. *International Journal of Behavioral Nutrition and Physical Activity* 8, 1 (2011), 1–9. https://doi.org/10.1186/1479-5868-8-103 pages #9

[110] David L Neumann, Robyn L Moffitt, Patrick R Thomas, Kylie Loveday, David P Watling, Chantal L Lombard, Simona Antonova, and Michael A Tremeer. 2018. A systematic review of the application of interactive virtual reality to sport. *Virtual Reality* 22, 3 (2018), 183–198. pages #3

[111] Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand, and Jon Howard Raphael. 2012. No room for reflexivity? Critical reflections following a systematic review of qualitative research. *Journal of Health Psychology* 17, 6 (2012), 866–885. https://doi.org/10.1177/1359105311427615 pages #14

[112] Niantic, Nintendo, and The Pokémon Company. 2016. *Pokémon Go*. Game [AR]. https://www.pokemongo.com/en-us/. pages #6

[113] Niels C Nilsson, Stefania Serafin, and Rolf Nordahl. 2013. Unintended positional drift and its potential solutions. In *2013 IEEE Virtual Reality (VR)*. IEEE, IEEE, New York City, NY, USA, 121–122. pages #13

[114] Rui Nouchi, Yasuyuki Taki, Hikaru Takeuchi, Atsushi Sekiguchi, Hiroshi Hashizume, Takayuki Nozawa, Haruka Nouchi, and Ryuta Kawashima. 2014. Four weeks of combination exercise training improved executive functions, episodic memory, and processing speed in healthy elderly people: evidence from a randomized controlled trial. *Age* 36, 2 (2014), 787–799. pages #2

[115] Anderson Souza Oliveira, Priscila Brito Silva, Morten Enemark Lund, Dario Farina, and Uwe Gustav Kersting. 2017. Balance training enhances motor coordination during a perturbed sidestep cutting task. *journal of orthopaedic & sports physical therapy* 47, 11 (2017), 853–862. pages #2

[116] Jorge Perez-Gomez and JA Calbet. 2013. Training methods to improve vertical jump performance. *The Journal of sports medicine and physical fitness* 53, 4 (2013), 339–357. pages #3, #4

[117] Dana P Piasecki, Kurt P Spindler, Todd A Warren, Jack T Andrish, and Richard D Parker. 2003. Intraarticular injuries associated with anterior cruciate ligament tear: findings at ligament reconstruction in high school and recreational athletes: an analysis of sex-based differences. *The American journal of sports medicine* 31, 4 (2003), 601–605. pages #3

[118] Jeffrey A Potteiger, Robert H Lockwood, Mark D Haub, Brett A Dolezal, Khalid S Almuzaini, Jan M Schroeder, and Carole J Zebas. 1999. Muscle power and fiber characteristics following 8 weeks of plyometric training. *Journal of strength and conditioning research* 13 (1999), 275–279. pages #3

[119] Michael E Powers. 1996. Vertical jump training for volleyball. *Strength and Conditioning* 18 (1996), 18–23. pages #2

[120] Harry Prapavessis and Peter J McNair. 1999. Effects of instruction in jumping technique and experience jumping on ground reaction forces. *Journal of orthopaedic & sports physical therapy* 29, 6 (1999), 352–356. pages #3

[121] Judy Robertson, Andrew Macvean, Samantha Fawkner, Graham Baker, and Ruth G. Jepson. 2018. Savouring our mistakes: Learning from the FitQuest project. *International Journal of Child-Computer Interaction* 16 (2018), 55–67. https://doi.org/10.1016/j.ijcci.2017.12.003 pages #2

[122] Sharon Stansfield, Daniel Shawver, Annette Sobel, Monica Prasad, and Lydia Tapia. 2000. Design and implementation of a virtual reality system and its application to training medical first responders. *Presence: Teleoperators & Virtual Environments* 9, 6 (2000), 524–556. pages #3

[123] Dejan Stošić, Slavoljub Uzunović, Saša Pantelić, Saša Veličković, Marko Đurović, and Danica Piršl. 2019. Effects of exercise program on coordination and explosive power in university dance students. *Facta Universitatis. Series: Physical Education and Sport* 17, 3 (2019), 579–589. pages #2

[124] Yusaku Sugiura, Tomoyuki Saito, Keishoku Sakuraba, Kazuhiko Sakuma, and Eiichi Suzuki. 2008. Strength deficits identified with concentric action of the hip extensors and eccentric action of the hamstrings predispose to hamstring injury in elite sprinters. *Journal of orthopaedic & sports physical therapy* 38, 8 (2008), 457–464. pages #2

[125] Ancret Szpak, Stefan Carlo Michalski, and Tobias Loetscher. 2020. Exergaming With Beat Saber: An Investigation of Virtual Reality Aftereffects. *J Med Internet Res* 22, 10 (23 Oct 2020), e19840. https://doi.org/10.2196/19840 pages #2, #4, #12

[126] Kazumoto Tanaka. 2017. 3D action reconstruction using virtual player to assist karate training. In *2017 IEEE Virtual Reality (VR)*. IEEE, IEEE, New York City, NY, USA, 395–396. pages #3

[127] Judith Tirp, Christina Steingröver, Nick Wattie, Joseph Baker, and Jörg Schorer. 2015. Virtual realities as optimal learning environments in sport-A transfer study of virtual and real dart throwing. *Psychological Test and Assessment Modeling* 57, 1 (2015), 57. pages #3

[128] Emanuel Todorov, Reza Shadmehr, and Emilio Bizzi. 1997. Augmented feedback presented in a virtual environment accelerates learning of a difficult motor task. *Journal of motor behavior* 29, 2 (1997), 147–158. pages #3

[129] Unity Technologies. 2020. Unity. Website. Retrieved October 26, 2020 from https://unity.com/. pages #6

[130] Stephen Uzor and Lynne Baillie. 2014. Investigating the Long-Term Use of Exergames in the Home with Elderly Fallers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2813–2822. https://doi.org/10.1145/2556288.2557160 pages #2

[131] Etienne van Wyk and Ruth de Villiers. 2009. Virtual Reality Training Applications for the Mining Industry. In *Proceedings of the 6th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa* (Pretoria, South Africa) *(AFRIGRAPH '09)*. Association for Computing Machinery, New York, NY, USA, 53–63. https://doi.org/10.1145/1503454.1503465 pages #3

[132] Athanasios Vanezis and Adrian Lees. 2005. A biomechanical analysis of good and poor performers of the vertical jump. *Ergonomics* 48, 11-14 (2005), 1594–1603. pages #3

[133] WanadevStudio. 2020. *Ragnaröck*. Game [VR]. WanadevStudio. Lyon, France. https://www.ragnarock-vr.com/home/. pages #7

[134] Peter Wilhelm and Dominik Schoebi. 2007. Assessing mood in daily life: Structural validity, sensitivity to change, and reliability of a short-scale to measure three basic dimensions of mood. *European Journal of Psychological Assessment* 23, 4 (2007), 258. https://doi.org/10.1027/1015-5759.23.4.258 pages #9

[135] Greg J Wilson and Aron J Murphy. 1996. The use of isometric tests of muscular function in athletic assessment. *Sports medicine* 22, 1 (1996), 19–37. pages #4

[136] Greg J Wilson, Aron J Murphy, and Anthony Giorgi. 1996. Weight and plyometric training: effects on eccentric and concentric force production. *Canadian Journal of Applied Physiology* 21, 4 (1996), 301–315. pages #4

[137] Dennis Wolf, Katja Rogers, Christoph Kunder, and Enrico Rukzio. 2020. *JumpVR: Jump-Based Locomotion Augmentation for Virtual Reality*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376243 pages #4

[138] World Health Organization. 2020. Physical activity. Website. Retrieved September 5, 2022 from https://www.who.int/news-room/fact-sheets/detail/physical-activity. pages #2

[139] XRWorkout Inc. 2021. *VRWorkout*. Game [VR]. Michael Gschwandtner, Wien, Austria.. pages #2

[140] Wenge Xu, Hai-Ning Liang, Qiuyu He, Xiang Li, Kangyou Yu, and Yuzheng Chen. 2020. Results and Guidelines From a Repeated-Measures Design Experiment Comparing Standing and Seated Full-Body Gesture-Based Immersive Virtual Reality Exergames: Within-Subjects Evaluation. *JMIR Serious Games* 8, 3 (27 Jul 2020), e17972. https://doi.org/10.2196/17972 pages #13, #14

[141] Yei-Yu Yeh and Louis D Silverstein. 1992. Spatial judgments with monoscopic and stereoscopic presentation of perspective displays. *Human Factors* 34, 5 (1992), 583–600. pages #3

[142] Soojeong Yoo and Judy Kay. 2016. VRun: Running-in-Place Virtual Reality Exergame. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (Launceston, Tasmania, Australia) *(OzCHI '16)*. Association for Computing Machinery, New York, NY, USA, 562–566. https://doi.org/10.1145/3010915.3010987 pages #2

[143] Frank TJM Zaal and Reinoud J Bootsma. 2011. Virtual reality as a tool for the study of perception-action: The case of running to catch fly balls. *Presence* 20, 1 (2011), 93–103. pages #3

9.3 MOVE, REACT, REPEAT! THE ROLE OF CONTINUOUS CUES IN IMMERSIVE EXERGAMES

[CorePub3] **Sukran Karaosmanoglu**, Moritz Wegner, Sebastian Cmentowski, and Frank Steinicke. Move, react, repeat! the role of continuous cues in immersive exergames. *Proceedings of the ACM on Human-Computer Interaction*, 8 (CHI PLAY), 2024. DOI 10.1145/3677091. URL https://doi.org/10.1145/3677091.

Supplementary materials for this article are available here: https://doi.org/10.1145/3677091.

Move, React, Repeat! The Role of Continuous Cues in Immersive Exergames

SUKRAN KARAOSMANOGLU, Human-Computer Interaction, Universität Hamburg, Germany
MORITZ WEGNER, Human-Computer Interaction, Universität Hamburg, Germany
SEBASTIAN CMENTOWSKI, HCI Games Group, Stratford School of Interaction Design and Business, University of Waterloo, Canada
FRANK STEINICKE, Human-Computer Interaction, Universität Hamburg, Germany

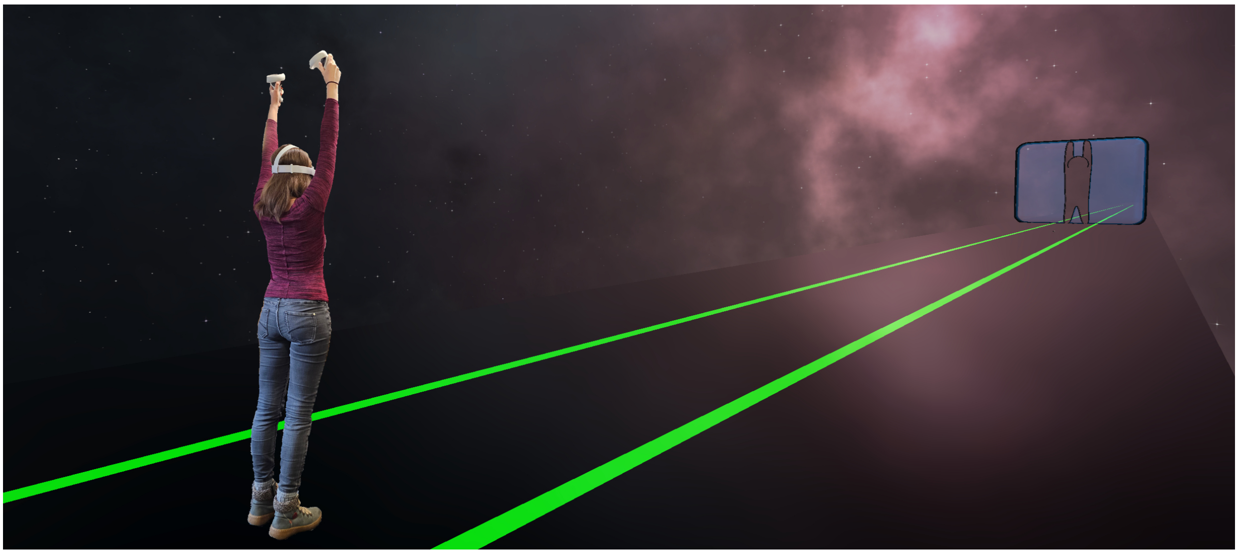

Fig. 1. Illustration of the gameplay with continuous cues: A player performs the featured pose on the approaching in-game wall and receives a visual cue (green lines) indicating that the execution was successful.

Immersive virtual reality (VR) exergames blend playful environments with physical activities. In such games, providing continuous cues, i.e., real-time indications of how the movements are being performed, can guide players and improve the execution of movements. However, research about how continuous cues impact player performance and experience in VR exergames remains sparse. To this end, we conducted a within-participants study (N=32) with four conditions: (i) no, (ii) audio, (iii) visual, and (iv) audiovisual cueing. The results show that both unimodal and bimodal cueing improve the player experience without causing cybersickness, while bimodal cueing improves pose accuracy. We provide three main design implications for VR exergames: (i) considering the granularity of cues, (ii) exploring different continuous cue designs, and (iii) underlining the use of continuous cues in other immersive experiences.

CCS Concepts: • **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**; • **Software and its engineering** → **Interactive games**.

Authors' Contact Information: Sukran Karaosmanoglu, Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, sukran.karaosmanoglu@uni-hamburg.de; Moritz Wegner, Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, moritz.wegner@studium.uni-hamburg.de; Sebastian Cmentowski, HCI Games Group, Stratford School of Interaction Design and Business, University of Waterloo, Waterloo, Canada, sebastian.cmentowski@uwaterloo.ca; Frank Steinicke, Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, frank.steinicke@uni-hamburg.de.

Additional Key Words and Phrases: virtual reality, exergames, feedback, cueing, audio, visual

ACM Reference Format:
Sukran Karaosmanoglu, Moritz Wegner, Sebastian Cmentowski, and Frank Steinicke. 2024. Move, React, Repeat! The Role of Continuous Cues in Immersive Exergames. *Proc. ACM Hum.-Comput. Interact.* 8, CHI PLAY, Article 326 (October 2024), 23 pages. https://doi.org/10.1145/3677091

1 Introduction

Physical activity improves well-being and greatly reduces the risk of severe common illnesses like dementia or cancer [71]. Despite these benefits, a report by the World Health Organization (WHO) [70] reveals that many people do not meet the recommendations for daily physical activity. However, sedentary behavior has profound effects on many aspects of our lives: it not only harms people's health but also has a significant impact on the economies and health systems of countries. This report suggests that insufficient physical activity is one of the most urgent problems in our society. To motivate people to exercise regularly, we need novel measures that provide "*affordable, enjoyable, and inclusive opportunities for all people to be active where they live, work and play*" [70].

Exergames incorporate physical effort as a fundamental aspect of gameplay and require players to play the games by moving [53]. These games are one way to help combat the problem of sedentary behavior. They offer a fun experience and motivate players to engage in physical activity. VR exergames, in particular, have unique features to motivate users to exercise [8]: they are immersive and captivating and typically build on spatial interactions using the VR system's six degrees of freedom (DOF) tracking capability [29]. Moreover, VR exergames are among the most played VR games in the commercial market (e.g., Beat Saber [19], OhShape [40]) and reach diverse user groups [12, 30, 31, 38]. Since these games aim to not only fulfill users' need for enjoyment but also deliver exerting workouts, they should incorporate informative feedback that motivates players to perform the movements optimally.

Delivering feedback (i.e., a response to one's actions [48]) is central to every interactive system used by humans and comes in various forms, serving a multitude of use cases. Our research focused on a specific form of feedback, which we refer to as *cues*. While the terms "feedback" and "cues" are sometimes used interchangeably in research [21, 41] and even referred to as "feedback cues" [47], we follow the common definitions [55, 58] and understand cues mainly as a type of feedback that provides players with low-complex and easily digestible information to initiate or continue to perform a movement. The use of cues for physical well-being is shown as a promising approach [55, 58]. A major use case of such cues is to guide users on how they need to adapt their actions. Typically, users receive cues based on their performance during individual tasks, giving them insights about the action they are performing. In exergames, providing immediate cues can potentially make players aware of their movement quality and thereby help them improve their exercise performance and avoid harmful situations. Accordingly, Mueller and Isbister [52] recommend providing moment-to-moment feedback on the players' movements but, at the same time, strongly emphasize the need to avoid giving excessive, overwhelmingly too much feedback. This highlights that balancing feedback and player experience is crucial but also challenging. To tackle this issue, Mueller and Isbister [52] suggest designers to allow players to self-reflect on their movements and learn from the provided feedback independently.

Research explored if and in which form exergames should provide feedback to users [60] (the authors did not consider visual feedback) and how different biofeedback visualizations benefit users [35]. Nevertheless, exploration of *continuous cues*, which we define as real-time indications of how movements are being performed, is limited. However, we believe continuous cues possess great potential for VR exergames by providing players with informative feedback. Delivering this information throughout the exercise in a continuous form (as opposed to a binary success /

failure experience) allows players to monitor and adapt their performance constantly (meeting the guidelines of Mueller and Isbister [52]). To our knowledge, there is no research on how to design these cues, which modality to use, and whether they are perceived as useful or distracting. Therefore, this paper answers two essential questions in VR exergame research:

RQ1: How does having continuous cues affect player experience and performance in VR exergames?

RQ2: How do different continuous cue modalities affect player experience and performance in VR exergames?

By answering these questions, we aim to understand how to improve the players' game experience as well as their performance. In our study, we focus on simple auditory and visual cues because those elements can be easily integrated into any VR exergame, and good audiovisual appeal is essential to the player experience [1]. As cues, we manipulated the color of an in-game visual element and the frequency of a repeatedly played sound. Both cues can be continuously changed in a controlled manner (similar to [3]). For our investigation, we implemented a VR exergame in which users performed a series of yoga-inspired static poses (similar to other games [8, 40], see Figure 1). In our game, the designed cues were presented to players during their movement execution (not after), guiding them to perform the pose of the currently approaching obstacle. Depending on the condition, players experienced no cues (*NO-CUE*), an audio cue (*AU-CUE*), a visual cue (*VI-CUE*), or a combination of audio and visual cues (*AV-CUE*) reflecting their moment-to-moment performance. We conducted a within-participants study (N=32) and assessed how continuous cues influenced player experience and performance. Our results indicate that continuous cues, regardless of modality, enhance player experience without causing cybersickness, and the combination of audio and visual cues improves pose accuracy. We discuss the potential of using continuous cueing and provide three main implications for future work: considering the granularity of cues, exploring different continuous cue designs, and underlining generalizability and relevance of continuous cues to other immersive experiences.

Overall, we summarize our contributions as follows:

- design and implementation of continuous cues in VR exergames,
- a first exploration of if and how different continuous cue modalities impact player experience and performance in VR exergames,
- design implications for VR exergame designers, developers, and researchers.

2 Related Work

This section summarizes prior research on feedback and cues in games, the benefits of VR exergames, and the limited exploration of continuous cues in VR exergames.

2.1 Games & Feedback

Feedback is a crucial element that gives meaning to every gameplay [18, 66]. In simple terms, feedback can be defined as a response to one's actions [48]. Inspired by education research [2], this concept can be divided into two types: summative (or outcome feedback [27]) and formative (or elaborative feedback [27]). The former concept concerns the post-assessment of performance, while the latter is related to real-time assessment. These concepts are translated into game research as well [20, 27]. In games, summative feedback is provided after completing certain tasks or overall gameplay (e.g., badges or leaderboards). This type of feedback provides an overall concrete assessment of the players' performance. On the other hand, formative feedback is given during the game in real-time. Such feedback allows players to adjust and improve their performance. Notable examples of this type of feedback include verbal narration or cues. While both types of feedback are valuable, formative feedback is particularly interesting as it can continuously inform players and allow them to make adjustments instantly. Especially, cues are subject to game [28, 41] and physical health research [55, 58] and generally have shown much promise.

Visual cues are a frequently used feature in non-VR [15] and VR games [41], as well as in its tutorials [28]. These cues can serve numerous purposes (e.g., guiding players or highlighting special objects) [15]. To help understand how visual cues are used to provide information to users, Dillman et al. [15] analyzed the use of visual interaction cues (i.e., guiding players' attention to specific interactive elements of the virtual world) in 49 video games and created a framework. Whereas Dillman et al. [15] only focus on visual interaction cues rather than audiovisual feedback cues, we, similar to others [41], can still use this closest-fitting framework to contextualize our cues within the larger design space. Overall, this framework categorizes visual cues according to task (i.e., purpose), markedness (i.e., design of the cue), and trigger (i.e., criteria to display the cue). Accordingly, we see our designed cues to fall into the following subcategories: Since our cues aim to help players by guiding their movements, we consider our cues to partially belong to the "go" subcategory of the task dimension. Thereby, we extend the original definition, focusing mainly on path navigation in video games, to indicating movement patterns and paths in exergames. In terms of markedness, our cues fall into the "integrated" subcategory of the markedness dimension, as they are additional virtual objects that can only be seen by players but are not part of the game world. Finally, our cues aim to provide consistent information about the player's movement and are therefore always visible to players (i.e., the "persistent" subcategory of the trigger dimension).

2.2 Exergames & Feedback

Motivation is one of the key challenges when performing regular training [51, 56]. Exergames [53] hold great potential to overcome this challenge by combining engaging experiences with simultaneous physical activity. Prior work indicated that exergames support psychological and physical well-being [23, 25, 39] among other benefits [36]. In particular, VR exergames offer immersive experiences and natural interactions to players. Players dive into computer-generated worlds and perform 6-DOF movements. Research showed that VR exergames are more motivating and engaging than non-immersive versions [8] and that VR exergame training can be used as an effective method (e.g., to improve table tennis performance [43]).

Like in real-world physical activities, feedback is at the heart of exergames [14, 45, 52]. There are many forms and purposes of feedback [45, 63]. For example, Stach and Graham [63] explored haptic feedback in non-immersive exergames for various purposes, such as balancing gameplay between people with asymmetric fitness levels or providing non-harmful interaction with a game. In a narrative review, Lyons [45] suggested that feedback has the potential to make exergames enjoyable and provide a sense of empowerment. However, feedback in exergames should be designed in a way that does not impose a high workload on players [45, 52]. To keep the cognitive load low, Clark and Mayer [11] recommend following the coherence principle, which refers to avoiding additional material that could be counterproductive to an instructional goal. The interference with learning can happen in three ways: distraction (i.e., taking attention away from the actual task), disruption (i.e., blocking the learning process for the actual task), or seduction (i.e., priming unrelated existing information). Following this principle, we keep our cues simple and make them responsive to the instructional goal of our game. Adhering to the recommendations of Mueller and Isbister [52] and Lyons [45] on the timing of cues, we focus on featuring moment-to-moment feedback, providing players with continuous information on their movement quality without overloading them with too much feedback [52]. With this approach, we hope to support players' experience and performance, thereby empowering them [45].

2.3 Feedback in VR

Incorporating feedback in virtual environments is an ongoing area of interest for the VR research community [6, 32]. However, most previous studies focused on fundamental interactions in VR, such as object selection [3, 4] or manipulation [9].

Canales and Jörg [9] reported that users preferred audio feedback (i.e., playing a sound when the task is completed) over visual feedback (i.e., changing an object's opacity) in a VR grasping task. At the same time, the audio feedback led to a decreased grabbing performance. Another study [4] found that applying adaptive audio feedback (i.e., adaptive pitch) based on the users' error rate benefits performance in an "as fast as possible" target pointing task in VR. Hu et al. [24] investigated the influence of the purposes (i.e., travel, selection, and manipulation) and timing of VR interaction cues on learning outcomes. For each purpose, the authors used different representations (e.g., boots, icons, controller models) with a semi-transparent green color continuously interpolated between the current and target state of the interaction (i.e., continuous cues). Their results indicate that immediately-provided cues promote faster learning and retention than those given with a delay. Similarly, Ariza et al. [3] investigated the impact of using auditory (i.e., frequency), visual (i.e., color), and haptic (i.e., intensity) forms of proximity-based feedback on user performance in a VR target selection task. The authors found that users preferred bimodal feedback to unimodal feedback and that binary feedback yielded better performance than continuous feedback. However, we argue that binary (discrete) feedback is not well suited for applications requiring continuous exertion because physical movements typically cannot be rated as good or bad with a binary decision. While the use of discrete feedback (e.g., red color for mistakes or point-based rewards for completing a task correctly) can provide high-level information about players' performance, it provides only a limited opportunity for players to self-reflect on their actions and does not guide them to improve their execution during the game. We believe continuous feedback is important when guiding players during the gameplay to improve their movements (aligning with the recommendations of Mueller and Isbister [52]). Therefore, we use continuous cueing in our research and inform the design of our cues (e.g., visual cue = color change) using the principles of Ariza et al. [3].

2.4 Feedback in VR Exergames

Movement science experts highlight the unique advantage of VR exergames: in these games, users can receive real-time feedback while feeling immersed in 3D worlds [12]. The unique possibilities of VR (e.g., immersion, full control) allow designers, developers, and researchers to integrate feedback into VR exergames in various forms, designs, and purposes.

For example, feedback can be used as a motivational tool. Xu et al. [72] used non-player audiences who cheered for players and opponents based on their performance; they found that this kind of feedback improved player performance, experience, and exertion. Similarly, Haller et al. [22] reported that using virtual spectator encouragement (i.e., rhythmic clapping) in a VR exergame based on high-intensity interval training can improve players' cycling speed and physical exertion. In line with research, some commercial VR games, such as *FitXR* [16], feature virtual trainers that guide and encourage players to perform better. Another way to use feedback is to create realistic experiences in VR exergames. In *DeceptiBike*, the authors used haptic cues (i.e., fan-based air stream) to create a realistic speed deception experience in a VR biking exergame [44].

Most importantly, feedback can be used to indicate whether players are performing well in a game. Yin et al. [73] used an environmental visual cue to show players' performance; the VR environment became darker if players' cycling speed did not match the music tempo. But, the authors did not evaluate this cue's effect on the player experience and performance. Shaw et al. [60] evaluated various forms of feedback in a VR cycling exergame; audio feedback (e.g., sounds if players failed to avoid an obstacle) outperformed wind and resistance feedback in terms of motivation and immersion, but the multimodal version yielded better results than any unimodal feedback. In a VR rowing exergame [35], players found different visualizations of constant biofeedback (e.g., lung animation) more helpful than not receiving feedback. However, these visualizations also led to a reduction in the correct respiration rate. She et al. [61] presented the VR rowing game *HIITCopter*, which provided players with a visual cue in three different colors (blue, green, red) to indicate when players are overexerting. They found that the game was successful at inducing high-intensity interval training; however, they did not evaluate the particular effect of the visual cue. Unlike prior research, we focus not only on providing information on players' performance through our implemented cues but also on offering guidance to improve players' execution of movements.

3 Virtual Reality Exergame

To answer our research questions, we created a VR exergame in Unity (v.2021.3.22f1) [65] and used models created in Blender [57]. Following the recommendations of Karaosmanoglu et al. [29], we report our VR game based on five factors: goals, people, exercises, design, and technologies. Our game aims to promote physical activity and provide physical training for younger adults. We drew inspiration from commercial and research-based VR games to design our single-player game (see below for more details). In the game, players perform static full-body poses that can be performed in a standing position, shown on 3D cut-out walls that approach them. Overall, the game has a sci-fi theme and features obstacle-based tasks that require players to fit into moving planes (see Figure 1). The game is adapted to the players' arm's length and reach. To experience our game, the players were equipped with a standalone headset and controllers (i.e., Meta Quest 2 [50]).

3.1 Poses

Our game features static body poses that players need to perform. We decided on this theme because such mechanics are used both in commercial games (e.g., OhShape [40]) and in research [8, 74]. For example, Zhang et al. [74] presented an approach to generate levels based on players' movement goals and validated their approach using a pose-based exergame. Our chosen style of gameplay focuses more on movement accuracy and isometric strength and less on high-intensity endurance training. Since our research mainly focuses on player experience and movement guidance, we deem this type of game more suitable for measuring the effectiveness of continuous cues.

We used various sources [40, 46, 54, 68] to decide on the selected in-game poses and incorporated movement patterns from established physical practices such as Tai Chi and yoga. Considering the limitations of the Meta Quest 2 (e.g., no leg tracking), we decided to use only movements that are detectable using the headset and controllers' positions. Specifically, our game features the following poses, which are repeated 2–5 times in a fixed, balanced order in the game (Figure 2): crescent moon (x4), five-pointed star with raised hands (x4), goddess pose (x4), mountain pose (x5), overhead squat (x4), sinking pose (x2), squat (x4), and five-pointed star (x3).

3.2 Walls

To indicate which pose players have to perform, they are confronted with 3D cut-out walls approaching them at a constant speed (see Figure 2 and Figure 1). We decided on this constant speed based on playtesting; each wall took six seconds to reach the player position. To account for individual differences of players, each wall is designed for a standardized size and scaled individually for each player before the game. The arms of the body poses in the walls were implemented considering that a person's arm span is close to their height [26]; each arm has 50% of the measured height. For the head, we performed informal playtests and decided on 28% of the body height as a head diameter measure.

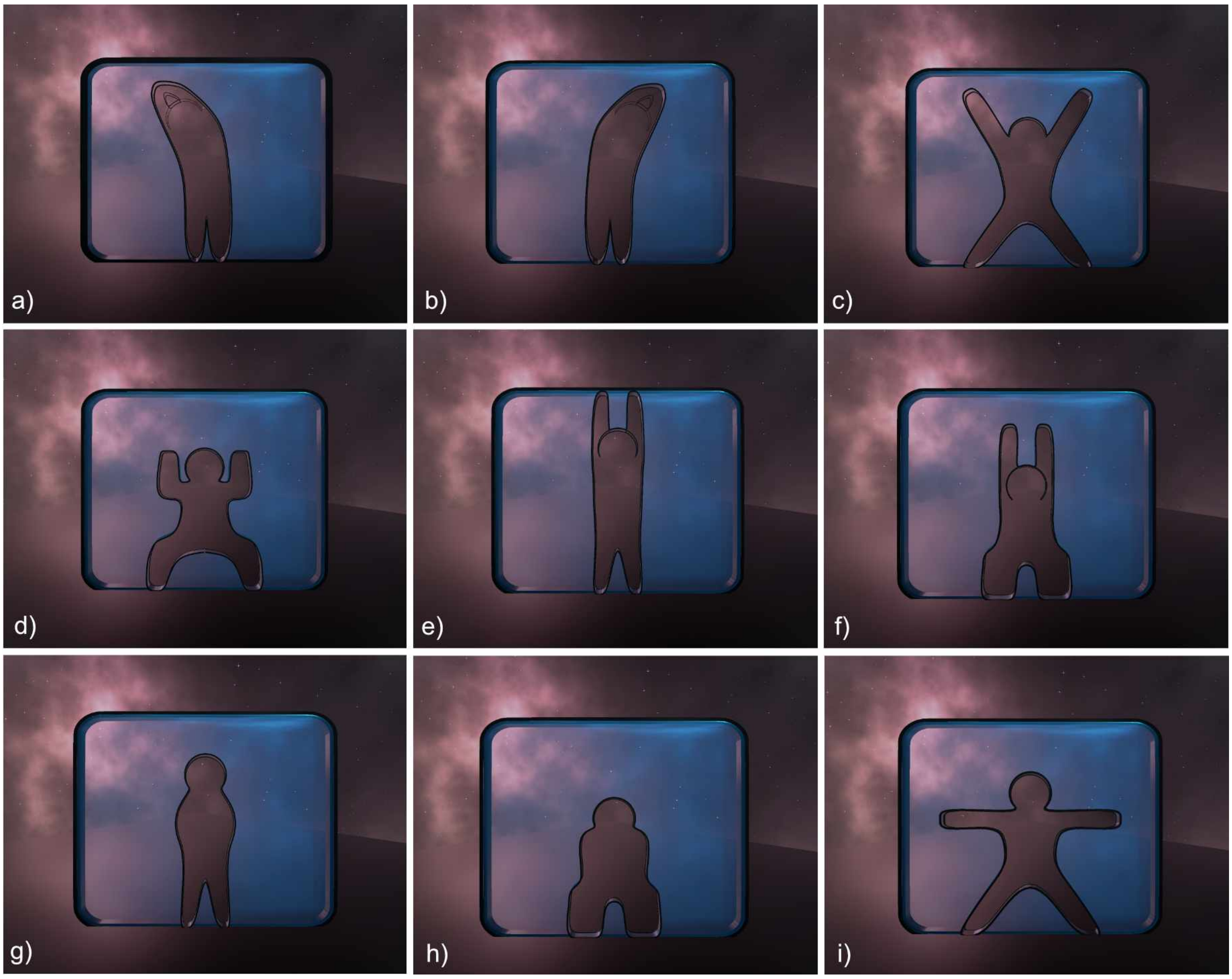

Fig. 2. The featured body poses in our game: a) crescent moon left, b) crescent moon right, c) five pointed star with raised hands, d) goddess pose, e) mountain pose, f) overhead squat, g) sinking pose, h) squat, and i) five pointed star.

3.3 Cues

We used two modalities to represent in-game cues in our game (see Table 1): auditory and visual. Our primary motivation behind this choice was the central importance of audio and visual effects for games. Additionally, the chosen cues are easy to integrate into almost every VR exergame. In contrast, providing haptic feedback in a VR exergame might have less external validity when applied to other VR exergames: Some games feature intense training and movement, and players might not feel the haptic feedback. Lastly, we also decided to focus only on two modalities to avoid fatigue and carryover effects. Please refer to our supplementary materials to see the sample of the used auditory cue, and a video of the gameplay.

Auditory Cue. The audio cue was designed as repeatedly playing sound guiding players to perform the pose of the currently approaching obstacle. If players did not match this pose at all, the sound would play at the minimal frequency of 0.87 Hz (every 1.15 seconds). Conversely, if players matched the required movement perfectly, the sound would play with a maximal frequency of 6.66 Hz (every 0.15 seconds). Depending on the players' performance, the frequency with which the sound is played gradually changes between those two extreme points (see Table 1). This approach is inspired by Ariza et al. [3]'s proximity-based feedback in VR. We also note that humans can

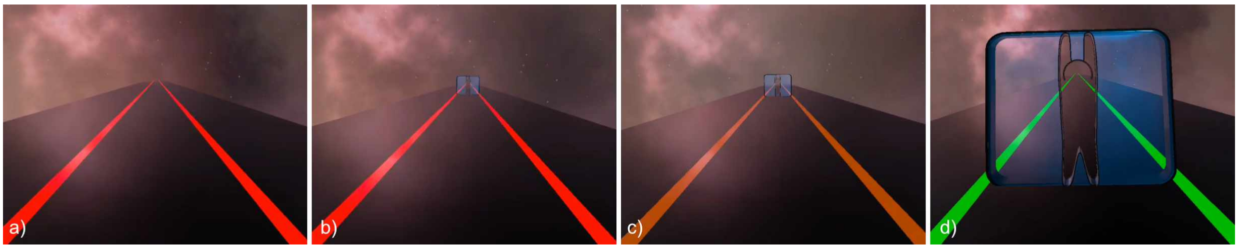

Fig. 3. VR in-game view of the visual cue presented in the game: a) The player waits for a wall to appear. b) The player sees the wall appear, and the visual cue (red) indicates that the movement has not yet been performed correctly. c) The player performs the movement, and the visual cue gradually changes color from red to green (orange). d) The player achieves the correct body posture shown in the game and receives a green visual cue that they are performing well.

Modality	Change	Minumum	Maximum
Auditory Cue	Frequency	1 per 1.15 seconds	1 per 0.15 seconds
Visual Cue	Color		

Table 1. Players received cues while performing each pose. The color and frequency of these cues gradually changed from minimum to maximum presentation values depending on the players' performance. The frequency refers to how often the sound was played.

distinguish differences in tempo between two repeating sounds as little as 8% [67]. Accordingly, our chosen frequency range could result in roughly 27 distinguishable levels; however, this value is based on a controlled lab experiment and likely does not fully transfer to a VR exergame session.

Visual Cue. Players experienced a visual cue in the form of two lines on the platform path. Similarly to the auditory cue, the lines' color changes depending on how accurately the player matches the pose of the approaching wall. If a player performs the movement with 0% percent success, the lines are colored in bright red (RGB = (255,0,0) — (minimum)). On the opposite, the lines are colored in green to indicate 100% success (RGB = (0,255,0) — (maximum)). Depending on the performance, the color of the two lines is interpolated between red and green (see Figure 3 and Table 1). Like in the *AU-CUE* condition, this approach was also informed by the work of Ariza et al. [3]. One disadvantage of the chosen scale is the limited number of distinguishable values and the limited contrast between different levels [62]. Alternative color maps, such as a rainbow scale or a linearized optimal color scale [42], would resolve this issue. However, users typically have more problems identifying the inherent ordering of such color maps [7, 62], and it might not be intuitively clear how the colors represent players' performance. Accordingly, we decided to prioritize understanding over level of detail. Interestingly, research on just-noticeable differences in colors with varying luminances [13] showed that humans can distinguish between 85 to 134 variances of each of the three primary colors on a standard display—more levels than with our auditory cue. But, these values are again theoretical and likely do not transfer to a physically and visually intense, immersive exergame, and exposure to visually intense systems with additional visual information may lead to a high cognitive workload [49] that may affect overall perception.

3.4 Movement Evaluation & Pilot Studies

Since we developed our game for the mobile Meta Quest 2, we could only rely on the headset and controllers' position to evaluate how players performed the different poses. While it is easy to

Table 2. List of which body parts are used to detect each body pose, provide cues, and calculate scores.

POSES	HEAD	ARMS
Crescent Moon (Figure 2a,b)	no	yes
Five Pointed Star with Raised Hands (Figure 2c)	yes	yes
Goddess Pose (Figure 2d)	yes	yes
Mountain Pose (Figure 2e)	no	yes
Overhead Squat (Figure 2f)	yes	yes
Sinking Pose (Figure 2g)	no	yes
Squat (Figure 2h)	yes	no
Five Pointed Star (Figure 2i)	yes	yes

determine the arm movements based on the controller motions, the torso and lower body require some estimation. Therefore, we first sampled the headset's vertical position to calibrate an estimate of the player's height. Then, we categorized the poses into four categories and determined how the head height changes for these types of poses. The first category encompasses standing poses, i.e., crescent moon (Figure 2a,b), mountain pose (Figure 2e), and sinking pose (Figure 2g), and requires players to stand fully upright. For the remaining categories, the legs are bent or spread, leading to a lower head height: squats (Figure 2f,h), goddess pose (Figure 2d), and spread leg body poses (Figure 2c,i). To understand how the head position changes for these poses, we conducted an informal pilot study (N=4, 2 women, 2 men) and asked them to perform these poses. Based on the evaluation, we determined the height changes for the body poses as follows: standing positions (no change), squat positions (-23.56%), goddess position (-13.63%), and spread legs positions (-7.12%).

Since the featured poses require a change in the positions of either the head, the hands, or a combination of both, we calculated the score and provided cues about players' performance based on how these involved body parts matched the target position (see Table 2). For example, in the mountain pose (Figure 2e), we only consider arm position changes, as players need to stay in their standing position and raise their arms. Moreover, we observed that linear mapping between the performed pose accuracy and the provided cue is affected by tracking errors. To minimize the noise in the tracking data, we decided to filter the raw data with an activation function. We tested the applicability of two different functions (easeOutSine and easeOutQuad)[1] in an informal study with 5 participants (3 women, 2 men). After evaluating the players' responses (7-point Likert item) on how their movement matched the provided cues, we decided to use the easeOutSine function.

3.5 Tutorial

To ensure that players can perceive and get familiar with the designed cues, we created an interactive VR tutorial that presents the cues and body poses featured in the game. First, participants calibrated their height and arm length (by performing T-pose and arm raise pose). Then, they were introduced to the featured poses and given text descriptions of how to perform these poses. Also, they were given a text explanation of the cue representations, shown both cue types at maximum and minimum success levels, and given the chance to perform the poses to see the changes in the cues.

3.6 Gameplay

The gameplay for each condition was 3 minutes long (like other research [8] featuring a similar game) to avoid fatigue effects. At the same time, it allowed players to experience all the featured cues (four conditions). In total, players experienced 30 walls per condition, each taking six seconds

[1] https://easings.net/

to reach them after spawning. Players were tasked with performing the movements featured on the walls and passing through them.

4 Method

To understand the impact of having performance cues on the players' experience and performance, we conducted a 2 (auditory cue: on vs. off) x 2 (visual cue: on vs. off) within-participants study.

4.1 Sample Characteristics

We recruited a total of 32 participants (9 women, 23 men). The mean age of participants was 25.81 years (*SD*=8.12). The majority of players were not regular VR HMD users: Only nine of them used VR 1−2 or more times per month. However, many reported exercising (at least 20 minutes) at least 1-2 times per month (*n*=28). The enjoyment of physical activity was rated as slightly positive (*M*=1.66, *SD*=1.29, range from -3 to 3).

4.2 Conditions

We used the following four conditions in our study:

- *NO-CUE - No Cueing:* Players performed featured body poses without receiving any cues.
- *AU-CUE - Audio Cueing:* Players experienced an audio cue guiding them to perform the featured in-game movements.
- *VI-CUE - Visual Cueing:* Players received a visual cue that guides them to perform the featured in-game movements.
- *AV-CUE - Audiovisual Cueing:* In this variant, both audio and visual cues were experienced.

4.3 Measures

Pre-Game Measures. We used a demographic questionnaire to gather information about the sample characteristics (e.g., age, gender). Additionally, we administered the Simulator Sickness Questionnaire (SSQ) [34] before the VR game to understand the severity of participants' latent cybersickness symptoms. This questionnaire consists of 16 items on a 4-point scale (*none/0 — severe/3*) and leads to Total SSQ, nausea, oculomotor disturbance, and disorientation scores.

Post-Game Measures. To assess the player experience, we used the Player Experience Inventory (PXI) [1]. It contains a total of 30 items in 10 subcategories, such as immersion and mastery. In our study, we also used the additional enjoyment subcategory of the PXI. Players rated their experience on a 7-point Likert scale (*strongly disagree/-3 — strongly agree/3*).

We measured the enjoyment of physical activity using the short version of the Physical Activity Enjoyment Scale (PACES-S) [10, 17]. It includes four items on a 5-point Likert Scale (*strongly disagree/1 — strongly agree/5*).

Lastly, we created additional custom questions to better understand the nuances of providing cues. To assess physical workload, players rated the single item "*I found this version of the game to be physically demanding*" on a 7-point Likert scale (*strongly disagree/-3 — strongly agree/3*). Similarly, we used the question "*I found this version of the game to be cognitively demanding*" (7-point Likert scale, *strongly disagree/-3 — strongly agree/3*) to measure the cognitive workload. With our third custom item, we asked for the perceived improvement: "*I found this version of the game helpful in improving my performance*" (7-point Likert scale, *strongly disagree/-3 — strongly agree/3*).

Final Measures. After participants completed all conditions, we asked a question to understand players' preferences regarding the used cues: "*Which version of the game did you like best?*". Finally, participants were presented with an open-ended question (i.e., "*Do you have any additional comments?*") where they could provide additional comments.

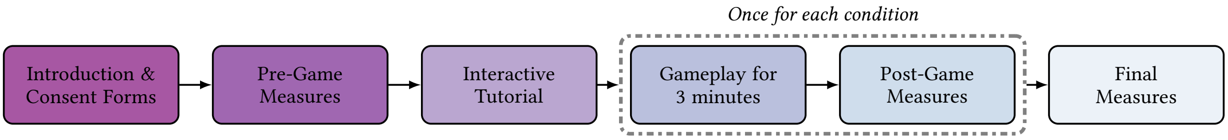

Fig. 4. We followed a within-participants design and primarily assessed questionnaire and performance data.

Game Performance. During the game, we collected the in-game performance by logging the matching accuracy and wall crashes for each pose. The pose accuracy captures the players' success in matching the required pose for each movement (i.e., headset height and controller position). After filtering tracking errors with the chosen activation function, we logged this accuracy at the time point at which players pass through the wall in the form of a percentage ranging from 0 to 100 (e.g., 72% successful performance). Additionally, a wall crash is registered whenever players fail to pass through the 3D hole in the wall. Given that our testbed game primarily features isometric strength exercises, we did not measure physiological measures, like heart rate, since these would likely not differ significantly. Instead, we focused exclusively on movement precision (targeted by our cue design) and player experience.

4.4 Procedure

We used a convenience sampling method to announce our study on various channels (e.g., university emailing lists, university recruiting system). Only participants with a minimum age of 18 were eligible. Additionally, they had to be able to perform full-body movements (e.g., squatting) with no uncorrected visual impairments.

Prior to the study, the participants filled out an informed consent form, followed by the demographic and pre-game questionnaire. Afterward, they were introduced to the Meta Quest 2 setup and performed the interactive tutorial.

Participants started with the tutorial and then played each condition for 3 minutes in a counterbalanced order (using a balanced Latin Square design). After each gameplay, they filled out the post-game questionnaires. Upon finishing all four conditions, players completed the final questionnaires. If applicable, participants received a confirmation (i.e., credit), which they need to collect during their studies. Completing the entire study (e.g., including an introduction to VR and signing the consent form) for each participant took approximately 90 minutes.

5 Analysis & Results

Our study included four conditions, featuring 2 (auditory cue: on vs. off) x 2 (visual cue: on vs. off) within-participants design. We confirmed the normality assumption with Shapiro-Wilk tests. If the normality was not violated, we used two-way repeated-measures ANOVAs (factors: auditory and visual) and paired t-tests for pairwise comparisons. Otherwise, we used non-parametric two-way repeated-measures ANOVAs[2] (factors: auditory and visual) based on Aligned Rank Transform [33, 69] (ART[3]) and Wilcoxon-signed rank tests for pairwise comparisons. Additionally, to analyze pre- and post-comparisons of SSQ scores, we conducted Friedman's ANOVAs and treated the measurement points as time sequences with five levels (pre, post-*NO-CUE*, post-*AU-CUE*, post-*VI-CUE*, and post-*AV-CUE*). Finally, we used an exact multinomial test to see the preference over game variants and exact binomial tests for pairwise comparisons.

[2]Using R's repeated measures model (aov)

[3]We use the subscript F_a to indicate the tests run based on this method.

Table 3. The table shows the descriptive values of the PXI, PACES-S, and SSQ subcategories.

Condition	PXI-mea. *Mean (SD)*	PXI-cur. *Mean (SD)*	PXI-mas. *Mean (SD)*	PXI-aut. *Mean (SD)*	PXI-imm. *Mean (SD)*	PXI-pro. *Mean (SD)*	PXI-aud. *Mean (SD)*	PXI-cha. *Mean (SD)*
NO-CUE	-0.21 (1.47)	0.11 (1.35)	0.72 (1.08)	-0.30 (1.51)	0.91 (1.30)	-2.02 (0.93)	1.31 (1.10)	0.68 (1.14)
AU-CUE	0.55 (1.40)	0.95 (1.42)	1.28 (1.02)	-0.54 (1.48)	1.69 (0.92)	0.94 (1.43)	1.65 (0.93)	1.07 (1.05)
VI-CUE	0.49 (1.27)	0.66 (1.30)	1.47 (1.03)	-0.30 (1.56)	1.51 (0.93)	0.80 (1.47)	1.67 (1.12)	0.70 (1.24)
AV-CUE	0.46 (1.45)	0.64 (1.24)	1.29 (1.08)	-0.63 (1.49)	1.80 (0.96)	1.16 (1.51)	1.58 (1.07)	1.17 (1.07)
Condition	PXI-eas. *Mean (SD)*	PXI-cla. *Mean (SD)*	PXI-enj. *Mean (SD)*	PACES-S. *Mean (SD)*	Total SSQ *Mean (SD)*	Nausea *Mean (SD)*	Oculomotor *Mean (SD)*	Disorientation *Mean (SD)*
Pre	–	–	–	–	13.32 (15.08)	12.52 (14.83)	10.42 (12.73)	12.18 (22.85)
NO-CUE	1.40 (1.28)	1.47 (1.39)	0.78 (1.26)	3.25 (0.83)	19.52 (22.51)	20.57 (17.82)	14.92 (19.59)	15.23 (30.69)
AU-CUE	1.72 (1.01)	2.07 (0.88)	1.50 (1.15)	3.53 (0.99)	17.53 (19.78)	19.68 (18.44)	12.55 (15.60)	13.49 (27.04)
VI-CUE	2.19 (0.76)	2.05 (1.09)	1.44 (0.97)	3.51 (0.95)	16.71 (18.07)	19.38 (16.52)	12.08 (14.90)	11.75 (27.41)
AV-CUE	2.08 (0.88)	1.93 (1.33)	1.63 (0.99)	3.59 (0.86)	18.82 (21.06)	19.68 (18.92)	14.21 (16.08)	15.23 (28.80)

Below, we report all pairwise tests with the Bonferroni correction method (i.e., p values are adjusted). Table 3 and 4 provide descriptive values of our measures. Example code snippets of our R analysis are provided in the supplementary materials.

5.1 Cybersickness

We did not observe a significant effect of time points on SSQ's categories: Total SSQ: $\chi^2(4)$=4.759, p=0.313, *Kendall's W*=0.037; SSQ-Nausea $\chi^2(4)$=8.860, p=0.065, *Kendall's W*=0.069; SSQ-Ocu: $\chi^2(4)$= 4.462, p=0.347, *Kendall's W*=0.035; SSQ-Dis: $\chi^2(4)$=3.399, p=0.493, *Kendall's W*=0.027.

5.2 Player Experience

Meaning. We found a significant main effect of having an auditory cue, $F_a(1,31)$=7.749, p=0.009, η_P^2 =0.200. Similarly, there was a significant effect of having a visual cue, $F_a(1,31)$=10.130, p=0.003, η_P^2 =0.246. There was also an interaction effect of both independent variables on the meaning ratings, $F_a(1,31)$=12.767, p=0.001, η_P^2 =0.292. Pairwise comparisons indicate that participants found *NO-CUE* less meaningful compared to *AU-CUE* (V=14, p=0.001), *VI-CUE* (V=26.5, p=0.003), and *AV-CUE* (V=37, p=0.002).

Curiosity. Having an auditory cue had a statistically significant effect on the curiosity scores, $F(1,31)$=7.860, p=0.009, η_P^2 =0.202. Nevertheless, we found no significant effect of having a visual cue, $F(1,31)$=0.514, p=0.479, η_P^2 =0.016. We observed an interaction effect between auditory and visual cues on the curiosity scores, $F(1,31)$=7.774, p=0.009, η_P^2 =0.201. Post-hoc tests indicated that only *NO-CUE* led to lower curiosity compared to *AU-CUE* ($t(31)$=-3.281, p=0.015).

Mastery. While the analysis of the mastery scores did not reveal any significant effect of having auditory cues, $F_a(1,31)$=2.360, p=0.135, η_P^2 =0.070, there was a significant effect of having visual cues, $F_a(1,31)$=12.698, p=0.001, η_P^2 =0.291. Similarly, we also observed an interaction effect between both independent variables, $F_a(1,31)$=11.744, p=0.002, η_P^2 =0.275. Pairwise comparisons showed that participants' feeling of mastery was higher for *VI-CUE* compared to *NO-CUE*, (V=62, p=0.005).

Autonomy. The autonomy scores did not differ significantly, regardless of the presence of auditory ($F_a(1,31)$=1.599, p=0.215, η_P^2 =0.049) or visual cues ($F_a(1,31)$=0.074, p=0.787, η_P^2 =0.002), nor did we observe an interaction effect ($F_a(1,31)$=0.145, p=0.706, η_P^2 =0.005).

Immersion. We found significant main effects on the immersion scores for the auditory cueing, $F_a(1,31)$=13.055, p=0.001, η_P^2 =0.296, and the visual cueing, $F_a(1,31)$=26.524, p<0.001, η_P^2 =0.461. Likewise, there was a significant interaction effect, $F_a(1,31)$=5.128, p=0.031, η_P^2 =0.142. Post-hoc tests revealed that all conditions had higher immersion scores compared to the *NO-CUE* condition: *AU-CUE* (V=48.5, p=0.008), *VI-CUE* (V=25.5, p=0.001), and *AV-CUE* (V=14, p<0.001).

Progressive Feedback. Unsurprisingly, the main effect of having auditory cues (F_a(1,31)=70.103, p<0.001, η_P^2 =0.693) and visual cues (F_a(1,31)=73.523, p<0.001, η_P^2 =0.703) was significant regarding the progressive feedback scores. Similarly, we also observed a significant interaction effect, F_a(1,31)=56.611, p<0.001, η_P^2 =0.646. Pairwise tests showed that *NO-CUE* was rated lower compared to *AU-CUE* (V=1, p<0.001), *VI-CUE* (V=6.5, p<0.001), and *AV-CUE* (V=0, p<0.001) conditions.

Audiovisual Appeal. While there was no significant effect of having auditory cues on the audiovisual appeal (F_a(1,31)=2.839, p=0.102, η_P^2 =0.084), we found a significant effect of the visual cues (F_a(1,31)=4.660, p=0.039, η_P^2 =0.131). There was also a significant interaction effect, F_a(1,31)=8.229, p=0.007, η_P^2 =0.210. Post-hoc comparisons showed that *NO-CUE* was found to be less audiovisually appealing than *AU-CUE* (V=31, p=0.021).

Challenge. There was a significant main effect of having auditory cues on the challenge scores, F_a(1,31)=9.158, p=0.005, η_P^2 =0.228. Neither the main effect of having visual cues (F_a(1,31)=0.166, p=0.687, η_P^2 =0.005) nor the interaction effect (F_a(1,31)=0.006, p=0.938, η_P^2 =0.000) were significant. Post-hoc tests indicate that *VI-CUE* was less challenging than *AV-CUE* (V=76, p=0.040).

Ease of Control. Regarding the perceived ease of control, we found no significant effect of the auditory cues (F_a(1,31)=0.367, p=0.549, η_P^2 =0.011) or the interaction effect (F_a(1,31)=2.418, p=0.130, η_P^2 =0.072). However, having visual cues had a significant effect (F_a(1,31)=41.535, p<0.001, η_P^2 =0.573). Post-hoc tests revealed that *NO-CUE* had lower ease of control scores compared to *VI-CUE* (V=30.5, p=0.001) and *AV-CUE* (V=53, p=0.004). Similarly, *AU-CUE* had lower ease of control scores ratings than *AV-CUE* (V=41, p=0.020).

Clarity of Goals. Significant main effects of both the auditory (F_a(1,31)=7.768, p=0.009, η_P^2 =0.200) and visual cues were found (F_a(1,31)=5.972, p=0.020, η_P^2 =0.162). Likewise, there was a significant interaction effect, F_a(1,31)=11.977, p=0.002, η_P^2 =0.279. Post-hoc tests showed that *NO-CUE* led to lower clarity of goal scores compared to *AU-CUE* (V=19, p=0.001) and *VI-CUE* (V=36.5, p=0.022).

Enjoyment. Similar to other subcategories, we found a significant main effect of having auditory cues (F_a(1,31)=10.630, p=0.003, η_P^2 =0.255) and visual cues (F_a(1,31)=10.334, p=0.003, η_P^2 =0.250), as well as an interaction effect (F_a(1,31)=5.409, p=0.027, η_P^2 =0.149) regarding the enjoyment scores. Post-hoc tests indicated that *NO-CUE* led to lower enjoyment when compared to *AU-CUE* (V=54, p=0.022), *VI-CUE* (V=85, p=0.044), and *AV-CUE* (V=39, p=0.003).

5.3 Physical Activity Enjoyment

We found a significant main effect of having auditory cues on physical activity enjoyment, F(1,31)=5.483, p=0.026, η_P^2 =0.150. However, there was neither a significant effect of having visual cues (F(1,31)=2.739, p=0.108, η_P^2 =0.081) nor an interaction effect (F(1,31)=1.224, p=0.277, η_P^2 =0.038) on the PACES-S scores. Post-hoc tests showed no significant difference between conditions.

5.4 Physical & Cognitive Workload

There was a significant main effect of having auditory cues on the physical workload scores, F_a(1,31)=12.627, p=0.001, η_P^2 =0.289. Neither the main effect of the visual cues (F_a(1,31)=0.105, p=0.748, η_P^2 =0.003) nor the interaction effect (F_a(1,31)=0.292, p=0.593, η_P^2 =0.009) were significant. Pairwise tests revealed that only *AU-CUE* led to higher physical workload scores compared to *NO-CUE* (V=5, p=0.022).

Table 4. The table shows the descriptives of physical and cognitive workload, perceived improvement, participants' game variant preferences, as well as, in-game performance metrics (pose accuracy and the number of crashes)

Condition	Phy. Workload *Mean (SD)*	Cog. Workload *Mean (SD)*	Per. Imp. *Mean (SD)*	Game Ver. Pref. (# Players)	Pose Accuracy (%) *Mean (SD)*	# of Crashes *Mean (SD)*
NO-CUE	0.06 (1.52)	-0.91 (1.65)	-1.47 (1.24)	0	89.39 (4.04)	9.34 (3.75)
AU-CUE	0.63 (1.52)	-0.50 (1.59)	1.47 (1.11)	6	91.19 (3.87)	8.94 (4.99)
VI-CUE	0.13 (1.50)	-0.94 (1.50)	0.72 (1.49)	6	90.22 (3.94)	8.81 (4.22)
AV-CUE	0.56 (1.41)	-0.69 (1.60)	1.69 (1.51)	20	91.58 (3.77)	8.16 (3.05)

We found no significant effect of having auditory (F_a(1,31)=2.573, p=0.119, η_P^2 =0.077) or visual cues (F_a(1,31)=0.126, p=0.725, η_P^2 =0.004), nor their interaction (F_a(1,31)=0.292, p=0.593, η_P^2 =0.009) on cognitive workload.

5.5 Perceived Improvement

We found a significant effect of the auditory cues (F_a(1,31)=75.234, p<0.001, η_P^2 =0.708), the visual cues (F_a(1,31)=30.172, p<0.001, η_P^2 =0.493), as well as an interaction effect (F_a(1,31)=43.935, p<0.001, η_P^2 =0.586) regarding the perceived improvement scores. Pairwise tests showed that participants experienced a lower perceived improvement in *NO-CUE* compared to *AU-CUE* (V=0, p<0.001), *VI-CUE* (V=0, p<0.001), and *AV-CUE* (V=2, p<0.001). Moreover, *VI-CUE* led to lower perceived improvement scores compared to *AV-CUE* (V=49.5, p=0.012).

5.6 Preference for the Game Variants

Overall, the majority of the participants preferred *AV-CUE* (n = 20) compared to the other game variants (*NO-CUE* =0, *AU-CUE* =6, *VI-CUE* =6). An exact multinomial test revealed that the proportions of the condition preference were significantly different from chance, p<0.001. Pairwise exact binomial test comparisons with Bonferroni correction indicated that proportions of *AV-CUE* preference differed significantly compared to *NO-CUE* (p<0.001).

5.7 Gameplay Metrics

Having auditory cues had a significant effect on the pose accuracy, F_a(1,31)=13.850, p=0.001, η_P^2 =0.309. But, there was neither a significant effect of having visual cues (F_a(1,31)=3.233, p=0.082, η_P^2 =0.094) nor an interaction effect. (F_a(1,31)=2.381, p=0.133, η_P^2 =0.071). Only *AV-CUE* led to significantly higher pose accuracy compared to *NO-CUE* (V=92, p=0.005).

Having auditory cues (F(1,31)=1.425, p=0.242, η_P^2 =0.044) or visual cues (F(1,31)=1.324, p=0.259, η_P^2 =0.041) did not have a significant effect on the number of wall crashes. Also, the interaction effect was not significant (F(1,31)=0.077, p=0.784, η_P^2 =0.002). To better understand how the conditions influenced the player performance, we plotted the cumulative wall crashes for the four conditions in Figure 5. The linear increase over time indicates that there was no learning effect. At the same time, the conditions involving cues show a flatter slope, which might hint towards a positive effect on game performance compared to *NO-CUE*, although the difference was not statistically significant.

5.8 Qualitative Feedback

Only 13 participants provided qualitative feedback. From those, four praised the designed audio-only cue: "*audio cues left a more impactful feeling on my performance*"-P_3. One participant highlighted the pros and cons of visual and audio cues, stating that "*audio cueing seemed to help me to immerse myself in the game, while visual cueing was better at informing me about my performance*"-P_{26}.

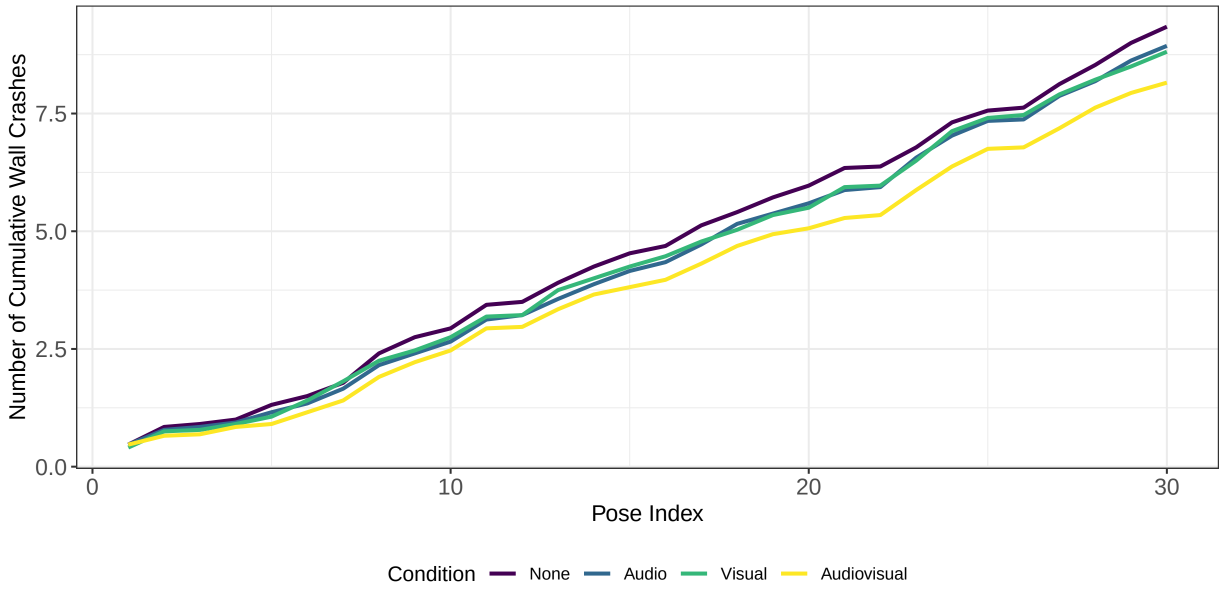

Fig. 5. Number of cumulative wall crashes by pose index across conditions.

Similarly, another player found the *VI-CUE* to provide a feeling of mastery: "*only visual cueing I felt more successful and confident in regards to being good at the game*"-P_{16}. But at the same time, they indicated that "*I feel like the version with both cueing and audio guidance is more successful in indicating what the perfect technique of the performance is*". Interestingly, two participants reported that "*the auditory cue and the visual cue sometimes felt contradictory because at some point it was probably no longer possible to perceive a greener green. However, a faster auditory signal would be possible*"-P_{29}. Finally, two participants wished for additional tips on how to better improve their movement execution.

6 Discussion

This section discusses our findings and provides a summary of our results with respect to our research questions. Based on the results, we then provide implications for including cues (i.e., formative feedback) when designing VR exergames.

6.1 No cybersickness.

Our findings show that players did not experience increased cybersickness symptoms after exposure to the VR exergame. Compared to prior studies on VR gaming in general [59] and on exergames in specific [12], our measured scores indicate that players did not experience notable cybersickness symptoms. This confirms that our designed game and the integrated cues can be used without negatively affecting participant's well-being. However, we also note the general criticism of using the SSQ to measure cybersickness in exergames, as there might be overlapping items between cybersickness symptoms or exercising outcomes (e.g., sweating) [12, 64].

6.2 Cues mostly do not influence workload.

In our study, adding cues did not lead to significantly higher cognitive workload. We attribute this to our adherence to the coherence principle (i.e., avoiding additional material) [11] in the design of our cues—our cues had a simple design. However, interestingly, audio cueing led to a higher physical workload compared to not receiving any cues. This might be due to players' higher enjoyment (thereby higher engagement) in *AU-CUE*, which may have encouraged them to move more. Alternatively, the repeatedly playing sound (i.e., audio cue) may have created a sense of urgency and pressure, which again may have resulted in more movement, leading players to feel a higher physical workload.

6.3 Cueing improves meaningfulness and clarity and provides progress information in VR exergames.

In terms of player experience, we observed that *NO-CUE* was found to be significantly less meaningful than all other conditions. Also, the quality of progress feedback was rated significantly higher for the conditions with cues compared to *NO-CUE*. These findings match our reasoning for providing cues to make the game respond to the players' movements and give them immediate feedback, thereby emphasizing the importance of correct movements.

Interestingly, the game's goal was perceived clearer when only unimodal cues (*AU-CUE* and *VI-CUE*) were provided. While this seems to contradict our finding that multimodal cues enhanced the progress feedback, qualitative feedback might hint towards a likely explanation. Two participants commented that the audio and visual feedback do not match perfectly with each other. In consequence, players might receive clearer instructions from individual cues, whereas the combination provides generally richer feedback. Given the limited knowledge of how different modalities influence each other, we emphasize the need for further research on this interplay.

Furthermore, we see that the meaningfulness and progress feedback scores were only slightly positive across conditions. Looking at the qualitative data, we speculate that this is due to our cues providing mainly high-level information; in the case of *VI-CUE*, players experienced a red color when they did not perform the movement correctly. But, they did not get additional information about which part of the body they should move differently to reach the green color. Hence, we recommend further exploration of using cues to provide detailed information.

6.4 Cueing adds to the immersion and enjoyment of VR exergames.

We see similar effects for immersion and enjoyment: Players found all conditions with cues more enjoyable and immersive than *NO-CUE*. We believe that adding new game elements (i.e., continuous cues) that react to the player makes the game more lively and, hence, more enjoyable and immersive. Interestingly, only the audio cues improved the game's audiovisual appeal significantly over the *NO-CUE* condition, although tendencies in the data hint towards the positive effect of all cueing conditions. The benefits towards immersion, enjoyment, and appeal match with Beaton et al. [5]'s lens of juiciness; providing continuous feedback (i.e., juicy feedback) supports player experience.

However, contrary to the common belief that more cues are better, our results highlight that users see no significant difference between unimodal and bimodal cues regarding immersion and enjoyment. Nevertheless, similar to previous studies in VR exergames [35], players perceive greater improvements in their performance when cues are used, particularly while playing the bimodal version. This observation is closely linked to the clear preference of the audiovisual round as the favorite condition (matching prior work [3]).

6.5 Unimodal cues have different effects on curiosity and mastery.

We observed contradictory findings for the players' feelings of curiosity and mastery. While only audio cueing made players significantly more curious compared to not receiving cues, for mastery, players rated having only a visual cue higher. For curiosity, this higher score for the audio condition can be explained twofold: (i) To understand the continuous change in audio, players must pay attention and figure out the mapping, which might stimulate curiosity. (ii) The visual cue is a more familiar concept from our daily lives (e.g., traffic lights) and, therefore, might trigger less curiosity. On the other hand, the lack of novelty of the visual cue likely contributed to gaining a sense of mastery in the game since it was easier to understand when performing well or poorly.

6.6 Bimodal cueing improves pose accuracy.

Players' pose accuracy was significantly better in *AV-CUE* compared to not receiving any cues. We conclude that bimodal cues are helpful for guiding players to perform correct movements. However, we see that while there was a tendency that cues helped players to reduce their number of crashes, this effect was not significant (see Figure 5). Therefore, we cannot make strong conclusions regarding performance improvements. We speculate that our game might have been too easy to detect significant differences across the game variants. Accordingly, we suggest further research with games involving dynamic difficulty adjustment to learn how cues might benefit players when performing individually tailored routines.

6.7 Takeaways

RQ1: How does having continuous cues affect player experience and performance in VR exergames? Overall, continuous cues (regardless of their modality) mostly improve player experience in VR exergames. In our study, players experienced more immersion and enjoyment when playing the conditions with cues and also felt the game was more meaningful and delivered better feedback. These benefits are also reflected in the participants' responses on the preferred game variant; no participant preferred not to have continuous cues. However, in terms of player performance, the effects of continuous cues are not apparent. Whereas cues did not have adverse effects (i.e., no cybersickness and similar physical activity enjoyment), only bimodal cueing improves pose accuracy significantly compared to no cueing. Additionally, purely descriptive differences in wall crashes and pose accuracies might hint towards the presence of a potential, undetected effect. However, in contrast to the positive influence on player experience, we cannot make definite conclusions regarding the effect of continuous cues on performance in exergames.

RQ2: How do different continuous cue modalities affect player experience and performance in VR exergames? The results reveal mixed effects; different modalities affect specific player experience constructs. For example, while audio cueing leads to higher curiosity compared to not receiving any cues, this is not the case for visual cueing. Conversely, players felt a higher mastery with the visual cues. Also, only combining both led to significantly improved player performance. A possible explanation lies within the players' prior experiences and sensorial differences between both modalities. Audio cues are less common, requiring players to learn the mapping, whereas visual cues are basic elements in many games, making them easier to grasp. At the same time, visual cues might diverge players' attention from other features in the virtual world (e.g., the obstacles). Together, the combination of both cues might help players learn and incorporate the feedback gained through both cues into their movements. Accordingly, our study confirms the importance of multimodality of exergame feedback.

6.8 Design Implications

Overall, our results support the idea that feedback is at the heart of games [14, 45, 52]. Based on our results, we provide three main implications and directions for future work when designing cues in VR exergames.

Consider Granularity of Cues. One of the main shortcomings of our visual cue is the possible granularity of feedback. Whereas it is relatively easy to notice changes in the sound's frequency, players had more difficulty distinguishing between two color variations. Our qualitative results hint that this may be related to the human perception of differentiating color changes, as interpolating between two colors (e.g., red and green) is naturally limited by these two extreme points—at some point, the green cannot become greener. Whereas our frequency-based audio cue offers a much lower range of feedback, according to controlled lab research on human perception [13, 67], we emphasize that the findings of these lab studies might not be directly applicable to VR exergames as they are physically and visually intensive. Accordingly, we call for future work to explore comparable cue modalities in VR exergames. Further, we emphasize choosing the optimal cue design for the intended feedback granularity in general. For a low to medium number of levels, visual color-coded elements offer the advantage of not requiring prior learning. Lastly, auditory cues could offer more fine-graded feedback —even with lower distinguishable levels— without cluttering the visual field of view.

Exploration of Different Continuous Cue Designs. For our game, we followed the implementation principles of Ariza et al. [3] to decide on the presented cues, i.e., using color and time of frequency as our visual and auditory modalities. To present these cues, we used two line representations (visual cue, colors) and constantly played medium-pitched sounds (auditory cue, time-frequency). We preferred this simple design as it could be easily integrated into many VR exergames. However, for some games, other designs could be more feasible and preferred. For example, in Yin et al. [73]'s game, darkening the environment is a better choice for an indirect cue since the environment already features a colorful race track. Similarly, games differ in the use of different channels. In games relying heavily on visual elements, visual cues force players to split their attention, whereas in rhythmic games, audio cues might be drowned out by the background music. Therefore, we suggest future work to explore the feasibility of different visual and auditory cue designs and their impact on player experience and performance.

Generalizability & Relevance to Other Immersive Experiences. Whereas our study focused on enhancing player experience and performance in VR exergames, our findings also apply to other types of immersive experiences. Almost every application provides at least limited cues to inform users of their performance. While our results might be transferable to other immersive applications, we note that compared to other types of applications, exergames naturally feature a higher task demand since players have to invest both physical and cognitive resources to succeed in the games. Accordingly, the requirements towards effective cues that do not overwhelm players with information are higher in this genre. Consequently, we believe that our central findings—that continuous audio cues complement the visual-centric gameplay with fine-graded feedback, whereas visual elements excel at easy learnability and that the combination of both is most effective at guiding players and preferred by users—confirm and extend best practices of general VR research [3, 60].

Furthermore, our results show that continuous cues can positively impact player experience and pose accuracy, indicating that cues can influence cognitive processing of performed movements in VR— supporting embodied cognition. Moreover, the use of continuous cues did not cause significantly higher cybersickness symptoms compared to not receiving any cues. This suggests that it is still possible to provide a comfortable experience for users while improving the player experience. Our results also align with rehabilitation research [55, 58] that emphasizes the potential and benefits of using cues in physical therapy. We confirm and extend their results, showing the benefits of using cues in the VR exergames realm. Accordingly, we recommend VR exergames targeting medical-focused training [29] to integrate cues to guide players in performing movements.

6.9 Limitations

In our game, we chose one of the most commonly used VR headsets (Meta Quest 2). While this choice has advantages (e.g., external validity), it also introduces drawbacks. In particular, the limited full-body tracking created an additional challenge for realizing our game featuring yoga-inspired poses. To ensure a good game experience, we conducted pilot tests to estimate the height changes for those poses and used these values for our user study. However, we cannot guarantee that minor differences in the personal physiology did not affect the individual game experience.

In terms of genre, our game featured a slow-paced VR exergame and did not require fast adjustments to the game tasks. Accordingly, we note that fast-paced exergames (e.g., *Beat Saber* [19]) and different game tasks [37] might yield different results regarding player experience and performance. Similarly, we note that the representation of cues and the design of the game can also impact the findings. As we stated in our implications, there are many ways to design cues, such as providing visual cues as a line or as the game's full background. Additionally, some games may include background music that may interfere with the provided audio cue. Hence, future studies should investigate the transferability of the results to other VR exergame genres, considering the interplay of different cue representations and the overall game design.

Our game focused on auditory and visual effects because (i) these modalities are crucial for games, (ii) the cues are easy to integrate into almost every VR exergame, and (iii) we aimed to avoid fatigue effects (duration of study was ca. 90 minutes). However, we encourage further exploration of other modalities (e.g., haptic feedback [60]) to enhance our understanding of the interplay between player experience, performance, and the modality of cues.

7 Conclusion

VR exergames have become widely used due to their fun and physical activity features. However, there is limited research on how to contribute to players' understanding of their movement execution and improvement. We investigated if and how providing continuous cues based on the in-game performance influences player experience and performance in VR exergames. We conducted a 2 (auditory: on vs. off) x 2 (visual: on vs. off) within-participants study (N=32) to investigate the effects of using continuous cues in a VR exergame. Our findings show that continuous cues, regardless of modality, enhance player experience, such as immersion and enjoyment, and bimodal cues increase pose accuracy. We provide three main design implications for using continuous cues in VR exergames, covering: (i) the granularity of cues, (ii) different continuous cue designs, and (iii) the generalizability of continuous cues to other immersive experiences.

Acknowledgments

This work was supported by the European Union, the German Federal Ministry of Education and Research (BMBF), the German Research Foundation (DFG), and partly funded under the Excellence Strategy of the Federal Government and the Länder. Additionally, this project was supported by the Provost's Program for Interdisciplinary Postdoctoral Scholars at the University of Waterloo.

References

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370. https://doi.org/10.1016/j.ijhcs.2019.102370 pages #3, #10

[2] Heidi Andrade and Gregory J Cizek. 2010. *Handbook of formative assessment.* Routledge. pages #3

[3] Oscar Ariza, Gerd Bruder, Nicholas Katzakis, and Frank Steinicke. 2018. Analysis of Proximity-Based Multimodal Feedback for 3D Selection in Immersive Virtual Environments. In *2018 IEEE Conference on Virtual Reality and 3D User*

Interfaces (VR). IEEE, Los Alamitos, CA, USA, 327–334. https://doi.org/10.1109/VR.2018.8446317 pages #3, #5, #7, #8, #16, #18

[4] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2021. The Effect of Pitch in Auditory Error Feedback for Fitts' Tasks in Virtual Reality Training Systems. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Los Alamitos, CA, USA, 85–94. https://doi.org/10.1109/VR50410.2021.00029 pages #5

[5] Dorcas E Beaton, Claire Bombardier, Francis Guillemin, and Marcos Bosi Ferraz. 2000. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 25, 24 (2000), 3186–3191. pages #16

[6] Joanna Bergström, Tor-Salve Dalsgaard, Jason Alexander, and Kasper Hornbæk. 2021. How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 533, 20 pages. https://doi.org/10.1145/3411764.3445193 pages #5

[7] David Borland and Russell M. Taylor Ii. 2007. Rainbow Color Map (Still) Considered Harmful. *IEEE Computer Graphics and Applications* 27, 2 (2007), 14–17. https://doi.org/10.1109/MCG.2007.323435 pages #8

[8] Felix Born, Sophie Abramowski, and Maic Masuch. 2019. Exergaming in VR: The Impact of Immersive Embodiment on Motivation, Performance, and Perceived Exertion. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, Los Alamitos, CA, USA, 1–8. https://doi.org/10.1109/VS-Games.2019.8864579 pages #2, #3, #4, #6, #9

[9] Ryan Canales and Sophie Jörg. 2020. Performance Is Not Everything: Audio Feedback Preferred Over Visual Feedback for Grasping Task in Virtual Reality. In *Proceedings of the 13th ACM SIGGRAPH Conference on Motion, Interaction and Games* (Virtual Event, SC, USA) *(MIG '20)*. Association for Computing Machinery, New York, NY, USA, Article 4, 6 pages. https://doi.org/10.1145/3424636.3426897 pages #5

[10] Cheng Chen, Susanne Weyland, Julian Fritsch, Alexander Woll, Claudia Niessner, Alexander Burchartz, Steffen C. E. Schmidt, and Darko Jekauc. 2021. A Short Version of the Physical Activity Enjoyment Scale: Development and Psychometric Properties. *International Journal of Environmental Research and Public Health* 18, 21 (2021), 15 pages. https://doi.org/10.3390/ijerph182111035 pages #10

[11] Ruth Colvin Clark and Richard E Mayer. 2016. Applying the Coherence Principle: Adding Extra Material Can Hurt Learning. e-Learning and the Science of Instruction. Hoboken. https://doi.org/10.1002/9781119239086.ch8 pages #4, #15

[12] Sebastian Cmentowski, Sukran Karaosmanoglu, Lennart E. Nacke, Frank Steinicke, and Jens Harald Krüger. 2023. Never Skip Leg Day Again: Training the Lower Body with Vertical Jumps in a Virtual Reality Exergame. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 786, 18 pages. https://doi.org/10.1145/3544548.3580973 pages #2, #5, #15

[13] Daniel D Desjardins and Patrick Gardner. 2015. Just noticeable color difference: implications for display systems. In *Display Technologies and Applications for Defense, Security, and Avionics IX; and Head-and Helmet-Mounted Displays XX*, Vol. 9470. SPIE, 116–127. pages #8, #18

[14] Heather Desurvire and Charlotte Wiberg. 2009. Game Usability Heuristics (PLAY) for Evaluating and Designing Better Games: The Next Iteration. In *Online Communities and Social Computing*, A. Ant Ozok and Panayiotis Zaphiris (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 557–566. pages #4, #17

[15] Kody R. Dillman, Terrance Tin Hoi Mok, Anthony Tang, Lora Oehlberg, and Alex Mitchell. 2018. A Visual Interaction Cue Framework from Video Game Environments for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173714 pages #4

[16] FitXR. 2019. *BoxVR*. Game [VR]. FitXR. London, England. https://fitxr.com/boxvr. Steam Reviews: https://store.steampowered.com/app/641960/BOXVR/. pages #5

[17] Julian Fritsch, Susanne Weyland, Katharina Feil, Alexander Burchartz, Steffen Schmidt, Alexander Woll, Ulrich Strauch, Benjamin Wienke, and Darko Jekauc. 2022. A Study on the Psychometric Properties of the Short Version of the Physical Activity Enjoyment Scale in an Adult Population. *International Journal of Environmental Research and Public Health* 19, 22 (2022), 11 pages. https://doi.org/10.3390/ijerph192215294 pages #10

[18] Tracy Fullerton. 2014. *Game design workshop: a playcentric approach to creating innovative games*. CRC press. pages #3

[19] Beat Games. 2019. *Beat Saber*. Game [VR]. Beat Games. Los Angeles, CA, USA. https://beatsaber.com/. Steam Reviews: https://store.steampowered.com/app/620980/Beat_Saber/. pages #2, #19

[20] Andrea Gauthier, Laura Benton, Leona Bunting, Elisabeth Herbert, Emma Sumner, Manolis Mavrikis, Andrea Revesz, and Asimina Vasalou. 2022. I Don't Usually Listen, I Read: How Different Learner Groups Process Game Feedback. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 88, 15 pages. https://doi.org/10.1145/3491102.3517480 pages #3

[21] Negar Haghbin and Marta Kersten-Oertel. 2021. Multimodal Cueing in Gamified Physiotherapy: A Preliminary Study. In *Proceedings of the 7th International Conference on Information and Communication Technologies for Ageing Well and e-Health - ICT4AWE*. INSTICC, SciTePress, Setúbal, Portugal, 137–145. https://doi.org/10.5220/0010396101370145 pages #2
[22] Joshua C. Haller, Young H. Jang, Jack Haller, Lindsay Shaw, and Burkhard C. Wünsche. 2019. HIIT The Road: Using Virtual Spectator Feedback in HIIT-based Exergaming. In *Proceedings of the Australasian Computer Science Week Multiconference* (Sydney, NSW, Australia) *(ACSW '19)*. Association for Computing Machinery, New York, NY, USA, Article 47, 9 pages. https://doi.org/10.1145/3290688.3290752 pages #5
[23] Günnur Hastürk and Müge Akyıldız Munusturlar. 2022. The effects of exergames on physical and psychological health in young adults. *Games for Health Journal* 11, 6 (2022), 425–434. pages #4
[24] Xinyu Hu, Alec Moore, James Coleman Eubanks, Afham Aiyaz, and Ryan P. McMahan. 2020. Evaluating Interaction Cue Purpose and Timing for Learning and Retaining Virtual Reality Training. In *Proceedings of the 2020 ACM Symposium on Spatial User Interaction* (Virtual Event, Canada) *(SUI '20)*. Association for Computing Machinery, New York, NY, USA, Article 5, 9 pages. https://doi.org/10.1145/3385959.3418448 pages #5
[25] Noor Azila Ismail, Hairul Anuar Hashim, and Hazwani Ahmad Yusof. 2022. Physical activity and exergames among older adults: A scoping review. *Games for Health Journal* 11, 1 (2022), 1–17. pages #4
[26] PF Jarzem and RB Gledhill. 1993. Predicting height from arm measurements. *Journal of Pediatric Orthopaedics* 13, 6 (1993), 761–765. pages #6
[27] Cheryl I. Johnson, Shannon K. T. Bailey, and Wendi L. Van Buskirk. 2017. *Designing Effective Feedback Messages in Serious Games and Simulations: A Research Review*. Springer International Publishing, Cham, 119–140. https://doi.org/10.1007/978-3-319-39298-1_7 pages #3
[28] Dominic Kao, Alejandra J. Magana, and Christos Mousas. 2021. Evaluating Tutorial-Based Instructions for Controllers in Virtual Reality Games. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 234 (oct 2021), 28 pages. https://doi.org/10.1145/3474661 pages #4
[29] Sukran Karaosmanoglu, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. 2024. Born to Run, Programmed to Play: Mapping the Extended Reality Exergames Landscape. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 309, 28 pages. https://doi.org/10.1145/3613904.3642124 pages #2, #6, #18
[30] Sukran Karaosmanoglu, Lucie Kruse, Sebastian Rings, and Frank Steinicke. 2022. Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 342, 7 pages. https://doi.org/10.1145/3491101.3519736 pages #2
[31] Sukran Karaosmanoglu, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. 2021. Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 252 (oct 2021), 27 pages. https://doi.org/10.1145/3474679 pages #2
[32] Sukran Karaosmanoglu, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. 2021. Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 443, 15 pages. https://doi.org/10.1145/3411764.3445492 pages #5
[33] Matthew Kay, Lisa A. Elkin, James J. Higgins, and Jacob O. Wobbrock. 2021. *ARTool: Aligned Rank Transform for Nonparametric Factorial ANOVAs*. The Comprehensive R Archive Network. https://CRAN.R-project.org/package=ARTool R package version 0.11.1. pages #11
[34] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220. https://doi.org/10.1207/s15327108ijap0303_3 pages #10
[35] Tanja Kojić, Lan Thao Nugyen, and Jan-Niklas Voigt-Antons. 2019. Impact of Constant Visual Biofeedback on User Experience in Virtual Reality Exergames. In *2019 IEEE International Symposium on Multimedia (ISM)*. IEEE, Los Alamitos, CA, USA, 307–3073. https://doi.org/10.1109/ISM46123.2019.00068 pages #2, #6, #16
[36] Bastian Kordyaka, Samuli Laato, Sebastian Weber, Juho Hamari, and Bjoern Niehaves. 2024. Exploring the association between engagement with location-based game features and getting inspired about environmental issues and nature. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 755, 15 pages. https://doi.org/10.1145/3613904.3642786 pages #4
[37] Bastian Kordyaka, Sebastian Weber, Gerhard Klassen, and Bjoern Niehaves. 2023. Decoding Gamification: A Randomized Within-Subjects Experiment on Affordances. (2023). https://aisel.aisnet.org/acis2023/113 pages #19
[38] Lucie Kruse, Sukran Karaosmanoglu, Sebastian Rings, and Frank Steinicke. 2022. Evaluating Difficulty Adjustments in a VR Exergame for Younger and Older Adults: Transferabilities and Differences. In *Proceedings of the 2022 ACM*

Symposium on Spatial User Interaction (Online, CA, USA) *(SUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 7, 11 pages. https://doi.org/10.1145/3565970.3567684 pages #2

[39] Samuli Laato, AKM Najmul Islam, and Teemu H Laine. 2022. Playing location-based games is associated with psychological well-being: an empirical study of Pokémon GO players. *Behaviour & Information Technology* 41, 9 (2022), 1888–1904. https://doi.org/10.1080/0144929X.2021.1905878 pages #4

[40] Odders Lab. 2019. *OhShape*. Game [VR]. Odders Lab. Seville, Spain. https://ohshapevr.com/. Steam Reviews: https://store.steampowered.com/app/1098100/OhShape/. pages #2, #3, #6

[41] Michael Lankes and Argenis Ramirez Gomez. 2022. GazeCues: Exploring the Effects of Gaze-based Visual Cues in Virtual Reality Exploration Games. *Proc. ACM Hum.-Comput. Interact.* 6, CHI PLAY, Article 237 (oct 2022), 25 pages. https://doi.org/10.1145/3549500 pages #2, #4

[42] Haim Levkowitz and Gabor T Herman. 1992. The design and evaluation of color scales for image data. *IEEE Computer Graphics and Applications* 12, 1 (1992), 72–80. pages #8

[43] Huimin Liu, Zhiquan Wang, Christos Mousas, and Dominic Kao. 2020. Virtual Reality Racket Sports: Virtual Drills for Exercise and Training. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 566–576. https://doi.org/10.1109/ISMAR50242.2020.00084 pages #4

[44] Markus Löchtefeld, Antonio Krüger, and Hans Gellersen. 2016. DeceptiBike: Assessing the Perception of Speed Deception in a Virtual Reality Training Bike System. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (Gothenburg, Sweden) *(NordiCHI '16)*. Association for Computing Machinery, New York, NY, USA, Article 40, 10 pages. https://doi.org/10.1145/2971485.2971513 pages #5

[45] Elizabeth J Lyons. 2015. Cultivating engagement and enjoyment in exergames using feedback, challenge, and rewards. *Games for health journal* 4, 1 (2015), 12–18. pages #4, #17

[46] Jake Mace. 2018. 7 Tai Chi Moves for Beginners | 15 Minute Daily Taiji Routine. https://www.youtube.com/watch?v=DmlsB_KOGtM pages #6

[47] Georgios Marentakis and Stephen A. Brewster. 2005. A comparison of feedback cues for enhancing pointing efficiency in interaction with spatial audio displays. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services* (Salzburg, Austria) *(MobileHCI '05)*. Association for Computing Machinery, New York, NY, USA, 55–62. https://doi.org/10.1145/1085777.1085787 pages #2

[48] B Jean Mason and Roger Bruning. 2001. Providing feedback in computer-based instruction: What the research tells us. 15 (2001), 2007. CLASS Research Report No.9. Center for Instructional Innovation, University of Nebraska-Lincoln. pages #2, #3

[49] Richard E. Mayer and Roxana Moreno. 2003. Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist* 38, 1 (2003), 43–52. https://doi.org/10.1207/S15326985EP3801_6 pages #8

[50] Meta Platforms, Inc. 2020. Meta Quest 2. https://www.oculus.com/quest-2/ Released: 13 October 2020. pages #6

[51] Anna Moschny, Petra Platen, Renate Klaaßen-Mielke, Ulrike Trampisch, and Timo Hinrichs. 2011. Barriers to physical activity in older adults in Germany: a cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity* 8, 1 (2011), 1–10. https://doi.org/10.1186/1479-5868-8-121 pages #4

[52] Florian Mueller and Katherine Isbister. 2014. Movement-Based Game Guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2191–2200. https://doi.org/10.1145/2556288.2557163 pages #2, #3, #4, #5, #17

[53] Florian 'Floyd' Mueller, Darren Edge, Frank Vetere, Martin R. Gibbs, Stefan Agamanolis, Bert Bongers, and Jennifer G. Sheridan. 2011. Designing Sports: A Framework for Exertion Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2651–2660. https://doi.org/10.1145/1978942.1979330 pages #2, #4

[54] Meng Ni, Kiersten Mooney, Anoop Balachandran, Luca Richards, Kysha Harriell, and Joseph F. Signorile. 2014. Muscle utilization patterns vary by skill levels of the practitioners across specific yoga poses (asanas). *Complementary Therapies in Medicine* 22, 4 (2014), 662–669. https://doi.org/10.1016/j.ctim.2014.06.006 pages #6

[55] A Nieuwboer, G Kwakkel, L Rochester, D Jones, E van Wegen, A M Willems, F Chavret, V Hetherington, K Baker, and I Lim. 2007. Cueing training in the home improves gait-related mobility in Parkinson's disease: the RESCUE trial. *Journal of Neurology, Neurosurgery & Psychiatry* 78, 2 (2007), 134–140. https://doi.org/10.1136/jnnp.200X.097923 pages #2, #4, #18

[56] Russell R Pate, Ruth P Saunders, Jennifer R O'Neill, and Marsha Dowda. 2011. Overcoming barriers to physical activity: helping youth be more active. *ACSM's Health & Fitness Journal* 15, 1 (2011), 7–12. https://doi.org/10.1249/FIT.0b013e318201c8ee pages #4

[57] Ton Roosendaal. 1994. *blender*. Blender Foundation. www.blender.org. pages #6

[58] Tamar C Rubinstein, Nir Giladi, and Jeffrey M Hausdorff. 2002. The power of cueing to circumvent dopamine deficits: a review of physical therapy treatment of gait disturbances in Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society* 17, 6 (2002), 1148–1160. https://doi.org/10.1002/mds.10259 pages #2, #4, #18

[59] Dimitrios Saredakis, Ancret Szpak, Brandon Birckhead, Hannah A. D. Keage, Albert Rizzo, and Tobias Loetscher. 2020. Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis. *Frontiers in Human Neuroscience* 14 (2020), 17 pages. https://doi.org/10.3389/fnhum.2020.00096 pages #15

[60] Lindsay Alexander Shaw, Burkhard Claus Wuensche, Christof Lutteroth, Jude Buckley, and Paul Corballis. 2017. Evaluating sensory feedback for immersion in exergames. In *Proceedings of the Australasian Computer Science Week Multiconference* (Geelong, Australia) *(ACSW '17)*. Association for Computing Machinery, New York, NY, USA, Article 11, 6 pages. https://doi.org/10.1145/3014812.3014823 pages #2, #5, #18, #19

[61] Harrison She, Jared Young, Ofek Wittenberg, Patrick Poole, Alexander Mercer, Alex Shaw, and Burkhard C. Wünsche. 2021. HIITCopter: Analysis of the Suitability of VR Exergaming for High-Intensity Interval Training. In *Proceedings of the 32nd Australian Conference on Human-Computer Interaction* (Sydney, NSW, Australia) *(OzCHI '20)*. Association for Computing Machinery, New York, NY, USA, 293–302. https://doi.org/10.1145/3441000.3441009 pages #6

[62] Samuel Silva, Joaquim Madeira, and Beatriz Sousa Santos. 2007. There is More to Color Scales than Meets the Eye: A Review on the Use of Color in Visualization. In *2007 11th International Conference Information Visualization (IV '07)*. 943–950. https://doi.org/10.1109/IV.2007.113 pages #8

[63] Tadeusz Stach and TC Nicholas Graham. 2011. Exploring Haptic Feedback in Exergames. In *Human-Computer Interaction – INTERACT 2011*. Springer, Springer, Berlin, Heidelberg, Germany, 18–35. pages #4

[64] Ancret Szpak, Stefan Carlo Michalski, and Tobias Loetscher. 2020. Exergaming With Beat Saber: An Investigation of Virtual Reality Aftereffects. *J Med Internet Res* 22, 10 (23 Oct 2020), e19840. https://doi.org/10.2196/19840 pages #15

[65] Unity Technologies. 2023. Unity Real-Time Development Platform. https://unity.com/ LTS Release 2021.3.22f1. 29 March 2023. https://unity.com/releases/editor/qa/lts-releases?version=2021.3&major_version=2021&minor_version=3&page=1. pages #6

[66] Katie Salen Tekinbas and Eric Zimmerman. 2003. *Rules of play: Game design fundamentals*. MIT press. pages #3

[67] Kim Thomas. 2007. Just noticeable difference and tempo change. *Journal of Scientific Psychology* 2 (2007), 14–20. pages #8, #18

[68] Luke Vecchio, Hays Daewoud, and Shannon Green. 2018. The health and performance benefits of the squat, deadlift. *and bench press. MOJ Yoga Phys. Ther* 3 (2018), 40–47. pages #6

[69] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963 pages #11

[70] World Health Organization. 2022. *Global status report on physical activity 2022*. World Health Organization, Geneva, Switzerland. https://www.who.int/publications/i/item/9789240059153 (visited 4 September 2023). pages #2

[71] World Health Organization. 2022. Physical Activity. https://www.who.int/news-room/fact-sheets/detail/physical-activity visited 10 Ferbuary 2024. pages #2

[72] Wenge Xu, Kangyou Yu, Xuanru Meng, Diego Monteiro, Dominic Kao, and Hai-Ning Liang. 2023. Exploring the effect of the Group Size and Feedback of non-player character spectators in virtual reality exergames. *Frontiers in Psychology* 14 (2023), 18 pages. https://doi.org/10.3389/fpsyg.2023.1079132 pages #5

[73] Betty Yin, Samuel Bailey, Emma Hu, Milinda Jayarekera, Alex Shaw, and Burkhard C. Wünsche. 2021. Tour de Tune 2 - Auditory-Game-Motor Synchronisation with Music Tempo in an Immersive Virtual Reality Exergame. In *Proceedings of the 2021 Australasian Computer Science Week Multiconference* (Dunedin, New Zealand) *(ACSW '21)*. Association for Computing Machinery, New York, NY, USA, Article 11, 10 pages. https://doi.org/10.1145/3437378.3437379 pages #5, #18

[74] Yongqi Zhang, Biao Xie, Haikun Huang, Elisa Ogawa, Tongjian You, and Lap-Fai Yu. 2019. Pose-Guided Level Design. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300784 pages #6

Received February 2024; revised June 2024; accepted July 2024

9.4 [Preprint] Unlocking the Potential: The Role of Game Elements on Player Motivation and Intention of Long-Term Use in Virtual Reality Exergames

[CorePre1] **Sukran Karaosmanoglu**, Bastian Kordyaka, and Frank Steinicke. Unlocking the potential: The role of game elements on player motivation and intention of long-term use in virtual reality exergames. 2025. DOI 10.25592/uhhfdm.16793. URL https://doi.org/10.25592/uhhfdm.16793.

Supplementary materials for this article are included after the full-text of the paper.

Unlocking the Potential: The Role of Game Elements on Player Motivation and Intention of Long-Term Use in Virtual Reality Exergames

Sukran Karaosmanoglu
Human-Computer Interaction
University of Hamburg
Hamburg, Germany
sukran.karaosmanoglu@uni-hamburg.de

Bastian Kordyaka
Åbo Akademi University
Turku, Finland
bastian.kordyaka@abo.fi

Frank Steinicke
Human-Computer Interaction
University of Hamburg
Hamburg, Germany
frank.steinicke@uni-hamburg.de

Figure 1: The impressions from the VR exergame *Beat Saber* (The images are retrieved from Steam Store; credit: Beat Games [5]).

Abstract

Virtual reality (VR) exergames provide enjoyable, immersive workouts, yet there is still a lack of knowledge about how to design these games effectively to unlock their potential. This paper investigates (i) which game elements (basic mechanics of games) contribute to player inspiration (a state-related form of motivation) and (ii) how inspiration influences players' long-term use. To answer these questions, we followed a mixed-methods design. In Study 1, we used an iterative approach to identify the most important VR exergames and their relevant game elements. Based on these findings, in Study 2, we created a survey that we sent to Beat Saber players (N=256). The results showed that nine game elements influence inspired by playing Beat Saber, and inspired to do physical exercise partially mediates the relationship between inspired by playing Beat Saber and long-term use. We conclude with implications that can inform future development in this field.

CCS Concepts

• **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**; • **Software and its engineering** → **Interactive games**.

Keywords

virtual reality, exergames, games, elements, gamification, motivation, Beat Saber

Conference'17, Washington, DC, USA

ACM Reference Format:
Sukran Karaosmanoglu, Bastian Kordyaka, and Frank Steinicke. 2025. Unlocking the Potential: The Role of Game Elements on Player Motivation and Intention of Long-Term Use in Virtual Reality Exergames. In . ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

Physical activity offers numerous health benefits, covering aspects of both physical health [2], psychological well-being [53], and aging [13]. However, we are currently facing the worldwide growing sedentary epidemic—people move less and less [72]. To combat insufficient physical activity, human-computer interaction (HCI) research is increasingly investigating and developing immersive virtual reality (VR) game artifacts, so-called *VR exergames*. These games have great potential to keep humans physically active while also offering enjoyable, immersive experiences.

VR exergames are booming [35]. They have gained traction both in academic research (e.g., [34, 36, 37, 73]) and digital gaming platforms (e.g., [5, 21, 33, 42, 52]). Recent research highlights a significant increase in the number of VR exergames; in a scoping review, Karaosmanoglu et al. [35] identified 186 VR exergames with distinct characteristics and gameplay. Additionally, commercial VR exergames tend to receive high praise from users on digital platforms. Notably, some VR exergames stand out more prominently than others, such as *Beat Saber* [5][1]. However, it remains a challenging question what makes well-selling exergames like *Beat Saber* so successful.

The central goal of every exergame is to motivate players to engage in physical exercise. Research indicates that immersive VR versions of exergames present even more motivating options for

[1]We note that *Beat Saber* requires players to perform rhythmic movements and has been explored as an exergame in many academic publications [35, 63, 74].

physical activity over traditional ones [8]. Similarly, players move more in VR exergames compared to two-dimensional setups [75]. However, it remains unclear what is particularly so motivating about these games. Also, motivation alone is not sufficient to holistically explain exergame play. To address the lack of physical activity, VR exergames have to motivate users over a longer period of time because without a commitment to regular exercise, individuals cannot fully experience the benefits. This long-term retention is a sought-after but rare achievement of most commercial exergames, with a few notable exceptions, such as *Beat Saber* which retain a large and vivid player community. Nevertheless, research [35] indicates a significant lack of exploration regarding the long-term use of VR exergames. Accordingly, we must understand which aspects of these successful exergames (i) motivate users in the first place and (ii) how they contribute to sustained long-term use.

VR exergames are complex systems; they include various design elements beyond mere gameplay mechanics, such as leaderboards and points. A large body of work in games research explores how game elements can be utilized to motivate users in non-gaming contexts—a concept known as gamification [16], motivational affordances [43], and game elements [45]. Since in our work we focus on exergaming environments, we refer to these design elements as *game elements* [16, 29]. Despite the growing success of certain VR exergames and the significant increase in the number of VR exergames, there is currently no comprehensive guidance on which game elements are crucial for enhancing player motivation or which elements developers and designers should incorporate to achieve long-term use in VR exergames.

To explore player motivation, we turn to inspiration theory—an approach that frames motivation as a state-based phenomenon rather than an innate mechanism [64, 65]. Unlike established motivational theories in HCI gaming research, such as self-determination theory (SDT) [58], inspiration theory focuses on external stimuli to evoke motivation. Given our specific interest in the influence of game elements in VR exergames (i.e., external stimuli) on player motivation, rather than the intrinsic desire that individuals might possess for playing these games, we use inspiration theory to investigate motivation (similar to other HCI works [44]). In simple terms, this theory suggests two components, positing that individuals can be *inspired by* external triggers—in our case a VR exergame—and *inspired to* take action—in our case to do physical exercise. Building on these components, our focus is to identify (i) which game elements of VR exergames can inspire players and (ii) how inspiration might impact their long-term to use of VR exergames. Therefore, we formulate our research questions as follows:

- **RQ1:** Which game elements influence inspiration in VR exergames?
- **RQ2:** How does inspiration in VR exergames influence long-term use?

To answer our research questions, we employed a mixed-design approach [14] and conducted two sequential studies. In *Study 1*, we aimed to identify relevant VR exergames and their elements. We did this in three steps: (i) creating a portfolio of VR exergames and game elements based on literature [35, 43] and available online sources [68], (ii) an expert online survey (N=5) to identify available and important game elements of VR exergames, and (iii) an expert interview session (N=2) to refine and provide example for each game element. The experts identified the most played VR exergame and gameplay elements, revealing a collective, anonymous consensus: *Beat Saber*.

In *Study 2*, building on the findings of *Study 1*, we carried out an online survey (N=256) about the most played (decided by our experts) and researched VR exergame [35]: *Beat Saber* [5]. In this survey, we asked players how specific game elements of *Beat Saber* influence their inspiration and long-term use of this game. Our results indicate that nine game elements significantly influence *inspired by playing Beat Saber*. Additionally, *inspired to do physical exercise* partially mediates the relationship between *inspired by playing Beat Saber* and *intention of long-term use*. These empirical insights suggest that certain game elements inspire players and, consequently, encourage them to do physical exercise. Moreover, this contributes to long-term use of VR exergames. We summarize our contributions as follows:

(1) identification of available and relevant game elements used in VR exergames,
(2) derivation of insights into how these game elements contribute to player inspiration in VR exergames,
(3) examination of insights into how player inspiration impacts the long-term intention of using a VR exergame,
(4) design implications for developers, designers, and researchers, such as which game elements to incorporate to enhance inspiration, physical activity, and long-term use.
(5) theoretical implications for researchers, such as the external validity of the theory of inspiration.

Our research enhances the understanding of how to design immersive VR exergames—games that encourage physical activity. We offer an initial exploration of how the fundamental components of VR exergames (i.e., game elements) influence player inspiration and their intent to use the games over the long term. Our findings can inform the development of more effective VR exergames that promote physical exercise and tackle the urgent problem of the sedentary lifestyle pandemic.

2 Related Work

This section provides a summary of research on the benefits of VR exergames and existing game elements, and highlights the importance of motivation and long-term use in the context of VR exergames.

2.1 Virtual Reality Exergames

With an aging population and increasingly sedentary lifestyles, the importance of physical activity in society has grown significantly over the last decades [72]. Research in HCI has focused on developing innovative solutions to promote physical activity [35, 47]. Among others, research investigated non-immersive [11, 25, 67] and immersive versions of exergames [8, 49, 73], which have shown promising effects. Due to their enjoyable and immersive qualities, VR exergames have emerged as one of the most popular genres in the VR gaming market (e.g., *Beat Saber* [5]) [62]. Since they do not require additional exercise equipment and offer a variety of workout options, ranging from dancing (e.g., [21]) to boxing (e.g., [52]), they became alternative, accessible options for exercising [36, 37].

These games actively engage players in physical activities while delivering enjoyable gameplay experiences [51]. Research shows that playing exergames through immersive (i.e., VR) head-mounted displays (HMDs) results in higher levels of motivation [8] and better game performance [73] compared to playing on TV displays.

VR offers endless possibilities for boosting motivation during physical workouts. For example, Michael and Lutteroth [49] leveraged players' past performances in a VR cycling exergame by depicting their previous and projected performances as ghost cyclists competing against their current efforts. Their longitudinal study findings indicated that this method not only enhanced players' physical performance but also increased their motivation. Another instance involves exaggerating players' virtual abilities [9, 34]. In the study of Born et al. [9], participants used a modified controller with a hand gripper to punch approaching characters. When players experienced a virtual augmentation (i.e., exaggeration of players' stroke impact through visuals), they engaged in strenuous activity for a longer duration.

Overall, these advantages have contributed to a growing interest in exergame research in the HCI field [35]. VR exergames have become motivational boosters to provide effective workouts for physical exercise—encouraging people to move more. Consequently, these games have emerged as alternative fitness tools for people of all ages [8, 12, 36–38, 73].

2.2 Game Elements

While general guidelines exist on what to consider when designing exergames—such as dual flow [61] and considering cognitive load [50]—the field lacks a comprehensive bottom-up approach. Therefore, we focus on the game elements (e.g., leaderboards, levels), which are building blocks of games; they shape the overall gaming experience. However, there is no definitive list of the existing game elements [16, 70].

In terms of identifying game elements, the majority of existing research has focused on their use in non-game contexts (e.g., [29, 30, 43]. For example, Hamari et al. [30] reviewed explored the effects of motivational affordances on psychological and outcomes. The authors identified 11 game elements, such as rewards and progress. In a scoping review, Hallifax et al. [29] found 15 game elements utilized in gamification research (e.g., social, feedback). Similarly, Koivisto and Hamari [43] conducted a review that provided an overview of empirical gamification studies, resulting in the identification of 46 game elements. Since not all games incorporate the same elements, certain components may be present in one game but absent in another [16]. As a result, the list of game elements could expand further depending on the specific applications or games being examined.

Games are complex systems composed of multiple elements. Research frequently assesses games as a whole rather than concentrating on individual game elements (e.g., [12, 37]). Although this holistic approach might be beneficial depending on the research objectives, it hinders the ability to determine the effects of specific game elements [15, 48] on outcomes. Understanding the influence of specific game elements on player experiences, however, can greatly enhance the design process for practitioners and researchers. Ultimately, this can lead to the creation of more effective VR exergames, rather than simply producing artifacts that do not target a specific goal.

2.3 Inspiration: A Motivational State

A large body of work in HCI has focused on understanding the role of motivation in games [58, 66]. Predominantly, game research studies draw on SDT to study the motivation of players [58, 66]. According to this theory, satisfying the three basic human needs (i.e., autonomy, competence, and relatedness) can contribute to intrinsic motivation— things that inherently interesting or enjoyable—and well-being [55, 58]. Specifically, autonomy refers to one feeling in control of one's choices, competence can be described as a feeling of mastery in a task, and relatedness considers the sense of connection with another being or group [58]. Although SDT is the most widely used theory in game research [66], its use in the field as a sole theory for addressing motivation-related research questions has been called into question recently [55]. Particularly, Poeller and Phillips [55] urges game scholars to engage in additional motivational theories that may be more suitable for their specific research questions.

Since we are specifically interested in the impact of VR exergames' game elements on player motivation, rather than the inherent motivation that individuals may have for playing these games, we used inspiration theory [64]. This theory views inspiration as a more state-based, situational form of motivation, in contrast to STD, which regards it as a trait-based, inherent construct. In simple terms, inspiration can be defined as a strong impulse—a motivational state—that occurs in response to a stimulus (i.e., trigger) and drives people to take action (i.e., target) [64, 65].

According to Thrash and Elliot [65], inspiration consists of three key characteristics: transcendence, evocation, and motivation. Transcendence involves becoming aware of new possibilities as a result of inspiration. Evocation means that inspiration is not initiated or controlled by the individual. Lastly, motivation pertains to the impulse to express or act upon a newly acquired idea or concept inspired by this experience. Inspiration can be classified into two components: *inspired by* and *inspired to*. The former concept pertains to the appreciation of the external trigger, while the latter encompasses the motivation to engage in actions and behaviors as a result of that evocative object [65]. Therefore, inspired by is considered as the antecedents of inspired to.

Overall, inspiration theory has been applied in research across various fields, including marketing [41], games [44], sports [20], and fitness contexts [60]. Kordyaka et al. [44] showed that games can serve as a source of inspiration, while Sandercock et al. [60] found that inspiration can result in increased levels of physical activity. Although previous research showed that several factors in VR exergames (e.g., emotional characters [26], augmentation [9]) can encourage more physical activity, there remains a gap in understanding whether the inspiration gained from playing VR exergames (i.e., inspired by) inspires people to do physical exercise (inspired to), and how this motivation subsequently influences users' long-term intentions to continue using VR exergames.

2.4 Long-Term Intention to Use

The World Health Organization (WHO) recommends that adults to engage in regular physical activity [71]; this can have positively affect physical and mental health [72]. For example, regular exercise can significantly reduce the likelihood of developing serious health conditions, including cardiovascular diseases [71, 72] and obesity [54, 69].

One of the primary goals of VR exergames is to provide benefits, such as improving physical training performance (e.g., [12]), or providing relaxation training (e.g., [3]). However, in line with WHO recommendations, these benefits are only possible through constant, long-term engagement. Since motivation can support long-term engagement in physical activity [59], VR exergames, due to their motivational nature, are believed to foster long-term engagement. To understand players' long-term intention to use VR exergames, several studies have asked participants about their long-term usage intentions regarding the games studied (e.g., [12, 74]). Specifically, Xu et al. [74] examined the acceptance of commercial VR exergames among older Chinese adults and identified several factors, such as perceived usefulness and enjoyment, that influence their intention to use these systems. However, to our knowledge, there has been no research investigating how the motivational nature of VR exergames may impact long-term usage intentions.

Overall, so far, there has been insufficient attention to the long-term use of VR exergames [35]. However, understanding long-term use intention is crucial for achieving the health objectives associated with VR exergames. Without insights into which game elements affect psychological constructs (i.e., inspiration) and their subsequent long-term use, we cannot effectively design VR exergames that provide benefits to players. Therefore, we echo the concerns expressed by Kayali et al. [40], who emphasizes the lack of clear guidelines for designing health-focused games for long-term use. While the authors provide considerations (e.g., relating to everyday habits and daily routines) using a research-through-design approach [40], we use a more quantitative lens and investigate how game elements and player inspiration influence the long-term intention to use VR exergames.

3 Research Focus & Mixed-Methods Design

VR exergaming offers opportunities for physical exercise and is recognized for its motivational appeal [8]. However, it remains unclear how specific game elements in these experiences impact player motivation, specifically regarding being *inspired by playing a VR exergame* and *inspired to do physical exercise*. Moreover, long-term regular use (like in every physical activity [71, 72] is crucial to have benefits from these games, yet we do not know if being *inspired by playing a VR exergame* and *inspired to do physical exercise* affects the *long-term intention of use*. Investigating these questions can lead to a more purposeful design of VR exergames and offer guidelines for developers and researchers alike to create effective games that target specific outcomes. For this, we hypothesize positive relationships between all three constructs that already were shown in previous work for another context [44]. Consequently, we outline our hypotheses in this research as follows:

H1: *Inspired by playing Beat Saber* is positively connected with *long-term intention of use.*

H2: *Inspired by playing Beat Saber* is positively connected with *inspired to do physical exercise.*

H3: *Inspired to do physical exercise Beat Saber* is positively connected with *long-term intention of use.*

To investigate our research questions and hypothesis, we used an exploratory sequential orientation of a mixed-methods study design [14]. This approach involves the collection of both qualitative and quantitative data, with qualitative data being gathered first for exploration to develop an instrument or identify variables, followed by quantitative data collection to test these instruments and variables. Our research consisted of two studies: (i) *Study 1:* identifying relevant VR exergames and meaningful game elements—qualitative—and *Study 2:* administering an online survey aimed at general users of the selected game from 110 VR exergame titles by experts in *Study 1*, *Beat Saber*—quantitative—. Overall, we illustrate our employed mixed-methods design approach for this study in Figure 2.

4 Study 1: Identification of VR Exergames & Game Elements

VR games offer inherently spatial interactions, meaning that most of them necessitate some degree of movement to play. According to a recent review [35], exergames can target different levels of physical activity, for example, based on the target user group. Similarly, there is no definitive list for game elements [16, 70], especially not for VR exergames. Moreover, every game contains certain types of gameplay elements [16]. Therefore, our goal in *Study 1* was to identify relevant VR exergames available in the commercial market along with their game design elements.

4.1 Method

In the following sections, we describe our data analysis and procedures, outline the data collection and measurements, and present the sample characteristics of our *Study 1*.

4.1.1 Data Analysis & Procedure. We approached our *Study 1*'s goal through three sequential steps: (i) creating a portfolio of VR exergames and game elements, (ii) conducting an expert survey and (ii) an expert interview session. Overall, we illustrate our steps of *Study 1* in Figure 3.

In the first step, we conducted searches to identify a list of VR exergames and game elements (*Step 1*). Next, in *Step 2*, we employed the identified lists of VR exergames and game elements, and invited experts to assess the availability and importance of each game element. Based on their evaluations, we streamlined our both lists and focused on a single game—selected by all of our experts—to evaluate its game elements: *Beat Saber*. Finally, in *Step 3*, we held an expert interview session to refine the game elements list (e.g., by incorporating examples for each element) from the previous step, resulting in a finalized collection of game elements.

4.1.2 Data Collection & Measurements. In the first step, we identified a large portfolio of VR exergames and game elements (*Step 1*) through literature and online database searches.

In *Step 2*, we recruited five experts via convenience sampling method and invited them to participate in an online survey. In the survey, we used identified VR exergames and design elements in

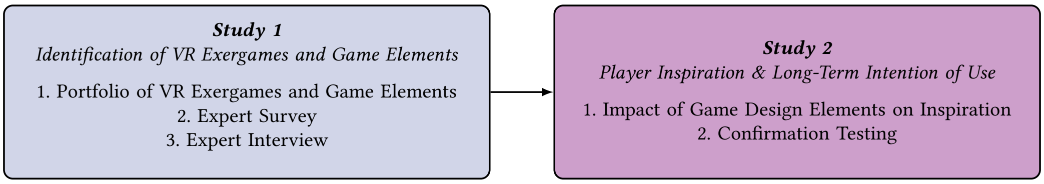

Figure 2: The mixed-methods design approach followed for this research.

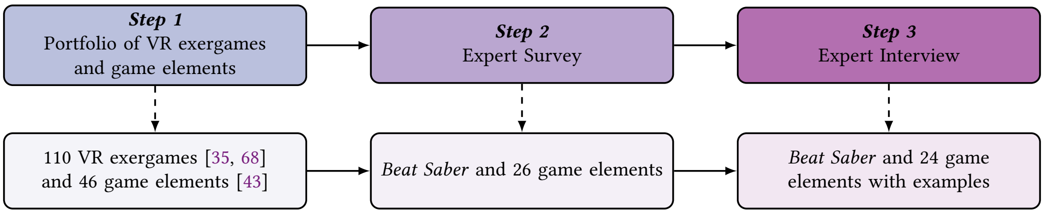

Figure 3: The upper part illustrates each step in *Study 1*, while the lower part presents the results obtained from each step.

Step 1, and asked experts to select their most played VR exergame and its available important elements (rating options: *important, not important, not sure, not applicable*). Since every game has unique design elements [16], we kept this question to be game-specific. Consequently, the participant selected only one game and answered the game elements questions related to this game. To ensure that we captured all relevant design elements specific to VR exergames, we also had an open-ended question for experts, asking them to note any important game design elements that were not included in our list.

In *Step 3*, we presented the lists of game elements from the previous step (i.e., Step 2) to the participants who are VR exergame researchers. They were prompted to discuss the identified elements, exclude or add elements, and provide examples for each element.

4.1.3 Sample Characteristics. Here, we describe our sample characteristics for the empirical parts of *Study 1*: *Step 2* and *Step 3*. For *Step 2*, we recruited experts who have access to VR hardware and play/have played exergames. In total, we reached out to five experts (3 women, 2 men) who work in the area of VR and play exergames. These experts had either a master's (n=4) or a higher degree (n=1), and their mean age was 31.40 (SD=1.82) years. On average, they reported having 8.4 (SD=2.61) years of contact with VR (*Min-Max*=4–10). Referring to this, they described their role in interacting with VR applications at various levels: user (n=1), player (n=4), developer (n=5), and designer (n=2). In *Step 3*, we recruited two experts; one from the research team and another from outside the team. Overall, this group comprised VR exergame researchers (1 woman, 1 man, M=30 years) who identified themselves as VR exergame developers, designers, and players (M=8 years of VR experience). They reported playing VR exergames on a weekly basis.

4.2 Results

In the following, we present the results of three steps of *Study 1*.

4.2.1 Step 1: Portfolio of VR Exergames and Game Elements. Mueller et al. [51] defines exergames as "*as a digital game[s] where the outcome of the game is predominantly determined by physical effort*". Since every VR game inherently includes spatial interaction, and thereby movement [35], it is challenging to decide what qualifies as an exergame and what does not. Therefore, to identify the most played VR exergames, we compiled a comprehensive list of 110 VR games that are considered as an exergame in research (in the scoping review of Karaosmanoglu et al. [35]) or in online sources [68]. The full list of VR exergames can be found in Table 4.

Games are complex and include a multitude of game elements. To generate a comprehensive list of design elements, we used the extensive list of Koivisto and Hamari [43] (similar to other works [44]) that includes 46 game design elements—the authors refer to them as affordances—identified in a literature review. Koivisto and Hamari [43] categorizes these design elements based on five categories, with each category having several game design elements: (i) achievement / progression (ten features (e.g., points, score, XP)), (ii) social (seven features (e.g, competition)), (iii) immersion (five features (e.g., avatar, character, virtual identity)), (iv) non-digital elements (eight features (e.g., real world / financial reward)), and (v) miscellaneous (sixteen features (e.g., warnings)).

4.2.2 Step 2: Expert Survey. All experts reported *Beat Saber* [5] as their most played VR exergame; we had one clear champion. Overall, 24 game design elements from our list were rated as important by at least one expert. As a result of this step, we excluded the 22 features: quizzes, social networking features, cooperation, collective voting, avatar, narrative, role play, real world / financial reward, check-ins, physical cards, physical cards, physical playboard, real world interactive objects, physical objects as game resources, physical dice, virtual currency, reminders, game slogans, funny movies, virtual pets, trading, making suggestions, and virtual objects as augmented reality. Additionally, experts reported the following game design elements that were not part of the initial list as important: music / rhythm (n=2) and pop-cultural integration (e.g., known songs) (n=1). As a result, this step of the analysis identified a total of 26

Table 1: List, examples, and sources of design elements remaining after *Study 1*; expert survey (ES) and interview (EI).

Design Elements	Examples
Achievement/progression [43]	
Points, score, XP	points based on cuts of cubes
Challenges, quests, missions, tasks, clear goals	cutting cubes of certain colors with certain colored beams
Badges, achievements, medals, trophies	badges like precision master
Leaderboards, rankings	rank A, rank B
Levels	expert level
Performance stats, performance feedback	text telling missed cubes, precision points
Progress, status bars, skill trees	campaign mode with certain missions to complete
Timer, speed	song duration, adjusting the speed
Increasing difficulty	modifiers, varying number of cubes at each difficulty level
Social [43]	
Competition	player vs. player competition
Customization, personalization	color customization of sabers
Multiplayer	playing with friends
Immersion [43]	
Virtual world, 3D world, game world	every song pack has a unique environment, dynamic environment
In-game rewards	combos
Non-digital elements [43]	
Motion tracking	playing with VR controllers
Miscellaneous [43]	
Assistance, virtual helpers	accessibility options, removing visual effects
Retries, health, health points	missing the cubes decreases health
Onboarding, benefits for beginners	tutorial, practice mode
Warnings	visual effects for collisions with blocks that must be avoided
Penalties	when you cut the wrong cube, the combo resets
Additional elements (ES & EI)	
Music / rhythm	gameplay accompanied with songs
Audiovisual congruence	visual elements match and change with music change
Downloadable contents / regular additions	music packs
Juicy effects and cues	visual embellishments, haptic feedback

game design elements and led us to focus exclusively on *Beat Saber* in the following steps.

4.2.3 Step 3: Expert Interview. Based on a discussion, the participants who are VR exergame researchers (one outside of the research team and one not) decided to exclude five game elements (e.g., full-game, game rounds) from *Step 2*'s game element list. For example, full-game was excluded because it does not describe a game element. However, considering *Beat Saber*, the participants also recognized the need to enhance the game elements list by incorporating new elements that could be important: juicy effects and cues, audiovisual congruence, and downloadable contents / regular additions. The resulting final list including a total of 24 game design elements and their examples can be found Table 1.

5 Study 2: Player Inspiration & Intention of Long-Term Use

Study 1 narrowed down the list of VR exergames to a single title, *Beat Saber* [5], confirming the popularity of this game in the commercial market [62] as well as VR exergame research [35]. Similarly, our experts in *Study 1* refined the game elements list of Koivisto and Hamari [43] to adapt it to the VR exergame context. After determining our VR exergame and game elements through a data-driven approach, we advanced to our *Study 2.* The objective of *Study 2* was to investigate two key aspects: (i) the specific game elements in *Beat Saber* that significantly contribute to *inspired by from playing Beat Saber*, and (ii) the extent to which this inspiration influences the *long-term intention of use.*

5.1 Method

In the following sections, we describe our stimuli, present our data analysis and procedure, outline the data collection and measurements, and detail the sample characteristics of our *Study 2.*

5.1.1 Stimuli. Beat Saber is a VR exergame developed by Beat Games and released in 2019 [5]. This music and rhythm-based game features a stylized, abstract theme with cubes and rectangles. Players are tasked with slicing through approaching items using sabers in each hand. Additionally, the game includes bombs and rectangles that must be avoided. Overall, *Beat Saber* receives high praise from users [5] and is considered one of the best-selling VR exergames available [62]. In addition to its commercial success, *Beat Saber* has been the focus of several academic studies—most researched commercial VR exergame [35]—due to its popularity [4, 27, 46, 63]. Researchers are exploring various aspects of the game, including the aftereffects of gameplay [63] and the effect of gameplay on cognitive performance [27].

5.1.2 Data Analysis & Procedure. To test our hypotheses in *Study 2*, we conducted an online cross-sectional survey to gather responses from *Beat Saber* players. We conducted all our analyses using R[2]. In summary, our data analysis involved two main steps: (i) testing

[2] https://www.r-project.org/

Table 2: The descriptive values of each game element in *Study 2*; expert survey (ES) and interview (EI).

Design Elements	N	Min	Max	Mean	SD
Achievement/progression [43]					
Points, score, XP	256	1	5	4.04	0.94
Challenges, quests, missions, tasks, clear goals	256	1	5	4.05	1.06
Badges, achievements, medals, trophies	256	1	5	3.75	1.14
Leaderboards, rankings	256	1	5	3.68	1.20
Levels	256	1	5	4.36	0.77
Performance stats, performance feedback	256	1	5	4.19	0.83
Progress, status bars, skill trees	256	1	5	3.82	1.13
Timer, speed	256	1	5	3.91	1.01
Increasing difficulty	256	1	5	4.14	0.79
Social [43]					
Competition	256	1	5	3.73	1.21
Customization, personalization	256	1	5	3.68	1.10
Multiplayer	256	1	5	3.73	1.30
Immersion [43]					
Virtual world, 3D world, game world	256	1	5	4.10	0.94
In-game rewards	256	1	5	4.04	0.92
Non-digital elements [43]					
Motion tracking	256	2	5	4.44	0.79
Miscellaneous [43]					
Assistance, virtual helpers	256	1	5	3.48	1.22
Retries, health, health points	256	1	5	3.91	0.94
Onboarding, benefits for beginners	256	1	5	3.87	1.09
Warnings	256	1	5	3.93	1.03
Penalties	256	1	5	3.56	1.05
Additional elements (ES & EI)					
Music / rhythm	256	2	5	4.45	0.82
Audiovisual congruence	256	1	5	4.25	0.86
Downloadable contents / regular additions	256	1	5	3.93	1.05
Juicy effects and cues	256	1	5	3.99	0.91

the exploratory potential of the identified relevant game elements on *inspired by* and (ii) examining a mediation model concerning the relationship between *long-term use* and inspiration.

For the exploratory testing, we specifically utilized the *stats* [56], *lm.beta* [6], *lmtest* [76], and *car* [23] packages to carry out stepwise backward multiple regression analysis and to evaluate the associated assumptions [19]. We chose the stepwise backward multiple regression analysis due to two reasons: (i) the presence of a large number of predictors (i.e., game elements) and (ii) the consideration of new design elements in our survey beyond those previously proposed in the literature by Koivisto and Hamari [43]. The backward approach of this regression model begins by considering all potential predictors in the model and iteratively refines it by removing predictors that do not significantly contribute to the model's explanatory power. The backward approach is recommended as it allows for the consideration of potential interactions among predictors while minimizing the risk of having a Type II error [19].

To test the mediation effect of *inspired to do physical exercise* on the relationship between *inspired by playing* Beat Saber and *intention of long-term use* (i.e., hypothesis testing), we used Hayes [31]'s PROCESS macro mediation model[3].

Sample codes for our analysis can be found in Appendix C.

[3]https://www.processmacro.org/download.html

5.1.3 *Data Collection & Measurements.* To reach out to people who are playing *Beat Saber*, we used the participant recruitment platform Prolific[4]. This platform was selected because in this platform, participants provide high-quality data [17, 18] and it has been used in previous HCI research [32, 44]. Additionally, it complies with the data privacy regulations established by the authors' university.

Following the best practices, we searched papers and identified studies that measured the same constructs relevant to our research [10, 22, 39, 43, 44, 64]. Using these findings, we developed an online survey incorporating empirically validated scales and items. All items used in the survey with the exact phrasings can be found Table 5.

Overall, our survey started with a brief introduction outlining the study's objectives and included an informed consent form. Following that, participants answered demographic questions about age, gender, education level, and usage of VR headsets and VR exergames. Then, participants rated 24 game elements with their examples relating to their importance on a 5-point Likert scale (from *not important at all/1* to *very important/5*). These elements were selected based on findings from *Study 1* (adapted from [43]). However, we did not present the broad categories of each game elements to participants. Afterwards, participants responded to items

[4]https://www.prolific.com/

Table 3: Stepwise backward multiple linear regression analysis; significant results for inspired by playing *Beat Saber*.

Design Elements	B	SE	t	β	p
Achievement/progression [43]					
Points, score, XP	0.18	0.08	2.346	0.129	0.020
Badges, achievements, medals, trophies	0.27	0.07	3.539	0.226	<0.001
Levels	0.22	0.09	2.333	0.124	0.020
Performance stats, performance feedback	-0.20	0.10	-2.073	-0.123	0.039
Progress, status bars, skill trees	-0.14	0.07	-2.024	-0.122	0.044
Social [43]					
Competition	0.21	0.06	3.326	0.191	0.001
Customization, personalization	0.21	0.07	3.118	0.176	0.002
Immersion [43]					
Virtual world, 3D world, game world	0.20	0.08	2.584	0.138	0.010
Miscellaneous [43]					
Assistance, virtual helpers	0.18	0.06	2.796	0.161	0.006

B = unstandardized regression coefficient, SE = Standard Error, t = t-value, β= standardized regression coefficient, p = p-value.

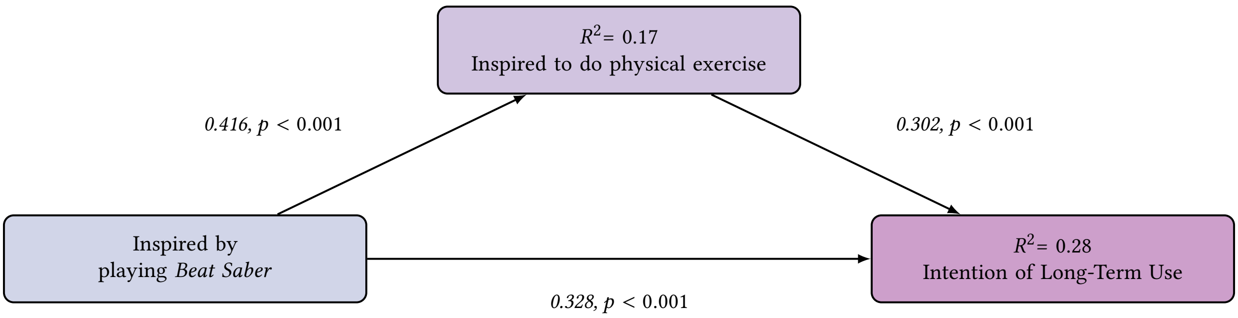

Figure 4: Mediation model describing the relationships between inspired by, inspired to, and intention of long-term use.

measuring the constructs inspired by *Beat Saber* (e.g., *"While playing Beat Saber, my imagination was stimulated"*; α= 0.91, *M*=5.06, *SD*=1.34) and inspired to do physical exercise (e.g., *"After playing Beat Saber, I was inspired to do physical exercise"*; α= 0.96, *M*=5.21, *SD*=1.48). These ten items were designed based on [10, 44, 64] and were rated on a 7-point Likert scale (*strongly disagree/1 — strongly agree/7*). Finally, participants rated four items related to their long-term intention to use *Beat Saber* (e.g., *I intend to play Beat Saber in the future*; α= 0.87, *M*=6.20, *SD*=0.85), also using a 7-point Likert scale (*strongly disagree/1 — strongly agree/7*); these items were from [22, 39].

5.1.4 Sample Characteristics. Overall, we received a total of 310 complete responses. However, after filtering out nonsensical responses (e.g., using Apple Vision Pro to play *Beat Saber*), spending unconsiderate time on the survey (i.e., less than four minutes), and those failing attention checks [28], we were left with 256 valid responses for our subsequent analysis. All participants received 1£ as compensation for their participation in the study.

Out of 256 participants, 97 reported their gender as women, 158 as men, and one as non-binary. The mean age was 31.51 years (*SD*=9.78). The majority of participants reported having completed a Bachelor's degree (*n*=126): followed by a high school (*n*=59), a Master's degree (*n*=56), a degree higher than a Master's (*n*=8), and middle school (*n*=7). Overall, participants reported playing *Beat Saber* for an average of 1.84 years (*SD*=1.50). Regarding frequency of play, the majority of participants (*n*=139) indicated that they played *Beat Saber* frequently (1-2 times per month), followed by often (1-2 times per week, *n*=66), rarely (1-2 times per year, *n*=26), every day (*n*=15), seldom (1-2 times total, *n*=10). To play *Beat Saber*, they have used the following headsets: Meta Quest headsets (*n*=204), PlayStation VR variants (*n*=107), HTC Vive headsets (*n*=23), Valve index (*n*=7), HP Reverb G2 (*n*=2), and Samsung HMD Odyssey (Mixed Reality) (*n*=1).

5.2 Results

In this section, we present our results regarding (i) the exploratory analysis of the identified 24 game elements' impact on *inspired by*, and (ii) the hypothesis testing related to our dependent variable, *long-term use*. Table 2 lists the descriptive statistics for each game design element.

5.2.1 Impact of Game Design Elements on Inspiration. According to Thrash and Elliot [64, 65], the state of *inspired by* serves as the antecedent to *inspired to*. In simpler terms, individuals get inspired by an evocative object, which then prompts them to take action as a result (i.e., inspired to). Therefore, we analyze the influence of game elements on *inspired by* in the following.

To analyze the impact of the identified game design elements (i.e., independent variables) on the *inspired by* construct (i.e., dependent variable), we initially examined the assumptions necessary for conducting a parametric multiple linear regression analysis.

Specifically, we verified the linearity with scatter plots (i.e., Q-Q residuals), checked for autocorrelation with the Durbin-Watson test (DW=2.074, p=0.728), and evaluated the multi-collinearity using variance inflation factors (all ≤ 1.896).

After confirming the assumptions, we ran a stepwise backward regression to identify significant predictors [19]. The final model ($F(13, 242) = 17.026$, $p<0.001$) explained 45% of the variance of the dependent variable inspired by playing *Beat Saber* (adjusted R^2=0.449). It contained nine significant predictor weights: *Points, score, XP* (β=0.129, p=0.020), *Badges, achievements, medals, trophies* (β=0.226, $p<0.001$), *Levels* (β=0.124 , p=0.020), *Performance stats, performance feedback* (β=-0.123, p=0.039), *Progress, status bars, skill trees* (β=-0.122, p=0.044), *Competition* (β=0.191, p=0.001), *Customization, personalization* (β=0.176, p=0.002), *Virtual world, 3D world, game world* (β=0.138, p=0.010), and *Assistance, virtual helpers* (β=0.161, p=0.006). However, all other five predictor weights did not reach to a significance: all p values ≥ 0.10. For readability reasons, we report all the significant results with their additional descriptives in Table 3.

5.2.2 Hypothesis Testing. To analyze the mediation effect of *inspired to do physical exercise* on the relationship between *inspired by playing Beat Saber* and *intention of long-term use*, we first tested the assumptions of conducting PROCESS macro mediation model [31]. Specifically, we confirmed the normality with scatter plots, and tested autocorrelation (DW=2.12, p=0.838) and multicollinearity (≤ 1.210).

Following the confirmation of assumptions, we conducted the mediation analysis. There was a significant direct effect of *inspired by playing* Beat Saber on *intention of long-term use* (β=0.328, $p<0.001$). Additionally, *inspired by playing Beat Saber* positively predicted *inspired to do physical exercise*, β=0.416, $p<0.001$. Similarly, *inspired to do physical exercise* positively predicted *intention of long-term use*, β=0.302, $p<0.001$. The significance of the indirect effect was assessed using bootstrapping with 5000 samples. The completely standardized indirect effect of *inspired to do physical exercise* on *intention of long-term use* was between 0.06 and 0.20 (95% CI). According to Hayes [31], this interval must not include zero. Therefore, our results indicate that *inspired to do physical exercise* significantly mediated the relationship between *inspired by playing Beat Saber* and *intention of long-term use*. These findings support our hypotheses, which proposed a positive relationship between *inspired by playing Beat Saber*, *inspired to engage in physical exercise*, and *intention of long-term use* (refer to the hypotheses in section 3).

6 Discussion

In our research, we used a bottom-up approach and disassembled a VR exergame into its building blocks. Accordingly, we examined the game elements that motivate—i.e., inspire—players in VR exergames, and consequently influence physical activity and long-term use. To achieve this, we conducted two sequential studies: *Study 1* focused on identifying VR exergames and their respective game elements, while *Study 2* quantitatively assessed the impact of these game elements on inspiration and long-term use. Here, we discuss our results by focusing on our research questions. Based on the insights we gathered, we provide implications for VR exergame researchers, designers, and developers to lay a groundwork for the purposeful design of VR exergames.

6.1 Which game elements influence inspiration in VR exergames?

In our *Study 1*, we identified 24 relevant game elements that seem particularly relevant in VR exergames. These game elements were identified as a result of three consecutive steps involving experts. Since we asked players about the importance of design elements without making a reference to inspiration, this list of game elements can be used by designers and developers alike to encourage VR exergame players to achieve their desired target states. Particularly, using these elements, researchers can investigate how they can various desired target states (e.g., physical exertion [7], several experiential outcomes of VR exergames through validated scales [1]).

Our findings in *Study 2* reveal that not all game elements significantly influence players' inspiration while playing *Beat Saber*. In particular, our results emphasize the prominence of the achievement / progression category over other categories identified by Koivisto and Hamari [43]. The achievement / progression category contributes five elements to the model, while the social category adds two and the immersion and miscellaneous categories contribute just one element each. However, we also highlight that in Koivisto and Hamari [43]'s list of game elements, the achievement / progression category contains more items than the other categories. Thus, this difference may have provided more nuanced insights for this category.

Furthermore, our identified nine significant game elements (*Points, score, XP*, *Badges, achievements, medals, trophies*, *Levels*, *Performance stats, performance feedback*, *Progress, status bars, skill trees*, *Competition*, *Customization, personalization*, *Virtual world, 3D world, game world*, *Assistance, virtual helpers*), as independent variables to explain *inspired by playing Beat Saber* in our *Study 2*, offer valuable insights for VR exergames. On a level of content, we find a nuanced picture showing that while two game elements (*Performance stats, performance feedback* and *Progress, status bars, skill trees*) negatively influence *inspired by playing Beat Saber*, the remaining seven elements (*Points, score, XP*, *Badges, achievements, medals, trophies*, *Levels*, *Competition*, *Customization, personalization*, *Virtual world, 3D world, game world*, *Assistance, virtual helpers*) affect this construct positively. Accordingly, we recommend considering these elements when implementing VR exergames to target player inspiration.

6.2 How does inspiration in VR exergames influence long-term use?

Our confirmatory hypothesis test showed that *inspired to do physical exercise* partially mediated the relationship between *inspired by playing Beat Saber* and *long-term use*. Specifically, we find empirical support for all of our hypotheses by showing that all variables are positively connected with each other. This result suggests that inspiration from playing *Beat Saber* may lead to a greater willingness to engage in physical activity and to use *Beat Saber* repeatedly over time.

In general terms, with our results, we show the external validity of the inspiration theory and transfer the results of other contexts (e.g., location-based games [44] and watching Olympics games [60]) to the VR exergames realm. Therefore, our results open up the possibility of using inspiration as a measure to investigate motivation in the context of the evocative object VR exergames.

Our mediation analysis revealed **(H1)** *inspired by playing Beat Saber* is positively connected with *long-term intention of use*. This relation suggests fostering inspiration through playing *Beat Saber* ensures players' engagement with a repeated play of the game. Moreover, our results support our **(H2)**: *Inspired by playing Beat Saber* is positively connected with *inspired to do physical exercise*. This validates the inspiration theory [64, 65] and shows that *inspired by* behaves like a predictor of *inspired to*. In simpler terms, if people got inspired by playing *Beat Saber*, they also got inspired to perform physical exercise. This aligns with the core goal of exergame research: promote physical activity [35]. Finally, we confirm our last hypothesis: **(H3)** *Inspired to do physical exercise Beat Saber* is positively connected with *long-term intention of use*. This implies that players who feel inspired to engage in physical exercise are also more likely to intend to use related resources—in our case *Beat Saber*–over the long term. Our confirmatory analysis provides an additional layer to existing approaches that identify predictors of long-term use of VR exergames, such as with Chinese older adults [74]. Our results strengthen the external validity of research efforts, particularly given that the majority of VR exergame users are young adults like in our sample.

In summary, our findings contribute to the ongoing development of innovative solutions addressing the challenges of insufficient physical activity and sedentary lifestyles [72]. These findings represent valuable information for developers and designers and suggest that integrating the specific game elements that inspire players could effectively address the problem of insufficient physical activity in today's society. Simply, we demonstrate that VR exergames have the potential to inspire players, making them an effective tool for tackling significant societal issues, which challenges the common assumption that games are solely for entertainment.

6.3 Implications for Design & Theory

Here, we present five implications for both design and theory to inform the development of future VR exergames.

6.3.1 Enhance player competence and minimize frustration. In our *Study 2*, the identified game elements such as *points*, *badges*, and *levels* help players feel accomplished and celebrate their successes. However, the negative influence of game elements *Performance stats, performance feedback* and *Progress, status bars, skill trees* may be explained by the feeling of potential frustration. For example, since *Performance stats, performance feedback* and *Progress, status bars, skill trees* emphasize the quantitative measurement of success or failure, players may fear poor performance and feel anxious. In *Beat Saber*, players particularly receive immediate feedback on the precision of their cuts of cubes. If their performance does not meet their expectations, seeing their low precision scores can lead to disappointment. Rather than encouraging inspiration, this can cause anxiety or stress. This focus on performance can distract from creative, exploratory or emotional engagement, that is often the source of inspiration [64, 65].

6.3.2 Employ competition. Competition is a fundamental aspect of numerous games across various genres (e.g., [57]), as well as a key element in most sports and exercise activities (e.g., swimming, football). This is also reflected in VR exergames. In the context of *Beat Saber*, competition can take various forms, including player vs. environment, player vs. player, and multilateral competition [24]. Our analysis indicates that the *Competition* element positively influences inspiration. However, we cannot fully conclude whether all specific forms of competition exert this effect. Therefore, we encourage further research to investigate how different types of competition affect inspiration.

6.3.3 Create or provide adjustment possibilities for the game environment. Audiovisual appeal plays a crucial in shaping player experiences [1]. Both *Customization, personalization* and *Assistance, virtual helpers* positively inspire players. *Beat Saber* offers a variety of features, including accessibility options and diverse color schemes for game objects. We believe that these customization options contribute to an appealing environment and satisfaction for players. Similarly, *Virtual world, 3D world, game world* inspired participants as well, which we attribute to *Beat Saber*'s abstract theme and vibrant neon color combinations, complemented by spatial visual effects. Therefore, we recommend developing visually appealing VR exergames (considering specific game element implementations of *Beat Saber*, such as dynamic environment) and providing customization options (e.g., color customization for game objects in *Beat Saber*) to accommodate players' personal preferences.

6.3.4 Develop a comprehensive list of game elements. We see our results as an indication that the original list of game elements used by Koivisto and Hamari [43] shows room for improvement, as we identified four additional elements (i.e., *Music / rhythm*, *Audiovisual congruence*, *Downloadable contents / regular additions*, and *Juicy effects and cues*) as a result of our expert survey and interview session that are relevant in the context of VR exergames. However, we note that the list of Koivisto and Hamari [43]'s game elements covers the gamification context, not games in particular. Consequently, we recommend that future work should strive to derive a comprehensive list of design features so that none of the elements are overlooked in empirical studies. On the other hand, none of the additional elements identified by experts showed a significant effect on *inspired by* in our confirmatory analyses. We think that this ambivalence happens because inspiration describes a state that goes beyond the gameplay experience (e.g., players get new ideas) [64, 65], and that the importance of a design element (e.g., *Music / rhythm*, *Audiovisual congruence*) stays within the game.

6.3.5 Understand why. Our primarily quantitative research identifies specific game elements that positively and negatively influence inspiration. Additionally, it demonstrates the connection between inspiration and long-term reuse. However, like many studies focused on quantitative methods, it falls short of explaining the underlying reasons for these findings. For instance, while we know that *Badges, achievements, medals, trophies* significantly enhances inspiration, we do not know for sure why this is the case. Is it because players enjoy earning badges, and/or because they like feeling competent? To address this gap, we advocate for future research that explores the reasons behind our quantitative results. We recommend employing qualitative methodologies to gain a more in-depth understanding of player inspiration and long-term usage patterns.

6.4 Limitations

In our *Study 2*, we employed *Beat Saber* to examine how specific game elements influence player inspiration and long-term use. While using a single commercial or self-implemented VR exergame for investigation (like in numerous other studies [36, 63]) limits the generalizability of the results, it also reduces the likelihood of creating confounds, since every game might have different, distinct game elements [16]. Still, this decision was guided by expert recommendations and the recognition that *Beat Saber* is one of the most successful commercial exergames [35, 62] and was played the most by all experts in *Step 2* of our *Study 1*. However, we also acknowledge that there are other successful VR exergames, such as Fit XR [21]. As with any game, these alternatives feature additional types of game elements. Consequently, our research may have overlooked certain game elements that are important in other games.

For data collection, we used Prolific. Although there are concerns about data quality associated with crowdsourcing platforms, numerous studies suggest that participants on Prolific tend to provide high-quality responses [17, 18]. Nonetheless, to further enhance data quality in our study, we implemented attention checks, set thresholds for the time spent on the survey, and excluded ambiguous responses. We also recommend that future research replicate this study using samples recruited through alternative methods.

7 Conclusion

VR exergames offer an innovative approach to physical exercise. However, our understanding of how to design them remains limited. In our paper, we addressed this issue by employing a bottom-up approach and conducted mixed-design two studies. We investigated (i) how specific game elements foster inspiration in the first place and (ii) how inspiration contributes to long-term use of VR exergames. In *Study 1*, we identified the most important VR exergames and their game elements through a three-step process: developing a portfolio of VR exergames and game elements, conducting an expert survey, and conducting an interview session with experts. At the end of this study, we identified 24 relevant design elements. Based on our findings, in *Study 2*, we designed and distributed a survey to *Beat Saber* players (N = 256). Our findings indicate that while seven game elements positively influence *inspired by playing Beat Saber*, two elements have a negative influence on this construct. Moreover, we found that *inspired to do physical exercise* partially mediates the relationship between *inspired by playing Beat Saber* and *long-term use*, providing empirical support for inspiration theory and paths to *long-term use*.

Acknowledgments

This project was supported by the European Union's Horizon Europe research and innovation program (grant number: 101135025) and the grant (number: 16SV8878) from the German Federal Ministry of Education and Research (BMBF).

References

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370. doi:10.1016/j.ijhcs.2019.102370

[2] Simon Adamson, Ross Lorimer, James N. Cobley, Ray Lloyd, and John Babraj. 2014. High Intensity Training Improves Health and Physical Function in Middle Aged Adults. *Biology* 3, 2 (2014), 333–344. doi:10.3390/biology3020333

[3] Dalya Al-Mfarej, Jennifer M. Vojtech, Serge H. Roy, Elise Townsend, Julie J. Keysor, Nancy Kuntz, Vamshi Rao, Joshua C. Kline, and Bhawna Shiwani. 2023. A Virtual Reality Exergame: Clinician-Guided Breathing and Relaxation for Children with Muscular Dystrophy. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 270–276. doi:10.1109/VRW58643.2023.00065

[4] Iannis Albert, Nicole Burkard, Dirk Queck, and Marc Herrlich. 2022. The Effect of Auditory-Motor Synchronization in Exergames on the Example of the VR Rhythm Game BeatSaber. *Proc. ACM Hum.-Comput. Interact.* 6, CHI PLAY, Article 253 (oct 2022), 26 pages. doi:10.1145/3549516

[5] Beat Games. 2019. *Beat Saber.* Game [VR]. Beat Games. Los Angeles, CA, USA. https://beatsaber.com/. Steam Reviews: https://store.steampowered.com/app/620980/Beat_Saber/.

[6] Stefan Behrendt. 2023. lm.beta: Add Standardized Regression Coefficients to Linear-Model-Objects. https://CRAN.R-project.org/package=lm.beta R package version 1.7-2.

[7] Gunnar Borg. 1998. *Borg's perceived exertion and pain scales.* Human kinetics.

[8] Felix Born, Sophie Abramowski, and Maic Masuch. 2019. Exergaming in VR: The Impact of Immersive Embodiment on Motivation, Performance, and Perceived Exertion. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, 1–8. doi:10.1109/VS-Games.2019.8864579

[9] Felix Born, Adrian Rygula, and Maic Masuch. 2021. Motivating Players to Perform an Optional Strenuous Activity in a Virtual Reality Exergame Using Virtual Performance Augmentation. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 225 (oct 2021), 21 pages. doi:10.1145/3474652

[10] Tim Böttger, Thomas Rudolph, Heiner Evanschitzky, and Thilo Pfrang. 2017. Customer Inspiration: Conceptualization, Scale Development, and Validation. *Journal of Marketing* 81, 6 (2017), 116–131. doi:10.1509/jm.15.0007

[11] Karina Caro, Mónica Tentori, Ana I. Martinez-Garcia, and Marina Alvelais. 2017. Using the FroggyBobby exergame to support eye-body coordination development of children with severe autism. *International Journal of Human-Computer Studies* 105 (2017), 12–27. doi:10.1016/j.ijhcs.2017.03.005

[12] Sebastian Cmentowski, Sukran Karaosmanoglu, Lennart E. Nacke, Frank Steinicke, and Jens Harald Krüger. 2023. Never Skip Leg Day Again: Training the Lower Body with Vertical Jumps in a Virtual Reality Exergame. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 786, 18 pages. doi:10.1145/3544548.3580973

[13] Stanley J. Colcombe, Arthur F. Kramer, Kirk I. Erickson, Paige Scalf, Edward McAuley, Neal J. Cohen, Andrew Webb, Gerry J. Jerome, David X. Marquez, and Steriani Elavsky. 2004. Cardiovascular fitness, cortical plasticity, and aging. *Proceedings of the National Academy of Sciences* 101, 9 (2004), 3316–3321. doi:10.1073/pnas.0400266101 arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.0400266101

[14] John W Creswell and Vicki L Plano Clark. 2017. *Designing and conducting mixed methods research.* Sage publications, Thousand Oaks, California, United States.

[15] Sebastian Deterding. 2011. Situated motivational affordances of game elements: A conceptual model. In *Gamification: Using game design elements in non-gaming contexts, a workshop at CHI*, Vol. 10.

[16] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From game design elements to gamefulness: defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (Tampere, Finland) *(MindTrek '11)*. Association for Computing Machinery, New York, NY, USA, 9–15. doi:10.1145/2181037.2181040

[17] Benjamin D. Douglas, Patrick J. Ewell, and Markus Brauer. 2023. Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLOS ONE* 18, 3 (03 2023), 1–17. doi:10.1371/journal.pone.0279720

[18] Peer Eyal, Rothschild David, Gordon Andrew, Evernden Zak, and Damer Ekaterina. 2021. Data quality of platforms and panels for online behavioral research. *Behavior research methods* (2021), 1–20. doi:10.3758/s13428-021-01694-3

[19] Andy Field, Zoe Field, and Jeremy Miles. 2012. Discovering statistics using R. (2012).

[20] Sean G. Figgins, Matthew J. Smith, Christopher N. Sellars, Iain A. Greenlees, and Camilla J. Knight. 2016. "You really could be something quite special": A qualitative exploration of athletes' experiences of being inspired in sport. *Psychology of Sport and Exercise* 24 (2016), 82–91. doi:10.1016/j.psychsport.2016.01.011

[21] FitXR. 2019. *FitXR - Box. HIIT. Dance. Sculpt. Combat.* Game [VR]. FitXR. London, England. Meta Reviews: https://www.meta.com/experiences/2327205800645550/.

[22] Lawrence Hoc Nang Fong, Long Wai. Lam, and Rob Law. 2017. How locus of control shapes intention to reuse mobile apps for making hotel reservations: Evidence from chinese consumers. *Tourism Management* 61 (2017), 331–342. doi:10.1016/j.tourman.2017.03.002

[23] John Fox and Sanford Weisberg. 2019. *An R Companion to Applied Regression* (third ed.). Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

[24] Tracy Fullerton. 2008. *Game Design Workshop-A Playcentric Approach to Creating Innovative Games.* Elsevier.

[25] Nadia Garcia-Hernandez, Karen Garza-Martinez, Vicente Parra-Vega, Armando Alvarez-Sanchez, and Laura Conchas-Arteaga. 2019. Development of an EMG-based exergaming system for isometric muscle training and its effectiveness to enhance motivation, performance and muscle strength. *International Journal of Human-Computer Studies* 124 (2019), 44–55. doi:10.1016/j.ijhcs.2018.11.010

[26] Linda Graf, Sophie Abramowski, Felix Born, and Maic Masuch. 2023. Emotional Virtual Characters for Improving Motivation and Performance in VR Exergames. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY, Article 417 (oct 2023), 21 pages. doi:10.1145/3611063

[27] Sidney Grosprêtre, Philémon Marcel-Millet, Pauline Eon, and Bettina Wollesen. 2023. How Exergaming with Virtual Reality Enhances Specific Cognitive and Visuo-Motor Abilities: An Explorative Study. *Cognitive Science* 47, 4 (2023). doi:10.1111/cogs.13278

[28] Tobias Gummer, Joss Roßmann, and Henning Silber. 2021. Using Instructed Response Items as Attention Checks in Web Surveys: Properties and Implementation. *Sociological Methods & Research* 50, 1 (2021), 238–264. doi:10.1177/0049124118769083

[29] Stuart Hallifax, Maximilian Altmeyer, Kristina Kölln, Maria Rauschenberger, and Lennart E. Nacke. 2023. From Points to Progression: A Scoping Review of Game Elements in Gamification Research with a Content Analysis of 280 Research Papers. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY, Article 402 (oct 2023), 21 pages. doi:10.1145/3611048

[30] Juho Hamari, Jonna Koivisto, and Harri Sarsa. 2014. Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In *2014 47th Hawaii International Conference on System Sciences.* 3025–3034. doi:10.1109/HICSS.2014.377

[31] Andrew F Hayes. 2017. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford publications.

[32] Teresa Hirzle, Maurice Cordts, Enrico Rukzio, Jan Gugenheimer, and Andreas Bulling. 2021. A Critical Assessment of the Use of SSQ as a Measure of General Discomfort in VR Head-Mounted Displays. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 530, 14 pages. doi:10.1145/3411764.3445361

[33] Holodia. 2021. *Holofit by Holodia.* Game [VR]. Holodia. Schiltigheim, France. https://www.holodia.com/. Meta Reviews: https://www.meta.com/experiences/2853664488067174/.

[34] Christos Ioannou, Patrick Archard, Eamonn O'Neill, and Christof Lutteroth. 2019. Virtual Performance Augmentation in an Immersive Jump & Run Exergame. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3290605.3300388

[35] Sukran Karaosmanoglu, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. 2024. Born to Run, Programmed to Play: Mapping the Extended Reality Exergames Landscape. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, 28 pages. doi:10.1145/3613904.3642124

[36] Sukran Karaosmanoglu, Lucie Kruse, Sebastian Rings, and Frank Steinicke. 2022. Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 342, 7 pages. doi:10.1145/3491101.3519736

[37] Sukran Karaosmanoglu, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. 2021. Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 252 (oct 2021), 27 pages. doi:10.1145/3474679

[38] Sukran Karaosmanoglu, Moritz Wegner, Sebastian Cmentowski, and Frank Steinicke. 2024. Move, React, Repeat! The Role of Continuous Cues in Immersive Exergames. *Proc. ACM Hum.-Comput. Interact.* 8, CHI PLAY, Article 326 (Oct. 2024), 23 pages. doi:10.1145/3677091

[39] Tuomas Kari and Markus Makkonen. 2014. Explaining the usage intentions of exergames. In *International Conference on Information Systems.* Association for Information Systems (AIS).

[40] Fares Kayali, Naemi Luckner, Peter Purgathofer, Katta Spiel, and Geraldine Fitzpatrick. 2018. Design considerations towards long-term engagement in games for health. In *Proceedings of the 13th International Conference on the Foundations of Digital Games* (Malmö, Sweden) *(FDG '18)*. Association for Computing Machinery, New York, NY, USA, Article 35, 8 pages. doi:10.1145/3235765.3235789

[41] Chung-Wha Chloe Ki, Sangsoo Park, and Youn-Kyung Kim. 2022. Investigating the mechanism through which consumers are "inspired by" social media influencers and "inspired to" adopt influencers' exemplars as social defaults. *Journal of Business Research* 144 (2022), 264–277.

[42] Kinemotik Studios. 2022. *Audio Trip.* Game [VR]. Kinemotik Studios. San Francisco, CA, USA. http://www.kinemotik.com/audiotrip/. Steam Reviews: https://store.steampowered.com/app/826540/Audio_Trip/.

[43] Jonna Koivisto and Juho Hamari. 2019. The rise of motivational information systems: A review of gamification research. *International Journal of Information Management* 45 (2019), 191–210. doi:10.1016/j.ijinfomgt.2018.10.013

[44] Bastian Kordyaka, Samuli Laato, Sebastian Weber, Juho Hamari, and Bjoern Niehaves. 2024. Exploring the association between engagement with location-based game features and getting inspired about environmental issues and nature. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 755, 15 pages. doi:10.1145/3613904.3642786

[45] Rui Leitão, Martin Maguire, Sarah Turner, and Laura Guimarães. 2022. A systematic evaluation of game elements effects on students' motivation. *Education and Information Technologies* (2022), 1–23. doi:10.1007/s10639-021-10651-8

[46] Amy Shirong Lu, Victoria Pelarski, Dar Alon, Aleksandra Baran, Emma McGarrity, Neha Swaminathan, and Caio Victor Sousa. 2023. The effect of narrative element incorporation on physical activity and game experience in active and sedentary virtual reality games. *Virtual Reality* 27, 3 (2023), 1607–1622. doi:10.1007/s10055-023-00754-7

[47] Joe Marshall, Florian 'Floyd' Mueller, Steve Benford, and Sebastiaan Pijnappel. 2016. Expanding exertion gaming. *International Journal of Human-Computer Studies* 90 (2016), 1–13. doi:10.1016/j.ijhcs.2016.02.003

[48] Elisa D. Mekler, Florian Brühlmann, Klaus Opwis, and Alexandre N. Tuch. 2013. Disassembling gamification: the effects of points and meaning on user motivation and performance. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 1137–1142. doi:10.1145/2468356.2468559

[49] Alexander Michael and Christof Lutteroth. 2020. Race Yourselves: A Longitudinal Exploration of Self-Competition Between Past, Present, and Future Performances in a VR Exergame. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–17. doi:10.1145/3313831.3376256

[50] Florian Mueller and Katherine Isbister. 2014. Movement-based game guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2191–2200. doi:10.1145/2556288.2557163

[51] Florian 'Floyd' Mueller, Darren Edge, Frank Vetere, Martin R. Gibbs, Stefan Agamanolis, Bert Bongers, and Jennifer G. Sheridan. 2011. Designing Sports: A Framework for Exertion Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2651–2660. doi:10.1145/1978942.1979330

[52] Odders. 2022. *Les Mills XR: BodyCombat.* Game [VR]. Les Mills XR: BodyCombat. Auckland, New Zealand. Meta Reviews: https://www.lesmills.com/bodycombat-vr/.

[53] Seong-Hi Park, Kuem Sun Han, and Chang-Bum Kang. 2014. Effects of exercise programs on depressive symptoms, quality of life, and self-esteem in older people: A systematic review of randomized controlled trials. *Applied Nursing Research* 27, 4 (2014), 219–226. doi:10.1016/j.apnr.2014.01.004

[54] Anatoli Petridou, Aikaterina Siopi, and Vassilis Mougios. 2019. Exercise in the management of obesity. *Metabolism* 92 (2019), 163–169.

[55] Susanne Poeller and Cody J. Phillips. 2022. Self-Determination Theory — I Choose You! The Limitations of Viewing Motivation in HCI Research Through the Lens of a Single Theory. In *Extended Abstracts of the 2022 Annual Symposium on Computer-Human Interaction in Play* (Bremen, Germany) *(CHI PLAY '22)*. Association for Computing Machinery, New York, NY, USA, 261–262. doi:10.1145/3505270.3558361

[56] R Core Team. 2023. R: A Language and Environment for Statistical Computing. https://www.R-project.org/

[57] Riot Games. 2009. *League of Legends.* Game. Riot Games. Los Angeles, CA, USA. https://www.leagueoflegends.com/.

[58] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion* 30 (2006), 344–360. doi:10.1007/s11031-006-9051-8

[59] Richard M. Ryan, Geoffrey C. Williams, Heather Patrick, and Edward L. Deci. 2009. Self-determination theory and physical activity: the dynamics of motivation in development and wellness. *Hellenic journal of psychology* 6, 2 (2009), 107–124.

[60] Gavin RH Sandercock, Chris Beedie, and Steve Mann. 2016. Is Olympic inspiration associated with fitness and physical activity in English schoolchildren? A repeated cross-sectional comparison before and 18 months after London 2012. *BMJ open* 6, 11 (2016), e011670. doi:10.1136/bmjopen-2016-011670

[61] Jeff Sinclair, Philip Hingston, and Martin Masek. 2007. Considerations for the design of exergames. In *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia* (Perth, Australia) *(GRAPHITE '07)*. Association for Computing Machinery, New York, NY, USA, 289–295. doi:10.1145/1321261.1321313

[62] SteamDB. 2024. SteamDB. https://steamdb.info/instantsearch/?refinementList%5Btags%5D%5B0%5D=VR Accessed: 21.08.2024 https://steamdb.info/instantsearch/?refinementList%5Btags%5D%5B0%5D=VR.

[63] Ancret Szpak, Stefan Carlo Michalski, and Tobias Loetscher. 2020. Exergaming With Beat Saber: An Investigation of Virtual Reality Aftereffects. *J Med Internet Res* 22, 10 (23 Oct 2020). doi:10.2196/19840

[64] Todd M Thrash and Andrew J Elliot. 2003. Inspiration as a psychological construct. *Journal of personality and social psychology* 84, 4 (2003), 871. doi:10.1037/0022-3514.84.4.871

[65] Todd M Thrash and Andrew J Elliot. 2004. Inspiration: core characteristics, component processes, antecedents, and function. *Journal of personality and social psychology* 87, 6 (2004), 957. doi:10.1037/0022-3514.87.6.957

[66] April Tyack and Elisa D. Mekler. 2020. Self-Determination Theory in HCI Games Research: Current Uses and Open Questions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3313831.3376723

[67] Daryoush Daniel Vaziri, Konstantin Aal, Yves J Gschwind, Kim Delbaere, Anne Weibert, Janneke Annegarn, Helios de Rosario, Rainer Wieching, David Randall, and Volker Wulf. 2017. Analysis of effects and usage indicators for a ICT-based fall prevention system in community dwelling older adults. *International Journal of Human-Computer Studies* 106 (2017), 10–25. doi:10.1016/j.ijhcs.2017.05.004

[68] Virtual Reality Institute of Health and Exercise. 2024. VR Ratings. https://vrhealth.institute/vr-ratings/ Accessed: 10.08.2024.

[69] Susanne B Votruba, Micah A Horvitz, and Dale A Schoeller. 2000. The role of exercise in the treatment of obesity. *Nutrition* 16, 3 (2000), 179–188.

[70] Kevin Werbach. 2014. (Re)Defining Gamification: A Process Approach. In *Persuasive Technology*, Anna Spagnolli, Luca Chittaro, and Luciano Gamberini (Eds.). Springer International Publishing, Cham, 266–272.

[71] World Health Organization. 2020. WHO guidelines on physical activity and sedentary behaviour: at a glance. https://iris.who.int/bitstream/handle/10665/337001/9789240014886-eng.pdf?sequence=1 Accessed: 07.09.2024.

[72] World Health Organization. 2022. Physical Activity. https://www.who.int/news-room/fact-sheets/detail/physical-activity Accessed: 18.08.2024.

[73] Wenge Xu, Hai-Ning Liang, Kangyou Yu, and Nilufar Baghaei. 2021. Effect of Gameplay Uncertainty, Display Type, and Age on Virtual Reality Exergames. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 439, 14 pages. doi:10.1145/3411764.3445801

[74] Wenge Xu, Hai-Ning Liang, Kangyou Yu, Shaoyue Wen, Nilufar Baghaei, and Huawei Tu. 2023. Acceptance of Virtual Reality Exergames Among Chinese Older Adults. *International Journal of Human–Computer Interaction* 39, 5 (2023), 1134–1148. doi:10.1080/10447318.2022.2098559

[75] Shengjie Yao and Gyoung Kim. 2019. The Effects of Immersion in a Virtual Reality Game: Presence and Physical Activity. In *HCI in Games*, Xiaowen Fang (Ed.). Springer International Publishing, Cham, 234–242.

[76] Achim Zeileis and Torsten Hothorn. 2002. Diagnostic Checking in Regression Relationships. *R News* 2, 3 (2002), 7–10. https://CRAN.R-project.org/doc/Rnews/

A VR Exergames List

Table 4: The list of VR exergames used in the expert survey. This list was created based on [35] and [68].

No	Game	No	Game
1	Audio Trip	2	Beat Saber
3	Black Box VR	4	Box VR
5	Carnival Games VR	6	Dance Central
7	First Steps	8	FitXR
9	Fruit Ninja VR	10	HoloFit
11	HoloPoint	12	NVIDIA VR Fun House
13	QuiVr	14	Snow Games VR
15	Sports Scramble	16	SyncSense
17	Thrill of the Fight	18	VirZoom
19	VRSports Challenge	20	VZFit
21	Supernatural	22	Until you Fall
23	Orc Hunter	24	Late For Work
25	ThunderGod	26	Animal Force
27	Music Escape	28	Mace and Grace
29	Flappy Flappy VR	30	Bardo
31	Singularity 5	32	Conveyor VR
33	Pavlov VR	34	Ninja Legends
35	Hellsplit: Arena	36	Holoception
37	DrumBeats VR	38	Synth Riders
39	Pistol Whip	40	Wave Circles
41	Hotel Transylvania Popstic	42	Blade and Sorcery
43	Hot Squat 2: New Glory	44	Dance Collider
45	Deus Vult	46	Big Breezy Boat
47	Everyday Golf VR	48	Flappy Arms
49	PowerBeatsVR	50	Song Beater: Quite My Tempo!
51	HoloBall	52	Chop It
53	Gates of Nowhere	54	Vector Velocity
55	Space Channel 5	56	SpaceFrog VR
57	To the Top	58	Blasters of the Universe
59	Audica	60	Windlands 2
61	Sweet Escape VR	62	Crystal Reign
63	Gorn	64	CYCOM: Cybernet Combat
65	Undead Development	66	Eclipse Cinequest
67	Smashbox Arena	68	Racket Fury: Table Tennis VR
69	Electronauts	70	Cricket Club
71	Cross Country Skiing VR	72	Final Soccer VR
73	Katana X	74	Dungeons & Treasure VR
75	Nightstar: Starfighter	76	Loco Dojo
77	Dash Dash Run!	78	The IOTA Project
79	OrbusVR	80	Detached
81	Acan's Call: Act 1	82	Karnage Chronicles
83	Sairento VR	84	The Wizards
85	VR Regatta	86	Soundboxing
87	Creed: Rise to Glory	88	Eleven Table Tennis
89	Audioshield	90	Sprint Vector
91	Vindicta	92	Goalie VR
93	Omni Arena	94	Fastest Fist
95	Knockout League	96	Drunkn Bar Fight
97	Raw Data	98	Rec Room
99	Superhot VR	100	Vanishing Realms
101	Tiltbrush	102	Job Simulator
103	Climbey	104	Racket: Nx
105	Arcade Saga	106	Space Pirate Trainer
107	Bitslap	108	Sword Master
109	GoalkeepVR	110	Hot Squat

B Study 2: Survey Items

Table 5: The list of used questionnaire items and their sources.

CONSTRUCT	#	ITEMS	REFS
Design Elements		*Please rate the importance of interacting with the design features listed below while playing Beat Saber on a scale ranging from 1 (not important at all) to 5 (very important).*	
	1	*Points, score, XP*	[43]
	2	*Challenges, quests, missions, tasks, clear goals*	[43]
	3	*Badges, achievements, medals, trophies*	[43]
	4	*Leaderboards, rankings*	[43]
	5	*Levels*	[43]
	6	*Performance stats, performance feedback*	[43]
	7	*Progress, status bars, skill trees*	[43]
	8	*Timer, speed*	[43]
	9	*Increasing difficulty*	[43]
	10	*Competition*	[43]
	11	*Customization, personalization*	[43]
	12	*Multiplayer*	[43]
	13	*Virtual world, 3D world, game world*	[43]
	14	*In-game rewards*	[43]
	15	*Motion tracking*	[43]
	16	*Assistance, virtual helpers*	[43]
	17	*Retries, health, health points*	[43]
	18	*Onboarding, benefits for beginners*	[43]
	19	*Warnings*	[43]
	20	*Penalties*	[43]
	21	*Music / rhythm*	ES & EI
	22	*Audiovisual congruence*	EI
	23	*Downloadable contents / regular additions*	EI
	24	*Juicy effects and cues*	EI
Inspired by		*Please indicate your answer on a scale ranging from 1 (strongly disagree) to 7 (strongly agree). Please think about your regular Beat Saber gameplay. While playing Beat Saber...*	
	25	...my imagination was stimulated	[10, 44, 64]
	26	...I was intrigued by a new idea	[10, 44, 64]
	27	...I unexpectedly and spontaneously got new ideas	[10, 44, 64]
	28	...my horizon was broadened	[10, 44, 64]
	29	...I discovered something new	[10, 44, 64]
Inspired to		*Please indicate your answer on a scale ranging from 1 (strongly disagree) to 7 (strongly agree). Please think about your regular Beat Saber gameplay. After playing Beat Saber...*	
	30	...I was inspired to do physical exercise	[10, 44, 64]
	31	...I felt a desire to do physical exercise	[10, 44, 64]
	32	...my interest to do physical exercise increased	[10, 44, 64]
	33	...I was motivated to do physical exercise	[10, 44, 64]
	34	...I felt an urge to do physical exercise	[10, 44, 64]
Long-Term Use		*Please rate the following items on a scale ranging from 1 (strongly disagree) to 7 (strongly agree).*	
	35	I intend to play *Beat Saber* in the future	[22, 39]
	36	I plan to play *Beat Saber* in the future	[22, 39]
	37	I am likely to play *Beat Saber* in the future	[39]
	38	I will always try to play *Beat Saber* in the future	[39]

C R Code Snippets

We report example R code snippets used to analyze quantitative data.

```
library(tidyverse)
library(tidyr)
library(dplyr)
library(lmtest)
library("lm.beta")
library(car)
options(scipen = 999)
options(digits = 16)

#linear regression
model <- stats::lm(Inspiration.by.Mean ~ . , data = data)
summary(model)

#stepwise regression
step.model <- stats::step(model, direction = "backward")
step.model
summary(step.model)

#Linearity
par(mfrow= c(2,2))
plot(step.model)

#Autocorrelation - Dubrin-Watson test
dw_test <- dwtest(step.model)
print(dw_test)

#Multi-collinearity - Variance Inflation Factors
vif_values <- car::vif(step.model)
max(vif_values)

#Standardized regression coefficient
standardized_coefs <- lm.beta(step.model)
print(standardized_coefs)

#Mediation Anaylsis
process(data = data, y = "Longterm.Intention", x = "Inspiration.by",
            m ="Inspiration.to", model = 4, stand =1, seed = 1)
```

10 IMMERSIVE VR GAMES FOR ACTIVE AGING

10.1 CANOE VR: AN IMMERSIVE EXERGAME TO SUPPORT COGNITIVE AND PHYSICAL EXERCISES OF OLDER ADULTS

[CorePub4] **Sukran Karaosmanoglu**, Lucie Kruse, Sebastian Rings, and Frank Steinicke. Canoe vr: An immersive exergame to support cognitive and physical exercises of older adults. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, (CHI EA '22), New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391566. DOI 10.1145/3491101.3519736. URL https://doi.org/10.1145/3491101.3519736.

Supplementary materials for this article are available here: https://doi.org/10.1145/3491101.3519736.

Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults

Sukran Karaosmanoglu
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
sukran.karaosmanoglu@uni-hamburg.de

Lucie Kruse
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
lucie.kruse@uni-hamburg.de

Sebastian Rings
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
sebastian.rings@uni-hamburg.de

Frank Steinicke
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
frank.steinicke@uni-hamburg.de

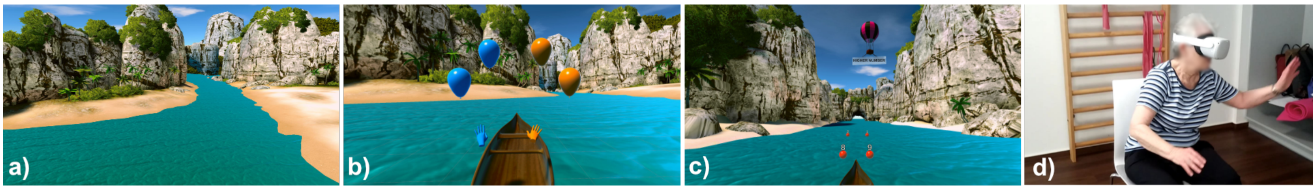

Figure 1: The figure shows the last version of *Canoe VR*: *a)* The game environment. *b)* Players calibrate their arm reach by pushing balloons away. *c)* Players pop soap bubbles with a higher number. *d)* An older adult plays the game.

ABSTRACT

Cognitive and physical exercises are important factors to support a healthy life, especially considering demographics change. Virtual reality (VR) exergames have great potential to support these activities in a more motivating way. However, regular usage of VR exergames by the older population is still limited. To address this issue, we designed and implemented a VR exergame: *Canoe VR*. We applied several prototyping sessions with older players and also report the results from interviews with physiotherapists. The results suggest that *Canoe VR* was very well received and can be used by older players as an additional fitness tool. We discuss the implications of extending a fitness routine with a VR exergame and using the game with players of different abilities.

CCS CONCEPTS

• **Software and its engineering** → **Interactive games**; • **Human-centered computing** → **Accessibility**.

KEYWORDS

Virtual reality, exergames, older adults

ACM Reference Format:
Sukran Karaosmanoglu, Lucie Kruse, Sebastian Rings, and Frank Steinicke. 2022. Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3491101.3519736

CHI '22 Extended Abstracts, April 29-May 5, 2022, New Orleans, LA, USA

1 INTRODUCTION

Given that the world population is aging [34], performing cognitive [35] and physical [8] exercises holds major importance for maintaining a healthy life. These exercises can help to reduce age-related decline in abilities [8, 21] and support well-being [7, 27]. Despite these benefits, having no interest, lack of opportunities for physical exercises, or COVID-19 restrictions can negatively affect participating in these activities [24, 29].

Immersive virtual reality (VR) systems are becoming commercially available and increasingly affordable for consumers (e.g., Meta Quest 2 [23]). In recent years, several VR exergames [37] (such as Beat Saber [13], OhShape [19], BoxVR [11]) have been introduced that incorporate physical exercises with cognitive aspects. Such training games can motivate users to exercise and continue training over longer periods. Especially given the effects of COVID-19, these systems allow exercising at home without contact with others or to reach exercise services without traveling. Nevertheless, VR technology is still novel and can expose some difficulties for users such as usability problems [1], cybersickness [20] or ergonomics [15]. At the same time, designing these games for older generations requires attention to the game design, age-related changes, and movements employed in the game.

In this paper, we implemented *Canoe VR* following a human-centered design (HCD) approach [12]. Older adults, fitness experts, and a health professional were involved in the process. We conducted eight prototyping sessions that considered nuances in the older population, such as older adults with and without dementia. The game was also tested in a gym with older adults and fitness

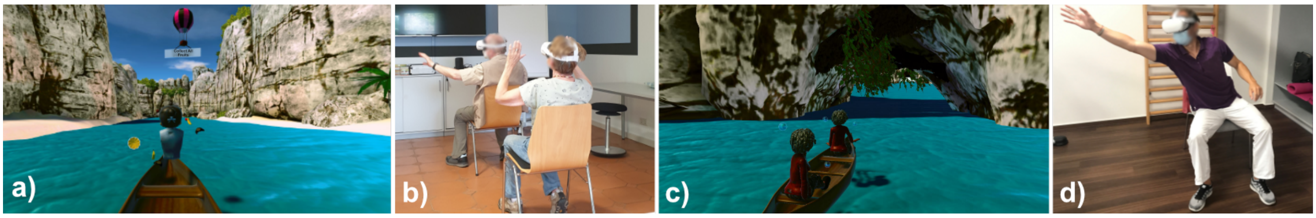

Figure 2: The figure shows *a)* an artificial agent playing the game together with the user. *b)* Two older adults in the multiplayer version. *c)* The representation of the two players in the multiplayer version. *d)* A physiotherapist testing the game.

experts. Based on the gathered feedback of play-testings and interviews, *Canoe VR* has been iteratively improved. We report the results of these sessions and lessons learned, which can guide future work on VR exergames.

The contribution of this paper are:

- Demonstration of a VR exergame designed for older users by considering nuances in the population as well as their training needs,
- Use of this VR exergame in the regular training sessions of older adults as an additional fitness tool.

2 RELATED WORK

Previous work showed that cognitive [35] and physical [8] activities are beneficial for healthy aging. For example, they can support cognitive functions [35] and reduce the effects of aging in the brain [8]. It is suggested that the combination of both interventions can lead to better outcomes [10], and in particular simultaneous training, can benefit cognitive functions of older population [32].

The World Health Organization recommends that older adults should be physically active (at least 150-300 minutes of moderate activity per week) [39]. However, COVID-19 imposed some challenges to obtain these activities due to necessary lockdown policies. While prior work identified some obstacles to physical activities (e.g., poor health conditions and missing company) [24] for older adults, the impacts of COVID-19 on physical activity is inevitable [29, 40]. Relating to our paper, we argue that exergames can offer simultaneous cognitive and physical training possibilities, in particular, during the COVID-19 pandemic. Thus, they can be a supportive tool for exercise routines of older adults.

The use of immersive VR is increasing [2] and it is becoming common for several applications ranging from business to gaming industry. VR supports unique experiences such as a potentially intense sense of presence [30], multisensory sensations [31], and natural interactions. For instance, in VR, older adults can experience outdoor world which might be limited due to COVID-19 and their health condition [17].

With VR exergames, players can benefit from both physical and cognitive stimulation while enjoying the game experience. In recent years, VR exergames were also explored for older adults with [16] and without dementia [18]. In an HCD process, Karaosmanoglu et al. [16] designed a VR exergame (*Memory Journalist VR*) for older players with dementia that included multiple stakeholders in the process, indicating positive player experiences for users. Similarly, a recent work [25] presented a VR exergame (*Seas the Day*) designed with input from various stakeholders for people with dementia. Kruse et al. [18] compared a VR exergame (*Maestro Game VR*) to a video-guided exercises in older adults. The findings suggested that both training options were comparably similar in terms of enjoyment and attention, and VR exergames can be a good addition for training. A recent study [6] indicated positive outcomes (e.g., usability, positive experiences) for the use of commercial VR games (e.g., BoxVR [11]) with four older adults during the pandemic.

In our work, we consider varying user groups (i.e., with and without dementia) in a single game and design for these groups, and use this VR exergame as an additional fitness tool. In this regard, users' needs, age-related changes, and health conditions should be considered [12, 14, 16]. Thus, in this paper, we used the HCD process [12] that considers needs and preferences of users by involving end-users as well as other stakeholders.

3 CANOE VR

In this section, we describe the technical properties and the last version of the game design. The changes applied to the game over the prototyping sessions can be found in Table 2.

The game was implemented in Unity (v. 2019.4.18f1) with the Universal Render Pipeline [33]. Since we used the Valve Index [36] in the early sessions, the game was run with an i9-9900K CPU and a GeForce RTX2080 Ti graphics card. Later, we used the Meta Quest 2 [23] as it is a standalone device and allows to conducting field studies without being bound to a computer. With this device, we were able to easily integrate hand tracking into our game.

The game play takes place on a tropical island scene (see Figure 1a), adapted from *BoatAttack* [22, 28]. Players are positioned inside a canoe that moves down a river. The game starts with a calibration phase. First, players perform a thumbs up gesture with both hands or they can press a button on both of their controllers to calibrate their seating position and comfortable view direction. Then, to obtain the maximum arm range of motion (ROM), the players need to touch four balloons (left, right, left up, right up) and push them away as far as possible (see Figure 1b).

After this calibration, users can select their game level. The game currently offers 22 levels with varying cognitive and physical tasks (see Table 1). Moreover, players can adjust the difficulty, length, and the number of game levels.

When the player presses a start button, the game provides written and audio instructions describing the level. Also as a reminder, a hot air balloon floats in the air in front of the players and displays the game level task in text form (see Figure 1c). Meanwhile, based on the ROM of each player, soap bubbles or other objects are spawned over the river. Using the ROM points, we create a 2D

Table 1: This table lists the game levels in *Canoe VR*. The asterisk indicates levels that can only be played in multiplayer mode.

Game Levels	Description	Game Levels	Description
(1) All Good	Touch as many bubbles as possible.	(12) Hand Change	Collect the bubbles by changing hands.
(2) Only Colored	Touch all the bubbles except the black ones.	(13) Fruit	Collect all fruits.
(3) Only Red	Collect only red bubbles.	(14) Music	Collect the bubbles with accompanying music.
(4) Two Simultaneously	Collect two bubbles at the same time.	(15) Dividable by 7	Collect all the bubbles showing a number dividable by seven.
(5) Color Change	Collect a red bubble and then a yellow bubble alternately.	(16) Lower Number	Collect the bubble with the lower number.
(6) Higher Number	Collect the bubble with the higher number.	(17) Vowels	Collect the bubbles showing a vowel.
(7) Maximum 5 Yellow	Collect a maximum of 5 yellow bubbles, but all red bubbles.	(18) Letters in Word	Collect all the bubbles showing letters contained in the word "canoe".
(8) Balloons Dodge	Avoid the balloons and collect all bubbles.	(19) Letter & Number	Collect a bubble with a letter and one with a number alternately.
(9) Butterflies	Free all the butterflies from the bubbles.	(20) Even Numbers	Collect the bubbles with even numbers.
(10) Colors for Hands	Collect the red and yellow bubbles with the corresponding red and yellow colored hand.	(21) Own Colors*	The front player collects all the yellow bubbles, the back player only collects the green bubbles.
(11) Cross Over	Collect the red and yellow bubbles with the corresponding red and yellow colored hand. This level includes cross over movements.	(22) Alternate Collect*	Always take turns collecting a bubble. One bubble by the back player, one by the front player, etc.

polygon between these outer limits and the canoe object to define possible spawn locations. We choose random, evenly distributed points inside the polygon to make sure that all generated points are within the player's reach. We can now dynamically modify the difficulty of a level by tuning multiple parameters, depending on the player's success. Different reach difficulty can be achieved by down-scaling the polygon and changing the distance between each of the generated bubbles in the direction the canoe travels. This enables easier, slow-paced, as well as harder levels that require fast decision-making. In the cross over level, we can additionally increase the percentage of bubbles that spawn on the opposite side of the hand that has to collect it.

Depending on the chosen level, different tasks are presented. For example, in the higher number level, players collect the soap bubble with a higher number relative to the bubble partner (see Figure 1c). While traveling along the river, players also pass through caves, which have several vines (see Figure 2c). To finish a level, players perform the following physical movements: torso, reaching, stretching, and leaning (see Figure 1d). At the end of the game, players get a scoreboard where they can see their score and compare it to other players.

4 RESULTS: PROTOTYPING SESSIONS & INTERVIEWS

We used different methods (e.g., observation and field notes, semi-structured interviews) to obtain data in our project. The full list of (translated) questions can be found in the supplementary material. For the game versions and changes used in each session, please see Table 2. All participants are inhabitants from the same local senior living facility. For sessions between I - IV, a healthcare professional was also present.

For the sessions that included interviews, the first author used a thematic analysis method (reflexive orientation) [3–5]. The researcher decided on deductive categories for each session[1]. Then inductively and iteratively coded the interviews under these categories and finally created themes. For the scope of this paper, we only present themes related to *Canoe VR*. We report quotes of the participants based on their specialty in the respective session[2].

Session I. We aimed to decide on an initial concept of an exergame. Two exergame ideas were implemented as early prototypes: *Canoe VR* and *Flower VR*. Two players tested both games (M=76 years, SD=0, *female*=2).

Canoe VR was found rich in terms of movement patterns. On a scale from 1 (not at all) to 7 (very much so), the mean enjoyment and exhaustion scores of the players were 6 (SD=0) and 1 (SD=0). The players and the healthcare professional liked the following aspects: "*left and right, recognizing color, head retraction, vines, upper body rotation, movement, points*"-*HP*, "*much more movement*"-P_1, and "*the overall picture of landscapes*"-P_2. One player experienced some technical problems due to accidentally pressing a button, which rotated their seating position. At the same time, the health professional reported concerns regarding the risk of falling. They suggested to improve the animal part of the game and add apples of different heights. They rated the enjoyment, presence, and exhaustion of the players as 6, 1, and 1 respectively. The health professional found *Flower VR* unsuitable for fit older adults, because the game flow included too many steps to remember easily. We decided to continue with *Canoe VR*, as it was well received by both the players and the health professional and offers more opportunities to create varying cognitive and physical exercises.

Session II. This session was conducted with two older adults (M=82.5 years, SD=7.78, *female*=2), who played *Canoe VR*.

The mean enjoyment, presence, and exhaustion scores of players were assessed as 7 (SD=0), 6 (SD=1.41), and 1 (SD=0), respectively. Both players enjoyed the game experience and accompanying music. They liked the following about the game: "*the whole game*"-P_1 and "*that you do something and that it also succeeds and the music was also good. The surroundings also look good, it was so nicely bright*"-P_2. They also reported some negative details: "*apples left and right sometimes hung too close, hard to tell which apples are in front and back*"-P_1. While for P_2, the instructions for feeding animals was too fast and not readable, for P_1, the feeding part was irritating because the basket was transferred to the land and the player stayed in the canoe.

[1]Session IV: physical activities, games, VR and VR games, and *Canoe VR*. The interviews with physiotherapists: games and VR, physical activities, and *Canoe VR*. Session VI: calibration phase, usability, agents and avatars, multiplayer, and *Canoe VR*.

[2]PH, PT, HP, P_1 represent a physiotherapist, trainer, and health professional and first player in a session, respectively.

Table 2: This table presents the changes and played game levels in each prototyping session.

Sessions	Game Version
Session I	In the first version of *Canoe VR*, the players were placed in a canoe on a river. Their task was to collect all red apples, and avoid the black apples and vines in tunnels (2 minutes). This was explained to them in a short tutorial. There was another level where they could collect fishes jumping out of the river, which had a duration of 45 seconds. At the end, they were asked to feed some animals with the collected food, by throwing the food towards them. The whole game took at least 200 seconds, including the tutorial. Background music accompanied the players. In *Flower VR*, players were placed in a garden scene. The task was to recreate a certain bouquet with different kinds of flowers that was shown to them. They had to pick the correct seeds, plant them in a flowerpot, water them, and afterwards place them in a vase.
Session II	We increased the length of the game to 220 meters (240 including the tutorial) and excluded the fishing part. At the end, the food was automatically given to the animals and did not have to be thrown anymore.
Session III	We introduced yellow and green apples in addition to red ones. The ending concept was changed to having a picnic with the collected apples. In three different levels, the players' task was to collect i) only red apples (60 seconds), ii) all red apples and 2-5 (fixed random number) yellow and 2-5 (fixed random number) green apples (90 seconds), and iii) the apple with a higher number (60 seconds).
Session IV	We made the second level easier by removing the green apples. The player played the levels i) only red apples, ii) all red and 2-5 (fixed random number) yellow apples, iii) the apple with the higher number, iv) alternatingly colored apples, and v) a level with soap bubbles. Each level was played for 60 seconds and the total game time was at least 280 seconds, including a tutorial.
Session V	This session featured *Canoe VR* on the Meta Quest 2 for the first time, with a new scene (see Figure 1a) and hand tracking instead of controllers. ROM calibration was added where the players sat straight and had to stretch their arms to the front as far as possible. The apple trees were exchanged for freely floating soap bubbles of different colors to enable more flexible positioning of the objects. Three levels were presented, with a length of 120 seconds each: all good, butterflies, and collecting red and yellow bubbles alternately. At the end of the level, a scoreboard was presented that showed how many bubbles were collected and how many vines were touched.
Session VI	Two game versions were presented: A single player version with a virtual agent as a second player, and a multiplayer version where two users could play together. The second player/agent was represented with two different avatars (changed during the game play) which had either i) only trunk and hands (see Figure 2a), or ii) full body (see Figure 2c). In the agent version, an agent was sitting in the canoe in front of the player, performing the tasks with them (see Figure 2a). The calibration was changed to a more playful solution, where the users had to push four balloons (two for each hand, one above them and one on the side) as far away as possible (see Figure 1b). Each balloon had to be touched at least once to finish the calibration. The following four levels were played for 60 seconds with a spawn distance of 1 meter between objects: fruits, butterflies, dodge balloons, and colors for hands. In the multiplayer version, the two players were sitting behind each other in the canoe (see Figure 2b,c). The played levels were fruit, own colors, alternate collect, all bubbles.
Session VII	The agent was removed in this version. The following four levels were played, with a duration of 60 seconds each, and a start distance of 1.5 m between collectible items: butterflies, colors for hands, cross over, and fruit. The distance between bubbles was evaluated every 10 bubbles and adjusted automatically by 0.2 m if more than 80 % of bubbles were correct or more than 40 % were wrong.
Session VIII	A remote was used to control the game flow. At the start, a tutorial asked the players to collect three bubbles while the canoe was still stationary to help them understand the task. We switched back from hand tracking to using controllers because of a more stable tracking. The game featured six levels, which were played in a random order and with different start distances between bubbles, 60 seconds each: all good (distance = 2 m), music (distance = 2 m), even numbers (distance = 1.2 m), letter and number (distance = 2 m), cross over (distance = 1.2 m), higher number (distance = 1.2 m). Distance between bubbles was evaluated every 5 meters and adjusted by 0.3 m if more than 80 % were correct or more than 20 % were wrong. The high score presented the top 10 scores and always included the players' score.

Session III. This session was conducted with older adults with dementia (*N*=3, *M*=81.33 years, *SD*=10.79, *female*=3).

Two players answered the enjoyment question with the following words: "*flawless*" and"*nice*". For presence, the answers were: "*yes*" and "*yes sure*", and "*no*". For exhaustion, a player answered with "*a little*", and the other with "*no*". Two players liked the task of collecting apples. The other player indicated that they liked the diversity with apples. Overall, the game level with all red, and a fixed amount of yellow and green apples was found too difficult. The health expert reported: for the enjoyment ("*well liked 5-6 [as a score]*") and exhaustion ("*no*"). They indicated that players had difficulties to understand and hear the instructions. They suggested to make the tutorial slower, avoid giving too much information, and to highlight the vines under the bridges more clearly or to use spiderwebs instead.

Session IV. Six older adults (*M*=78.83 years, *SD*=3.06, *female*=5, *male*=1) were recruited. We aimed to gain a deeper understanding about our user group, and, therefore, conducted semi-structured interviews with them.

Theme: ***Canoe VR is very well received, but the preferences of players for game objects and levels varied.*** Most players enjoyed the game: "*Yes, I like it very much*"-P_5. The cognitive and physical value of the game was appreciated by some players: "*That you just always keep moving and always have to pay attention, then duck, again pay attention to the apples etc. [...] I thought it was very nice, very interesting*"-P_1. But also: "*That really challenged me [smirks]*"-P_4. All players found the length of the overall game appropriate, however, P_2 player commented on having shorter individual levels and quick changes between them. Some players mentioned that they liked the apples more compared to soap bubbles, while only one reported a reason: "*Yeah, was more representational somehow*"-P_3. We speculate that the preference for apples could be due to the variety in apple tasks. There was subjectivity for the best task: all red and exactly [2-5] yellow apples, the higher number, combination of all these different things, and collecting red and yellow apples alternately. All players were positive to play the game regularly with some task variations: "*Yes, I can imagine that*"-P_3.

Interview with Physiotherapists. After the positive feedback in the last session, we decided to present the game to physiotherapists. This semi-structured interview was conducted in a video call attended by two experts (*M*=46 years, *SD*=7.07, *female*=1, *male*=1). During the interview, we presented a video from the last prototyping session with an older adult playing the game and showed them some concept pictures.

Theme: ***Canoe VR features rich movements and can offer valuable options (being mobile and social value), but it should be realistic and familiar.*** Both therapists found the featured movements very good: "*That you can also lower the arm and work with the trunk, raise the arm again, that I also have pause phases in between and not constantly a tension, a contraction of a muscle group. Because that is much too strenuous in the long run. [...] a game like that is really super*"-PH_2. The idea of having the game portable is found interesting because users can easily try out the game. A multiplayer mode to facilitate social interaction was also well received ("*when we interact as a team. [...] There comes then also again an added value to it*"-PH_2). Both experts strongly emphasized the need for graphical realism, in particular, consistency with the real world: "*Not that the graphics have to be designed in a great way, but I wouldn't keep it abstract. [...] But I wouldn't overburden [...] the participants with disturbing things [...] where everything is unrealistic. Or where the apple is not red and yellow, but blue and purple... So I would go very close to reality, especially with the graphics*"-PH_2.

Session V. This session was conducted in the gym of the facility with an attendance of a physiotherapist and a trainer (*N*=2, *M*=41.5 years, *SD*=0.71, *female*=1, *male*=1). We asked the older adults (*N*=7, *M*=83.71 years, *SD*=3.25, *female*=6, *male*=1) in the gym if they would like to play *Canoe VR*. The experts also played the game to give more precise feedback.

Canoe VR was found enjoyable by the players (*M*=3.81, *SD*=0.50, KIM [38] enjoyment/interest sub-scale, *not at all true/0 — very true/4*). The players felt low levels of exhaustion (*M*=1.57, *SD*=0.98, a single custom item, *not exhausted at all/1 — very exhausted/5*). Most players also gave positive comments: "*overall, an interesting, movement-intensive game*"-P_6. However, P_4 reported technical issues such as the font becoming blurry and quickly disappearing. Also, for P_7, the game was not something they would play every day. For the experts, the use of VR for older adults' exercises is considered positive: "*very good approach to apply digital movement offers in training therapy, diversified, innovative, motivating for all difficulty levels*"-*PH*. Both also rated the enjoyment of the players positively (*M*=4.5, *SD*=0.71, a single custom item, *very bad/1 — very good/5*). The trainer reported that the VR headset "*[...] size partly too big*", however, for the therapist the usability of the system was "*very good, [...] could be individualized if necessary, otherwise good flow, smooth gameplay, pleasant design ambiance*". The movements (*M*=4, *SD*=0) and their daily functionality for older adults (*M*=4, *SD*=0) were rated high (a single custom item, *very bad/1 — very good/5*). While the trainer found the bubbles a bit too fast for the players, the physiotherapist suggested to have "*shorter sequences, lower movements [...], dynamic, sitting "actively" at the front of the chair, cross-over movements (of the hands oriented to the center of the body)*".

Session VI. In this session, we focused on the social value of the game. We tested a version with older adults (*M*=78.75 years, *SD*=4.71, *female*=6, *male*=2) playing with a virtual agent (*N*=4), and a multiplayer version with two players playing together (*N*=4). We took observation, field notes, and interviewed the players in groups of one, three, and four players.

Theme: ***The sitting configuration and the agent were not received well, but overall players reported positive player experience with consideration of some aspects.*** All players playing with the agent noticed the agent, however, it faded out after some time: "*I only saw them for a short moment and then I didn't see them anymore*"-P_3. P_1 explicitly described the changes in the appearance of the avatars, however, most of these players did not see a need for an agent: "*actually, they can go away*"-P_2. Similarly, P_4 also realized the changes, but they focused more on the task: "*At the beginning I saw that, but then I was so concentrated on [task] what concerns me [...]*"-P_4. In the multiplayer mode, players mostly indicated that they enjoyed to play in pairs, while a player, who played with the agent negatively, commented about playing with a real person "*[...] you would concentrate on that person and [...] would be too distracted*"-P_2. However, a front player felt like they played alone: "*I did not notice my [partner]*"-P_7.

Overall, many players indicated positive player experience: "*Somehow you had the feeling that you were sitting in the boat. Then the waves moved. So I thought that was actually quite nicely done*"P_6 and "*You couldn't just sit there and watch, but it was a challenge to [perform the tasks]. So I found that quite good*"-P_4. Many players preferred the avatar with the full body representation while for P_4, it did not matter as the task was more important. Regarding the design, a few players had difficulties to see the butterflies, required some time to understand the thumbs up gesture, or had issues with the behavior of the calibration balloons: "*So I had the impression, the balloons do not really react when you push them away*"-P_8. Some players suggested to increase the length of the gameplay.

Session VII. This session was held in the same gym as Session V, and with the same fitness experts. Three older adults (*M*=84.33 years, *SD*=1.53, *female*=2, *male*=1) and the experts played *Canoe VR* (see Figure 2d).

The players rated their enjoyment of the game as high (*M*=3.89, *SD*=0.19, KIM). However, their exhaustion was low (*M*=1, *SD*=0). For P_3, "*overall, the game is interesting, challenging, concentration and movement enhancing*". While P_1 suggested to have the texts visible for a longer time, P_2 suggested to improve the calibration sequence: "*Slightly change the setting at the start so that the balloons do not fly so far (away)*". According to the experts, the older players enjoyed the game (*M*=4.5, *SD*=0.71). For usability, the physiotherapist commented on a safety issue: "*possibly too wide cross movements (inclinations to fall)*". The evaluation of the movements patterns (*M*=4.5, *SD*=0.71) and their functionality (*M*=4.5, *SD*=0.71) for daily life of older adults were high. The butterfly level caused a little challenge: "*visibility of the "butterflies" sometimes a bit difficult*"-*PH*, as also mentioned by P_1. Both experts found the game valuable: "*variety is good, timing is ideal, cross movements is extremely valuable therapeutically, good trunk activity (rotation/trunk extension)*"-*PH*.

For further improvements, they suggested: "*more stretching movements of the torso (arms stretching upwards)*"-*PH*, "*reduce the range of motion of insecure people if necessary [...], fun of movements in the game is given, if necessary sound as confirmation of the achieved movement*"-*PH*, and "*Bonus round! "Special effect" if you made the quest (e.g., fireworks or a lovely melody), optional music*"-*PT*.

Session VIII. We conducted this session with people with dementia (*N*=5, *M*=81.2 years, *SD*=8.58, *female*=4, *male*=1). We focused on three points: interaction with the game after the improvements, understanding how challenging the game is, and finding a suitable starting difficulty. To determine the difficulty, we adjusted the distance between bubbles using a modified version of the staircase method [9]. The levels started with either 1.2 meters or 2 meters distance between bubbles and adjusted itself every 5 meters by 0.3 meters, depending on game performance within this section.

The players played an average of 3.8 (*SD*=1.79, *Min-Max*=1 − 6) levels, based on their own wishes. During the game, some difficulties were observed. For example, P_1 had challenges to complete the tutorial. Since we did not implement a way to skip the tutorial on the remote control, it was turned off for the other four players. Also, none of the participants was able to finish the calibration sequence. When they had tried their best and did not make any more progress, we skipped the calibration on the remote control, playing with the default values. With varying physical and cognitive abilities, their success in the game also changed from person to person, e.g., for P_5, the game worked quite smoothly while P_1 had some difficulties. With the dynamic difficulty, we found that the mean distance between bubbles (mean of all data points) approached 2.38 meters (*SD*=0.91). We note that players played different levels with varying cognitive and physical tasks and a different amount of levels, therefore, the average distance serves as a reference start point for the future, not for comparison purposes.

5 DISCUSSION & CONCLUSION

In this work, we presented and evaluated *Canoe VR*, which enables cognitive and physical training for older adults. We considered cognitive variances as well as physical needs of older adults. In an HCD process [12], we gathered feedback from older adults with and without dementia, a health professional, and fitness experts. We iteratively improved the game to provide positive player experience and offer both cognitive and physical training opportunities.

We used various methods in this study for data collection. Although, this can be seen as a limitation in terms of not obtaining homogeneous findings, it also offers rich results. Our sample size is low and the results should be interpreted with this in mind. Also, the reflexive orientation of thematic analysis requires the deep engagement of the researcher with data, therefore does not require more than one coder [3–5]. Following qualitative research practices [26], we report on reflexivity that might have led to bias in the analysis, such as the main author's prior research experience with older adults, VR game experience, and a background in psychology and cognitive systems. Yet, these also stands as a strong ground to understand and interpret the data.

Overall, the results indicate that the game was very well received by users and the movements were found rich by the experts. Although participants suggested improvements or pointed to unsuccessful points as well, they generally expressed positive experiences. Our results, aligning with prior work [6], show that older adults' VR exergame experiences were positive, in particular during COVID-19. For this specific game scenario, the used virtual agent was not found to be helpful. In contrast, a multiplayer mode was a good addition, but the front players also desired more interactions with their partner, leaving a gap for improvement.

Canoe VR was developed through a design process with multiple stakeholders and several prototyping sessions. We also showed that this game can be used in regular gym sessions of older adults. By building on previous work [18], we show that a VR exergame can be suitable as an additional fitness tool for this user group, yet further scientific explorations are required to generalize this result. Although the game features a variety of cognitive tasks and movements, more research is needed to examine the cognitive and physical benefits (e.g., using subjective and objective measures) of this designed game.

In terms of varying abilities in the population, our findings indicate that *Canoe VR* offers opportunities for both, players with and without dementia. The game settings provide options to change the game tasks, elements, length, and difficulty. Further, dynamic difficulty integration allowed the players with varying abilities to experience the game and exercise. Therefore, based on the sessions with players with dementia, we argue that this game also matches their abilities. The importance of this is also emphasized in [16]. However, we particularly plan to explore this game in more detail with both groups in our future work. We plan to test this game with older players with and without dementia in a long-term study to understand which aspects of the game are suitable for which group and how these affect the game and their cognitive abilities.

ACKNOWLEDGMENTS

We would like to thank our participants and Hospital zum Heiligen Geist Hamburg for supporting our research and Daniel Apken for his contribution to the implementation of the game. This research was supported by the German Federal Ministry of Education and Research (BMBF).

REFERENCES

[1] Vero Vanden Abeele, Brenda Schraepen, Hanne Huygelier, Celine Gillebert, Kathrin Gerling, and Raymond Van Ee. 2021. Immersive Virtual Reality for Older Adults: Empirically Grounded Design Guidelines. *ACM Trans. Access. Comput.* 14, 3, Article 14 (aug 2021), 30 pages. https://doi.org/10.1145/3470743

[2] ARtillery Intelligence. 2021. Virtual reality (VR) headset unit sales worldwide from 2019 to 2024 (in million units). https://arinsider.co/ The Statistic Portal, Statista from https://www.statista.com/statistics/677096/vr-headsets-worldwide/. Accessed: 10.01.2022.

[3] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[4] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47. https://doi.org/10.1002/capr.12360

[5] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology* 18, 3 (2021), 328–352. https://doi.org/10.1080/14780887.2020.1769238

[6] Pablo Campo-Prieto, Gustavo Rodríguez-Fuentes, and José Mª Cancela-Carral. 2021. Immersive virtual reality exergame promotes the practice of physical activity in older people: an opportunity during COVID-19. *Multimodal Technologies and Interaction* 5, 9 (2021), 52. https://doi.org/10.3390/mti5090052

[7] Joyce Y C Chan, Tak Kit Chan, Timothy C Y Kwok, Samuel Y S Wong, Allen T C Lee, and Kelvin K F Tsoi. 2020. Cognitive training interventions and depression in mild cognitive impairment and dementia: a systematic review and meta-analysis of randomized controlled trials. *Age and Ageing* 49, 5 (05 2020), 738–747. https://doi.org/10.1093/ageing/afaa063

[8] Stanley J. Colcombe, Kirk I. Erickson, Naftali Raz, Andrew G. Webb, Neal J. Cohen, Edward McAuley, and Arthur F. Kramer. 2003. Aerobic Fitness Reduces Brain Tissue Loss in Aging Humans. *The Journals of Gerontology: Series A* 58, 2 (02 2003), M176–M180. https://doi.org/10.1093/gerona/58.2.M176

[9] Tom N Cornsweet. 1962. The staircase-method in psychophysics. *The American journal of psychology* 75, 3 (1962), 485–491. https://www.jstor.org/stable/1419876

[10] Claudine Fabre, Karim Chamari, Patrick Mucci, J Masse-Biron, and C Prefaut. 2002. Improvement of cognitive function by mental and/or individualized aerobic training in healthy elderly subjects. *International Journal of Sports Medicine* 23, 6 (2002), 415–421. https://doi.org/10.1055/s-2002-33735

[11] FitXR. 2019. *BoxVR.* Game [VR]. FitXR. London, England. https://fitxr.com/boxvr. Steam Reviews: https://store.steampowered.com/app/641960/BOXVR/. Accessed: 30.12.2021.

[12] International Organization for Standardization. 2019. ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. https://www.iso.org/standard/77520.html International Organization for Standardization. Geneva, Switzerland.

[13] Beat Games. 2019. *Beat Saber.* Game [VR]. Beat Games. Los Angeles, CA, USA. https://beatsaber.com/. Steam Reviews: https://store.steampowered.com/app/620980/Beat_Saber/. Accessed: 30.12.2021.

[14] Kathrin Gerling, Jonas Schild, and Maic Masuch. 2010. Exergame Design for Elderly Users: The Case Study of SilverBalance. In *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology* (Taipei, Taiwan) *(ACE '10)*. Association for Computing Machinery, New York, NY, USA, 66–69. https://doi.org/10.1145/1971630.1971650

[15] Teresa Hirzle, Maurice Cordts, Enrico Rukzio, Jan Gugenheimer, and Andreas Bulling. 2021. A Critical Assessment of the Use of SSQ as a Measure of General Discomfort in VR Head-Mounted Displays. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 530, 14 pages. https://doi.org/10.1145/3411764.3445361

[16] Sukran Karaosmanoglu, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. 2021. Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 252 (oct 2021), 27 pages. https://doi.org/10.1145/3474679

[17] Lucie Kruse, Sukran Karaosmanoglu, Sebastian Rings, Benedikt Ellinger, Daniel Apken, Thandiwe Feziwe Mangana, and Frank Steinicke. 2021. A Long-Term User Study of an Immersive Exergame for Older Adults with Mild Dementia during the COVID-19 Pandemic. In *ICAT-EGVE 2021 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments* (Sankt Augustin, Germany), Jason Orlosky, Dirk Reiners, and Benjamin Weyers (Eds.). The Eurographics Association, Geneve, Switzerland, 9–18. https://doi.org/10.2312/egve.20211322

[18] Lucie Kruse, Sukran Karaosmanoglu, Sebastian Rings, Benedikt Ellinger, and Frank Steinicke. 2021. Enabling Immersive Exercise Activities for Older Adults: A Comparison of Virtual Reality Exergames and Traditional Video Exercises. *Societies* 11, 4 (2021), 134. https://doi.org/10.3390/soc11040134

[19] Odders Lab. 2019. *OhShape.* Game [VR]. Odders Lab. Seville, Spain. https://ohshapevr.com/. Steam Reviews: https://store.steampowered.com/app/1098100/OhShape/. Accessed: 30.12.2021.

[20] Joseph J. LaViola. 2000. A Discussion of Cybersickness in Virtual Environments. *SIGCHI Bull.* 32, 1 (jan 2000), 47–56. https://doi.org/10.1145/333329.333344

[21] Ting Li, Ye Yao, Yan Cheng, Bing Xu, Xinyi Cao, David Waxman, Wei Feng, Yuan Shen, Qingwei Li, Jijun Wang, et al. 2016. Cognitive training can reduce the rate of cognitive aging: a neuroimaging cohort study. *BMC Geriatrics* 16, 1 (2016), 1–12. https://doi.org/10.1186/s12877-016-0194-5

[22] Andre McGrail, Alex Best, and Stintah. 2020. *Boat Attack.* https://github.com/Unity-Technologies/BoatAttack. Accessed: 06.01.2022.

[23] Meta Platforms, Inc. 2020. Meta Quest 2. https://www.oculus.com/quest-2/ Released: 13 October 2020.

[24] Anna Moschny, Petra Platen, Renate Klaaßen-Mielke, Ulrike Trampisch, and Timo Hinrichs. 2011. Barriers to physical activity in older adults in Germany: a cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity* 8, 1 (2011), 1–10. https://doi.org/10.1186/1479-5868-8-121

[25] John Muñoz, Samira Mehrabi, Yirou Li, Aysha Basharat, Laura E Middleton, Shi Cao, Michael Barnett-Cowan, and Jennifer Boger. 2022. Immersive Virtual Reality Exergames for Persons Living With Dementia: User-Centered Design Study as a Multistakeholder Team During the COVID-19 Pandemic. *JMIR Serious Games* 10, 1 (19 Jan 2022), e29987. https://doi.org/10.2196/29987

[26] Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand, and Jon Howard Raphael. 2012. No room for reflexivity? Critical reflections following a systematic review of qualitative research. *Journal of Health Psychology* 17, 6 (2012), 866–885. https://doi.org/10.1177/1359105311427615

[27] Seong-Hi Park, Kuem Sun Han, and Chang-Bum Kang. 2014. Effects of exercise programs on depressive symptoms, quality of life, and self-esteem in older people: A systematic review of randomized controlled trials. *Applied Nursing Research* 27, 4 (2014), 219–226. https://doi.org/10.1016/j.apnr.2014.01.004

[28] Prefrontal Cortex. 2020. *Boat Attack.* https://github.com/prefrontalcortex/BoatAttack/tree/boat-attack-vr. Accessed: 13.01.2022.

[29] Hamilton Roschel, Guilherme G. Artioli, and Bruno Gualano. 2020. Risk of Increased Physical Inactivity During COVID-19 Outbreak in Older People: A Call for Actions. *Journal of the American Geriatrics Society* 68, 6 (2020), 1126–1128. https://doi.org/10.1111/jgs.16550

[30] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

[31] Frank Steinicke. 2016. *Being Really Virtual: Immersive Natives and the Future of Virtual Reality.* Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-43078-2

[32] Jamie L. Tait, Rachel L. Duckham, Catherine M. Milte, Luana C. Main, and Robin M. Daly. 2017. Influence of Sequential vs. Simultaneous Dual-Task Exercise Training on Cognitive Function in Older Adults. *Frontiers in Aging Neuroscience* 9 (2017), 368. https://doi.org/10.3389/fnagi.2017.00368

[33] Unity Technologies. 2021. Unity Real-Time Development Platform. https://unity.com/ LTS Release 2019.4.18f1. 13 January 2021. https://unity3d.com/unity/qa/lts-releases?version=2019.4&page=2. Accessed: 30.12.2021.

[34] United Nations, Department of Economic and Social Affairs, Population Division. 2020. World Population Ageing 2019 (ST/ESA/SER.A/444). https://doi.org/10.18356/6a8968ef-en

[35] Michael Valenzuela and Perminder Sachdev. 2009. Can cognitive exercise prevent the onset of dementia? Systematic review of randomized clinical trials with longitudinal follow-up. *The American Journal of Geriatric Psychiatry* 17, 3 (2009), 179–187. https://doi.org/10.1097/JGP.0b013e3181953b57

[36] Valve Corporation. 2019. Valve Index. https://www.valvesoftware.com/en/index Released: 28 June 2019.

[37] Joeke Van Santen, Rose-Marie Dröes, Marije Holstege, Olivier Blanson Henkemans, Annelies Van Rijn, Ralph De Vries, Annemieke Van Straten, and Franka Meiland. 2018. Effects of exergaming in people with dementia: results of a systematic literature review. *Journal of Alzheimer's Disease* 63, 2 (2018), 741–760. https://doi.org/10.3233/JAD-170667

[38] Matthias Wilde, Katrin Bätz, Anastassiya Kovaleva, and Detlef Urhahne. 2009. Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften* 15 (2009), 31–45. https://archiv.ipn.uni-kiel.de/zfdn/pdf/15_Wilde.pdf

[39] World Health Organization. 2020. Physical Activity. https://www.who.int/news-room/fact-sheets/detail/physical-activity Accessed: 09.01.2022.

[40] Minoru Yamada, Y Kimura, D Ishiyama, Y Otobe, M Suzuki, S Koyama, T Kikuchi, H Kusumi, and H Arai. 2020. Effect of the COVID-19 epidemic on physical activity in community-dwelling older adults in Japan: a cross-sectional online survey. *The Journal of Nutrition, Health & Aging* 24 (2020), 948–950. https://doi.org/10.1007/s12603-020-1501-6

10.2 Lessons learned from a human-centered design of an immersive exergame for people with dementia

[CorePub5] **Sukran Karaosmanoglu**, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. Lessons learned from a human-centered design of an immersive exergame for people with dementia. *Proceedings of the ACM on Human-Computer Interaction*, 5 (CHI PLAY), 2021. Association for Computing Machinery. ISBN 9781450391566. DOI 10.1145/3474679. URL https://doi.org/10.1145/3474679.

Supplementary materials for this article are available here: https://doi.org/10.1145/3474679.

Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia

SUKRAN KARAOSMANOGLU, Human-Computer Interaction, Universität Hamburg, Germany
SEBASTIAN RINGS, Human-Computer Interaction, Universität Hamburg, Germany
LUCIE KRUSE, Human-Computer Interaction, Universität Hamburg, Germany
CHRISTIAN STEIN, gamelab.berlin, Humboldt Universität zu Berlin, Germany
FRANK STEINICKE, Human-Computer Interaction, Universität Hamburg, Germany

Cognitive-physical exercises can reduce the progression of dementia. However, traditional methods often induce problems (e.g., lack of motivation), whereas the success of recent virtual reality (VR) exergames such as Beat Saber may provide a playful, motivational, and immersive alternative. Yet, until now, it remains unclear which game mechanics, concepts, and designs work best for people with dementia, and how to implement exergames for and with this user group. In this paper, we adapted a human-centered design approach to address the specifics of developing VR exergames for people with dementia. This includes semi-structured interviews with stakeholders and contextual inquiries to better analyze the user requirements. Based on our analysis, we present *Memory Journalist VR* - a novel VR exergame specifically designed for people with dementia in a participatory design process. We report the qualitative evaluation based on the feedback gathered in five focus group sessions. Finally, we discuss the lessons learned, which provide important insights for the design of future VR exergames for people with dementia: (i) creating social gaming activities with a focus on shared aspects, (ii) support of an inverse game flow channel addressing decline and variance in cognitive-physical abilities, and (iii) ensuring a safe VR exergame experience.

CCS Concepts: • **Human-centered computing → Accessibility**; • **Applied computing → Life and medical sciences**; • **Software and its engineering → Interactive games**.

Additional Key Words and Phrases: serious games; VR exergames; people with dementia; human-centered design approach; health

ACM Reference Format:
Sukran Karaosmanoglu, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. 2021. Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 252 (September 2021), 27 pages. https://doi.org/10.1145/3474679

1 INTRODUCTION

According to the World Health Organization (WHO), the number of people with dementia worldwide has increased to more than 50 million. Due to demographic change, the figure continues to increase substantially every year [93]. Dementia causes extreme changes in someone's cognitive-physical abilities [7, 22] and impacts the economic and social resources of families, health systems, and governments [13, 94]. Although there is no cure for dementia yet [7], previous work has shown that cognitive and physical exercises can reduce the risk and progression of dementia [38, 95]. The combined interventions highlighted more positive effects compared to separate use [21, 60]. These

Authors' addresses: Sukran Karaosmanoglu, Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, sukran.karaosmanoglu@uni-hamburg.de; Sebastian Rings, Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, sebastian.rings@uni-hamburg.de; Lucie Kruse, Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, lucie.kruse@uni-hamburg.de; Christian Stein, gamelab.berlin, Humboldt Universität zu Berlin, Berlin, Germany, christian.stein@hu-berlin.de; Frank Steinicke, Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, frank.steinicke@uni-hamburg.de.

traditional exercise methods contain some challenges (e.g., lack of motivation, distraction from external stimuli, and missing supervision [72]). Innovative technology in the area of serious games has the potential to overcome these challenges.

Virtual reality (VR) can address some of the mentioned limitations. The current VR technology such as head-mounted displays (HMDs) is used in many areas ranging from entertainment, education, and training to health. VR systems can provide fully-immersive environments, which stimulate multiple human sensory channels [28, 78]; thus it can minimize the distraction from actual tasks. The possibility to display virtual content everywhere around the user enables full body motions while ensuring that users still see instructions or information on the HMD. Also, VR allows for customization of virtual environments (VEs) [28], which can help to serve the specific needs of user groups; for example, enabling worlds that are no longer accessible for people with dementia living in psychiatric hospitals [83]. However, this technology also exposes some disadvantages, such as a limited spatial range of movement, ergonomics, and nausea [39, 79]. Especially for people with dementia, the use of VR requires special attention to avoid unethical and undesirable effects (e.g., emotional stress) because VR can be realistic and create a sense of presence [75, 76]. Nevertheless, prior work showed that VR experiences can be engaging and hold promise for future exploration, especially when the interests of this user group are considered [40].

Serious games that combine physical exercises with cognitive stimulation [89] can motivate people, in particular people with dementia [17], to perform physical exercises due to their playful nature. VR *exergames* have enormous potential as a training tool when designed around their effectiveness and attractiveness [73]. These games can adapt VEs and game tasks to individuals' ability levels and create safety-controlled exercises. While previous works showed the significant potential of VEs [40, 83] and the motivational aspects of VR exergames for people with dementia [17], a comprehensive understanding of their needs when designing VR exergames remains sparse. Particularly, it is largely unknown, which game mechanics, concepts, and designs work best for people with dementia, and how to implement exergames for and with this user group.

While there are several established methods for developing interactive systems/games in a human-centered design (HCD) approach to ensure effective, efficient, and satisfactory usage [24], these standard procedures might not be appropriate for people with dementia [31]. This is due to the fact that people with dementia are characterized by very individual, often daily form-dependent manifestations of dementia, which limits their cognitive-physical capabilities, and also sometimes hinders their ability to express feelings or thoughts [49, 93]. Hence, it is absolutely necessary that caregivers and people who have close relationships with people with dementia (significant others) are taken into account as well as the routines of care activities and the living conditions of this user group. To address these limitations of current VR exergames and HCD approaches, our work is driven by the following research questions:

- RQ_1: "*Does the designed VR exergame provide positive and good player experiences, usability, and accessibility for people with dementia?*"
- RQ_2: "*Which aspects of the HCD approach should be adapted in the context of VR exergames for people with dementia?*"

Therefore, we introduce an HCD approach for developing VR exergames adapted to the special requirements and context of dementia. In particular, this approach contains (i) semi-structured interviews with stakeholders as well as contextual inquiries to analyze the user requirements of people with dementia. Furthermore, we report about the remaining iterative steps of the HCD apporach [24]: (ii) conceptual design of the VR exergame, (iii) implementation of the prototype, and (iv) evaluation of the prototype. In **Step 1**, we interviewed clinical professionals, a physiotherapist, technical experts, and significant others. Following that, we conducted contextual inquiry sessions with older adults with and without dementia. We employed thematic analysis on the contextual inquiry sessions and interview data to extract user requirements, which we applied to our exergame. In **Step 2**, based on the findings, we designed a VR exergame prototype to gain a better understanding of varying preferences and needs of people with dementia, which was implemented in **Step 3**. In **Step 4**, we evaluated this VR exergame with five focus group sessions by creating an iterative design circle, which allowed us to improve and adapt the game to the needs of people with dementia.

The contributions of this work can be summarized as follows:

- an adaptation of the HCD approach to design VR exergames for people with dementia: semi-structured interviews with stakeholders and contextual inquiry sessions.
- conceptual design and iterative development of a VR exergame based on the findings from the user requirement analysis: *Memory Journalist VR*,
- qualitative evaluation of the VR exergame prototype across five focus groups,
- provision of design implications of VR exergames for people with dementia to allow novel insights for game designers and developers.

2 RELATED WORK

Previous studies have shown the positive effects of physical-cognitive training on people with dementia. While there are prior works indicating exergames can create positive experiences, immersive games for them have rarely been studied yet. Moreover, a detailed exploration of the needs and requirements of this user group as well as the question on how an HCD approach should be undertaken with them remains sparse in this context.

2.1 Physical & Cognitive Activities

Neurodegenerative conditions such as dementia can cause a decline in people's cognitive and physical abilities [93], significantly impacting their daily life. Thus, non-pharmacological research focuses on physical [16, 69] or cognitive intervention techniques [71, 77] to reduce and prevent dementia.

Cognitive activities (e.g., reading books and solving puzzles) have been shown to serve as a technique to reduce the risk of dementia [71], as they increase cognitive reserve through stimulation in the brain [80, 81]. Spector et al. [77] showed an improvement in quality of life scores for people with dementia who received cognitive stimulation therapy (questions related to reminiscence, money, and games) compared to a control group. Drawing from these studies, the human-computer interaction (HCI) area also employed elements of reminiscence therapy in their technologies [33, 48].

The findings of physical activity studies also highlighted a decreased risk of dementia [16, 69] and improved cognitive and physical well-being [38, 96]. Rolland et al. [68] reported that physical exercises can help to reduce the progression of people with dementia's performance decline in everyday activities such as walking. Moreover, prior work [96] reported an increase in cognitive functions (e.g., attention, visual memory, and working memory) of users who completed 6 weeks of exercise training compared to a control group; thus, prompting us to explore this in our research.

To increase both interventions' benefits, research investigated the effects of combined cognitive-physical training with older adults; the findings indicated higher cognitive [21] and lowered depression scores [60]. Similarly, the combined interventions improved the cognitive functions, daily living activities, and mood of people with dementia [47]. These findings motivated us to explore the combination of cognitive-physical activities in our immersive game, as VR offers a medium to create simultaneous cognitive-physical exercises by serving multiple sensory channels [28, 78] and exergames support mental and physical stimulation [45].

2.2 Exergames

Games have a great potential to increase older adults' enjoyment and motivation [26, 29]; this also applies for people with dementia [17]. Exergames can help to overcome some of the difficulties of traditional exercise methods such as loss of motivation and time constraints of caregivers/players [72, 89]. However, the decline in cognitive-physical abilities of older adults –especially for people with dementia– requires consideration to make games accessible to these users. Gerling et al. [30] suggested designing exergames for older adults, taking into account the changing cognitive-physical abilities of players due to aging.

Research has explored exergaming for dementia in many contexts. Padala et al. [61] reported an improved balance and reduced fear of falling for people with dementia in a home-based caregiver supervised *Wii-Fit* [58] program. Concerning social well-being, Unbehaun et al. [86] investigated the social potential of video game-based exergames on people with dementia, their formal/family caregivers; their findings indicated an improved social interaction for this user group and their relatives. To our knowledge, however, there are no studies or recommendations on how to design or make exergames accessible for this user group. However, it can be assumed that the design requirements of people with dementia will differ from those without dementia.

Due to its fully-immersive nature allowing people to dive with their full body and mind into a virtual world [78], VR provides an interesting medium for exergames. It serves multiple sensory channels and thereby supports physical actions and cognitive stimulation. By evoking the feeling of presence [76], VR can increase attention to tasks and reduce external cues which might distract from the actual task. Moreover, VR offers the opportunity to customize the entire 3D VEs and tasks [28], allowing people with dementia to visit places they cannot visit in the real world [83] and to have enriching experiences [40]. These customizable features of VR can help cater to users' varying abilities by tailoring physical movement patterns to their needs and provide them a better representation in 3D; leading to the use of VR as a training and evaluation tool in various studies [59, 66]. Although VR technology is promising and used in many areas, this technology also presents many drawbacks such as limited physical tracking space [79] and cybersickness [52]. Exposure to VR can lead to some adverse effects on humans such as nausea, disorientation, eye strain, and neck pain [39, 54]. In particular, major conflicts between visual and vestibular senses and discrepancies between users' head movements and visual VR scene movements were shown as common causes of these symptoms [52, 54]. Thus, we decided to avoid creating such scenarios (e.g., big discrepancies between senses) to reduce the potential negative effects of VR on our users.

VR exergames are becoming increasingly popular in the commercial market for entertainment and exercise purposes (e.g., Beat Saber [27], OhShape [51], Ragnaröck [92]). This trend also took place in academic research. Unlike the commercial market, research has also focused on the potential of VR exergames for people with dementia [17, 18]. For instance, Eisapour et al. [17] designed a VR exergame program to promote physical exercise, resulting in an enjoyable experience. Similarly, a recent work [62] reported preliminary results of a VR exergame study, in which people with dementia played the role of a conductor. Their findings indicated that VR exergames can be seen as a motivator for performing physical activities. Eisapour et al. [18] compared a VR exergame-based program with traditional exercises guided by a therapist. They found comparable results on the values of enjoyment, comfort, and difficulty. With respect to reduced physical abilities due to aging, Rings et al. [65] presented seated VR exergame exercises designed with the participation of physiotherapists.

While there is little research showing the effectiveness and motivational aspects of VR exergames [17, 18, 62] for people with dementia, previous work has neither focused on their requirements in VR exergames nor on the actual HCD approach, which could guide the development of future immersive applications for this user group.

2.3 Human-Centered Design Approach

The HCD aims to ensure high accessibility, usability, and user experience by providing methods to understand users' needs, capabilities, and limitations. This approach prioritizes the users and offers opportunities to engage with them and stakeholders throughout the entire design process [24]. It has been used in the context of health [37] and games [56, 90]. For instance, Harte et al. [37] applied a three-step HCD methodology to design a fall prevention system for older adults, including end-users and experts. Dementia affects everyone differently and with varying severity [93]. Therefore, familiarization with this user group, observation of their interactions with the designed systems, and gaining perspective of stakeholders holds major importance.

The HCI field is increasingly exploring dementia in design, for example in the context of art therapy [53], everyday sound [42], and VR [40]. While standard HCD approaches suggest including end-users and stakeholders, designing for people with dementia poses often some challenges that cannot be addressed with a typical HCD approach. For example, the decline in cognitive-physical abilities might limit their capabilities of expression and reasoning skills [49]. Yet, the changes in their abilities should not be seen as a constraint to the design, but as an opportunity to create new and meaningful experiences by focusing on their current capabilities [23, 53, 91]. Therefore, people with dementia must undoubtedly be included in the process, but stakeholders such as significant others or caregivers could provide insight on behalf of or together with the end user group as they are the ones who know best about people with dementia's daily routines and behavior patterns.

Furthermore, the design of VR exergames for people with dementia requires the expertise and the input of people from various fields. As the field covers health, physical activities, VR, and game research, it is important to include different experts who can enrich the design process and inform other parties (e.g., about health conditions). For example, health professionals might simply not be aware of technological advancements while game researchers might have difficulties understanding the abilities –and thereby requirements– of people with dementia to create appropriate game mechanics. The lack of empathy by mostly young product developers is often mentioned as a possible problem [19, 25]. Exergames should also be designed considering the effectiveness of the physical activities [73]; thus making the involvement of physiotherapists inevitable. Therefore, it is important to first get a better understanding of the needs and requirements of people with dementia from stakeholders.

Another issue that should be considered is how the developed systems are used and what purpose they serve; we hope to support people with dementia in clinical settings by first demonstrating the feasibility of VR exergames for them. Many studies are showing that long-term training is required to achieve these outcomes [68, 77]. Regarding this, systems should be easy to use for caregivers/significant others, as they often motivate people with dementia to use the technology, especially in long-term use [67]. From another angle, this feature can also help to reduce the workload for caregivers and increase social interaction [86]. Therefore, many studies have involved caregivers, physiotherapists, and hospital staff to serve these purposes when designing systems for people with dementia [17, 55, 86, 91].

To develop a VR exergame, which provides high usability, player experience, and accessibility, we formed a research team consisting of a motion scientist, game designers, developers, and a health professional to serve the entire design process. In addition, we enriched this research with additional stakeholders (e.g., clinical experts, physiotherapist, technicians, significant others) to better understand the people with dementia's needs.

3 STEP 1: UNDERSTANDING THE NEEDS AND REQUIREMENTS OF PEOPLE WITH DEMENTIA

The main purpose of the study was to understand the needs and requirements of people with dementia and to apply this knowledge in the context of VR exergames, thereby providing engaging experiences. As a secondary goal, we were interested in adapting the HCD approach in this context. To extract the user requirements, our research team conducted semi-structured interviews [2] and contextual inquiry sessions [6]. We chose these methods to understand the needs and experience (for which there are no validated questionnaires) of the participating older adults, which were from the same senior living facility (Hospital zum Heiligen Geist Hamburg[1]). The participants of our study were recruited mainly with the help of the senior living facility. The overall study procedure is illustrated in Figure 1.

In order to consider ethical implications, we conducted an ethics workshop at the first phase of our large project prior to the described study in this paper and invited an ethics expert (female, age=57, ethics in IT) to observe our process. We applied the lessons learned to our study (e.g., termination of experience in case of any psychological or physical stress) and closely collaborated with the senior living facility (e.g., preparation of suitable questions and presence of a caregiver during each session).

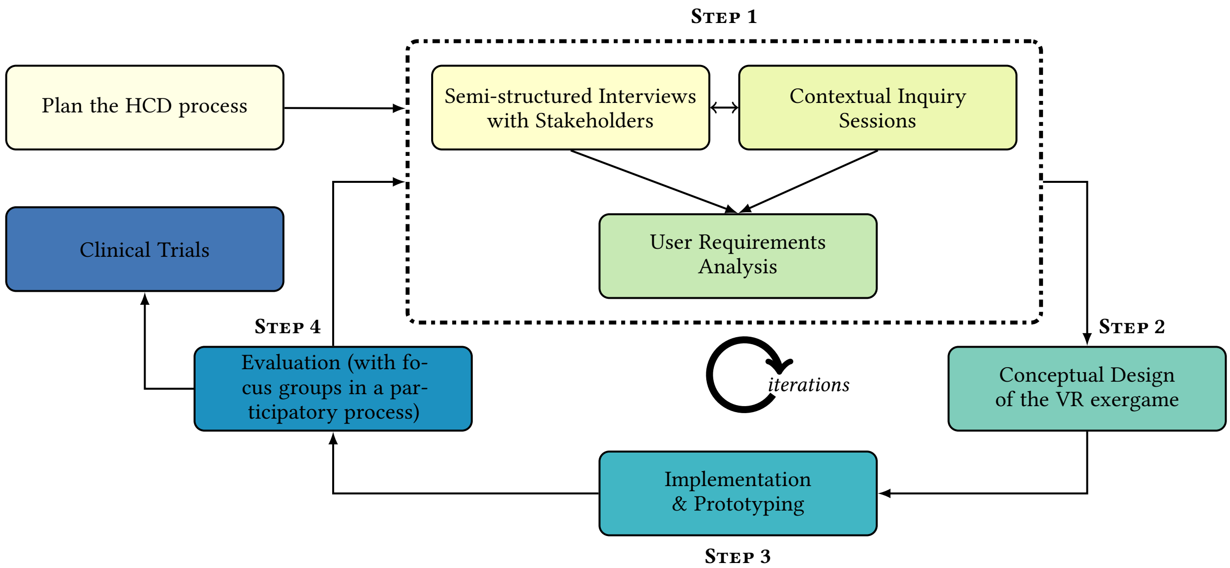

Fig. 1. In Step 1, we conducted semi-structured interviews with stakeholders and the contextual inquiry sessions with people with dementia to analyze the user requirements. We further followed the remaining steps of the HCD approach [24]: conceptual design of a VR exergame (Step 2), implementation of the VR exergame (Step 3), and evaluation the VR exergame with five focus group sessions (Step 4). We iterated the findings over requirements, design, and implementation steps.

3.1 Semi-structured Interviews with Stakeholders

We conducted semi-structured interviews with the stakeholders to explore their attitudes towards VR technology, conditions and needs of people with dementia, the possible physical movements, and

[1] https://www.hzhg.de/

game scenarios that can be implemented in VR. As game researchers/developers, the project was our first experience with this user group, and, therefore, we placed importance on understanding them and how to interact with them before our first face-to-face encounter. The full list of interview questions can be found in the supplementary materials.

Participants. The interviews were conducted with 11 stakeholders: three significant others (n=3 female, age: M=61.33, SD=2.89 years), five clinical dementia experts[2] (n=4 female, n=1 male, age: M=43.6, SD=10.55 years), a physiotherapist (male, age: 39 years), and two technicians (n=2 male, age: M=50, SD=15.56 years).

Procedure. While we followed our interview guideline with pre-determined questions, we addressed follow-up questions to our participants. All interviews were performed separately with each stakeholder. The interviews were audio-recorded, except for two participants due to privacy concerns. In addition to audio-recordings, field notes were taken by the researchers.

3.2 Contextual Inquiry Sessions

In these sessions, the research team aimed to observe people with dementia's cognitive-physical conditions, interactions, and attitudes from a first-person perspective (see Figure 2a,b). The followed guideline for this step can be found in the supplementary materials.

Participants. We included both people with and without dementia to observe a wider interaction of this age group with the gameplay in VR. While the design requirements of individuals with dementia are different from those without dementia, they may show some commonalities due to shared culture, similar age range, and shared aesthetic values. The sessions were conducted separately for groups of people with and without dementia. We recruited eight participants (7 female, 1 male; n=3 mild dementia, n=5 people without dementia) with an average of 81 years (SD=6.05). A health professional from the research team (female, age: 51, 26 years of experience working with people with dementia) also accompanied the players in these sessions.

Procedure. Following the contextual inquiry method, we have enriched the familiar social space used by older adults for reading, board, or video games (i.e., Kinect games) with VR hardware. We audio-recorded all sessions.

First, the participants were introduced to the purpose of our project and gave their consent (and legal representatives for people with dementia) to participate in our study. This document included information on the planned project, what the player will experience in VR, and the symptoms of cybersickness. Following that, we observed their experience while they played Kinect games, which is part of their usual daily routine in this senior living facility. Before the VR session, all participants were verbally informed about their upcoming experience, warned of possible negative symptoms of VR, and asked to tell us if they feel uncomfortable or experience any difficulties. The research team members observed the whole experience to terminate in case of any problems. The participants first experienced the SteamVR [87] home environment, which shows a large terrace with a view of mountains, trees, and birds. After some time of familiarization (ca. 5 minutes for each participant), we used the VR game *Tilt Brush* [32] to provide a more interactive gaming experience. We observed their interaction with the controllers in VR while decorating a snowman. We then used the NVIDIA VR arcade application [82] only for people without dementia (due to time constraints of people with dementia and the health professional in their daily routine) to see a larger variety in types of interaction.

[2] All clinical experts had different specialties: E_1 Responsible expert for working with people with dementia, E_2 Head of inpatient and short-term care at a senior living facility, E_3 Responsible person for social care at a senior living facility, E_4 Head of a senior living facility, E_5 Head of nursing at a senior living facility.

Fig. 2. Impressions from the HCD approach in which we introduced VR technology to people with dementia (a) and created a social environment between people with dementia, spectators, and caregivers (b). We further designed a VR exergame for people with dementia using a 3D printed camera controller (c) and included people with dementia with varying cognitive-physical abilities in these experiences (d).

3.3 Data Analysis & Results

To analyze qualitative data, we transcribed the audio recordings of the contextual inquiry sessions and the interviews (a total of 13:06 hours). Afterwards, the first author applied the reflexive approach of inductive thematic analysis, which does not require multiple coders [9–11]; the reflexive approach of thematic analysis involves deep engagement of researchers and their interpretations on data [10]. Our approach consisted of the steps introduced by Braun and Clarke [10]: familiarization of data, inductive coding of data, creation of initial themes, iteratively reviewing and shaping the themes, and final theme generation phase.

We identified four themes through our inductive coding, which serve to determine the user requirements. We report quotes based on the participants' specialty, and number: people with (P_{PWD}) and without dementia (P_{PW}), the health professional from the research team (P_{HP}), clinical experts (P_E), technicians (P_T), physiotherapist (P_{PH}), and significant others (P_S); for example, P_{PWD_1} represents the quotes from the first player with dementia in the contextual inquiry session.

Theme 1: ***Novelty effects of VR technology***. Only one player reported having prior VR experience, which resulted in a negative one: "*I had to leave the [place] because I got sick*"-P_{PW_2}. For the rest VR was unknown, yielding conversations between the research team and players: "*[...] iPad, I play a lot with it, and that is already VR [laughs]*"-P_{PW_1}. This unfamiliarity created some hesitation to play VR games for older adults without dementia (e.g., "*I don't even know if I can play it, because I have a fear of heights and am prone to balance disorder*"-P_{PW_2}), while people with dementia warmly welcomed the new technology. Yet, after trying VR games, the majority of participants reported highly positive experiences: "*Oh, how beautiful!*"-P_{PWD_1} and "*exhausted from the excitement, but it was positive excitement*"-P_{PW_3}. Especially after the players experienced and were informed about the possibilities of VR, they reported two environmental themes they wished to see in VR exergames: exposure to nature (e.g., beach, sea) and sports activities (e.g., skiing, swimming). Moreover, these themes were always associated with the players' memories and outside environments: "*I was a passionate swimmer, was also in the swimming club. [...] We always went to the North Sea*"-P_{PWD_1}.

Theme 2: ***The role and requirements of social gaming activities***. The results indicated a trend towards designing social environment-based VR exergames with the involvement of spectators and caregivers. Generally, the participants saw VR exergaming activities as a potential to create a social environment among all parties, which can improve psychological well-being (e.g., "*These [social activities] can help to prevent frustration and reactivate skills [...] from the past*"-P_{E_5}). The experts reported the group dynamics that emerged from other game activities (e.g., board games, Kinect games) and how supportive they were to improve the psychological health of older adults: "*when the older adults say that they would like to be better and are on the last places again, the group can* *normally catch them well and cheer them up*"-P_{E_3} and "*In a game, where all people with dementia can participate, they are always integrated*"-P_{E_2}. We also observed these social interactions during our contextual inquiry sessions: spectators to a player with some vision problems: "*But you were great today!*" P_{PWD_1}: "*Really? Well, that's how it is, you develop tricks to compensate and that you can see again*".

However, the stakeholders found shared cognitive-physical conditions, interests, hobbies, and biographies essential to create positive experiences for people with dementia in a social setting: "*biography work has to be done beforehand to understand which game would be the most suitable one*"-P_{E_1}, and "*we need [...] to look at what the residents liked to do, right?*"-P_{E_2}. Otherwise, these aspects can turn out to be a potential danger, which can impact players' motivation and psychological well-being: "*You have to make sure that the characters fit together and also the severity of dementia, because if there are people who can still do it [an exercise] really well and others need a long time and guidance, then the others may get impatient or make a comment that is not so constructive to the group mood*"-P_{E_3}. As a notable challenge, the stakeholders reported reduced social interest of people with dementia in interacting with family members (e.g., "*she is rather apathetic, does not take part in conversation and is just sitting there*"-P_{S_3}) and with peers (e.g., "*I don't think she has contact with others*"-P_{S_2}).

Theme 3: ***Challenges due to reduced cognitive-physical abilities****.* All stakeholders pointed out how dementia affects cognitive abilities in an increasing way: "*she [person with dementia] always asks the same about relatives. They are all dead [...]. Always the same things and that repeats every five to seven minutes. [...] at the start, it was 10-15 minutes, but now it gets shorter*"-P_{S_2}. This also impacts how people with dementia perceive and react to communication (e.g., "*they [people with dementia] need a moment to understand it, and then it might be better to talk in slow motion. For them, it is normal tempo*"-P_{E_3}). These challenges were also observed during the contextual inquiry sessions. Particularly, we identified differences between people with and without dementia in their interactions with controllers, the need to give detailed game instructions, and to repeat tasks due to forgetfulness. We had to remove the HMD for a moment to create an association between real and virtual hands for people with dementia in VR. Further, we repeated the game details multiple times for them compared to people without dementia: "*Here? [showing the controllers] (Research team: Yes, there! Just up there, yes? Do you see that?) No.*"-P_{PWD_3}. However, people with dementia also showed differences within the group in terms of guidance needed in VR (e.g., instructing them and turning their chairs to look in different directions in the VEs).

The reduced cognitive abilities affected people with dementia's physical skills: "*the problem here is the cognitive limitations, because movement always means that I have to instruct someone, right?*"-P_{E_1}. However, the conditions of people with dementia can differ within the group: "*there are people who already have problems performing movements. Everyone makes a movement differently [...] and has their own movement patterns*"-P_{PH}. Thus, the physiotherapist explicitly suggested employing different levels of difficulty (e.g, "*I would always do these [exercises] from easy to difficult because even someone with problems can do an easy exercise well at first, then [...] have the feeling that [they] can do it well*"-P_{PH}.) and seated exercises (e.g., "*I [referring to people with dementia] do not have to use the torso. I can also use my arm while sitting, my body center of gravity shifts*"-P_{PH}.) In this way, a feeling of accomplishment should be given to motivate people with dementia: "*If I [physiotherapist] show them [people with dementia] a little bit of success, that would be an unbelievable psychological component*"-P_{PH}. Yet, one should at first consider the everyday effectiveness "*[...] to maintain their [people with dementia] independence and autonomy and all the things that affect everyday life*"-P_{PH}.

Theme 4: ***The Priority of Safety***. All participants always prioritized the safety of people with dementia. This important issue was further specified in terms of the technical setup, feeling safe in company, and psychological well-being.

Regarding the technical setup, the main outcome was to provide a safe environment while players were performing physical movements in VR: "*Fixed installations are preferred*"-P_{T_2} and "*I would never use tripods here*"-P_{T_1}. Particularly, the majority of the participants mentioned the risk of VR to be an isolating experience due to immersion and thereby losing the association between the real and virtual world for players: "*they wouldn't realize if a fire breaks out or something else happens while people with dementia wear the VR headset*"-P_{T_2}. The safety of the physical world was further discussed by taking into account the changing cognitive-physical abilities of people with dementia: "*the height of any fixtures, so people with wheeled walkers do not collide with it*"-P_{T_1}. Wearing an HMD was discussed on many occasions with the consideration of people with dementia's reaction to it (e.g.,"*the head is one of our most intimate areas*"-P_{E_1}), which can potentially create danger. Nevertheless, we did not encounter any complaints about wearing an HMD from our players in the contextual inquiry sessions.

Another notable topic was sharing VR exergame play between player and caregiver, which was partially attributed to a social factor but most importantly to control the gaming experience to ensure the safety of players: "*See, if she [person with dementia] is selecting the game, the program does not react how she expects and is confused and thinks she is doing something wrong*"-P_{HP} and "*they [all participants] can then share what they have experienced or achieved*"-P_{PH}. Also, being alone created some hesitance for the players to play VR exergames: "*As long as someone is there, everything is good! Yes, it gives me support and says: So you stand right or you can not fall down anywhere. [...] You can't practically play it alone at home*"-P_{PW_2}. These concerns were amplified for people with dementia: "*it could happen that they suddenly want to run and get lost in the truest sense of the word because they have forgotten that they are wearing their glasses*"-P_{PW_1}.

Finally, a main safety characteristic was preserving the psychological well-being of our participants. VR holds enormous potentials for creating many VR exergaming activities, but not every game is suitable for the well-being of people with dementia: for example, rejuvenation effects (e.g.,"*Even though people with dementia often think they are around 30 years old, it would cause a crisis if they saw themselves young while in a phase of clarity and in knowledge of their real age*"-P_{E_1}.), creating completely new encounters (e.g., "*That is what is in their everyday life anyway, they always see something new. With the VR glasses, we would only intensify this*"-P_{E_1}.), creating unrealistic experiences ("*you have to expect that the person who is in the forest [...] wants to pick up a stone*"-P_{E_1}.), and enclosed environments ("*[...] in a forest, the path is a natural boundary, but you can still see into the distance. Also at the beach, [...] a dock should not invite you to step on it [because it could be dangerous] and rather only be in the distance*"-P_{E_1}). Likewise, the health professional of the research team explicitly suggested that "*the people with dementia's experience [exergame and the time before and after] should follow a predictable procedure when possible, so that the people with dementia keep a sense of security*"-P_{HP}.

3.4 User Requirements Analysis

We discuss below the results of the user requirements analysis and summarize the recommendations for the VR exergames.

Create social gaming environments with the consideration of shared aspects. Our findings highlight the need to create a group-based VR exergame experience with the participation of spectators and caregivers. Most of the participants attributed this mainly to its potential to create social interaction between and among the parties. This feature was further observed in our contextual inquiry sessions as a source of motivation to encourage players. Nevertheless, our results also indicate the importance of shared cognitive-physical conditions, hobbies, interests, and biographies among older adults with dementia. This was associated with becoming more insecure and frail in behavior and loss of interest. Therefore, game designers/developers avoid these potentially discouraging factors to motivate people with dementia to play exergames.

Support of an inverse game flow channel addressing decline and variance in cognitive-physical abilities. Our evaluation shows that dementia has an extreme impact on the cognitive-physical abilities of older adults, which must be addressed when developing game mechanics and concepts. All players with dementia had difficulty understanding the use of VR controllers. This showcases that people with dementia have difficulties learning new interaction techniques. This requires designers consider providing intuitive techniques for this user group to interact with the virtual world. Second, people with dementia required multiple and detailed instructions to complete their gameplays due to forgetfulness. Hence, the designers should introduce continuous hints, explanations, reminders, or narratives into the gameplay.

Dementia can lead to varying cognitive-physical effects on older adults: requiring more repetitions of instructions or more physical support. Beyond this, every person with dementia can show this variance in their abilities over time [93]. We suggest game designers to consider the possibility of creating an inverse game flow channel to support the decline and variance in cognitive-physical abilities: for example, by creating different levels of difficulty or adaptivity in exergames. These can help a wide range of people with dementia to become a part of the VR exergame and to provide a sense of accomplishment.

Provision of a safe VR exergame experience. The results emphasize the importance of creating a safe VR exergaming play for people with dementia. First, developers must install fixed VR systems to avoid any physical injury, taking into account the effects of dementia. Second, caregivers should be part of the users' VR exergame experience to support them with their possible cognitive-physical and psychological difficulties; for example, by reacting to situations with empathy in a timely manner (might be challenging for young developers [19, 25]). Also, for most players, the presence of the caregiver was the connection to the physical world; they expressed their concerns about falling while wearing an HMD and performing movements without a caregiver. This indicates the need for a caregiver and to reduce the risk of falls in VR exergames for people with dementia. Third, game designers must pay attention to the details of in-game scenarios to preserve the psychological well-being of people with dementia and avoid creating scenarios that might cause emotional stress for this user group.

4 STEP 2: CONCEPTUAL DESIGN OF THE VR EXERGAME

Based on the findings of the user requirements analysis, we designed a VR exergame: *Memory Journalist VR*. In the early stage of this research, the implementation of the exergame was presented in [63, 64]. We note that the immersive exergame was designed in the HCD approach as a result of the iterative design process; we incorporated our focus group findings into the game design. The improvements and changes applied to the VR exergame prototype through five focus group sessions can be seen in Figure 5. Below, we report the final conceptual design of the VR exergame.

Prior work showed that VEs have the potential to create reminiscence effects on people with dementia [83]. We were inspired by this finding and hoped to invoke the past memories of people with dementia while they were playing the VR exergame: providing both cognitive and physical stimulation. Therefore, we created *Memory Journalist VR* in which the player explores a 360° 3D recording of famous landmarks and their surroundings in a VE. The 3D recordings presented

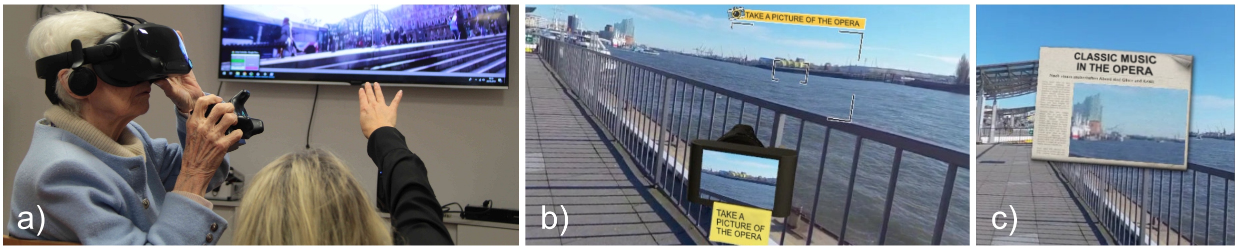

Fig. 3. The player explores the game environment with a 3D printed camera (a), communicates with a health professional to find a landmark in the scene (b), and takes a correct picture of the landmark which is shown in a newspaper (c).

the capital and city of residence of the older adults to provide a meaningful experience [34] (see Figure 3).

We aimed to provide an intuitive way for this user group to interact with the virtual world without them having to learn and repeat how to use new interaction devices. Thus, we chose to design a 3D printed camera device and mapped this physical version of the device to a virtual replica in the game.

We enabled the option of playing the exergame as a seated experience to reduce the risk of falling [65] and to reach more target users with varying cognitive-physical abilities. We created the recordings with the consideration of the eye level of participants in a sitting position. We also avoided creating scenarios that could conflict between the participants' visual and vestibular senses (e.g., by not including forward movement in the seated VR exergame) [54] or cause emotional stress (e.g., unrealistic experiences). The players mainly performed torso, arm, stretching, and head movements that could assist in some of their daily life tasks: grabbing objects, reaching something, and following someone with their head while holding conversations (see Figure 2d).

To ensure the safety and well-being of the participants, we gave control of the VR exergame to the health professional/caregivers by building a browser-based application. In addition, we shared the VR exergame experience of HMD players with the spectators (both older adults and health professionals) on the TV screen, providing them with the opportunity to communicate and interact with each other in a social environment (see Figure 2b).

5 STEP 3: IMPLEMENTATION OF THE VR EXERGAME

In this section, we describe the technical equipment used and the final implementation of the game. The changes applied to the VR exergame prototype through the HCD process can be seen in Figure 5.

5.1 Technical Setup

Memory Journalist VR was implemented using C# in Unity (v.2018.4.18f1) [84]. To ensure high visual fidelity, the game was run using an i9-9900 CPU and an Nvidia RTX 2080 Ti graphics card. The players were equipped with a HMD (HTC Vive Pro Eye [43] or Valve Index [88]: depending on the availability) and a custom 3D-printed camera to interact with *Memory Journalist VR*. We created the 3D device based on the size and visuals of a *Leica M4* camera [20], which was likely to be known by the target users. For this, we built a 3D model in Maya [3] and combined it with a universal Vive tracker and a button (see Figure 2c) to give a realistic feeling of an actual camera.

Furthermore, we used an *Insta360 Pro* device [44] together with a tripod at 120 cm height to create 360° 3D scenes for *Memory Journalist VR*. Moreover, an application, *Remote App*, was implemented on top of an express (v4) [41] web server running on NodeJS (v8.9.4) [15] (see Figure 4b). We chose

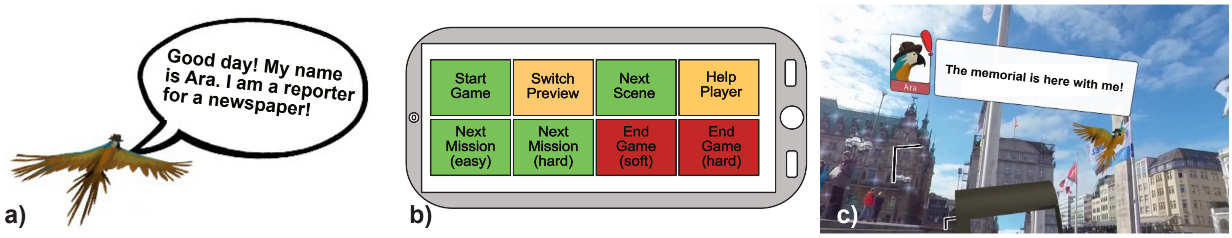

Fig. 4. The player meets with the virtual guide (parrot) at the beginning of the game (a). The health professional uses the *Remote App* which is shown on a mobile phone (b) and sends a hint using the *Remote App* (c).

to employ a safety-first approach: therefore, instead of using a tripod, we fixed the VR system's lighthouses to the wall and concealed all unnecessary cables.

5.2 *Memory Journalist VR*

The player has the role of a reporter, who takes photographs of landmarks (see Figure 3). The game provides an audio-visual interactive introduction of the player role and tasks considering varying cognitive-physical abilities of the target users (e.g., reduced hearing and seeing abilities). Afterwards, participants are introduced to the 3D-printed camera controller and are given time to get accustomed.

At the same time, the *Remote App* shows eight buttons that allow the health professional to control the VR exergame experience: Start Scene, Switch Preview, Next Scene, Help Player, Next Mission (easy), and Next Mission (hard) and two red buttons (soft or hard) to terminate the game due to urgency (see Figure 4b). Next Mission (easy) focuses on popular landmarks that can be detected easily by performing head movements, while Next Mission (hard) additionally requires zooming with the camera by performing arm movements (i.e., stretching out and retracting), or the rotation of the camera to get an upright photo. This enables to control the difficulty level of the VR exergame. For example, when players are given the task of photographing a landmark on their left side, they perform arm, head, torso, and stretching movements to complete the task. While we note that this game is not a typical exertion heavy exercise game, we focused on physical activities that people with dementia can perform [53] and tried to tailor the technology to their current interaction abilities.

After the player is introduced to the game mechanics, the health professional can activate the missions. Accordingly, the game guides the player by using a parrot as a virtual guide explaining the missions via a short story (see Figure 4a). Once the explanation is finished, the player is presented with a yellow sticky note attached to the camera and a text above the view frame of the camera to remind them of the current mission (see Figure 3b). At the same time, the target locations are highlighted by using sparkles to make people with dementia pay more attention to the given task. Additionally, if the player cannot find a landmark or does not know it, the health professional can send a hint with the parrot using the help button in the *Remote App* (see Figure 4b). When the camera is triggered while pointing at the requested landmark, a newspaper article with the image appears in front of the player to provide feedback on the accomplished mission (see Figure 4c). However, if the player takes photographs of other areas, the taken pictures will appear without a newspaper in their view to provide constructive feedback. Further, the health professional can offer bonus scenes to give the players the possibility of exploration without asking for the completion of any task. At the end of the VR exergame play, the players and health professionals can review the photographs taken in each session on the *Remote App* to share the feeling of success together.

I	The research team provides verbal instructions on what to photograph in what format (i.e. landscape, portrait).
II	In this version, we created a virtual guide that explains tasks (parrot), texts which show all objectives, smooth virtual camera movement, and a floating frame which shows the field of view of the camera.
III	The player is introduced to the *reporter narrative*. Also, the missions are shown as sticky-notes with large text always in view and are given individually by following after each other. The hints are represented 30 and 60 seconds after the missions start.
IV	The *Remote App* is introduced to control mission succession and hints. The missions can contain zooming tasks. The virtual camera turns transparent when it is close to the player's face to encourage players to stop searching for a location through the lens. We further added bonus scenes that show more of the world without adding tasks.
V	In the game, the objectives are highlighted with sparkles to foster more attention. We randomized the order of the missions. The zooming tasks are adapted to the players' arm reach. The *RemoteApp* shows the fictive newspaper after the game finishes.

Fig. 5. The design and implementation changes/improvements applied in *Memory Journalist VR* through five focus group sessions.

6 STEP 4: EVALUATION OF THE VR EXERGAME WITH FOCUS GROUPS

We created an iterative design circle for *Memory Journalist VR*. This demonstrates a robust design process to present the validity of the VR exergame. Here, we aimed to better observe and address the needs of people with dementia while they were playing the exergame. Therefore, we tested it with five focus group sessions to ensure the best player experience.

6.1 Participants

We conducted focus group sessions on five different dates, for which a total of 11 people with (n=10 mild dementia, n=1 moderate dementia; n=8 female, n=3 male, age: M=85.82, SD=9.02 years) and 6 without dementia (n=6 female, age: M=82.67, SD=4.46 years) were recruited. Some participants took part in multiple sessions. People with dementia and those without did not participate in the same focus group sessions (but some on the same dates). None of the participants in our focus groups reported any previous VR experience before our study. The demographic information of the players can be seen in Table 1.

Based on our findings, we involved a health professional (female, age: 51, 26 years of experience working with people with dementia) in all sessions (except the last one) as a part of the research team. The health professional helped by guiding the players through the sessions. In the last session, three caregivers (n=2 female, n=1 male, age: M=54.3, SD=4.73 years, experience working with people with dementia: M=12, SD=3 years) were observed in a gameplay with people with dementia.

6.2 Procedure

We chose to test the VR exergame with people with and without dementia (only in the early sessions) to enrich the data collection due to their shared values. Further, we note that it was not always possible to recruit large numbers of participants with dementia, due to their established daily routines (e.g., exercise activities, and doctor appointments). The given feedback from each session was integrated into *Memory Journalist VR*. We performed focus groups separately (people with and without dementia) to ensure group harmony (based on the findings of the user requirements analysis). Each focus group took place approximately one month after each other. During the sessions, we made observations and took field notes.

Focus Group Sessions	Dementia Status With/Without	Gender Female	Gender Male	Age Mean	Age SD	Duration of Stay Mean	Duration of Stay SD	Number of Players Participated Before
Focus Group Session I	with dementia	3	1	82	10.23	1.82	1.25	0
	without dementia	6	0	82.67	4.46	3.72	3.06	0
Focus Group Session II	with dementia	1	1	74.5	9.19	1.83	2.14	2
	without dementia	2	0	78.5	4.95	5.28	4.59	2
Focus Group Session III	with dementia	1	1	91.5	0.71	1.56	0.50	0
	without dementia	2	0	78.5	4.95	5.19	4.60	2
Focus Group Session IV	with dementia	4	1	86.2	9.36	0.97	0.89	1
Focus Group Session V	with dementia	3	0	77.3	13.65	1.58	1.83	2

Table 1. The demographic information of the players across five focus groups: status of dementia, gender, age, and duration of stay at the senior living facility in years. The last column depicts the amount of people that participated in prior focus groups.

In *Focus Groups I-II-III-IV*, we started the sessions by creating a relaxed atmosphere and inviting the players to a table with tea and cake. First, the players were introduced to the HMD and game concept. They and their legal representatives were asked to sign an informed consent form which included information on the planned project, what the player will experience in VR, and the symptoms of cybersickness. Also, all participants were verbally briefed about their upcoming experience and warned of possible negative symptoms of VR. There was no fixed order in which the participants played. While one player was playing the exergame, the health professional was at their side and working with them to complete the objectives. One researcher was taking notes of their activity, while the second researcher was taking notes and talking with the remaining participants at the table. The last researcher was responsible for the game experience and instructions The spectators could follow the gameplay on a large TV screen behind the HMD player. When a participant removed the HMD, they were able to talk about their experience in a brief interview. The questions were about the overall gaming experience, enjoyment, scenes in VR, and causes of discomfort. The overall duration of the study was approximately 30 minutes for each participant, 20 minutes in VR.

In *Focus Group V*, we aimed to finalize the evaluation of the exergame with the additional involvement of different caregivers to see the generalizability of the design process and the exergame. We wanted to observe the usability of the VR exergames for the case of long-term use without the help of the research team. We observed the interaction between caregivers and people with dementia while they played the VR exergame together, led for the first time by the caregivers. We presented an extensive instruction sheet to the caregivers to explain the *Remote App*, the game mechanics, and the usage of the HMD[3]. For each pair, we prepared the caregiver by reading out the instructions and answering open questions. We also gave them a printed copy to assist the caregivers during the session. Afterwards, the caregivers were instructed to start and lead a gaming session with the people with dementia up to completion. The pairs played the VR exergame one after the other without the other pairs' presence. During this process, we stayed in the room to observe the interactions of people with dementia and caregivers. The overall duration of the study was approximately 60 minutes for each participant, 20-30 minutes in VR.

6.3 Data Analysis & Results

To analyze qualitative data, we gathered field and observation notes taken by the members of the research team. Here again, the first author employed the reflexive orientation of inductive thematic

[3]The instruction sheet explained the following points: mounting the HMD, devices in the room, the reason for using the VR exergame, game flow, *Remote App*, difficulty level between the missions, and exit strategies.

analysis [9–11] by following the same steps as stated in subsection 3.3. We derived four themes through inductive coding. We report quotes based on the participants' group and number: focus group players with (F_{PWD}) and without dementia (F_{PW}), the health professional (F_{HP}), and focus group caregivers (F_C). For example, $F5_{C_1}$ represents the quotes from the first caregiver in the fifth session.

*Theme 1: **Player experience in*** **Memory Journalist VR**. This theme captures the players' experience of people with and without dementia using the designed VR exergame. **For people without dementia**, while VR and the game initially showed novelty effects, in the long-run, the game was found to be repetitive: "*I liked the recognition value in the scenes, but it gets boring very quickly*"-$F2_{PW_1}$. Particularly, in the third focus group, one of the older adults without dementia performed the game tasks without waiting for the explanation of the parrot because "*it [the exergame] is boring*"-$F3_{PW_1}$. This trend was also observed among the other spectators. Hence, $F3_{PW_2}$ preferred to not play *Memory Journalist VR* after watching the HMD player; leading us to conclude the focus groups with people without dementia.

For people with dementia, *Memory Journalist VR* provided a positive player experience (e.g., "*Can't find words right now. It was beautiful*"-$F4_{PWD_3}$ and "*I enjoy it, all the time*"-$F4_{PWD_2}$). Most players with dementia perceived *Memory Journalist VR* as an immersive and realistic game play. For example, people with dementia made attempts to talk with pedestrians and to touch objects in the scene: "*I felt right in the middle of it*"-$F5_{PWD_3}$. Particularly, for one player with dementia, $F4_{PWD_5}$, we observed the plausibility effect of VR; he tried touch the soap bubbles in the VE with his tongue. These positive experiences were mainly attributed to the opportunity to explore an outdoor environment in VR. These opportunities were often difficult to reach for people with dementia: "*I found it nice to discover so much, especially the buildings [...] I used to go for walks there. Since I have been in care here, that does not work so well anymore*"-$F4_{PWD_3}$. The players took their time to discover the buildings and the objects in the scene, which is further supported with adding bonus scenes to our exergame (e.g., "*I like both, special tasks [missions] and the possibility to discover for myself*"-$F4_{PWD_4}$). However, in some cases, the feeling of autonomy through exploration led the players to ignore the virtual guide (the parrot) or be angry at it for interrupting the game by explaining tasks: "*Ugly bird, I don't like him [...]. I don't want to hear that [the hint]! I don't need a helper*"-$F3_{PWD_4}$. This led us to switch hints from the automatic time-based version to the need-based caregiver controlled version, which was found useful for both, the players and caregivers (e.g., "*The parrot was good for help*"-$F5_{C_1}$).

*Theme 2: **Emerging social dynamics**.* This theme focuses on instances of social interactions created by *Memory Journalist VR*. For the majority of people with dementia, the exergame elicited a variety of social interactions between players and caregivers/the health professional. The effects of the game were observed both during and after the focus groups.

During gameplays, communications were raised due to familiarity with VEs and the desire to complete *Memory Journalist VR* successfully. Since the game includes VEs that the participants often recognized from their real lives, we observed situations where both the players and the caregivers benefit from each other's knowledge (e.g., "*girl [the health professional], help me out!*"-$F4_{PWD_2}$). While our game design holds a challenge in case that both, the caregiver and the player, do not know the landmarks, the *Remote App* helped both sides by providing hints through the parrot: "*Do you see the arcades? Let's take a picture of them!*"-$F5_{C_2}$. The players shared –when they remembered– their memories of the places they visited in *Memory Journalist VR* with the research team, the spectators, and the caregivers: "*I hate that building, way too expensive!*"-$F4_{PWD_2}$. Particularly, this enabled long conversations between parties.

Furthermore, the need for a sense of accomplishment was observed. **After their sessions**, most of the players asked the research team and caregivers how well they completed their tasks. They shared their success in the application by viewing the photographs they took together: "*the shots are quite good*"-F5$_{C_1}$. The players were instantly happy when they received positive feedback. After the game sessions, some players shared their experiences about *Memory Journalist VR* with their relatives, which can be seen as an indicator of the game that enables communication: "*The player proudly told her daughter all the things she had experienced today, that she had been out in town*"-F5$_{C_1}$.

Theme 3: ***Impact of cognitive-physical abilities on the VR exergame.*** This theme addresses how the VR exergame and the cognitive-physical abilities of the participants interact with each other. **For people without dementia**, the exergame's tasks were not found to be cognitively challenging; all the players completed their missions and used the 3D camera without requiring explanation. Physically, they all preferred to play the game while standing.

For people with dementia, every focus group session showed signs of further improvement in the game mechanics and design. We encountered two **cognitive factors**, which required special attention: forgetfulness and problem-solving. The players' level of forgetfulness varied. Some players could not remember the game in the end (e.g., "*I took photos?*"-F3$_{PWD_1}$), while some recalled the missions and features of *Memory Journalist VR* after almost one month (e.g., "*Here I am again, ah parrot!*"-F4$_{PWD_2}$. To serve these varying levels of forgetfulness, we created (i) sticky notes fixed to the camera as a reminder of current tasks (a familiar concept from daily life), (ii) 3D sparkles to highlight the tasks, and (iii) random sequencing of tasks to avoid repetition.

The players also presented differing degrees in problem-solving skills, with the majority showing a deficiency. For example, when some received a mission, they could not comprehend the instructions to find the target (e.g., turning left and right to follow the parrot). Another notable example was regarding the use of the 3D camera. While we only expected that players would understand the placement and the use of the button, the affordance of this old timely device prompted the players to hold the device right to their face, expecting a viewfinder, which was not present. As a result, players could not see anything and we initially had to ask them to move the camera away from their eyes. This shows that people with dementia expected all the usual means of interacting with the device, not only some of the provided affordances with the camera model. To facilitate the problem-solving process, we introduced (i) the reporting theme gives the player a purpose and to present the missions in an easily understandable context, (ii) a transparent 3D camera when the camera is close to the players' head, and (iii) a *Remote App* to show hints depending on the needs of players.

In particular, the *Remote App* helped the caregivers to easily guide the players without any interruption from outside. This allowed caregivers to control the game flow dynamically to match the needs of people with dementia (e.g., frequency of sending help and easy or hard missions). This also helped the caregivers to enjoy *Memory Journalist VR* because it was "*a good mix of conversation and guided play in my opinion*"-F5$_{C_2}$. As a result, we observed that *Memory Journalist VR* was helpful to cognitively stimulate long-term memories of people with dementia: "*it was nice by the water, in the cafe down there*"-F5$_{PWD_2}$.

Regarding **physical abilities**, the players with dementia always chose to play the exergame in a sitting position. In the beginning, we held our sessions with a fixed chair. However, dementia showed some impact on the physical abilities of the players. While most basic arm movements were performed by the players, for torso movements, the research team members had to physically assist players to turn left and right. Hence, we decided to use a fixable swivel chair. The use of a swivel chair enabled even people with low physical abilities to explore the 360° VEs. The zoom feature was explicitly wished from a player who used to be a professional photographer: "*I want* *to get closer!*"-F3$_{PWD_2}$. The creation of a level of physical difficulty in this intuitive task was well conveyed to the other participants. All players intuitively performed stretching out and retracting movements without a need for instructions. Especially with an adaptation to their individual arm reach, the players became competent and motivated to perform the zooming tasks: "*she makes a physical effort for zoom*"-F5$_{C_1}$.

Theme 4: ***Encountered barriers for the use of VR and the VR exergame.*** This theme outlines the potential challenges observed in the use of VR and *Memory Journalist VR* for people with dementia and their caregivers.

Regarding technical aspects, some caregivers reported concerns about the HMD's weight, but we encountered only one occasion where the HMD caused neck pain for the player (F4$_{PWD_2}$), leading us to terminate the experience immediately. Despite using a fixed swivel chair, we did not observe any safety issues other than stretching the HMD cable, which was controlled by the research team. **For caregivers of people with dementia**, the usability of the system was "*overwhelming at the beginning but later quite interesting*"-F5$_{C_2}$. The reasons were explained as the required time to become familiar with the VR technology (e.g., placing the HMD, sharpening the view of the players in VR) and understanding how the players interact with *Memory Journalist VR* (e.g., sending the parrot for help). Nevertheless, all caregivers were able to complete the exergame without requiring any help from the research team.

Although *Memory Journalist VR* was found to be an immersive experience, this sometimes led to isolation from the real world. A player accidentally touched the foot of the caregiver who was around the player since the beginning of the experience, but the player got irritated ("*who's on my foot*"-F4$_{PWD_2}$). Using the 3D camera as an intuitive way to interact with *Memory Journalist VR* was well received by our players. People with dementia did not require an explanation on the purpose of the device or how the 3D camera would be used to photograph landmarks. However, in a few cases, some players had difficulties finding the camera button to take pictures, creating a feeling of insecurity for them.

7 OVERALL DISCUSSION

We first discuss our findings by focusing on our research questions. Then, we refine our preliminary implications described in subsection 3.4 based on the emerged nuances in the focus group sessions.

7.1 RQ_1: Does the designed VR exergame provide positive and good player experiences, usability, and accessibility for people with dementia?

Overall, the findings show that it is feasible to develop VR exergames for people with dementia and integrate them into their daily life and routines. The VR exergames can provide positive and good player experiences, usability, and accessibility when game mechanics, aspects, and concepts are specifically targeted to the needs and abilities of this user group. This is in line with studies that emphasize the importance of understanding the abilities of this user group and how they interact with the designed systems, rather than seeing dementia as a barrier to interacting with technology [23, 53, 91].

In terms of **player experience**, our results highlight how changes in cognitive-physical abilities of older adults can affect the needs and requirements of VR exergame design. First of all, *Memory Journalist VR* was no interesting experience for adults without dementia, whereas it was highly enjoyable for people with dementia even after multiple sessions. We attribute this mainly to the exergame not offering a good **game flow** for older adults without dementia [14]; the game tasks simply were too easy and their skill level was too high. Inversely, the exergame provided a good game flow for people with dementia by offering them a concrete clear goal (i.e., photographing), which matches their cognitive-physical abilities (e.g., arm, head, and torso movements) [4].

We further observed **novelty effects** of VR. This positively affected the experience of both player groups. Although the exergame became repetitive for the players without dementia, interestingly the novelty effects did not disappear for the majority of people with dementia; they were always excited to play. VR game designers should consider the power and implications of this effect and explore this while creating novel VR experiences/games for people with dementia.

The **enjoyment** of people with dementia was associated with the possibilities of exploration of outdoor environments. Our results reflect on Tabbaa et al. [83]'s work on the use of outdoor VEs for people with dementia living in psychiatric hospitals. Further, we build on their work by creating a 360° VEs-based exergame and supporting this with a meaningful experience (e.g., focusing on reminiscence) [34]. This exploration concept of the game demonstrated that this user group still holds the need for autonomy [70] and when violated this can lead to frustration. Therefore, game designers should ensure that the player feels autonomy to avoid a potential conflict of people with dementia with the designed games.

While *Memory Journalist VR* did not explicitly require any communication between both parties, the game elicited a variety of very well received **social interactions** between people with dementia and their caregivers/the health professional; they helped each other and shared their long-term memories, relating to people with dementia's need for relatedness [70]. We attribute these to our design choice to integrate the shared biographies of our end-users and even caregivers into a VR exergame. Despite the risk that players and caregivers might not know these landmarks, this became an opportunity for further discussion and exploration, as it offers common ground for both sides. Thus, it might be in the interest of game researchers to explore social potentials of VR exergames (similar to exploration of social potentials in video-based exergames [85, 86]).

People with dementia showed several signs that *Memory Journalist VR* was a **realistic and immersive** gameplay due to the high visual fidelity achieved through the 3D 360° videos. The players showed both place and plausibility illusion effects of VR [75] by trying to interact with people, and touch and taste objects. Interestingly, on some occasions, this user group was immersed in VR so much that they lost their connection to the physical world (in contrast to typical users [75]). While it may be considered a good aspect of the game, VR researchers should keep in mind that illusions of place and plausibility can lead to potential ethical concerns and dangers in the experience of people with dementia. For example, places should not show the past (e.g. their hometown in their 20s) but reflect the present, and any depiction of the individuals using the device should reflect reality. If a mirror is used in VR, it should show the user exactly as they are, not younger or in different attire since this could disrupt the sense of reality for the people with dementia. In particular, seeing a false reality of something they may have experienced before can have negative psychological consequences (e.g., anxiety), thus it can be unethical.

With respect to **usability** and **accessibility**, wearing an HMD was not physically demanding for people with dementia except for one player in the short-term exposure. VR controllers were simply too complex to use; the layout of the buttons on the controllers and the instructions on how to use them were cognitively too demanding. We argue that it is difficult to teach new game interaction techniques in VR exergames to people with dementia due to their decline in cognitive abilities. Our results indicate that with the use of the 3D-printed camera, the participants were able to interact in the game without requiring high cognitive demand yet they expected all affordances of the camera (i.e., viewfinder). We recommend using or creating interaction devices that they are familiar with from real-life use. Also, game designers should consider providing all the affordances of the device or offer solutions to overcome familiarities that people with dementia may not find on the device (e.g., by making the camera transparent).

We had to use a fixed swivel chair to increase the accessibility of the exergame. This allowed the game to be explored with a wider user group, whereas reducing the intensity and effectiveness of body movements. The accessibility of VR games is also related to the usability of the system by **caregivers**. In this regard, our results are the first step (to our knowledge) towards understanding the possibility of caregivers using VR exergames for people with dementia in hospitals without any help from the game development team. The caregivers' experience highlights that novelty effects [50] of VR may occur due to the first use of the new technology, which can be overwhelming. Nevertheless, the results indicate that caregivers can use the VR exergame in their daily routine without support if given proper instructions.

7.2 RQ_2: Which aspects of the HCD approach should be adapted in the context of VR exergames for people with dementia?

Our findings demonstrate that standard HCD approaches [24] should be adapted to the specific situation of developing novel innovative technology for people with neurological diseases, which limit their cognitive as well as physical abilities.

Classical HCD approaches emphasize the benefits of including end-users and other stakeholders [24, 37]. However, for developing VR exergames for people with dementia in a clinical setting it is essential to include a variety of stakeholders. First of all, while the players must be included in the HCD approach, they could be limited in expressing their thoughts, needs, and feelings [49]. Our results support this assumption by showing that they cannot always explain the rationale behind their reactions. For example, some players did not like the parrot without being able to specify the reason for it. Thus, making it crucial to include stakeholders, who can add comments, and translate thoughts and feelings of people with dementia.

Considering our experience throughout the design process, we as game designers/developers realized how little we knew about the effects of dementia, and the daily life and routines of people affected by it. In fact, based on the stakeholders' involvement with **semi-structured interviews**, we had to change some of the traditional tools for the game design; for example, we always provided positive feedback and did not implement unpredictability and challenges involving cognitive-physical mechanics [8]. Thus, we suggest including stakeholders such as significant others, clinical dementia experts with different specialties, physiotherapists, and technicians to gain a better understanding of this specific user group.

Furthermore, people with dementia in clinical settings pose significant challenges for the successful development of VR exergames. Especially, **conducting contextual inquiries** is of major importance because they as well as the involved stakeholders typically follow strict routines to organize their daily work and life. For example, we could not complete the planned gameplays in the contextual inquiry due to time constraints of people with dementia and caregivers in daily work. However, a successful setup into the daily life of people with dementia requires (i) caregivers, who play exergames together with this user group, (ii) technicians that install and maintain the VR technology, and (iii) physicians who organize the physical training program with consideration of the amount and intensity of the VR exergames. Therefore, to identify the potential constraints (e.g., daily routine, ethical, and legal issues), we recommend game designers conduct contextual inquiries.

Our focus group findings show the importance of involving a caregiver in VR exergames to avoid potential safety issues. For example, reacting in a timely manner when players are faced with the psychological consequences of VR games: feeling insecure due to the inability to perform tasks or loss of connection with the physical world. Similarly, when players encounter problems that could threaten their physical health (e.g., cable problems due to the use of a swivel chair). Yet not only for those but also for positive effects: to give a feeling of competence to the players by supporting them with positive feedback [70] and help to transfer their thoughts to young game developers who might have a lack of empathy to understand older adults easily [19, 25]. Therefore, stakeholders must be involved in the user requirements analysis step. However, even after the initial analysis, especially formal or informal caregivers must also be incorporated in the remaining iteration steps. We argue that the inclusion of stakeholders and contextual inquiries are not optional aspects of the HCD, but should be mandatory for the development of VR exergames for people with dementia.

7.3 Revisiting the Initial Implications for Design

Here, we visit our initial implications and discuss them by lensing through the results of focus groups.

Create social gaming environments with the consideration of shared aspects. The findings partially support this initial implication. The VR exergame can create a social environment for people with dementia with their caregivers/the health professional, but we did not observe any increased social interaction among older adults with dementia. We speculate that this might be due to the feeling of insecurity to initiate conversations and the loss of interest in socially interacting with peers [74]. However, future research should explore whether and how all people with dementia can be integrated with each other in such a play environment. For example, positive results are increasingly reported for creating multiplayer games and employing asymmetries to promote communication between players [35, 46]. While creating interdependent asymmetric games can be considered promising (among players with dementia or between these users and caregivers), this research may also face some challenges in providing positive experiences for all users due to differing abilities.

Support of an inverse game flow channel addressing decline and variance in cognitive-physical abilities. Contrary to traditional games, one cannot always expect an improvement or learning effects during the phase of gameplay, when people with dementia are involved, especially in such short-term games. In fact, we observed reoccurring novelty effects during the process. Also, researchers should not forget the progressive nature of dementia, which over time further affects abilities [93]. These can impose some changes to known elements of the game design (e.g., challenge and introducing new concepts with the progress), which are often based on game flow concepts in which the difficulty of the game increases over time when users improve their skills [4, 14]. We recommend that the cognitive-physical difficulty of VR exergames (like other games involving people with dementia) should be adaptable to the abilities of this user group, which declines over time so that the games are still playable and enjoyable for them. The findings emphasize that we supported the inverse game flow in our exergame by introducing the *Remote App*; caregivers were able to control the game's tasks at certain difficulty levels to tailor the exergame to the needs of players. For short-term use, this helped caregivers to address the decline and variance in the cognitive-physical abilities of the users. However, this leaves a research gap for future studies to explore dynamic inverse game flow adaptivity in VR exergames for people with dementia, in particular for long-term use.

Provision of a safe VR exergame experience. Our study strongly emphasizes the importance of safety in VR exergames. While we always preferred the safety-first approach (e.g., using wall-mounted systems and involving a health professional in the entire design process), we also faced some difficulties related mainly to the technical setup to ensure the safety of the participants. For example, we used a fixed swivel chair and this did not cause serious problems, but we speculate that it could happen. Yet, following the concerns of our caregivers about the weight of HMDs, we chose not to use wireless HMD systems for people with dementia. However, we also observed a player having neck pain due to the weight of the VR-HMD. This was also shown as one of the discomfort factors of using VR technology in a recent study [39]. We recommend that experimenters and relevant healthcare professionals observe players during the experience and immediately terminate the experience if they encounter or suspect any problems (e.g., physical and psychological). Future VR research should test the possible use of wireless HMD systems and commercial companies should consider providing accessible VR hardware solutions for this user group.

7.4 Limitations

We performed our study with a small number of participants (*n*=11). While our results serve as the first attempt to understand the needs of people with dementia in VR exergames, future work should test, verify, and demonstrate the long-term results of this using a higher sample size. However, one should be aware that the possibilities of reaching such user groups with specific needs (e.g., due to legal, ethical, and social issues) is limited in overall accessibility research.

We were only able to reach participants in the same senior living facility. Also, these participants had previous experience with Kinect games. We speculate that these could bias the findings in two ways: (i) enabling more social interaction among players and (ii) having positive attitudes towards the use of emerging technology.

While our focus was on the process and design of VR exergames for people with dementia, we also note that we did not compare this designed game across different technologies. Future work should consider evaluating and comparing the applicability of different technologies with this user group (e.g., augmented reality).

In our research, we did not use surveys to evaluate player experiences and usability. This may be considered as a limitation, however, prior work [31] suggests that using common testing approaches might not be suitable for people with dementia. Also, there is currently no validation of the use of these questionnaires (e.g., Player Experience Inventory [1], Nasa-TLX [36], System Usability Scale [12]) for this user group. Finally, we note on the reflexivity [5, 57] of qualitative findings; the main researcher has a background in psychology, cognitive systems, and VR gaming experience, which might introduce bias into interpretation of the results.

8 CONCLUSION

This work explored the feasibility of an immersive exergame for people with dementia following an HCD approach consisting of four steps. First, we conducted (i) semi-structured interviews with several stakeholders and contextual inquiries with people with dementia. Based on the analysis of user requirements, we (ii) conceptually designed, (iii) implemented *Memory Journalist VR*, and (iv) evaluated this exergame with five focus group sessions involving our research team members (i.e., a motion scientist, game designers/developers, and a health professional) through the processes.

The results indicate that VR exergames can provide positive and good player experiences (e.g., enjoyment, presence, and social interaction), usability, and accessibility for this user group, especially when the interests and abilities of people with dementia are considered. The lessons learned also strongly emphasize the necessity to adapt the standard HCD approach to design VR exergames for this specific user group by including the stakeholders and contextual inquires. Finally, we provide design implications for future work: (i) creating social gaming environments with the consideration of shared aspects, (ii) the support of an inverse game flow to address the decline and variance in cognitive-physical abilities, and (iii) offering a safe VR exergame experience. Thus, this work provides important insights for game designers by demonstrating a solid design process in the development of VR exergames for people with dementia.

ACKNOWLEDGMENTS

We thank our participants and Hospital zum Heiligen Geist Hamburg for making this research possible with their help, Daniel Apken for his work on the implementation, and the anonymous reviewers for their constructive feedback. This research was funded by the German Federal Ministry of Education and Research (BMBF), the German Research Foundation (DFG), the European Union's Horizon 2020 research and innovation program, and the Federal Ministry for Economic Affairs and Energy (BMWi). The authors also acknowledge the support of the Cluster of Excellence "Matters of Activity. Image Space Material" funded by the DFG under Germany's Excellence Strategy - EXC 2025 - 390648296.

REFERENCES

[1] Vero Vanden Abeele, Katta Spiel, Lennart E. Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370. https://doi.org/10.1016/j.ijhcs.2019.102370

[2] William C. Adams. 2015. *Conducting Semi-Structured Interviews.* John Wiley & Sons, Ltd, New Jersey, United States, USA, Chapter 19, 492–505. https://doi.org/10.1002/9781119171386.ch19

[3] Alias Systems Corporation. 1998. Maya. https://www.autodesk.com/products/maya/

[4] Sean Baron. 2012. *Cognitive flow: the psychology of great game design.* Gamasutra-The Art & Business of Making Games. https://www.gamasutra.com/view/feature/166972/cognitive_flow_the_psychology_of_.php

[5] Roni Berger. 2015. Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative research* 15, 2 (2015), 219–234. https://doi.org/10.1177/1468794112468475

[6] Hugh Beyer and Karen Holtzblatt. 1997. *Contextual Design: Defining Customer-Centered Systems.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[7] Christiaan G Blankevoort, Marieke JG Van Heuvelen, Froukje Boersma, Helga Luning, Jeltsje De Jong, and Erik JA Scherder. 2010. Review of effects of physical activity on strength, balance, mobility and ADL performance in elderly subjects with dementia. *Dementia and geriatric cognitive disorders* 30, 5 (2010), 392–402. https://doi.org/10.1159/000321357

[8] Brenda Brathwaite and Ian Schreiber. 2008. *Challenges for Game Designers.* Charles River Media, Inc., USA.

[9] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[10] Virginia Braun and Victoria Clarke. 2020. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21 (2020), 37–47. https://doi.org/10.1002/capr.12360

[11] Virginia Braun and Victoria Clarke. 2020. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology* 0, 0 (2020), 1–25. https://doi.org/10.1080/14780887.2020.1769238

[12] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[13] Kimberly McClure Cassie and Sara Sanders. 2008. Chapter 12; familial caregivers of older adults. *Journal of Gerontological Social Work* 50, S1 (2008), 293–320. https://doi.org/10.1080/01634370802137975

[14] Mihaly Csikszentmihalyi and Mihaly Csikzentmihaly. 1990. *Flow: The psychology of optimal experience.* Vol. 1990. Harper & Row, New York, USA.

[15] Ryan Dahl. 2009. NodeJS. https://nodejs.org/en/

[16] Renée FAG de Bruijn, Elisabeth MC Schrijvers, Karen A de Groot, Jacqueline CM Witteman, Albert Hofman, Oscar H Franco, Peter J Koudstaal, and Mohammad Arfan Ikram. 2013. The association between physical activity and dementia in an elderly population: the Rotterdam Study. *European journal of epidemiology* 28, 3 (2013), 277–283. https://doi.org/10.1007/s10654-013-9773-3

[17] Mahzar Eisapour, Shi Cao, Laura Domenicucci, and Jennifer Boger. 2018. Participatory Design of a Virtual Reality Exercise for People with Mild Cognitive Impairment. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18).* Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3170427.3174362

[18] Mahzar Eisapour, Shi Cao, Laura Domenicucci, and Jennifer Boger. 2018. Virtual reality exergames for people living with dementia based on exercise therapy best practices. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 62, 1 (2018), 528–532. https://doi.org/10.1177/1541931218621120

[19] Astrid Elsbernd, Sonja Lehmeyer, and Ulrike Schilling. 2014. *So leben ältere und pflegebedürftige Menschen in Deutschland - Lebenslagen und Technikentwicklung.* Jacobs Verlag, Detmold, Germany.

[20] Ernst Leitz GmbH. 1966. Leica M4 Camera. Ernst Leitz GmbH.

[21] Claudine Fabre, Karim Chamari, Patrick Mucci, J Masse-Biron, and C Prefaut. 2002. Improvement of cognitive function by mental and/or individualized aerobic training in healthy elderly subjects. *International journal of sports medicine* 23, 06 (2002), 415–421. https://doi.org/10.1055/s-2002-33735

[22] Howard H Feldman, Bart Van Baelen, Shane M Kavanagh, and Koen EL Torfs. 2005. Cognition, function, and caregiving time patterns in patients with mild-to-moderate Alzheimer disease: a 12-month analysis. *Alzheimer Disease & Associated Disorders* 19, 1 (2005), 29–36. https://doi.org/10.1097/01.wad.0000157065.43282.bc

[23] Sarah Foley, Daniel Welsh, Nadia Pantidi, Kellie Morrissey, Tom Nappey, and John McCarthy. 2019. *Printer Pals: Experience-Centered Design to Support Agency for People with Dementia.* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300634

[24] International Organization for Standardization. 2019. ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. https://www.iso.org/standard/77520.html International Organization for Standardization. Geneva, Switzerland.

[25] Wolfgang Friesdorf and Achim Heine. 2006. *sentha-seniorengerechte Technik im häuslichen Alltag: Ein Forschungsbericht mit integriertem Roman.* Springer Science & Business Media, Heidelberg, Germany.

[26] Luciano Gamberini, M Fabregat, A Spagnolli, L Prontu, B Seraglia, M Alcaniz, A Zimmerman, Tapani Rontti, J Grant, R Jensen, et al. 2008. Eldergames: videogames for empowering, training and monitoring elderly cognitive capabilities. *Gerontechnology* 7, 2 (2008), 111.

[27] Beat Games. 2019. *Beat Saber.* Game [VR]. Beat Games. Los Angeles, CA, USA. https://beatsaber.com/.

[28] Rebeca I Garcia-Betances, Viveca Jiménez-Mixco, Maria T Arredondo, and María F Cabrera-Umpiérrez. 2015. Using virtual reality for cognitive training of the elderly. *American Journal of Alzheimer's Disease & Other Dementias®* 30, 1 (2015), 49–54. https://doi.org/10.1177/1533317514545866

[29] Kathrin Gerling, Ian Livingston, Lennart E. Nacke, and Regan Mandryk. 2012. *Full-Body Motion-Based Game Interaction for Older Adults.* Association for Computing Machinery, New York, NY, USA, 1873–1882. https://doi.org/10.1145/2207676.2208324

[30] Kathrin Maria Gerling, Jonas Schild, and Maic Masuch. 2010. Exergame Design for Elderly Users: The Case Study of SilverBalance. In *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology* (Taipei, Taiwan) *(ACE '10).* Association for Computing Machinery, New York, NY, USA, 66–69. https://doi.org/10.1145/1971630.1971650

[31] Aideen Gibson, Claire McCauley, Maurice Mulvenna, Assumpta Ryan, Liz Laird, Kevin Curran, Brendan Bunting, Finola Ferry, and Raymond Bond. 2016. Assessing Usability Testing for People Living with Dementia. In *Proceedings of the 4th Workshop on ICTs for Improving Patients Rehabilitation Research Techniques* (Lisbon, Portugal) *(REHAB '16).* Association for Computing Machinery, New York, NY, USA, 25–31. https://doi.org/10.1145/3051488.3051492

[32] Google. 2016. *Tilt Brush.* Game [VR]. https://www.tiltbrush.com/

[33] Gary Gowans, Jim Campbell, Norm Alm, Richard Dye, Arlene Astell, and Maggie Ellis. 2004. Designing a Multimedia Conversation Aid for Reminiscence Therapy in Dementia Care Environments. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria) *(CHI EA '04).* Association for Computing Machinery, New York, NY, USA, 825–836. https://doi.org/10.1145/985921.985943

[34] Barbara J. Harmer and Martin Orrell. 2008. What is meaningful activity for people with dementia living in care homes? A comparison of the views of older people with dementia, staff and family carers. *Aging & Mental Health* 12, 5 (2008), 548–558. https://doi.org/10.1080/13607860802343019

[35] John Harris, Mark Hancock, and Stacey D. Scott. 2016. Leveraging Asymmetries in Multiplayer Games: Investigating Design Elements of Interdependent Play. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16).* Association for Computing Machinery, New York, NY, USA, 350–361. https://doi.org/10.1145/2967934.2968113

[36] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, Amsterdam, Holland, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[37] Richard Harte, Liam Glynn, Alejandro Rodríguez-Molinero, Paul MA Baker, Thomas Scharf, Leo R Quinlan, and Gearóid ÓLaighin. 2017. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. *JMIR human factors* 4, 1 (2017), e8. https://doi.org/10.2196/humanfactors.5443

[38] Patricia Heyn, Beatriz C Abreu, and Kenneth J Ottenbacher. 2004. The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. *Archives of physical medicine and rehabilitation* 85, 10 (2004), 1694–1704. https://doi.org/10.1016/j.apmr.2004.03.019

[39] Teresa Hirzle, Maurice Cordts, Enrico Rukzio, Jan Gugenheimer, and Andreas Bulling. 2021. A Critical Assessment of the Use of SSQ as a Measure of General Discomfort in VR Head-Mounted Displays. In *Proceedings of the 2021 CHI*

Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, Article 530, 14 pages. https://doi.org/10.1145/3411764.3445361

[40] James Hodge, Madeline Balaam, Sandra Hastings, and Kellie Morrissey. 2018. Exploring the Design of Tailored Virtual Reality Experiences for People with Dementia. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174088

[41] TJ Holowaychuk, StrongLoop, et al. 2010. NodeJS Express. http://expressjs.com/

[42] Maarten Houben, Rens Brankaert, Saskia Bakker, Gail Kenning, Inge Bongers, and Berry Eggen. 2020. The Role of Everyday Sounds in Advanced Dementia Care. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376577

[43] HTC and Valve Corporation. 2019. HTC Vive Pro Eye. https://www.vive.com/eu/product/vive-pro-eye/overview/

[44] Insta360. 2017. Insta360 Pro. https://www.insta360.com/product/insta360-pro

[45] Dennis L Kappen, Pejman Mirza-Babaei, and Lennart E. Nacke. 2019. Older adults' physical activity and exergames: a systematic review. *International Journal of Human–Computer Interaction* 35, 2 (2019), 140–167. https://doi.org/10.1080/10447318.2018.1441253

[46] Sukran Karaosmanoglu, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. 2021. Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 443, 15 pages. https://doi.org/10.1145/3411764.3445492

[47] EGA Esther Karssemeijer, JA Justine Aaronson, WJ Willem Bossers, T Tara Smits, RPC Roy Kessels, et al. 2017. Positive effects of combined cognitive and physical exercise training on cognitive function in older adults with mild cognitive impairment or dementia: A meta-analysis. *Ageing research reviews* 40 (2017), 75–83. https://doi.org/10.1016/j.arr.2017.09.003

[48] Basel Kikhia, Josef Hallberg, Johan E Bengtsson, Stefan Savenstedt, and Kare Synnes. 2010. Building digital life stories for memory support. *International journal of Computers in Healthcare* 1, 2 (2010), 161–176. https://doi.org/10.1504/IJCIH.2010.037460

[49] Jacqueline Kindell, Karen Sage, John Keady, and Ray Wilkinson. 2013. Adapting to conversation with semantic dementia: using enactment as a compensatory strategy in everyday social interaction. *International Journal of Language & Communication Disorders* 48, 5 (2013), 497–507. https://doi.org/10.1111/1460-6984.12023

[50] Chen-Lin C. Kulik and James A. Kulik. 1991. Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior* 7, 1 (1991), 75 – 94. https://doi.org/10.1016/0747-5632(91)90030-5

[51] Odders Lab. 2019. *OhShape*. Game [VR]. Odders Lab. Seville, Spain. https://ohshapevr.com/.

[52] Joseph J. LaViola. 2000. A Discussion of Cybersickness in Virtual Environments. *SIGCHI Bull.* 32, 1 (Jan. 2000), 47–56. https://doi.org/10.1145/333329.333344

[53] Amanda Lazar, Caroline Edasis, and Anne Marie Piper. 2017. A Critical Lens on Dementia and Design in HCI. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2175–2188. https://doi.org/10.1145/3025453.3025522

[54] Christopher H Lewis and Michael J Griffin. 1997. Human factors consideration in clinical applications of virtual reality. In *Virtual reality in neuro-psycho-physiology: Cognitive, clinical and methodological issues in assessment and rehabilitation*, Giuseppe Riva (Ed.). Vol. 44. IOS Press, Amsterdam, Netherlands, 35–56. https://doi.org/10.3233/978-1-60750-888-5-35

[55] Kellie Morrissey, John McCarthy, and Nadia Pantidi. 2017. *The Value of Experience-Centred Design Approaches in Dementia Research Contexts*. Association for Computing Machinery, New York, NY, USA, 1326–1338. https://doi.org/10.1145/3025453.3025527

[56] John Edison Muñoz, Afonso Gonçalves, Élvio Rúbio Gouveia, Mónica S. Cameirão, and Sergi Bermúdez i Badia. 2019. Lessons Learned from Gamifying Functional Fitness Training Through Human-Centered Design Methods in Older Adults. *Games for Health Journal* 8, 6 (2019), 387–406. https://doi.org/10.1089/g4h.2018.0028

[57] Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand, and Jon Howard Raphael. 2012. No room for reflexivity? Critical reflections following a systematic review of qualitative research. *Journal of Health Psychology* 17, 6 (2012), 866–885. https://doi.org/10.1177/1359105311427615

[58] Nintendo EAD Group. 2007. *Wii Fit*. Game [Wii]. https://www.nintendo.com/wiifit/launch/

[59] Gabriele Optale, Cosimo Urgesi, Valentina Busato, Silvia Marin, Lamberto Piron, Konstantinos Priftis, Luciano Gamberini, Salvatore Capodieci, and Adalberto Bordin. 2010. Controlling memory impairment in elderly adults using virtual reality memory training: a randomized controlled pilot study. *Neurorehabilitation and neural repair* 24, 4 (2010), 348–357. https://doi.org/10.1177/1545968309353328

[60] Wolf D Oswald, Thomas Gunzelmann, Roland Rupprecht, and Bernd Hagen. 2006. Differential effects of single versus combined cognitive and physical training with older adults: the SimA study in a 5-year perspective. *European journal of ageing* 3, 4 (2006), 179. https://doi.org/10.1007/s10433-006-0035-z

[61] Kalpana P Padala, Prasad R Padala, Shelly Y Lensing, Richard A Dennis, Melinda M Bopp, Paula K Roberson, and Dennis H Sullivan. 2017. Home-based exercise program improves balance and fear of falling in community-dwelling older adults with mild Alzheimer's disease: a pilot study. *Journal of Alzheimer's disease* 59, 2 (2017), 565–574. https://doi.org/10.3233/JAD-170120

[62] Sebastian Rings, Sukran Karaosmanoglu, Lucie Kruse, Daniel Apken, Tobias Picker, and Frank Steinicke. 2020. Using Exergames to Train Patients with Dementia to Accomplish Daily Routines. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Canada) *(CHI PLAY '20)*. Association for Computing Machinery, New York, NY, USA, 345–349. https://doi.org/10.1145/3383668.3419883

[63] Sebastian Rings, Frank Steinicke, Tobias Picker, and Caspar Prasuhn. 2020. Enabling Patients with Neurological Diseases to perform Motor-Cognitive Exergames under Clinical Supervision for Everyday Usage. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, New York, USA, 360–364. https://doi.org/10.1109/VRW50115.2020.00078

[64] Sebastian Rings, Frank Steinicke, Tobias Picker, and Caspar Prasuhn. 2020. Memory Journalist: Creating Virtual Reality Exergames for the Treatment of Older Adults with Dementia. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, New York, USA, 686–687. https://doi.org/10.1109/VRW50115.2020.00194

[65] Sebastian Rings, Frank Steinicke, Tobias Picker, and Caspar Prasuhn. 2020. Seated Immersive Exergaming for Fall Prevention of Older Adults. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, New York, USA, 289–290. https://doi.org/10.1109/VRW50115.2020.00063

[66] Giuseppe Riva. 1997. Virtual reality as assessment tool in psychology: Virtual reality in neuro-psycho-physiology. In *Virtual reality in neuro-psycho-physiology: Cognitive, clinical and methodological issues in assessment and rehabilitation*, Giuseppe Riva (Ed.). Vol. 44. IOS Press, Amsterdam, Netherlands, 71–79. https://doi.org/10.3233/978-1-60750-888-5-71

[67] Philippe Robert, Alexandra König, Hélene Amieva, Sandrine Andrieu, François Bremond, Roger Bullock, Mathieu Ceccaldi, Bruno Dubois, Serge Gauthier, Paul-Ariel Kenigsberg, et al. 2014. Recommendations for the use of Serious Games in people with Alzheimer's Disease, related disorders and frailty. *Frontiers in aging neuroscience* 6 (2014), 54. https://doi.org/10.3389/fnagi.2014.00054

[68] Yves Rolland, Fabien Pillard, Adrian Klapouszczak, Emma Reynish, David Thomas, Sandrine Andrieu, Daniel Rivière, and Bruno Vellas. 2007. Exercise program for nursing home residents with Alzheimer's disease: A 1-year randomized, controlled trial. *Journal of the American Geriatrics Society* 55, 2 (2007), 158–165. https://doi.org/10.1111/j.1532-5415.2007.01035.x

[69] Suvi Rovio, Ingemar Kåreholt, Eeva-Liisa Helkala, Matti Viitanen, Bengt Winblad, Jaakko Tuomilehto, Hilkka Soininen, Aulikki Nissinen, and Miia Kivipelto. 2005. Leisure-time physical activity at midlife and the risk of dementia and Alzheimer's disease. *The Lancet Neurology* 4, 11 (2005), 705–711. https://doi.org/10.1016/S1474-4422(05)70198-8

[70] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68. https://doi.org/10.1037/0003-066X.55.1.68

[71] Christine Sattler, Pablo Toro, Peter Schönknecht, and Johannes Schröder. 2012. Cognitive activity, education and socioeconomic status as preventive factors for mild cognitive impairment and Alzheimer's disease. *Psychiatry research* 196, 1 (2012), 90–95. https://doi.org/10.1016/j.psychres.2011.11.012

[72] Karen A Schutzer and B Sue Graves. 2004. Barriers and motivations to exercise in older adults. *Preventive medicine* 39, 5 (2004), 1056–1061. https://doi.org/10.1016/j.ypmed.2004.04.003

[73] Jeff Sinclair, Philip Hingston, and Martin Masek. 2007. Considerations for the Design of Exergames. In *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia* (Perth, Australia) *(GRAPHITE '07)*. Association for Computing Machinery, New York, NY, USA, 289–295. https://doi.org/10.1145/1321261.1321313

[74] David Singleton, Naaheed Mukadam, Gill Livingston, and Andrew Sommerlad. 2017. How people with dementia and carers understand and react to social functioning changes in mild dementia: a UK-based qualitative study. *BMJ Open* 7, 7 (2017). https://doi.org/10.1136/bmjopen-2017-016740

[75] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

[76] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of Presence in Virtual Environments. *Presence: Teleoper. Virtual Environ.* 3, 2 (Jan. 1994), 130–144. https://doi.org/10.1162/pres.1994.3.2.130

[77] Aimee Spector, Lene Thorgrimsen, BOB Woods, Lindsay Royan, Steve Davies, Margaret Butterworth, and Martin Orrell. 2003. Efficacy of an evidence-based cognitive stimulation therapy programme for people with dementia: randomised controlled trial. *The British Journal of Psychiatry* 183, 3 (2003), 248–254. https://doi.org/10.1192/bjp.183.3.248

[78] Frank Steinicke. 2016. *Being really virtual*. Springer, Berlin, Germany. https://doi.org/10.1007/978-3-319-43078-2

[79] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. 2010. Estimation of Detection Thresholds for Redirected Walking Techniques. *IEEE Transactions on Visualization and Computer Graphics* 16, 1 (2010), 17–27. https://doi.org/10.1109/TVCG.2009.62

[80] Cindy Stern and Zachary Munn. 2010. Cognitive leisure activities and their role in preventing dementia: a systematic review. *International Journal of Evidence-Based Healthcare* 8, 1 (2010), 2–17. https://doi.org/10.1111/j.1744-1609.2010.00150.x

[81] Yaakov Stern. 2012. Cognitive reserve in ageing and Alzheimer's disease. *The Lancet Neurology* 11, 11 (2012), 1006–1012. https://doi.org/10.1016/S1474-4422(12)70191-6

[82] Lightspeed Studios. 2016. *NVIDIA VR Funhouse.* Game [VR]. NVidia, California, CA, USA. https://store.steampowered.com/app/468700/NVIDIA_VR_Funhouse/.

[83] Luma Tabbaa, Chee Siang Ang, Vienna Rose, Panote Siriaraya, Inga Stewart, Keith G. Jenkins, and Maria Matsangidou. 2019. Bring the Outside In: Providing Accessible Experiences Through VR for People with Dementia in Locked Psychiatric Hospitals. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300466

[84] Unity Technologies. 2020. Unity Real-Time Development Platform. https://unity.com/

[85] David Unbehaun, Konstantin Aal, Daryoush Daniel Vaziri, Rainer Wieching, Peter Tolmie, and Volker Wulf. 2018. Facilitating Collaboration and Social Experiences with Videogames in Dementia: Results and Implications from a Participatory Design Study. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–23. https://doi.org/10.1145/3274444

[86] David Unbehaun, Daryoush Daniel Vaziri, Konstantin Aal, Rainer Wieching, Peter Tolmie, and Volker Wulf. 2018. *Exploring the Potential of Exergames to Affect the Social and Daily Life of People with Dementia and Their Caregivers.* Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3173574.3173636

[87] Valve Corporation. 2016. *SteamVR.* [VR]. https://store.steampowered.com/steamvr

[88] Valve Corporation. 2019. Valve Index. https://www.valvesoftware.com/en/index

[89] Joeke Van Santen, Rose-Marie Dröes, Marije Holstege, Olivier Blanson Henkemans, Annelies Van Rijn, Ralph De Vries, Annemieke Van Straten, and Franka Meiland. 2018. Effects of exergaming in people with dementia: results of a systematic literature review. *Journal of Alzheimer's Disease* 63, 2 (2018), 741–760. https://doi.org/10.3233/JAD-170667

[90] Vero Vanden Abeele and Veerle Van Rompaey. 2006. Introducing Human-Centered Research to Game Design: Designing Game Concepts for and *with* Senior Citizens. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI EA '06)*. Association for Computing Machinery, New York, NY, USA, 1469–1474. https://doi.org/10.1145/1125451.1125721

[91] Jayne Wallace, Anja Thieme, Gavin Wood, Guy Schofield, and Patrick Olivier. 2012. Enabling Self, Intimacy and a Sense of Home in Dementia: An Enquiry into Design in a Hospital Setting. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2629–2638. https://doi.org/10.1145/2207676.2208654

[92] WanadevStudio. 2020. *Ragnaröck.* Game [VR]. WanadevStudio. Lyon, France. https://www.ragnarock-vr.com/home/.

[93] World Health Organization (WHO). 2020. Dementia. https://www.who.int/news-room/fact-sheets/detail/dementia Accessed: 28.05.2021.

[94] Anders Wimo, Linus Jönsson, John Bond, Martin Prince, Bengt Winblad, and Alzheimer Disease International. 2013. The worldwide economic impact of dementia 2010. *Alzheimer's & Dementia* 9, 1 (2013), 1–11. https://doi.org/10.1016/j.jalz.2012.11.006

[95] Bob Woods, Elisa Aguirre, Aimee E Spector, and Martin Orrell. 2012. Cognitive stimulation to improve cognitive functioning in people with dementia. *The Cochrane Database of Systematic Reviews* 15, 2 (2012). https://doi.org/10.1002/14651858.CD005562.pub2

[96] Lidia Yágüez, Kendra N Shaw, Robin Morris, and David Matthews. 2011. The effects on cognitive functions of a movement-based intervention in patients with Alzheimer's type dementia: a pilot study. *International journal of geriatric psychiatry* 26, 2 (2011), 173–181. https://doi.org/10.1002/gps.2510

Received February 2021 ; revised June 2021 ; accepted July 2021

10.3 [PREPRINT] EVALUATING AUGMENTED LOCOMOTION AND RANGE OF REACHABLE OBJECTS FOR OLDER ADULTS IN A VIRTUAL REALITY EXERGAME

[CorePre2] **Sukran Karaosmanoglu**, Sebastian Finnern, Frank Steinicke, and Katja Rogers. Evaluating augmented locomotion and range of reachable objects for older adults in a virtual reality exergame. 2025. DOI 10.25592/uhhfdm.16795. URL https://doi.org/10.25592/uhhfdm.16795.

Supplementary materials for this article are included after the full-text of the paper.

Evaluating Augmented Locomotion and Range of Reachable Objects for Older Adults in a Virtual Reality Exergame

Sukran Karaosmanoglu, Sebastian Finnern, Frank Steinicke, Katja Rogers

***Abstract*—Virtual reality (VR) exergames are increasingly used to support user motivation and retention. Similarly, virtual *augmentation*—virtually adjusting players' abilities in virtual environments—of user abilities can elicit motivation and empowerment. However, research on augmented VR interaction with older adults is limited. We designed an exergame for older adults with augmented interaction, specifically augmented locomotion and augmented reachable range for object manipulation. In two user studies, we explored how augmentation affects player experience and performance with younger adults (aged 18–35 years, N=29) and older adults (aged 69–89 years, N=24), respectively. Our study with younger adults indicates that augmentation (primarily locomotion) significantly contributed to intrinsic motivation, physical activity enjoyment, and game performance without increasing cybersickness or diminishing physical activity. However, augmentation did not yield the same effects for older adults, and even decreased physical activity. Our work suggests exploration of different augmented interaction types and carefully considering age group differences when implementing augmented interaction in VR exergames.**

***Index Terms*—virtual reality, exergames, older adults, augmentation, locomotion, swimming.**

I. Introduction

Insufficient physical activity is a major problem that threatens public health [1]–[3]. The reasons for not engaging in physical activities can be various [4], [5]. For example, maintaining regular physical exercise habits can be challenging for many. However, exergames, defined as "*digital game[s] where the outcome [...] is predominantly determined by physical effort*" [6], can be motivating to support physical exercise and psychological well-being [7] among other benefits [8].

People are increasingly using virtual reality (VR) for physical training. VR exergames are amongst the most popular types of VR games; they can provide immersive experiences, fulfill entertainment needs of people, and offer new ways to perform physical activities. VR's unique opportunity to mediate realistic experiences holds great and varied potential in different contexts, such as for people with accessibility needs [9], [10], for medical training [11], or physical performance [12]. *Unrealistic* VR experiences even increase the possibilities of this technology: Exaggerated interaction can be used to improve efficiency and accessibility [13], motivation [14], as well as the user's feeling of empowerment [15].

Both for games and exercising, motivation plays a crucial role, and exaggerated interactions can be exploited to increase a user's motivation. Past studies have used different terminologies to describe enhanced abilities of users [16]: "*virtual performance augmentation*" [14], [17], "*mixed reality empowerment*" [18], "*exaggeration*" [19], "*magic interaction techniques*" [16], and "*hyper- and super-natural interaction techniques*" [20]. In this paper, similar to prior work [14], [17], [21], we use the term *augmented* interaction (i.e., the ways in which users engage with a virtual environment) to express the concept of virtually adjusting players' abilities and enhancing their capabilities in VR. By augmenting their abilities, users can fly [22] or overcome great distances with teleportation or jumping [21], which can also help them explore virtual worlds more flexibly. Such adjustments to players' abilities can positively affect their motivation [14], [15]. Despite these apparent benefits, investigation of using these elements in VR exergames [14], [15], [17], [19], [23] is sparse. Existing work mostly includes participants who were adults and general population [14], [17], [23]. However, none of these studies have considered or have been adapted for older adults in the VR exergames context. Research into exergames for older adults has shown promising results (e.g., improved performance [24] and physical functions [25]), but also consistently emphasizes the need for adapting factors like speed (e.g., [26]) and a preference for scenarios familiar from the real world or everyday life [27]. We thus need empirical data to learn if augmented interaction in a VR exergame can positively affect the experience of older adults.

The objective of this work is to understand effects of augmented interaction in a VR exergame for *older adults*, defined as aged 65+ years [28]. We also explored effects for *younger adults* (defined as aged 18–39 years [29]) to connect with prior work on augmented interaction with younger people [14], [17], [23], and to understand the difference and similarities in augmentation experiences between young and older adults. To do so, we designed an accessible exergame in which older adults can perform seated swimming arm movements by using a mobile VR: *ExerSwimVR*. We chose this theme because swimming is a popular activity and a relatively gentle exercise (especially when approximated in VR) that may be more inclusive for older adults, and this theme has been used [30] and requested [9] by older adults in prior VR studies.

Sukran Karaosmanoglu, Sebastian Finnern, and Frank Steinicke are with Human-Computer Interaction group, Universiät Hamburg, Germany. Katja Rogers is with Digital Interactions Lab, University of Amsterdam, the Netherlands.

Corresponding author e-mail: sukran.karaosmanoglu@uni-hamburg.de
Manuscript submitted July 16, 2024.

Augmentation was designed and applied separately in two ways: how users were able to move—i.e., swim—in the VR world (*augmented locomotion*) and how users were able to interact with virtual objects (*augmented object range*). We focused on these two aspects because (i) movement is a fundamental aspect of exercising and virtual locomotion, and (ii) objects are the basic elements that we interact with in VR. Thus, understanding augmentation for these two aspects can provide higher generalizability to other VR exergames and applications, since almost all involve movement and/or some form of interaction with objects.

In this work, we followed a mixed-methods research design [31] to assess how well augmentation was received by users in a VR exergame. We conducted two within-participants studies to investigate the effects of augmentation on player experience (PX): *User Study 1* (*N*=29) with younger adults and *User Study 2* (*N*=24) with older adults. In both studies, we assessed PX on the following constructs: (i) intrinsic motivation, (ii) cybersickness, (iii) immersion, (iv) physical activity enjoyment, and (v) performance. Our findings show that for younger adults, primarily augmentation of locomotion led to significantly higher intrinsic motivation, perceived competence and choice, and physical activity enjoyment, without causing cybersickness or less physical activity. However, for older adults, augmentation of any type of interaction largely did not significantly affect their experience, but decreased physical activity. We discuss potentials of augmented interaction for younger adults as well as older adults, and similarities and differences in their experience. Finally, we also provide takeaways to guide the future design of augmented interactions in VR exergames: highlighting the need to consider augmentation of different interactions and age group differences. We summarize our contributions as follows:

- a demonstration of an accessible VR exergame with a swimming theme designed for older adults,
- a first exploration of how older people experience and perform with augmented interaction for locomotion and/or object range in a VR exergame,
- an exploration of the same game with younger adults to connect our results with previous works, and
- a discussion of how the same augmented interaction differs in effect when experienced by older adults.

II. Background

This section provides our rationale on the design choice for a swimming theme, and considerations on the potential of enhanced VR interactions, how VR exergames can benefit from augmentation, and exergame design for older adults.

A. Swimming Theme in VR

The theme of our exergame is swimming—a common way to exercise [32]. Swimming activates many muscle structures [33] and positively affects mental and physical health [34]. It thus hold a positive association for many. Further, general exposure to aquatic environments can positively impact mental health [35]. However, we also note that these advantages might not occur in VR (e.g., no actual water exposure). Yet, prior VR studies have used underwater VR worlds as stimuli for relaxation [36], the commercial market offers swimming themed experiences (e.g., [37]–[39]), swimming-themed VR has been used to test technology for older adults [30], and a swimming theme was explicitly requested by older adults in a prior VR study [9].

Swimming is a challenging activity to represent in VR due to its water environment and horizontal whole-body movement. Yet VR research features artifacts with swimming-themed experiences. A few prototypes can be used underwater [40], [41], but most work aims to provide a VR swimming experience without getting into the water. While a few prototypes were designed with a harness to enable flexible movement from a recumbent in-air position [42], [43], most swimming-themed VR applications are experienced from a standing or seated position [37]–[39]. To our knowledge, swimming in VR has not yet been tested with a seated user position for older adults. However, seated VR experiences in general have shown much promise [44]. We thus draw on these previous explorations to implement a seated swimming-themed experience, thus leveraging accessibility and safety for our target group of older adults [9].

B. Enhancement of User Abilities in VR

Enhancement of users' abilities can result in improved user experience and there are various ways to achieve this. For example, to overcome the spatial boundary limitations of VR, a large body of locomotion research has explored redirected walking, which allows users to walk endlessly along a curved path in virtual environments [45], [46]. Similarly, gesture-based techniques (e.g., walk-in-place [47], run-in-place [14]) are widely used to provide more natural experiences and improved presence compared to less natural locomotion techniques (e.g., hand pointing) [47], which inspires us to use a swim-in-place technique for augmented locomotion.

Various researchers have implemented "*superpowers*" via altered interaction options in VR. Some report positive feedback from designers/developers [48]. One technique gave users the ability to navigate large VR environments by switching between first- and third-person views [49]; while in third-person view, users could choose the targets from a vantage point to move their virtual avatar. Another example [21] demonstrated a VR technique of augmented jump abilities. Most conditions with augmented jumping ability led to higher intrinsic motivation compared to the non-augmented one.

Enhanced VR abilities have also been explored to allow users to interact with distant objects, such as via ray-casting techniques (a virtual laser beam emanates from users' hands for remote object interaction [50]) or stretching avatars' arms [51]–[53]. Ray-casting can have disadvantages in situations with high object occlusion [54] or required precision [52], yet early VR research suggests it may be preferred for grabbing objects compared to arm-extension techniques [51]. As it is an established technique in VR [55], and well suited to grabbing objects (thus increasing the chances of success), we use this to explore virtual augmentation for grabbing objects in our VR exergame.

C. VR Exergames and Augmentation in Exergames

VR exergames have been extensively researched recently [56]. Unlike other systems, these games often include built-in six degrees-of-freedom motion tracking that allows users to physically interact with spatial worlds. Users experience these worlds through stereoscopic immersive displays. Research on VR exergames yields promising results [9], [57]–[59]. For example, VR exergames support intrinsic motivation more than their non-VR versions [58]. Similarly, a previous study [24] showed that players' exergame performance was higher in VR than when playing on a large screen. While VR systems can offer more motivating exergame experiences, their use also holds challenges, such as cybersickness symptoms [60], usability issues [61], or disconnection from the real world [9]. Thus, VR exergames require careful considerations in game design, pilot testing, and the involvement of experts, which we follow in our research.

For some years now, enhancing the virtual abilities of players has been of interest for exergames research. Though the literature uses heterogeneous terminology to express the same concept [16], augmented interactions especially in exergames share a common goal: to motivate players [14], [15], [17]. Previous works have supported the link between virtual augmentation in exergames and motivation [14], [15], and intrinsic motivation was found to be associated with long-term exercise adherence [62]; yet no studies considered older adults.

Most research on augmentation in exergames has augmented players' ability to move in the game world. Jumping was a notable topic, with augmentation showing a positive effect (e.g., motivation, enjoyment, competence) in mixed-reality trampolining [15], [23], and a jump and run VR game [14]. Moving beyond jumping, Granqvist et al. [19] examined the effect of VR avatar flexibility on user performance and experience in three conditions (realistic, moderately exaggerated, and highly exaggerated). They found that moderate exaggeration led to better performance and higher competence compared to the realistic variant. Born et al. [17] found that augmentation via knock-back effects in VR punching can motivate players to engage in optional strenuous activity for longer compared to other motivational conditions (points and no motivator). Nevertheless, research into other types of augmented interaction in exergames remains limited.

Many exergame studies compare an augmented version of interaction to a non-augmented version. We note that these non-augmented versions are of course still not fully realistic. For instance, Wolf et al. [21] detect players' vertical jumps using the VR HMD position, but did not track other body parts or muscle activity to provide more realistic jump detection. Born et al. [17] developed a custom hand trainer-turned-controller that players pressed to punch enemies. Thus while they need to make a fist, the arm does not need to travel to the target to perform a punching motion, and one feels no haptic feedback upon impact. In line with these studies [17], [21], our non-augmented versions provide an approximation of swimming arm motions (locomotion) and grasping objects (range); the augmented versions for these activities (locomotion speed and reachable range, respectively) augment the user's abilities.

D. Exergames for Older Adults

For older adults, the use of exergames can reduce depression [63] and support physical functions [25]. Nevertheless, many researchers highlight design factors that need to be considered particularly for older adults. For example, in immersive VR exergames, researchers emphasize the need to consider health conditions (e.g., dementia [9]), need for audiovisual congruence [61], and the use of natural movements [64]. Reviews on using immersive VR for older adults' physical health and rehabilitation discuss potential adverse effects (e.g., cybersickness [65]), but also benefits [66]. For instance, older adults found exercising on a treadmill less motivating than a VR game (i.e., Beat Saber [67]) [64], and preliminary work [57] reported promising results on an arm-extension VR exergame. We draw on the lessons learned and guidelines of designing immersive VR exergames for older adults in *ExerSwimVR*: for example, by creating a seated experience, and requiring only gentle activity through breaststroke arm movements for locomotion.

So far, research on augmentation in VR exergames has primarily focused on children [15] and adults in the general population [14], [17], [23]. There are some indications that age differences can affect exergame experiences or performances in general [24], [68], yet there is no empirical data on how older adults respond to such augmentation approaches. Understanding the perspectives of older adults on these aspects holds major importance: the aging world population [28] may be more likely to use exergames and VR technology, yet the field's understanding of older adults' experience of VR and games is still limited. Further, previous research on exergames for older adults has often suggested a preference for "realistic" [69] scenarios [27]. Thus, it is possible augmented interaction—by being unrealistic—would not improve older adults' experience, prompting a need for empirical data with older adults that we answer with this paper.

III. Research Questions

We contribute to the field's knowledge by investigating whether and how certain augmented interaction affects PX and performance in a VR exergame. We do so in a study with younger adults to see how our results compare with previous studies exploring augmented interaction in VR [21] or exergames with a similar age group [14], [17], [23]. We then conduct a study with the same stimuli (VR game and augmented interaction techniques) with older adults. While our key focus is on older adults, we included a study with younger adults for three reasons: (i) to reduce the chance that our results would be specific to our artifact, (ii) to see similarities or differences in the experience of these user groups using the same stimuli, and (iii) to replicate the findings of previous studies with a younger adult population [14], [17], [21], [23].

Building on the outlined research gap, this work focuses on two main interaction types that are essential to exergames as well as VR in general: augmenting (i) *locomotion* and (ii) *the range of reachable objects*. Specifically, we augmented the players' virtual locomotion speed (i.e., movement manipulation) and/or the range at which they could reach objects (i.e.,

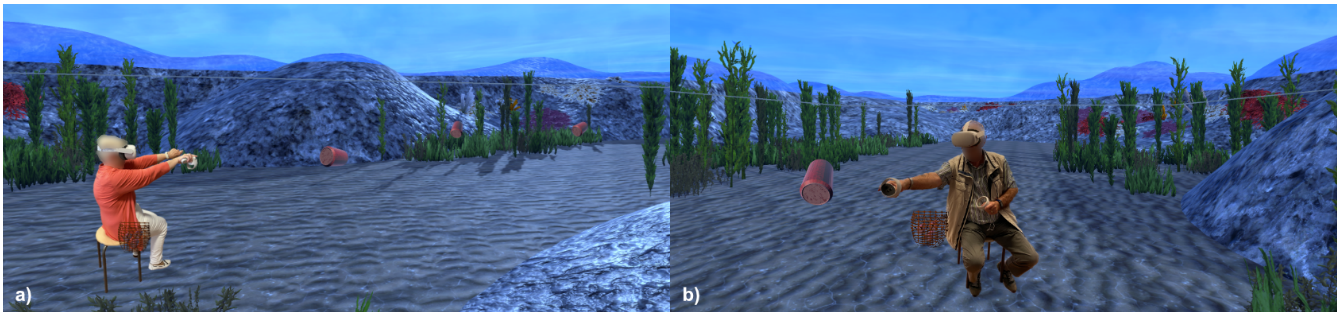

Fig. 1. In *ExerSwimVR*, players explore a virtual underwater world in VR. They perform breaststroke arm movements to swim (a) and collect pollution objects to clean the ocean (b).

object range manipulation). Our research was driven by the following research questions:

RQ1: How does augmented interaction (i.e., (i) augmented locomotion and (ii) augmented range of reachable objects) in a VR exergame affect intrinsic motivation, immersion, physical activity enjoyment, performance, and cybersickness of players?

RQ2: How does the same augmented interaction differ in effect on those same PX components and performance when experienced by older adults?

IV. ExerSwimVR

To answer our research questions, we designed and implemented *ExerSwimVR* (see design rationale in the background and game design sections). We report our VR exergame following existing guidelines [56]. In this seated single-player VR exergame (i.e, player mode and exercise position [56]), players perform physical activities in the form of breaststroke arm movements (i.e., stretching arms and hands out in front with hands at a 45-degree outward angle, pulling them slightly down and outward until shoulder level [70]) and reaching motions for object collection (i.e., upper-body exercises, and collecting/catching- and path-based tasks [56]). Our game was implemented in Unity (v.2019.4.21f1) [71][1]. We primarily used the Meta Quest 2 [75] (and in the first study its predecessor Meta Quest 1 [76] in a few cases when no more Meta Quest 2 devices were available or when participants joined with their own device) as the VR head-mounted display (HMD) devices, because they can run the game stand-alone without a computer (i.e, mobile setup [56]).

A. Game Design and Pilot Testing

ExerSwimVR is a swimming-themed VR game (i.e., theme with everyday familiarity [56]) in which people experience an underwater world, as desired by older adults in a previous VR exergame study [9]. The underwater game scene features sandy soil, marine plants, corals, anemones and algae, and includes ambient underwater sounds, as audiovisual congruence in VR is recommended for older adults [61]. These elements were used as decoration but sparingly, to not cause distraction or increase players' cognitive load (see Figure 1). We did not use any sea animals in the game scene to prevent potential zoophobia. Canyon rocks delineated the game path that players should follow. In addition, the game was populated with rusted red cans (as object targets) along the path; the players were tasked with cleaning up this pollution by collecting these objects. All targets (cans) were positioned approximately 7.5 meters from each other (see Figure 2); these positions were identical in all game variants (with different augmented interaction). The game was implemented as a seated VR exergame to account for safety, disability, and age-related physical conditions. This is in line with prior work [9] on VR applications for older adults (i.e., target age range [56]) and also to meet recommendations by a physiotherapist that we interviewed—see below). With this game, we aim to provide physical training and promote physical activity [56].

During the design process, we conducted two informal pilot tests to carefully decide on optimal manipulation values: one with younger adults (*N*=5; age: *M*=28, *SD*=1.22 years; 1 female, 4 male), and one with older adults (*N*=4; age: *M*=67.25, *SD*=0.5 years; 2 female, 2 male). Through these, we gained a preliminary understanding of player perspectives on different exaggeration variants and magnitudes, i.e., how much locomotion and object range should be augmented in the main studies. Based on these informal studies, we chose the augmentation values on movement and object manipulation that are described in the next section.

Additionally, we conducted a semi-structured interview with a physiotherapist with 28 years of work experience to assess potential benefits and safety aspects of this game (questions provided in supplementary materials); consulting a physiotherapist when designing an exergame for older adults is recommended [9], [77]. The physiotherapist considered the game design safe for older adults. Based on her recommendations, we made sure to use a stable chair (without wheels) during gameplay. Further, we chose to use a chair without back or armrests so that they would not injure themselves when performing the swimming arm movements. Regarding the duration of each game session, the expert recommended to keep it at one minute to avoid fatigue among the older adults: "*I would not do it longer, that's for sure, not in one part. [...] I think, If it's for older people, I would stay in [...] 1 minute. Even less maybe. The first level even less.*" Given the

[1]In addition, we used: Oculus XR Plugin (v.1.11.2) [72], Oculus Integration Package (v.29.0) [73], and XR Interaction Toolkit (v.0.10.0.preview.7) [74].

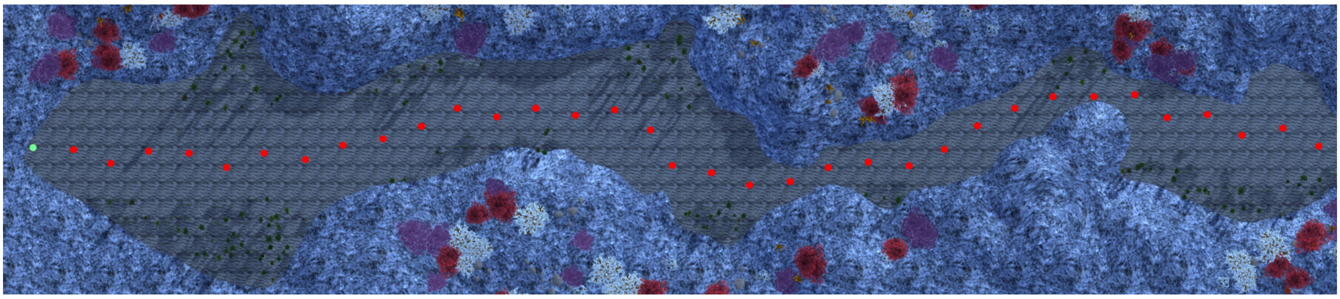

Fig. 2. The figure illustrates part of the top view of the game environment. The green dot represents the start position of players; the red dots represent the positions of objects that the players need to pick up, thus forming the path they need to follow.

expert's strong emphasis on brief game levels for older adults, we kept the play duration at one minute for each game level as going against their advice would be unethical. Similar levels of exposure time have been used in previous VR exergames studies [78], [79] and other VR works [80], [81], and short VR exposure times are recommended for older adults [82]. Finally, we implemented an interactive VR tutorial level for each game variant for both user groups (and softened the time limit for older adults), to give them enough time to become familiar with *ExerSwimVR*.

B. Implementing Approximated Underwater Swimming in VR

In *ExerSwimVR*, players perform breaststroke arm movements in a seated position to navigate the virtual world. To implement an approximated swimming experience in VR, we calculated players' velocity based on acceleration (*thrust*) and deceleration (*drag*) in relation to their arm motions.

1) Acceleration and Thrust: The acceleration of players was calculated based on drag and lift forces [83] as well as additional factors ($thrust = (dragForce + liftForce) \cdot wingspanFactor \cdot forceDirectionFactor$). Effective acceleration is created when using the hands with a downward-outward angle of about 45 degrees [70]. To integrate this aspect, we multiplied the controllers' velocity by the hand orientation of players (*dragForce*). *DragForce* was calculated as the dot product of the 45-degree angle and the orientation of the controller normalized to range from zero to one. Deviations from a 45-degree angle result in less force.

Lift is the upward force and behaves perpendicularly to the drag force [83]. To calculate the *liftForce*, we multiplied the angle between the players' hand orientation and the drag force direction (horizontal). The hand orientation factor was obtained by the dot product of a vector perpendicular to the players' hand orientation and a global vector pointing up. Consequently, lift force was maximized when the players' hands were parallel to the ground.

Further, we took into account how players perform the arm movements. We measured the players' wingspan at the beginning of the game; the players were asked to stretch their arms out sideways. Based on this, we implemented a *wingspanFactor* calculated by dividing the distance between controller and body center (estimated based on HMD position) by half of the maximum wingspan. For example, an arm extended half way (i.e., a quarter of the maximum wingspan) results in a wingspan factor of 0.5. This ensured that forces produced at near maximum wingspan (further away from the players' center) led to greater thrust; in contrast, moving their hands closer to their body reduces the distance between controller and body, and the resulting force is decreased.

For the computation of the composite force, it was ensured that only force perpendicular to the players' arm movement was considered (*forceDirectionFactor*). To achieve this, we calculated the dot product of a vector drawn between the players' center and the respective controller, and the movement direction of the controller. Therefore, force was maximized when swimming movements were performed in circular motions around the body.

2) Deceleration and Drag: To calculate deceleration, we considered the drag force and its dependence on players' velocity (v). Since Unity [71] does not offer drag force calculation for liquids, we calculated the deceleration forces using the following formula [84], [85]: $drag = 0.5 \cdot C \cdot \rho \cdot A \cdot v^2$. For the drag coefficient (*C*), we used the value 0.2 (approximated from [86]). The water density value (ρ=995) based on a temperature of around 30 degrees Celsius [87] was chosen due to the tropical look of the underwater environment. The players' surface area (*A*) was computed based on average height values of the national population [88]. We used the height values of people aged older than 18 for the *User Study 1*, and the height values of people older than 65 years for *User Study 2*. The surface area of the upper body for the two studies was then calculated as half of the body height multiplied by the quarter of the total body height (approximated from [89]), resulting in 0.37 m^2 for *User Study 1* and 0.36 m^2 for *User Study 2*, respectively.

C. Gameplay and Augmented Interaction

ExerSwimVR starts with a user interface that lets players or experimenters configure the game, e.g., to select an identification number (allowing us to run the study in a counterbalanced fashion), and indicate their hand preference (left or right handedness). Players calibrate an arm avatar (see Figure 4a) to match their physical arm size, giving us a measurement of their wingspan and allowing us to adapt the range of motion [56] based on players' arm reach.

Players' task and main interaction is to clean up the water by collecting pollutants in the underwater world (see Figure 4). Before each level, players completed a tutorial consisting

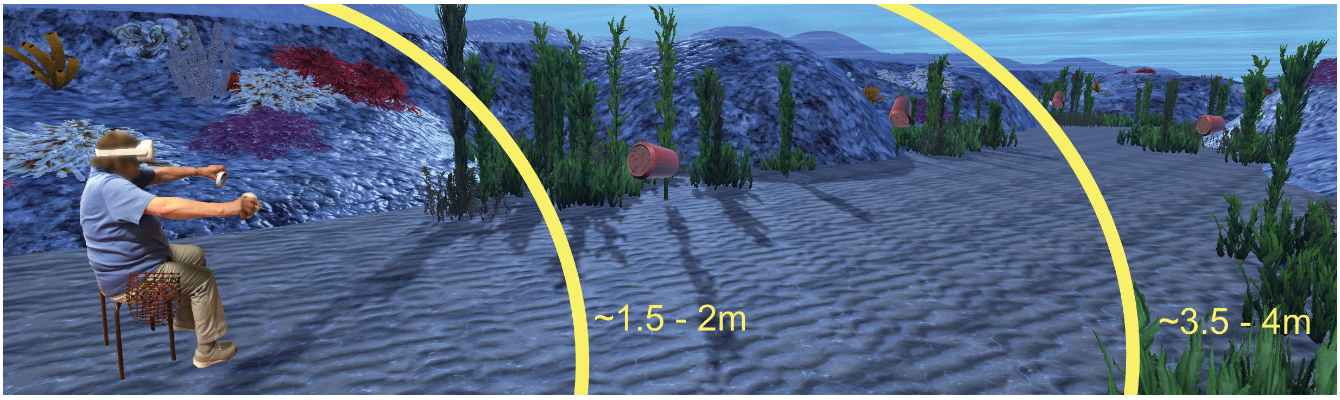

Fig. 3. In the game, when players did not experience locomotion augmentation (*NO-AUG* and *OBJ-AUG*), they moved 1.5-2 meters with one arm stroke in VR and reaching an object approximately took 4-5 strokes. However, when they experienced locomotion augmentation (*LOC-AUG* and *BOTH-AUG*), one arm stroke made them to move 3.5-4 meters in VR and reaching an object approximately took 2-3 strokes. We note that these values are only an estimate based on trials by one researcher and can be affected by factors considered in the calculation of locomotion (e.g., arm length and angle of hands).

Fig. 4. In the game, the players were tasked to collect pollution objects (a). In *NO-AUG* and *LOC-AUG*, the players had to swim close to the objects and collect them by pressing a grip button of a controller (b). In *OBJ-AUG* and *BOTH-AUG*, the players were able to collect objects from three meters away using a laser beam (c). After grabbing the objects, they put the objects in their basket (d).

of the same world as the actual game level, but with only three placed pollutants, giving them a chance to try out the relevant interaction technique. For *User Study 1*, the tutorial ended after 30 seconds or upon successful collection of any target. We applied this time constraint for younger adults because the *User Study 1* was conducted remotely, and we wanted to prevent users exerting themselves in tutorial levels and testing the experimental levels with already increased exertion. For the second study, we wanted to give the older adults sufficient time to get familiar with the task and VR technology, following guidelines for exergames for older adults [90]. Hence, we did not constrain the tutorial duration in *User Study 2* and only used the end criteria of successfully collecting a target, while the experimenter monitored to interrupt in the case of over-exertion. This was not necessary; participants' completed the successful collection with a median of 30.34 seconds duration. After the tutorial, players experienced the main game level (in each respective game variant of the study condition) for a 1-minute duration as recommended by the physiotherapist.

1) Locomotion: Players performed breaststroke arm movements to reach the objects (see Figure 3). In the game, the avatar's hands followed the position of the controllers; the position and movements of the remaining arm portions were calculated using inverse kinematics. In the no-augmentation variant, locomotion was implemented to represent a close-to-lifelike response to swimming arm motions (see implementation details above). In the augmented variant, the same arm movements in real life were translated into virtual movement at an exaggerated velocity by a factor of six, resulting in further virtual distance for the same physical movement.

2) Range of Reachable Objects: Once in reaching distance of the objects, players could collect them and place them inside a basket positioned either at their side (matching their handedness, see Figure 4d). In the no-augmentation variant, the range at which objects could be collected resembled natural object manipulation: Players had to press their controller against the object and hold the trigger button on the controller to grab them (see Figure 4b). In the augmented variant, pressing the controller triggered the emission of a laser pointer from the player's virtual hand (see Figure 4c). When targeting an object at a distance with this laser, the object would be transported to the player's hand in a snapping motion. The laser pointer was set to enable players to target objects located at up to three meters distance from the center of the avatar hand to the outer surface of the object (3.4 meters from the center of the avatar hand to the center of the object).

V. User Studies

We conducted two within-participants user studies to evaluate PX of augmented interactions in the VR exergame *ExerSwimVR*. The first was conducted remotely (due to COVID-19) and recruited young adults in the general population, in line with prior studies on augmented interaction techniques in VR [14], [17], [23]. A few months later, we then conducted

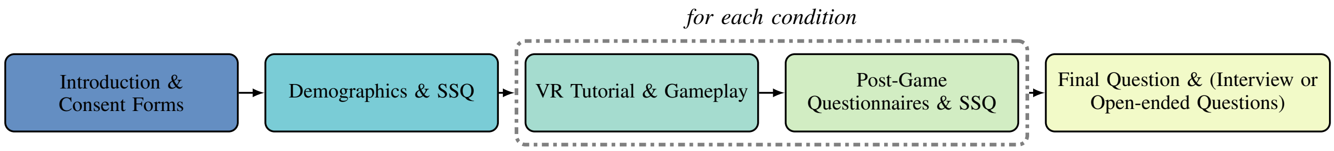

Fig. 5. The studies followed a within-participants design (two factors: augmented locomotion and augmented reachable range for object manipulation) and assessed questionnaire, performance, and qualitative data.

User Study 2 with older adults as an in-person study at a living facility[2]. Ethics assessments were conducted for both studies; due to the specifics of the user group in *User Study 2*, we additionally applied for and received ethics approval by the local ethics committee of Universität Hamburg (#006/2021).

A. Conditions

Both studies consisted of four conditions, and a video of these interaction variants can be found in the supplementary:

- *NO-AUG - No Augmentation:* Players performed swimming motions without any manipulation of their virtual velocity. They collected targets without any manipulation of object range (see Figure 4b).
- *LOC-AUG - Augmentation of Locomotion:* Players' virtual velocity was increased by a factor of six. However, the reachable range for object manipulation was identical to *NO-AUG* (see Figure 4b).
- *OBJ-AUG - Augmentation of Reachable Range of Objects:* Players' movement was not augmented, but they were able to grab the target objects from up to three meters away (see Figure 4c).
- *BOTH-AUG - Augmentation of Locomotion and Reachable Range of Objects:* In this variant, both virtual locomotion and reachable range of objects (see Figure 4c) were augmented.

B. Participants

User Study 1 had young adults (defined as under 40 [29]) whereas *User Study 2* focused on older adults (65 years or older [28]). As sample characteristics can impact results (e.g., negative correlations between gaming experience and cybersickness [91]), we provide further sample details (e.g., prior VR and gaming experience) in the supplementary.

a) Sample Characteristics: User Study 1: We initially recruited 34 remote participants. After excluding two participants who did not complete all experimental conditions, and three more who were older than 39 years, our sample consisted of 29 participants. They were aged between 18 and 35 years (*M*=24.69, *SD*=3.91); 19 identified as female and 10 as male.

b) Sample Characteristics: User Study 2: Twenty-five participants were recruited with the help of the senior living facility. One withdrew, leaving a total of 24 older adults as participants in the study analysis (female: 19; male: 5). Their average age was 81 (*SD*=4.85, *Min-Max*= 69–89). One player reported a visual impairment (i.e., decreased vision) and another reported a bone-related impairment that affected finger movements / touch sensitivity, however neither felt that it impeded their VR experience severely.

[2]COVID-19 measures were then relaxed and this allowed us to provide more support to our participants if needed. Older adults assigned a care level were not recruited, as a further safety precaution due to COVID-19.

C. Measures

In both studies, we collected data using the same questionnaires, gameplay metrics, and interview questions. Questionnaire items and interview questions are provided in the supplementary materials.

a) Pre-Game Questionnaires: The pre-game demographics questionnaire captured participants' age, gender, gaming habits, VR experience, and physical activity. We also employed the Simulator Sickness Questionnaire (SSQ) [92] (German version [93]) to assess the intensity of cybersickness symptoms. It consists of 16 items on a 4-point scale (*none/0 — severe/3*), yielding a total SSQ score, and nausea, oculomotor disturbance, and disorientation sub-categories.

b) Post-Game Questionnaires: After each gameplay, we employed the SSQ again to assess the effect of gameplay on cybersickness symptoms. Additionally, we used three subcategories of the Kurzskala intrinsischer Motivation (KIM) questionnaire [94], which is the shorter validated German version of the Intrinsic Motivation Inventory (IMI) [95]. The interest/enjoyment, perceived competence, and perceived choice categories of KIM (a total of 12 items) are collected on a 5-point Likert scale (*not at all true/0 — very true/4*). Players rated their immersion on the immersion construct of the Player Experience Inventory (PXI) [96] (German version [97]) on a 7-point Likert scale (*strongly disagree/-3 — strongly agree/3*). To measure their enjoyment of movements, we used the physical activity enjoyment scale (PACES-8) [98], which consists of eight bipolar items rated on a 7-point scale (e.g, *It's no fun at all; It's a lot of fun*, ranging from 1 to 7). The items of this scale were translated, discussed, and finalized by three German speakers. To assess players' perceived exertion level, we used a custom-single item on a 7-point Likert scale (*strongly disagree/1 — strongly agree/7*): "*I found the game to be physically demanding*".

c) Final Phase: After all conditions, we asked a final question to understand players' preference regarding the game variants: "*Which version of the game did you like best?*" Following this, we used five questions to gather qualitative feedback from participants in the form of open-ended questions (in *User Study 1*) and semi-structured interviews (*User Study 2*), respectively. In *User Study 1*, we opted for open-ended questions instead of interviews to avoid difficulties in scheduling video calls with remote participants. In *User Study*

2, we conducted semi-structured interviews instead to more easily capture the older adults' views on the game variants. The questions targeted their experience of the augmented interaction techniques and reasons for their perspective on them; the questions are listed in the supplementary.

d) Gameplay Metrics: We logged physical activity as a vector length calculated based on the controller positions[3] (using Unity unit metrics). We summed up the values for left and right hand distance traveled to obtain a total physical activity metric in each condition. Additionally, the number of collected targets in each game variant was recorded.

D. Procedure of User Studies

Both studies followed a similar user study procedure as illustrated in Figure 5; differences are detailed as follows:

a) Procedure of User Study 1: This study took place in a remote public study week of Universität Hamburg. The study week was announced using a convenience sampling method (e.g., student mailing lists). During this study week, if required, we provided participants with an HMD that they took home to participate in user studies.

At the beginning of the study, participants were provided a digital information letter and consent form, a procedure document that explained the study, and a VR usage document. After giving their consent, they completed the demographics and initial SSQ (administered online). Once they wore the HMD, they were guided to calibrate their virtual wingspan by matching the size of a virtual arm avatar by stretching their arms out to the sides in a T shape. Following this, in a counterbalanced fashion (using a balanced Latin Square [99]), participants completed a tutorial (30 seconds or until they collected a target) and then 1 minute of gameplay for each condition. We note that both in written forms and in the textual explanation of the tutorials, participants were informed that they will experience alteration of their virtual movement or object range aspects. However, we did not use the word "augmentation" or synonyms to avoid priming effects. After each condition, the participants filled out the post-game questionnaires. At the end of the study, participants completed the final questionnaire about game version preferences and the additional open-ended questions capturing their experiences with different augmentation variants. Finally, if applicable, those participants who were students of Universität Hamburg received a participation certificate, which they need to collect to complete their studies. No remuneration was provided otherwise. We estimated the duration of the entire study procedure for young adults to have been ˜90 minutes, though we did not ask them to record the time.

b) Procedure of User Study 2: This study followed a similar procedure as the first one, but was conducted face-to-face in a senior living facility. To decrease the risk associated with COVID-19, we used disinfection wipes and a CleanBox system [100] to clean the VR HMDs and controllers. The participants gave their consent and were provided with a VR usage information document. After this, the players were given a verbal explanation of their tasks and shown the VR HMD and controllers. Following this, like in the first study, the participants filled out the demographics and initial SSQ (administered on paper), calibrated their virtual arms, completed the four counterbalanced conditions of tutorial and gameplay, and filled out the post-game questionnaires after each condition. As a final step, we conducted semi-structured interviews (instead of open-ended questions) to gain an understanding of the older adults' experience of the interaction variants. Overall, the study took ˜90 minutes.

E. Quantitative Analysis Methods

First, we tested the normality assumption using Shapiro-Wilk tests. When data were normally distributed, we ran two-way repeated-measures ANOVAs (factors: locomotion and object range manipulation) and paired t-tests for pairwise comparisons. Otherwise, we used non-parametric two-way repeated-measures ANOVAs[4] based on Aligned Rank Transform [101], [102] (ART[5]) and Wilcoxon-signed rank tests for pairwise comparisons. Additionally, to analyze pre- and post-comparisons of SSQ scores, we conducted Friedman's ANOVAs and treated the measurement points as time sequences with five levels (pre, post-*NO-AUG*, post-*LOC-AUG*, post-*OBJ-AUG*, and post-*BOTH-AUG*). We note that in *User Study 2*, we had a single missing data point in one of the SSQ item ratings. For this, we employed a k-nearest neighbour (kNN) imputation method [103], [104] to determine a replacement value of this single data point considering all SSQ items (k=3). We performed exact multinomial tests to analyze players' preferences for the game variants and exact binomial tests for pairwise comparisons. All pairwise test values were adjusted and reported with Bonferroni correction. Example code snippets of our R analysis scripts are provided in the supplementary materials.

As our quantitative data were generally non-normally distributed, we report median (*Mdn*) and interquartile range (*IQR*) of our measures [105] (see Tables I and III). For readability, we only report the significant results in the text; Table II shows the statistical results of all dependent variables.

F. Qualitative Analysis Method

The interviews were transcribed using Dovetail [106] and checked by German speakers against any mistakes. The German answers to open-ended questions and interviews were translated using DeepL Pro Translator [107] into English and were again rechecked by German speakers.

We chose to use a reflexive orientation of thematic analysis [108], [109], which requires in-depth engagement of researchers with the data, and is commonly used in HCI [69], [110], [111]. Since the purpose of the method is not to reach consensus between multiple coders (unlike coding reliability approaches [109]) but to show richness in the data, having one coder for analysis is valid and "*indeed good practice*" [112]. The analysis was shaped by the first author's background and perspective (positionality) [109], [112]: They have previous experience with conducting VR user studies with both

[3]https://docs.unity3d.com/ScriptReference/Vector3-magnitude.html

[4]Using R's repeated measures model (aov)

[5]We use the subscript F_a to indicate the tests run based on this method.

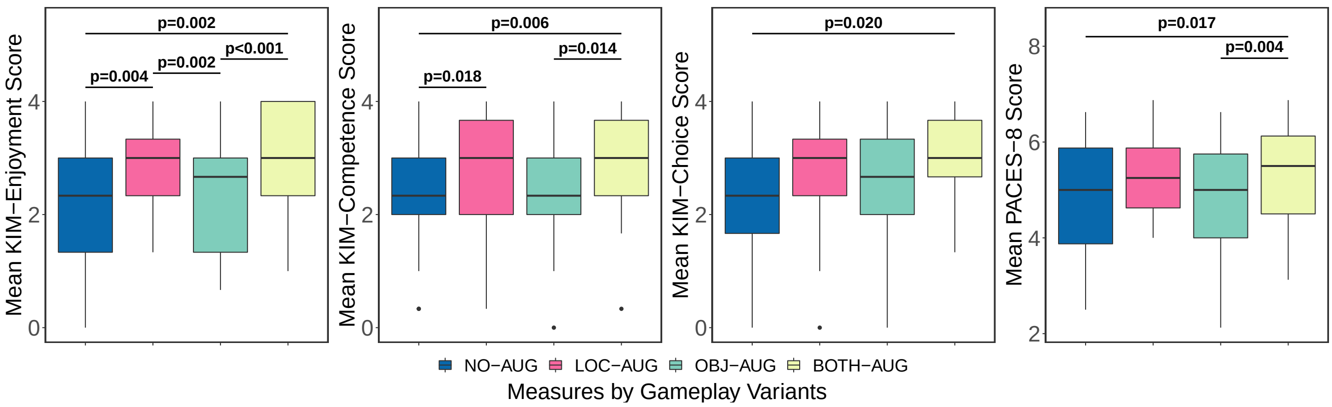

Fig. 6. The boxplots illustrate pairwise comparisons in the KIM sub-categories and PACES-8 in *User Study 1*. Comparisons with significant test results (Bonferroni correction applied) are indicated via p values.

younger and older generations and work on the intersection of accessibility, VR, and exergames research. They have a background in cognitive systems and psychology. We believe that this background is well suited to understand and analyze the qualitative data; we discuss potential downsides to this in the limitations.

Our qualitative analysis consisted of multiple steps. First, the lead author developed deductive categories based on the research questions: *reasons for preferring a game variant*, *impacts of movement augmentation*, *impacts of object range augmentation*, and *general feedback & improvements*. They then iteratively coded the data (with the broad deductive categories as a guide) using inductive descriptive codes (e.g., *increased speed feels natural*, *increased object range is fun*). Afterwards, focusing on the deductive categories relevant to the research question, the primary author used affinity mapping to organize the codes and quotes, and derived initial themes. The researcher iterated on the themes by going over the codes and quotes from the participants multiple times. After this step, the first author discussed the created themes with the last author and finalized them.

In the reporting, dominance of aspects is reported based on both the frequency and depth of the responses. We report the quotes of participants in the format PX_{UY}; e.g., $P3_{U1}$ refers to the third participant from the *User Study 1*.

VI. Results: *User Study 1* with Younger Adults

a) ***SSQ: Pre- / Post-Comparison:*** We did not observe a significant effect of time points on the total SSQ or most of its categories (with total SSQ: $\chi^2(4)$=3.10, p=0.541, *Kendall's W*=0.027; SSQ-Ocu: $\chi^2(4)$=8.96, p=0.062, *Kendall's W*=0.077; SSQ-Dis: $\chi^2(4)$=3.00, p=0.558, *Kendall's W*=0.026). There was a significant effect of time of measurement on SSQ-Nausea scores, $\chi^2(4)$=10.83, p=0.029, *Kendall's W*=0.093. However, pairwise tests indicated that there were no significant differences between times of measurement.

b) ***Post-Game Questionnaires:*** We found a significant main effect of locomotion augmentation on **KIM interest/enjoyment** values, F_a(1,28)=44.20, p<0.001, η^2_P =0.612. Participants enjoyed the *LOC-AUG* (V=20.5, p=0.004) and *BOTH-AUG* (V=36.5, p=0.002) variants more compared to the *NO-AUG*. Similarly, they felt significantly higher interest/enjoyment in *LOC-AUG* (V=338.5, p=0.002) and *BOTH-AUG* (V=20.5, p<0.001) compared to the *OBJ-AUG* condition. We observed a significant main effect of the locomotion augmentation on **KIM perceived competence** ratings, F_a(1,28)=17.05, p<0.001, η^2_P =0.378. Perceived competence was rated significantly lower for *NO-AUG* when compared to *LOC-AUG* (V=46, p=0.018) and *BOTH-AUG* (V=45.5, p=0.006). Moreover, for the same measure, *BOTH-AUG* was rated significantly higher than the *OBJ-AUG* variant (V=62, p=0.014). For the **KIM perceived choice** scores, again a significant main effect of locomotion augmentation was observed, F_a(1,28)=8.28, p=0.008, η^2_P =0.228. Perceived choice was rated significantly higher for *BOTH-AUG* compared to *NO-AUG*, V=47, p=0.020. There was no significant main effect of independent variables or interaction effect on **PXI-immersion** scores. For **PACES-8**, however, we found a significant main effect of locomotion augmentation, F(1,28)=15.66, p<0.001, η^2_P =0.359. Post-hoc tests showed that participants' physical activity enjoyment was higher for *BOTH-AUG* when compared to *NO-AUG* (t(28)=-3.26, p=0.017) and *OBJ-AUG* (t(28)=-3.84, p=0.004). The single-item **exertion** scores yielded a significant main effect of locomotion augmentation, F(1,28)=11.37, p=0.002, η^2_P =0.289—yet post-hoc tests did not indicate a significant difference between conditions.

c) ***Gameplay Metrics:*** There was a significant main effect of locomotion augmentation (F(1,28)=140.21, p<0.001, η^2_P =0.834) on the number of **collected targets**. The targets collected were significantly higher in *LOC-AUG* compared to *NO-AUG* (t(28)=-9.56, p<0.001) and *OBJ-AUG* (t(28)=8.38, p<0.001). A similar pattern was observed for the same conditions when compared to *BOTH-AUG*: *NO-AUG* (t(28)=-10.07, p<0.001) and *OBJ-AUG* (t(28)=-11.03, p<0.001). There was neither a significant main effect nor an interaction effect of the independent variables on physical activity.

d) ***Preference of Augmentation Variant:*** Participants preferred the conditions with locomotion augmentation applied: *BOTH-AUG* (n=16), *LOC-AUG* (n=11), in comparison to *NO-AUG* (n=2) and *OBJ-AUG* (n=0). An exact multinomial

TABLE I
THE TABLE SHOWS THE DESCRIPTIVE VALUES OF THE DEPENDENT VARIABLES IN *User Study 1* AND *User Study 2*.

	CONDITIONS	TOTAL SSQ		SSQ-NAU.		SSQ-OCU. DIS.		SSQ-DIS.	
		Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
STUDY 1	*PRE*	7.48	0 — 18.7	0	0 — 0	15.16	0 — 22.74	0	0 — 13.92
	NO-AUG	3.74	3.74 — 18.7	0	0 — 9.54	7.58	0 — 15.16	0	0 — 27.84
	LOC-AUG	3.74	0 — 14.96	0	0 — 9.54	7.58	0 — 15.16	0	0 — 13.92
	OBJ-AUG	7.48	0 — 18.7	0	0 — 9.54	7.58	0 — 15.16	0	0 — 27.84
	BOTH-AUG	7.48	0 — 14.96	0	0 — 9.54	7.58	0 — 15.16	13.92	0 — 27.84
STUDY 2	*PRE*	11.22	2.81 — 19.64	0	0 — 11.93	7.58	0 — 15.16	6.96	0 — 17.4
	NO-AUG	0	0 — 12.16	0	0 — 9.54	0	0 — 3.79	0	0 — 13.92
	LOC-AUG	0	0 — 7.48	0	0 — 2.39	0	0 — 0	0	0 — 3.48
	OBJ-AUG	1.87	0 — 14.96	0	0 — 2.39	0	0 — 9.48	0	0 — 13.92
	BOTH-AUG	0	0 — 11.22	0	0 — 9.54	0	0 — 0	0	0 — 13.92

	CONDITIONS	KIM-ENJ.		KIM-PER. COM.		KIM-PER. CHO.		PACES-8	
		Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
STUDY 1	*NO-AUG*	2.33	1.33 — 3	2.33	2 — 3	2.33	1.67 — 3	5	3.88 — 5.88
	LOC-AUG	3	2.33 — 3.33	3	2 — 3.67	3	2.33 — 3.33	5.25	4.63 — 5.88
	OBJ-AUG	2.67	1.33 — 3	2.33	2 — 3	2.67	2 — 3.33	5	4 — 5.75
	BOTH-AUG	3	2.33 — 4	3	2.33 — 3.67	3	2.67 — 3.67	5.5	4.5 — 6.13
STUDY 2	*NO-AUG*	4	3 — 4	2.5	1.67 — 3.42	3.17	2.25 — 4	6.56	5.31 — 7
	LOC-AUG	3.83	3 — 4	2.33	1.25 — 3.33	3	1.92 — 4	6.69	5.22 — 7
	OBJ-AUG	4	2.92 — 4	2.17	1.92 — 3.33	3	2 — 4	6.19	5.28 — 6.56
	BOTH-AUG	3.67	3.25 — 4	2.5	1.58 — 3.33	3	2.33 — 4	6.44	5.22 — 7

	CONDITIONS	PXI-IMM.		EXERTION		COLLECTED TARGETS		PHYSICAL ACTIVITY	
		Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
STUDY 1	*NO-AUG*	1.67	1 — 2.67	5	4 — 6	5	4 — 6	212.98	171.76 — 241.16
	LOC-AUG	2	1.33 — 2.67	5	3 — 5	10	7 — 12	186.43	157.04 — 210.62
	OBJ-AUG	2	1.33 — 2.67	5	4 — 6	6	5 — 7	199.34	170.99 — 221.89
	BOTH-AUG	2	1.33 — 2.67	5	3 — 5	11	9 — 13	182.7	153.85 — 209.7
STUDY 2	*NO-AUG*	2	1.33 — 2.67	1	1 — 3	2	1 — 3	107.96	76.71 — 129.63
	LOC-AUG	2	1 — 2.67	1	1 — 2	3	2 — 4	98.79	78.98 — 110.15
	OBJ-AUG	2.33	1.33 — 3	1	1 — 2	2	1 — 3	78.87	64.40 — 106.78
	BOTH-AUG	2.17	1.33 — 3	1	1 — 2	4	2 — 6	70.53	57.63 — 97.92

test revealed the proportions of preferred conditions were significantly different from chance, $p<0.001$. Pairwise comparisons with exact binomial tests showed that proportions of *BOTH-AUG* preference differed significantly compared to *NO-AUG* (p=0.008) and *OBJ-AUG* ($p<0.001$). *LOC-AUG* preference differed significantly compared to *OBJ-AUG* (p=0.006).

e) ***Theme 1: Augmentation preferences are primarily based on a sense of achievement and secondarily on experiential factors****:* Regarding preferences, our qualitative findings show a slightly different distribution than the single-item responses, yet overall the same tendencies: *BOTH-AUG* (n=17), *LOC-AUG* (n=11), and *NO-AUG* (n=1). Overall, participants attributed their preferences for augmentation variants to several factors. For example, one based their decision on experiencing nausea in some conditions, and therefore preferring *NO-AUG*. For the most part, however, the strongest underlying factor contributing to a variant preference was how it supported their sense of achievement. For most players, the locomotion augmentation (in *LOC-AUG* and *BOTH-AUG*) was preferred and an important feature. This primarily related to how it supported them in achieving their goal. For example, they emphasized their resulting "*real sense of achievement by moving faster because [they] got to the trash faster and could collect more*"-*P9*$_{U1}$.

No participants solely preferred the object range augmentation, yet for those that mentioned it positively in the context of the combined variant, it also related to how it aided their progress: "*I also found steering left and right a bit difficult, so I was glad for the range extension*"-*P14*$_{U1}$. However we note that augmentation aiding progress is not always a unanimously liked feature. For one participant, object range augmentation reduced the challenge, making it less favorable: "*Having to row all the way to the cans was also more fun, because you not only have to move fast, but you also have to navigate more accurately, and I like the challenge of that more than the lack of it*"-*P19*$_{U1}$.

As a second underlying cause for preference, experiential factors like enjoyment were also common—albeit less so than achievement (and in one case, *through* the sense of achievement: "*[it] allows for faster collection, so that was more fun for me as well*"-*P5*$_{U1}$). For example, the speed at which locomotion augmentation allowed participants to move brought fun to their experience: "*I could swim faster and had more fun*"-*P25*$_{U1}$. For some, interestingly, their positive experience of augmentation was because it made "*movement fe[el] the most natural*"-*P20*$_{U1}$, and "*uninhibited*"-*P21*$_{U1}$.

f) ***Theme 2: Augmentation increased motivation, but more so for locomotion augmentation than object range augmentation****:* The majority commented that locomotion augmentation positively affected their motivation in several ways. First, relating to the previous theme, players were motivated by feeling that they can or could achieve more with this

TABLE II
THIS TABLE LISTS THE RESULTS OF EMPLOYED STATISTICAL TESTS ON DEPENDENT VARIABLES OF *User Study 1* AND *User Study 2*. THE RESULTS ARE REPORTED BASED ON THE MAIN EFFECT OF LOCOMOTION (LOC.) AND OBJECT RANGE AUGMENTATION (OBJ.), AND THE INTERACTION EFFECT OF BOTH INDEPENDENT VARIABLES (INT.). THE BOLD TEXT INDICATES SIGNIFICANT RESULTS.

		KIM-ENJ.	KIM-PER. COM.	KIM-PER. CHO.	PACES-8
STUDY 1	Loc.	**F_a(1,28)=44.20, p<0.001, η_P^2 =0.612**	**F_a(1,28)=17.05, p<0.001, η_P^2 =0.378**	**F_a(1,28)=8.28, p=0.008, η_P^2 =0.228**	**F(1,28)=15.67, p<0.001, η_P^2 =0.359**
	Obj.	F_a(1,28)=1.39, p=0.248, η_P^2 =0.047	F_a(1,28)=0.05, p=0.819, η_P^2 =0.002	F_a(1,28)=3.29, p=0.081, η_P^2 =0.105	F(1,28)=0.22, p=0.646, η_P^2 =0.008
	Int.	F_a(1,28)=0.51, p=0.481, η_P^2 =0.018	F_a(1,28)=2.12, p=0.157, η_P^2 =0.070	F_a(1,28)=0.50, p=0.486, η_P^2 =0.017	F(1,28)=0.00, p=1.000, η_P^2 =0.000
STUDY 2	Loc.	F_a(1,23)=0.02, p=0.881, η_P^2 =0.001	F_a(1,23)=0.32, p=0.578, η_P^2 =0.014	F_a(1,23)=0.01, p=0.905, η_P^2 =0.001	F_a(1,23)=2.89, p=0.102, η_P^2 =0.112
	Obj.	F_a(1,23)=0.02, p=0.896, η_P^2 =0.001	F_a(1,23)=0.03, p=0.861, η_P^2 =0.001	F_a(1,23)=0.09, p=0.765, η_P^2 =0.004	F_a(1,23)=3.30, p=0.082, η_P^2 =0.125
	Int.	F_a(1,23)=0.03, p=0.860, η_P^2 =0.001	F_a(1,23)=0.57, p=0.458, η_P^2 =0.024	F_a(1,23)=0.06, p=0.811, η_P^2 =0.003	F_a(1,23)=1.56, p=0.224, η_P^2 =0.063

		PXI-IMM.	EXERTION	COLLECTED TARGETS	PHYSICAL ACTIVITY
STUDY 1	Loc.	F_a(1,28)=0.94, p=0.341, η_P^2 =0.032	**F_a(1,28)=11.37, p=0.002, η_P^2 =0.289**	**F(1,28)=140.21, p<0.001, η_P^2 =0.834**	F_a(1,28)=3.83, p=0.060, η_P^2 =0.120
	Obj.	F_a(1,28)=1.51, p=0.229, η_P^2 =0.051	F_a(1,28)=1.49, p=0.233, η_P^2 =0.050	F(1,28)=3.18, p=0.085, η_P^2 =0.102	F_a(1,28)=1.30, p=0.265, η_P^2 =0.044
	Int.	F_a(1,28)=0.76, p=0.389, η_P^2 =0.027	F_a(1,28)=0.04, p=0.840, η_P^2 =0.001	F(1,28)=0.11, p=0.742, η_P^2 =0.004	F_a(1,28)=0.14, p=0.707, η_P^2 =0.005
STUDY 2	Loc.	F_a(1,23)=0.29, p=0.597, η_P^2 =0.012	F_a(1,23)=1.43, p=0.243, η_P^2 =0.059	**F_a(1,23)=18.07, p<0.001, η_P^2 =0.440**	F(1,23)=3.61, p=0.070, η_P^2 =0.136
	Obj.	F_a(1,23)=2.81, p=0.107, η_P^2 =0.109	F_a(1,23)=2.71, p=0.113, η_P^2 =0.105	**F_a(1,23)=4.87, p=0.038, η_P^2 =0.175**	**F(1,23)=8.61, p=0.007, η_P^2 =0.272**
	Int.	F_a(1,23)=0.26, p=0.613, η_P^2 =0.011	F_a(1,23)=0.93, p=0.345, η_P^2 =0.039	F_a(1,23)=3.22, p=0.086, η_P^2 =0.123	F(1,23)=0.13, p=0.718, η_P^2 =0.006

feature, e.g., "*The motivation was increased because I felt I could do more in less time*"-*P23*$_{U1}$. Second, they expressed it as a factor that motivated them to put them more effort: "*Seeing the adjustment of my speed and power reflected in the movement gave me the impetus to collect more cans and move faster*"-*P1*$_{U1}$. Third, their experience of augmented locomotion as an enjoyable element coincided with greater motivation: "*with the adjustments it was more fun and the motivation was definitely higher*"-*P26*$_{U1}$. In contrast, some players emphasized that without locomotion augmentation, their experience was negatively impacted because "*it felt like you would never reach your destination*"-*P22*$_{U1}$.

When asked how object range augmentation affected their motivation, players also related this to how it aided their progress and achievement. Object range augmentation offered opportunities to move strategically in the game: "*[it] also motivated me, as I was able to keep the swum course straight and did not have to adjust my orientation as often*"-*P4*$_{U1}$ and "*you could take the most efficient routes and no longer had to zigzag to each can*"-*P2*$_{U1}$. Participants appreciated this feature because it helped them in their task, e.g., "*it made me collect more cans [...]*"-*P29*$_{U1}$. For some, this feature led to positive experiences (e.g., fun, excitement). However, a few players found it difficult to estimate at which distance objects could be collected, causing them to lose time: "*I found it rather hindering, because I never knew exactly whether I would reach the object with the "blue beam"*"-*P22*$_{U1}$. A small number of players pointed out that with this feature they "*didn't feel like [they] actually "earned" the individual points*"-*P21*$_{U1}$ or experienced "*no challenge*"-*P10*$_{U1}$. Further, a few did not even notice its existence: "*I honestly didn't notice much in the levels*"-*P27*$_{U1}$. Overall, most participants gave the impression that their motivation was more strongly influenced by augmentation of locomotion than of object range, e.g., "*In fact, the virtual [object] adjustment was less important to me than the movement adjustment*"-*P29*$_{U1}$.

TABLE III
THIS TABLE SHOWS THE DESCRIPTIVE VALUES OF TIME SPENT AND PHYSICAL ACTIVITY IN THE CONDITION TUTORIALS.

	CONDITIONS	TIME (IN SECONDS) *Mdn*	*IQR*	PHYSICAL ACTIVITY *Mdn*	*IQR*
STUDY 2	*NO-AUG*	42.72	25.86 — 69.64	48.01	34.05 — 71.28
	LOC-AUG	27.92	19.02 — 60.80	31.69	18.48 — 51.38
	OBJ-AUG	29.11	18.76 — 39.94	22.06	19.11 — 27.42
	BOTH-AUG	18.9	15.63 — 63.27	14.85	10.50 — 30.34
	All variants	30.34	18.74 — 61.39	27.43	16.82 — 45.73

VII. RESULTS: *User Study 2* WITH OLDER ADULTS

*a) **Interactive Tutorials**:* Unlike the younger adults, the older adults did not have a time limit to complete the interactive VR tutorials before each game variant. Although most older adults reported prior VR experience (n=18), they were not regular users. By providing tutorials, we gave the older adults enough time to become familiar with the VR system and the interaction techniques. Descriptive data on time spent and physical activity in the tutorials can be found in Table III.

*b) **SSQ: Pre- and Post-Comparison**:* We observed a significant effect of time points on total SSQ scores, $\chi^2(4)$=22.39, p<0.001, *Kendall's W*=0.233. Participants reported significantly higher total SSQ in the pre-compared to the post-*LOC-AUG* measure (V=140, p=0.027). We found a significant effect of time points on SSQ-Nausea ($\chi^2(4)$=10.69, p=0.030, *Kendall's W*=0.111), SSQ-Ocu ($\chi^2(4)$=16.26, p=0.003, *Kendall's W*=0.169), and SSQ-Dis ($\chi^2(4)$=9.96, p=0.041, *Kendall's W*=0.104); pairwise tests with Bonferroni correction did not indicate significant differences between time points.

*c) **Post-Game Questionnaires**:* Neither a main effect of the independent variables nor their interaction effect was observed on the post-game measures.

*d) **Gameplay Metrics**:* There were significant main effects of locomotion augmentation (F_a(1,23)=18.07, p<0.001, η_P^2 =0.440) and object range augmentation (F_a(1,23)=4.87, p=0.038, η_P^2 =0.175) on the number of collected targets.

Participants collected a significantly higher number of targets when they were in the *BOTH-AUG* condition in comparison to the *NO-AUG* (V=14.5, p=0.003) and the *OBJ-AUG* playthroughs (V=10, p=0.001).

We observed a significant main effect of object range augmentation on physical activity of players, F(1,23)=8.61, p=0.007, η_P^2 =0.272. They performed significantly less physical activity in the *BOTH-AUG* condition compared to the *NO-AUG* condition (t(23)=3.40, p=0.015).

*e) **Preference of Game Variants**:* Most participants preferred the *BOTH-AUG* condition (n=12). However, some preferred the other game variants: *LOC-AUG* (n=6), *NO-AUG* (n=4), *OBJ-AUG* (n=2). An exact multinomial test indicated the proportions of preferred conditions were significantly different from chance, p=0.034. However, pairwise exact binomial test comparisons with Bonferroni correction did not indicate significant differences between game variants.

*f) **Theme 1: Augmentation preferences were primarily based on a sense of achievement, but also learning effects and experiential factors**:* For the older adults, being successful played the most dominant role in choosing a preferred condition: "*because that's where you see the best success*"-*P7*$_{U2}$. This led to a feeling of achievement (e.g., "*the sense of achievement for me was the biggest thing*"-*P3*$_{U2}$).

Experiential factors like enjoyment were also mentioned as playing a secondary important role, e.g., "*it was fun*"-*P4*$_{U2}$. In a few cases, this enjoyment traced back to the sense of achievement ("*you enjoyed that [*BOTH-AUG*] more because you accomplished more there*"-*P14*$_{U2}$); in others, it related to the enjoyment of moving ("*[having] more movement in it*"-*P17*$_{U2}$, referring to the *LOC-AUG* condition).

Another secondary factor unique to this set of interviews related to the important role of learning effects and understanding the mechanics required of them to accomplish their task. For example: "*I understood how to do it [swimming and grabbing objects]. And then I was happy that I understood it. And then that was fun for me.*"-*P18*$_{U2}$.

A few participants attributed their preference of a specific gameplay variant to feeling that it was more realistic or tangible, e.g., "*I just thought it was better [...] because it was closer to reality*"-*P2*$_{U2}$. However, interestingly, answers with this reasoning occurred referencing all of the different variants.

*g) **Theme 2: Augmentation was viewed positively but connected less to motivation**:* When asked about whether or how augmentation affected their motivation, older adults responded with positive feedback. Both locomotion and object range augmentation were connected positively to motivation. For example, for locomotion augmentation: "*in the game it motivated me more*"-*P21*$_{U2}$. In particular, it motivated some players to collect more objects: "*You wanted even more cans than you achieved*"-*P10*$_{U2}$ and was seen as a motivating element "*because you also see successes more quickly*"-*P7*$_{U2}$. Occasionally, this increased motivation was also talked about in terms of feeling greater ambition: "*the ambition was awakened*"-*P10*$_{U2}$ and "*one had also the ambition to participate properly*"-*P17*$_{U2}$. For object range augmentation, this was similar: "*And there is the motivation that I can then also grab more cans, because I can reach them faster*"-*P11*$_{U2}$.

However, participants also sometimes evaded the connection to motivation and responded with feedback that did not explicitly relate to an effect on motivation: "*I don't know if it motivated me that much. I just thought it was better without the laser because it was closer to reality*"-*P2*$_{U2}$. Further, a number of players explicitly reported that augmentation did *not* impact their motivation, e.g., "*it didn't change my motivation in any way*"-*P22*$_{U2}$ and "*motivated one way or the other, right? It didn't particularly encourage me or anything*"-*P15*$_{U2}$. For one player, using the augmented technique (in this instance, object range augmentation) required more effort, thus leading them to take longer: "*[without the augmentation] I don't have to think long and hard about whether I can get it in on the right or the left*"-*P18*$_{U2}$. Finally, a few players did not notice that locomotion augmentation had occurred: "*I didn't really notice that anything had changed*"-*P16*$_{U2}$.

VIII. Discussion

Here, we discuss the importance of augmented interaction for younger vs. older adults in these VR exergame studies, and differences and similarities between these user groups' experiences. Based on this, we highlight key takeaways for HCI researchers as well as developers in the domain of VR exergames. Finally, we address study limitations.

A. Augmented Interaction for Younger Players: Positive Effect of Augmented Locomotion

The *User Study 1* findings show a positive effect of augmented interaction on younger adults' experience—one that was mainly driven by augmented locomotion: Younger players enjoyed the game experience more when locomotion was augmented virtually. This aligns with previous work [14], [15] exploring how virtual augmentation in exergames can support intrinsic motivation, and extends them by showcasing how they apply to another type of augmented movement, i.e., swimming. Perceived competence and choice also follow a similar trend: like motivation, these were increased with augmented locomotion or the combination of locomotion and object range augmentation. This together with the previous studies lends support to the supposition that augmentation of locomotion contributes to fulfillment of autonomy and competence—two of the three basic human needs defined by self-determination theory [95]. While other theories of motivation may also explain these positive effects and should be considered in future work, this theory has been dominant in games research [113] including work on augmentation in exergames. Our qualitative results also reflect on these concepts; participants explained their reasoning for preferring augmented locomotion with how it supported feelings of achievement (relating to competence), supporting enjoyment.

The participants also enjoyed physical activity more in the condition featuring both types of augmented interaction compared to no augmentation or only object range augmentation. This complements work by Lehtonen et al. [15]—in which players enjoyed physical activity in a multiplayer mixed reality trampoline game featuring augmentation—but also provides a comparison between augmented and non-augmented variants. It further shows physical activity enjoyment can be achieved through augmented interaction in a single-player VR exergame.

While we implemented *ExerSwimVR* as a seated experience to account for accessibility and safety, this did increase our concerns about potential cybersickness risks due to the mismatch between players' real and virtual movements—particularly so for the augmented locomotion. However, our findings show that players did not experience a significant increase in their SSQ scores. This matches with prior work [21] showing that locomotion can be augmented in some cases without causing cybersickness; our findings transfer these results to seated VR exergames. This paves the way to using augmented locomotion in seated VR exergame experiences, which can reduce injury risk.

The augmented interaction techniques provide advantages in our exergame tasks: it allows players to cover more virtual distance and collect objects more easily. This was unsurprisingly reflected in the task performance of our participants; they collected significantly more targets in the conditions with augmented locomotion (although not the augmented object range condition). The greater ease with which participants were able to complete their task could have led to participants then overall putting less effort into their physical activity. However, despite the help of these augmented techniques, the participants did not perform significantly less physical activity based on their logged movements. This suggests that augmented interaction can be used without significantly decreasing physical activity. However, it also did not *increase* physical activity—so virtual augmentation did not work to induce more movement. This contrasts with Born et al. [17]'s work, in which virtual augmentation led to more strenuous activity for a longer duration. Yet we note that nevertheless greater motivation could lead to better retention long-term.

In contrast to the augmented locomotion, our results for object range augmentation indicate that it did not significantly affect participants' experience. This might be because players prefer more granular object manipulation over more simplified or abstract implementations; this would be in line with previous work making such a comparison with object manipulation tasks in a more general VR application [114]. Alternatively, since most players had prior VR experience, they may not have perceived this feature as an exciting superhuman ability, but rather something familiar like a teleportation feature that they might know from other VR applications [115].

B. Augmented Interaction for Older Adults

The primary goal of this research was to understand experiences of older adults with augmented interaction. This contributes to ongoing research in the HCI field to design VR games to support older adults' physical activity [9], [57], [64], [68]. Exergames designed based on empirical evidence can hopefully motivate and support our aging world population to engage in more physical activity, which can benefit their physical and mental health [3].

Our older adult participants rated all game variants positively, indicating that overall *ExerSwimVR* was received very well. Intrinsic motivation scores of older adults remained high regardless of type of game variant—even higher than younger adults for example for interest/enjoyment. This is in line with the observations of the study conductor who noted that most participants were quite enthusiastic and happy to experience all variants. For example, even when told they could simply grab objects with a single controller, a few insisted on picking up the objects with both hands. However, in contrast to the previous study, our quantitative findings for *User Study 2* showed no significant difference between conditions; augmentation did not significantly increase the positive experience for this user group.

We look to our qualitative data to better understand why the effect of augmented interaction may differ for older adults. For both *User Study 1* and *User Study 2*, participants reported a sense of achievement as the dominant reason for preferring a certain game variant. However, unlike the younger adults in *User Study 1*, some of the older adults also talked about learning effects: pointing out that they had felt happy to have understood the required interaction. They also had less frequent VR experience. It is possible that learning the interaction took more effort from these participants, thus hindering positive effects on their experience. Interestingly, for a few, preferences for game variants were linked to the realism or tangibility of the experience, but overall there was no consensus on what was realistic (echoing general ambiguity in games research surrounding realism [69]). Given that preferences for realism crop up in exergame research with older adults [9], [27], [116], [117], this suggests that more care should be taken in understanding what older adults mean when they talk about wanting or liking realistic aspects. They may have had preconceptions about swimming that our VR approximation did not meet, yet they gave no such feedback.

Further, fewer older adults seemed to connect augmentation with their motivation. Whether this is due to a difference in how augmentation inherently affects motivation for this user group, or a different contributing factor remains to be seen. For example, another difference can be observed in the game metrics, specifically regarding physical activity. While *User Study 1* participants showed no difference in physical activity across conditions, *User Study 2* participants performed *less* physical activity when both types of augmentation were present compared to the no-augmentation condition. Overall, this could be seen as one kind of augmentation being linked to decreased physical activity for older adults—i.e., a loss of effectiveness of the exergame. This will need to be explored in the future. Nevertheless, we note that physical activity was overall much higher in *User Study 1* than *User Study 2*. Matching this, perceived exertion remained moderately high in *User Study 1*, and low across all conditions in *User Study 2*. Considering the difference in physical activity, it is not surprising that younger adults felt higher exertion compared to older adults.

Game performance differed between the two studies: while sometimes augmentation helped both groups to collect more targets, regardless of conditions, the older adults collected less targets than the younger adults. This is in line with previous works that have explored game performance across different age groups [24], [118]. However, we also observed similarities in the experience of both groups: In contrast to prior work [21], the augmented interaction did not contribute to higher immersion in either study.

The older adults reported lower total SSQ scores after experiencing the locomotion augmentation (*LOC-AUG*) compared to the initial measurement. While this is surprising at first glance, we speculate that this was due to their excitement in taking part in the study, which can potentially elicit symptoms that intersect with the SSQ, such as difficulty focusing, or concentrating. Overall, however, this suggests that seated VR exergames may be a useful and safe option for older adults without unduly increasing cybersickness.

C. Main Takeaways

We provide takeaways for consideration in the design of VR exergames with augmented interaction.

a) Consider augmentation of different interactions: Aligning with prior work, our results largely suggest that augmented locomotion elicited positive experiences. However, augmentation of the range of object manipulation had less of an impact on players' experience compared to augmentation of locomotion. Exergame research has dominantly explored augmented interaction in movement [14], [15], [23]. As movement is an essential element in exergames, this is understandable. However, there are other augmented interaction possibilities in VR exergames that could be explored. While the object range augmentation was less successful in increasing motivation or encouraging more physical activity, other alternative augmentations (e.g., [17]) may be promising. Therefore, we recommend greater variety in the augmented interaction types we explore in VR exergames and their effect on motivation.

b) Consider age group differences: Similar to Xu et al. [24]'s work on different age groups (younger vs. middle-aged adults) in VR games, our findings suggest an effect of age differences (younger vs. older adults) on the player experience: These age differences can play a role in how augmented interaction are perceived. While the augmentation was coupled with positive experiences for younger adults, this largely was not upheld for older adults. This necessitates care when considering augmentation in VR exergames: If designing one for the general population, the addition of virtual augmentation may be useful to increase motivation. However, when the user group consists only of older adults, it may not be worth the effort to implement this. In such a case, while virtual augmentation may not negatively impact experience, it seems it can reduce physical activity and thus be counterproductive to the goal of most exergames. Alternatively, if the game is multiplayer, designers could consider employing asymmetric game design (augmented vs. non-augmented) to address differences in users' experiences [110], [119].

D. Limitations

Our target group was older adults, and hence, this game was designed with their requirements in mind (e.g., seated experience, not too rapid pace). For example, we limited the experimental gameplay time to one minute, at the recommendation of the physiotherapist, thus upholding safety as a primary design goal for older adults. Yet relying on our qualitative analysis, we speculate that more exposure could have resulted in augmentation affecting older adults' motivation. Future studies may explore safe ways to enable longer exposure to virtual augmentation with older adults. Though the design targeted older adults, we also evaluated the game with younger adults. The requirements that shaped the exergame's design to be suitable for older adults could have affected how well it suited the younger adults. Yet this effect appears small as the experience was rated highly by the younger adults; motivation in *User Study 1* matched results of previous studies' effects of virtual augmentation.

The studies had some differences that could affect comparability. We provided an interactive tutorial with no time limit before each game variant for older adults, as prior work has emphasized that exergames for older adults should give them time to become acquainted with interactive systems [90]. While the older adults did not require much more time to collect one object than the younger ones in *User Study 2*, it is still possible that they would have needed more time to become comfortable with the augmentation techniques, and with VR itself. Because of COVID-19 regulations, *User Study 1* was conducted remotely, while *User Study 2* was performed in-person. This difference could have affected the comparability between studies, although research suggests results should nevertheless be at least similar in task performance [120]. Additionally, the methods used to gain qualitative feedback from participants differed (open-ended survey questions in *User Study 1* and semi-structured interviews in *User Study 2*). This may have affected the level of response detail—although, unexpectedly, the survey responses were often more detailed than the interview responses.

In terms of game design, we implemented the object range augmentation in a way that players were able to grab an object from up to three meters away. However, if players wanted, they could still grab the objects only when at a closer distance. We did this to avoid frustration and not break the flow of the game by requiring players to precisely estimate the distance to the object. Nevertheless, this could have led some participants to not engage in this feature, although we only observed this in a few remote participants' logs.

Regarding realism, we again note that our game's "non-augmented" version only offers an approximation of a swimming experience (prioritizing the safety of older adult players via the seated design, but integrating more complex physical aspects of swimming—like the angle of the hands and the density of the water to reflect the game's tropical underwater scene [87]—where possible).

In line with prior work on augmented interaction [17], [21], our study design compares augmented interaction variants with enhanced user abilities to a non-augmented version that does not enhance user abilities. As technology progresses and becomes more affordable, we hope that future projects will compare additional increasingly "realistic" non-augmented variants (e.g., recumbent swimming position, or an underwater HMD), even though this may be difficult with older adults.

Finally, the qualitative analysis was shaped by the researchers' involvement [109], [121] and our methodological choices: The lead author has experience designing VR games for different user groups and playing augmentation-based VR games. Similarly, the last author, who was involved in the discussion of themes in the final step, has explored realistic and augmented interaction in VR games. With these backgrounds, the analysis may have focused more on the potential connections with dimensions of realism and virtual augmentation than other perspectives, e.g., in-depth theories of motivation.

IX. Conclusion

In this paper, we implemented an accessible swimming VR exergame, *ExerSwimVR*, for older adults. In two user studies, we investigated how different types of augmented interaction—(i) augmented locomotion and (ii) augmented range of object manipulation—impact motivation, immersion, physical activity enjoyment, performance, and cybersickness of players. For younger adults, we found that primarily the augmented locomotion elicits significantly higher intrinsic motivation, perceived competence and choice, and physical activity enjoyment without causing cybersickness or less physical activity. However, for older adults, augmentation (of either sort) mostly did not induce significant differences, and in fact augmentation in one case *decreased* physical exertion. These findings suggest a need for more nuanced exploration of augmentation techniques, and to carefully consider age group differences when designing exergames as takeaways for VR researchers and designers.

Acknowledgments

We would like to thank our participants and Hospital zum Heiligen Geist Hamburg for supporting our research. This work was supported by the European Union, the German Federal Ministry of Education and Research (BMBF), and the German Research Foundation (DFG). Additionally, we acknowledge financial support from the Open Access Publication Fund of Universität Hamburg.

References

[1] D. Burrows, "Sedentary pandemic threatens eu health," October 2017. [Online]. Available: https://euobserver.com/health-and-society/139651

[2] J. H. Park, J. H. Moon, H. J. Kim, M. H. Kong, and Y. H. Oh, "Sedentary lifestyle: overview of updated evidence of potential health risks," *Korean journal of family medicine*, vol. 41, no. 6, pp. 365–373, 2020.

[3] World Health Organization, "Physical activity," November 2020. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/physical-activity

[4] A. Moschny, P. Platen, R. Klaaßen-Mielke, U. Trampisch, and T. Hinrichs, "Barriers to physical activity in older adults in germany: a cross-sectional study," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 8, no. 1, pp. 1–10, 2011. [Online]. Available: https://doi.org/10.1186/1479-5868-8-121

[5] R. R. Pate, R. P. Saunders, J. R. O'Neill, and M. Dowda, "Overcoming barriers to physical activity: helping youth be more active," *ACSM's Health & Fitness Journal*, vol. 15, no. 1, pp. 7–12, 2011.

[6] F. F. Mueller, D. Edge, F. Vetere, M. R. Gibbs, S. Agamanolis, B. Bongers, and J. G. Sheridan, "Designing sports: A framework for exertion games," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 2651–2660. [Online]. Available: https://doi.org/10.1145/1978942.1979330

[7] A. N. I. Samuli Laato and T. H. Laine, "Playing location-based games is associated with psychological well-being: an empirical study of pokémon go players," *Behaviour & Information Technology*, vol. 41, no. 9, pp. 1888–1904, 2022.

[8] B. Kordyaka, S. Laato, S. Weber, J. Hamari, and B. Niehaves, "Exploring the association between engagement with location-based game features and getting inspired about environmental issues and nature," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, ser. CHI '24. New York, NY, USA: Association for Computing Machinery, 2024. [Online]. Available: https://doi.org/10.1145/3613904.3642786

[9] S. Karaosmanoglu, S. Rings, L. Kruse, C. Stein, and F. Steinicke, "Lessons learned from a human-centered design of an immersive exergame for people with dementia," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, no. CHI PLAY, oct 2021. [Online]. Available: https://doi.org/10.1145/3474679

[10] L. Tabbaa, C. S. Ang, V. Rose, P. Siriaraya, I. Stewart, K. G. Jenkins, and M. Matsangidou, "Bring the outside in: Providing accessible experiences through vr for people with dementia in locked psychiatric hospitals," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–15. [Online]. Available: https://doi.org/10.1145/3290605.3300466

[11] S. G. Izard, J. A. Juanes, F. J. García Peñalvo, J. M. Estella, M. Ledesma, and P. Ruisoto, "Virtual reality as an educational and training tool for medicine," *Journal of Medical Systems*, vol. 42, no. 3, pp. 1–5, 2018.

[12] J. Koulouris, Z. Jeffery, J. Best, E. O'Neill, and C. Lutteroth, "Me vs. super(wo)man: Effects of customization and identification in a vr exergame," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–17. [Online]. Available: https://doi.org/10.1145/3313831.3376661

[13] R. J. Jacob, A. Girouard, L. M. Hirshfield, M. S. Horn, O. Shaer, E. T. Solovey, and J. Zigelbaum, "Reality-based interaction: A framework for post-wimp interfaces," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: Association for Computing Machinery, 2008, p. 201–210. [Online]. Available: https://doi.org/10.1145/1357054.1357089

[14] C. Ioannou, P. Archard, E. O'Neill, and C. Lutteroth, "Virtual performance augmentation in an immersive jump & run exergame," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–15. [Online]. Available: https://doi.org/10.1145/3290605.3300388

[15] L. Lehtonen, M. D. Kaos, R. Kajastila, L. Holsti, J. Karsisto, S. Pekkola, J. Vähämäki, L. Vapaakallio, and P. Hämäläinen, "Movement empowerment in a multiplayer mixed-reality trampoline game," in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, ser. CHI PLAY '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 19–29. [Online]. Available: https://doi.org/10.1145/3311350.3347181

[16] B. Dewitz, S. Karaosmanoglu, R. W. Lindeman, and F. Steinicke, "Magic, superpowers, or empowerment? a conceptual framework for magic interaction techniques," in *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2023, pp. 807–808.

[17] F. Born, A. Rygula, and M. Masuch, "Motivating players to perform an optional strenuous activity in a virtual reality exergame using virtual performance augmentation," *Proc. ACM Hum.-Comput. Interact.*, vol. 5, no. CHI PLAY, oct 2021. [Online]. Available: https://doi.org/10.1145/3474652

[18] P. Hämäläinen, J. Marshall, R. Kajastila, R. Byrne, and F. F. Mueller, "Utilizing gravity in movement-based games and play," in *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, ser. CHI PLAY '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 67–77. [Online]. Available: https://doi.org/10.1145/2793107.2793110

[19] A. Granqvist, T. Takala, J. Takatalo, and P. Hämäläinen, "Exaggeration of avatar flexibility in virtual reality," in *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, ser. CHI PLAY '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 201–209. [Online]. Available: https://doi.org/10.1145/3242671.3242694

[20] M. Nabioyuni and D. A. Bowman, "An evaluation of the effects of hyper-natural components of interaction fidelity on locomotion performance in virtual reality," in *Proceedings of the 25th Inter-*

national Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments, ser. ICAT - EGVE '15. Goslar, DEU: Eurographics Association, 2015, p. 167–174.

[21] D. Wolf, K. Rogers, C. Kunder, and E. Rukzio, *JumpVR: Jump-Based Locomotion Augmentation for Virtual Reality*. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–12. [Online]. Available: https://doi.org/10.1145/3313831.3376243

[22] J. Yoon, S. Lee, and T. Park, "Jediflight: Design and evaluation of wing-based flying experience in virtual reality," in *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, ser. CHI PLAY '18 Extended Abstracts. New York, NY, USA: Association for Computing Machinery, 2018, p. 309–320. [Online]. Available: https://doi.org/10.1145/3270316.3273043

[23] R. Kajastila, L. Holsti, and P. Hämäläinen, "Empowering the exercise: a body-controlled trampoline training game," *International Journal of Computer Science in Sport*, vol. 1, no. 13, 2014. [Online]. Available: https://users.aalto.fi/~hamalap5/publications/IJCSS_2014_personal_copy_Kajastila.pdf

[24] W. Xu, H.-N. Liang, K. Yu, and N. Baghaei, "Effect of gameplay uncertainty, display type, and age on virtual reality exergames," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445801

[25] N. Skjæret, A. Nawaz, T. Morat, D. Schoene, J. L. Helbostad, and B. Vereijken, "Exercise and rehabilitation delivered through exergames in older adults: An integrative review of technologies, safety and efficacy," *International Journal of Medical Informatics*, vol. 85, no. 1, pp. 1–16, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1386505615300514

[26] K. M. Gerling, J. Schild, and M. Masuch, "Exergame design for elderly users: The case study of silverbalance," in *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology*, ser. ACE '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 66–69. [Online]. Available: https://doi.org/10.1145/1971630.1971650

[27] A. Barenbrock, M. Herrlich, and R. Malaka, "Design lessons from mainstream motion-based games for exergames for older adults," in *2014 IEEE Games Media Entertainment*, 2014, pp. 1–8.

[28] United Nations, Department of Economic and Social Affairs, Population Division, "World population ageing 2019 (st/esa/ser.a/444)," October 2020. [Online]. Available: https://www.un-ilibrary.org/content/books/9789210045544

[29] J. P. Ziegelmann, S. Lippke, and R. Schwarzer, "Adoption and maintenance of physical activity: Planning interventions in young, middle-aged, and older adults," *Psychology & Health*, vol. 21, no. 2, pp. 145–163, 2006, pMID: 21985115. [Online]. Available: https://doi.org/10.1080/14768320500018891

[30] W. Zhao, S. Baker, and J. Waycott, "Challenges of deploying vr in aged care: A two-phase exploration study," in *32nd Australian Conference on Human-Computer Interaction*, ser. OzCHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 87–98. [Online]. Available: https://doi.org/10.1145/3441000.3441018

[31] J. W. Creswell and V. L. P. Clark, *Designing and conducting mixed methods research*. Thousand Oaks, California, United States: Sage publications, 2017.

[32] H. Tanaka, "Swimming exercise," *Sports Medicine*, vol. 39, no. 5, pp. 377–387, 2009.

[33] M. Yfanti, A. Samara, P. Kazantzidis, A. Hasiotou, and S. Alexiou, "Swimming as physical activity and recreation for women," *Tims. Acta: naučni časopis za sport, turizam i velnes*, vol. 8, no. 2, 2014.

[34] I. Cumming, "The health & wellbeing benefits of swimming," 2017. [Online]. Available: https://www.swimming.org/swimengland/health-and-wellbeing-benefits-of-swimming

[35] E. Britton, G. Kindermann, C. Domegan, and C. Carlin, "Blue care: a systematic review of blue space interventions for health and wellbeing," *Health Promotion International*, vol. 35, no. 1, pp. 50–69, 12 2018. [Online]. Available: https://doi.org/10.1093/heapro/day103

[36] F. Soyka, M. Leyrer, J. Smallwood, C. Ferguson, B. E. Riecke, and B. J. Mohler, "Enhancing stress management techniques using virtual reality," in *Proceedings of the ACM Symposium on Applied Perception*, ser. SAP '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 85–88. [Online]. Available: https://doi.org/10.1145/2931002.2931017

[37] Blitz Games, 505 Games, "*Michael Phelps: Push the Limit*," Game [Kinect, Xbox 360], October 2011, blitz Games, 505 Games. https://www.blitzgames.com/games/mpptl.

[38] Unknown Worlds Entertainment, "*Subnautica*," Game [VR], San Francisco, CA, USA., January 2018, unknown Worlds Entertainment. San Francisco, CA, USA. https://unknownworlds.com/subnautica/. https://store.steampowered.com/app/264710/Subnautica/.

[39] Llyr ap Cenydd, "*Ocean Rift*," Game [VR], May 2019, picselica Ltd. https://www.oculus.com/experiences/quest/2134272053250863/.

[40] D. Hatsushika, K. Nagata, and Y. Hashimoto, "Scuba vr: Submersible-type virtual underwater experience system," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 962–963.

[41] S. Yamashita, X. Zhang, and J. Rekimoto, "Aquacave: Augmented swimming environment with immersive surround-screen virtual reality," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, ser. UIST '16 Adjunct. New York, NY, USA: Association for Computing Machinery, 2016, p. 183–184. [Online]. Available: https://doi.org/10.1145/2984751.2984760

[42] D. Jain, M. Sra, J. Guo, R. Marques, R. Wu, J. Chiu, and C. Schmandt, "Immersive terrestrial scuba diving using virtual reality," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 1563–1569. [Online]. Available: https://doi.org/10.1145/2851581.2892503

[43] S. Fels, Y. Kinoshita, T. pei Grace Chen, Y. Takama, S. Yohanan, A. Gadd, S. Takahashi, and K. Funahashi, "Swimming across the pacific: a vr swimming interface," *IEEE Computer Graphics and Applications*, vol. 25, no. 1, pp. 24–31, 2005.

[44] J. P. Freiwald, O. Ariza, O. Janeh, and F. Steinicke, "Walking by cycling: A novel in-place locomotion user interface for seated virtual reality experiences," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–12. [Online]. Available: https://doi.org/10.1145/3313831.3376574

[45] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe, "Estimation of detection thresholds for redirected walking techniques," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 1, pp. 17–27, 2010.

[46] N. C. Nilsson, T. Peck, G. Bruder, E. Hodgson, S. Serafin, M. Whitton, F. Steinicke, and E. S. Rosenberg, "15 years of research on redirected walking in immersive virtual environments," *IEEE Computer Graphics and Applications*, vol. 38, no. 2, pp. 44–56, 2018.

[47] M. Slater, M. Usoh, and A. Steed, "Taking steps: the influence of a walking technique on presence in virtual reality," *ACM Trans. Comput.-Hum. Interact.*, vol. 2, no. 3, p. 201–219, sep 1995. [Online]. Available: https://doi.org/10.1145/210079.210084

[48] S. Sadeghian and M. Hassenzahl, "From limitations to "superpowers": A design approach to better focus on the possibilities of virtual reality to augment human capabilities," in *Designing Interactive Systems Conference 2021*, ser. DIS '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 180–189. [Online]. Available: https://doi.org/10.1145/3461778.3462111

[49] S. Cmentowski, A. Krekhov, and J. Krüger, "Outstanding: A multi-perspective travel approach for virtual reality games," in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, ser. CHI PLAY '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 287–299. [Online]. Available: https://doi.org/10.1145/3311350.3347183

[50] M. R. Mine, "Virtual environment interaction techniques," *UNC Chapel Hill CS Dept*, 1995.

[51] D. A. Bowman and L. F. Hodges, "An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments," in *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, ser. I3D '97. New York, NY, USA: Association for Computing Machinery, 1997, p. 35–ff. [Online]. Available: https://doi.org/10.1145/253284.253301

[52] I. Poupyrev and T. Ichikawa, "Manipulating objects in virtual worlds: Categorization and empirical evaluation of interaction techniques," *Journal of Visual Languages & Computing*, vol. 10, no. 1, pp. 19–35, 1999. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1045926X98901124

[53] J. McIntosh, H. D. Zajac, A. N. Stefan, J. Bergström, and K. Hornbæk, "Iteratively adapting avatars using task-integrated optimisation," in *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '20. New York, NY, USA:

Association for Computing Machinery, 2020, p. 709–721. [Online]. Available: https://doi.org/10.1145/3379337.3415832

[54] K. P. Herndon, A. van Dam, and M. Gleicher, "The challenges of 3d interaction: A chi '94 workshop," *SIGCHI Bull.*, vol. 26, no. 4, p. 36–43, oct 1994. [Online]. Available: https://doi.org/10.1145/191642.191652

[55] F. Argelaguet and C. Andujar, "A survey of 3d object selection techniques for virtual environments," *Computers & Graphics*, vol. 37, no. 3, pp. 121–136, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0097849312001793

[56] S. Karaosmanoglu, S. Cmentowski, L. E. Nacke, and F. Steinicke, "Born to run, programmed to play: Mapping the extended reality exergames landscape," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, ser. CHI '24. New York, NY, USA: Association for Computing Machinery, 2024. [Online]. Available: https://doi.org/10.1145/3613904.3642124

[57] S. Karaosmanoglu, L. Kruse, S. Rings, and F. Steinicke, "Canoe vr: An immersive exergame to support cognitive and physical exercises of older adults," in *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3491101.3519736

[58] F. Born, S. Abramowski, and M. Masuch, "Exergaming in vr: The impact of immersive embodiment on motivation, performance, and perceived exertion," in *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, 2019, pp. 1–8.

[59] S. Cmentowski, S. Karaosmanoglu, L. E. Nacke, F. Steinicke, and J. H. Krüger, "Never skip leg day again: Training the lower body with vertical jumps in a virtual reality exergame," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3544548.3580973

[60] T. Hirzle, M. Cordts, E. Rukzio, J. Gugenheimer, and A. Bulling, "A critical assessment of the use of ssq as a measure of general discomfort in vr head-mounted displays," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445361

[61] V. V. Abeele, B. Schraepen, H. Huygelier, C. Gillebert, K. Gerling, and R. Van Ee, "Immersive virtual reality for older adults: Empirically grounded design guidelines," *ACM Trans. Access. Comput.*, vol. 14, no. 3, aug 2021. [Online]. Available: https://doi.org/10.1145/3470743

[62] P. J. Teixeira, E. V. Carraça, D. Markland, M. N. Silva, and R. M. Ryan, "Exercise, physical activity, and self-determination theory: a systematic review," *International journal of behavioral nutrition and physical activity*, vol. 9, no. 1, pp. 1–30, 2012.

[63] D. Rosenberg, C. A. Depp, I. V. Vahia, J. Reichstadt, B. W. Palmer, J. Kerr, G. Norman, and D. V. Jeste, "Exergames for subsyndromal depression in older adults: a pilot study of a novel intervention," *The American Journal of Geriatric Psychiatry*, vol. 18, no. 3, pp. 221–226, 2010.

[64] W. Xu, H.-N. Liang, K. Yu, S. Wen, N. Baghaei, and H. Tu, "Acceptance of virtual reality exergames among chinese older adults," *International Journal of Human–Computer Interaction*, vol. 0, no. 0, pp. 1–15, 2022. [Online]. Available: https://doi.org/10.1080/10447318.2022.2098559

[65] E. R. Høeg, T. M. Povlsen, J. R. Bruun-Pedersen, B. Lange, N. C. Nilsson, K. B. Haugaard, S. M. Faber, S. W. Hansen, C. K. Kimby, and S. Serafin, "System immersion in virtual reality-based rehabilitation of motor function in older adults: A systematic review and meta-analysis," *Frontiers in Virtual Reality*, vol. 2, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frvir.2021.647993

[66] P. Campo-Prieto, J. M. Cancela, and G. Rodríguez-Fuentes, "Immersive virtual reality as physical therapy in older adults: present or future (systematic review)," *Virtual Reality*, vol. 25, no. 3, pp. 801–817, 2021.

[67] B. Games, "*Beat Saber*," Game [VR], Los Angeles, CA, USA, May 2019, beat Games. Los Angeles, CA, USA. https://beatsaber.com/. Steam Reviews: https://store.steampowered.com/app/620980/Beat_Saber/.

[68] L. Kruse, S. Karaosmanoglu, S. Rings, and F. Steinicke, "Evaluating difficulty adjustments in a vr exergame for younger and older adults: Transferabilities and differences," in *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*, ser. SUI '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3565970.3567684

[69] K. Rogers, S. Karaosmanoglu, M. Altmeyer, A. Suarez, and L. E. Nacke, "Much realistic, such wow! a systematic literature review of realism in digital games," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3491102.3501875

[70] M. Young, "How to do breaststroke arms," 2022. [Online]. Available: https://www.swim-teach.com/how-to-do-breaststroke-arms.html

[71] U. Technologies, "Unity real-time development platform," Brighton, United Kingdom, UK, February 2021, lTS Release 2019.4.21f1. 25 February 2021. https://unity3d.com/unity/qa/lts-releases?version=2019.4&page=2. [Online]. Available: https://unity.com/

[72] ——, "Oculus xr plugin," 2022, https://assetstore.unity.com/packages/tools/integration/oculus-integration-82022.

[73] Oculus, "Oculus integration," 2022, https://docs.unity3d.com/Packages/com.unity.xr.oculus@1.11/manual/index.html.

[74] U. Technologies, "Xr interaction toolkit," 2022, https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@0.10/manual/index.html.

[75] Meta Platforms, Inc, "Meta quest 2," October 2020, released: 13 October 2020. [Online]. Available: https://www.oculus.com/quest-2/

[76] Oculus, "Oculus quest 1," May 2019, released: 21 May 2019. [Online]. Available: https://www.oculus.com/quest-2/

[77] J. Muñoz, S. Mehrabi, Y. Li, A. Basharat, L. E. Middleton, S. Cao, M. Barnett-Cowan, J. Boger *et al.*, "Immersive virtual reality exergames for persons living with dementia: User-centered design study as a multistakeholder team during the covid-19 pandemic," *JMIR Serious Games*, vol. 10, no. 1, p. e29987, 2022.

[78] A. Elor, M. Powell, E. Mahmoodi, N. Hawthorne, M. Teodorescu, and S. Kurniawan, "On shooting stars: Comparing cave and hmd immersive virtual reality exergaming for adults with mixed ability," *ACM Trans. Comput. Healthcare*, vol. 1, no. 4, sep 2020. [Online]. Available: https://doi.org/10.1145/3396249

[79] A. Elor, S. Lessard, M. Teodorescu, and S. Kurniawan, "Project butterfly: Synergizing immersive virtual reality with actuated soft exosuit for upper-extremity rehabilitation," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019, pp. 1448–1456.

[80] Y. Shiban, I. Schelhorn, P. Pauli, and A. Mühlberger, "Effect of combined multiple contexts and multiple stimuli exposure in spider phobia: A randomized clinical trial in virtual reality," *Behaviour Research and Therapy*, vol. 71, pp. 45–53, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0005796715000972

[81] M. F. Gago, D. Yelshyna, E. Bicho, H. D. Silva, L. Rocha, M. Lurdes Rodrigues, and N. Sousa, "Compensatory Postural Adjustments in an Oculus Virtual Reality Environment and the Risk of Falling in Alzheimer's Disease," *Dementia and Geriatric Cognitive Disorders Extra*, vol. 6, no. 2, pp. 252–267, 06 2016. [Online]. Available: https://doi.org/10.1159/000447124

[82] M. Washburn, A. Hagedorn, and S. Moore, "Creating virtual reality based interventions for older adults impacted by substance misuse: Safety and design considerations," *Journal of Technology in Human Services*, vol. 39, no. 3, pp. 275–294, 2021.

[83] H. Toussaint, P. Hollander, C. Berg, and A. Vorontsov, "Biomechanics of swimming," *Exercise and sport science*, pp. 639–660, 2000.

[84] lumenlearning.com, "Drag forces," 2022, https://courses.lumenlearning.com/physics/chapter/5-2-drag-forces/.

[85] P. P. Urone and R. Hinrichs, *College Physics*, 2012, https://openstax.org/details/books/college-physics.

[86] M. H. Koo and A. S. M. Al-Obaidi, "Calculation of aerodynamic drag of human being in various positions," *Department of Mechanical Engineering, Taylor's University, Malaysia*, 2013.

[87] J. B. Patterson and E. C. Morris, "Measurement of absolute water density, 1 °c to 40 °c," *Metrologia*, vol. 31, no. 4, pp. 277–288, jan 1994. [Online]. Available: https://doi.org/10.1088/0026-1394/31/4/001

[88] S. B. (Destatis), *Statistisches Jahrbuch: Deutschland und Internationales*. Statistisches Bundesamt (Destatis), 2019. [Online]. Available: https://www.destatis.de/DE/Themen/Querschnitt/Jahrbuch/statistisches-jahrbuch-2019-dl.pdf?__blob=publicationFile

[89] www.cansonstudio.com, "Drawing: Human proportions," 2022, https://www.cansonstudio.com/drawing-human-proportions.

[90] A. Velazquez, A. I. Martinez-Garcia, J. Favela, A. Hernandez, and S. F. Ochoa, "Design of exergames with the collaborative participation of older adults." IEEE, June 2013. [Online]. Available: https://doi.org/10.1109/cscwd.2013.6581016

[91] P. Kourtesis, R. Amir, J. Linnell, F. Argelaguet, and S. E. MacPherson, "Cybersickness, cognition, & motor skills: The effects of music,

gender, and gaming experience," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 5, pp. 2326–2336, 2023.

[92] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The International Journal of Aviation Psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[93] A. Hösch, "Simulator sickness in fahrsimulationsumgebungen-drei studien zu human factors," Ph.D. dissertation, Technische Universität Ilmenau, 2018. [Online]. Available: https://www.db-thueringen.de/receive/dbt_mods_00037865

[94] M. Wilde, K. Bätz, A. Kovaleva, and D. Urhahne, "Überprüfung einer kurzskala intrinsischer motivation (kim)," *Zeitschrift für Didaktik der Naturwissenschaften*, vol. 15, pp. 31–45, 2009. [Online]. Available: https://archiv.ipn.uni-kiel.de/zfdn/pdf/15_Wilde.pdf

[95] R. M. Ryan, C. S. Rigby, and A. Przybylski, "The motivational pull of video games: A self-determination theory approach," *Motivation and emotion*, vol. 30, no. 4, pp. 344–360, 2006, the IMI items are available online under this link: http://selfdeterminationtheory.org/intrinsic-motivation-inventory/. [Online]. Available: https://doi.org/10.1007/s11031-006-9051-8

[96] V. V. Abeele, K. Spiel, L. Nacke, D. Johnson, and K. Gerling, "Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences," *International Journal of Human-Computer Studies*, vol. 135, p. 102370, 2020.

[97] L. Graf, M. Altmeyer, K. Emmerich, M. Herrlich, A. Krekhov, and K. Spiel, "Development and validation of a german version of the player experience inventory (pxi)," ser. MuC '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3543758.3543763

[98] S. P. Mullen, E. A. Olson, S. M. Phillips, A. N. Szabo, T. R. Wójcicki, E. L. Mailey, N. P. Gothe, J. T. Fanning, A. F. Kramer, and E. McAuley, "Measuring enjoyment of physical activity in older adults: invariance of the physical activity enjoyment scale (paces) across groups and time," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 8, no. 1, pp. 1–9, 2011.

[99] J. V. Bradley, "Complete counterbalancing of immediate sequential effects in a latin square design," *Journal of the American Statistical Association*, vol. 53, no. 282, pp. 525–528, 1958. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/01621459.1958.10501456

[100] CleanBox Technology, "*CleanBox*," Nashville, TN USA, 2022, cleanBox Technology. Nashville, TN, USA. https://cleanboxtech.com/.

[101] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins, "The aligned rank transform for nonparametric factorial analyses using only anova procedures," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 143–146. [Online]. Available: https://doi.org/10.1145/1978942.1978963

[102] M. Kay, L. A. Elkin, J. J. Higgins, and J. O. Wobbrock, "Package 'artool'," pp. 1–13, 2021, package 'ARTool' v.0.11.1. [Online]. Available: https://cran.r-project.org/web/packages/ARTool/ARTool.pdf

[103] M. Templ, A. Kowarik, A. Alfons, , G. de Cillia, B. Prantner, and W. Rannetbauer, "Package 'vim'," 2021, package 'ARTool' v.6.2.2. [Online]. Available: https://cran.r-project.org/web/packages/VIM/VIM.pdf

[104] A. Kowarik and M. Templ, "Imputation with the r package vim," *Journal of Statistical Software*, vol. 74, no. 7, p. 1–16, 2016. [Online]. Available: https://www.jstatsoft.org/index.php/jss/article/view/v074i07

[105] F. Habibzadeh, "Common statistical mistakes in manuscripts submitted to biomedical journals," *European Science Editing*, vol. 39, no. 4, pp. 92–94, 2013.

[106] Dovetail Research PTY. LTD., "Dovetail," Surry Hills, Australia, 2016, surry Hills, Australia. https://dovetailapp.com/.

[107] DeepL SE, "Deepl translator," Cologne, Germany, 2017, deepL Pro version. Cologne, Germany. https://www.deepl.com/pro?cta=header-pro/.

[108] V. Braun and V. Clarke, "Reflecting on reflexive thematic analysis," *Qualitative Research in Sport, Exercise and Health*, vol. 11, no. 4, pp. 589–597, 2019. [Online]. Available: https://doi.org/10.1080/2159676X.2019.1628806

[109] ——, "Can i use ta? should i use ta? should i not use ta? comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches," *Counselling and Psychotherapy Research*, vol. 21, no. 1, pp. 37–47, 2021. [Online]. Available: https://doi.org/10.1002/capr.12360

[110] S. Karaosmanoglu, K. Rogers, D. Wolf, E. Rukzio, F. Steinicke, and L. E. Nacke, "Feels like team spirit: Biometric and strategic interdependence in asymmetric multiplayer vr games," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445492

[111] R. P. Gauthier, M. J. Costello, and J. R. Wallace, ""i will not drink with you today": A topic-guided thematic analysis of addiction recovery on reddit," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3491102.3502076

[112] V. Clarke and V. Braun, "Thematic analysis: A practical guide," *Thematic Analysis*, pp. 1–100, 2021.

[113] A. Tyack and P. Wyeth, "Exploring relatedness in single-player video game play," in *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, ser. OZCHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 422–427. [Online]. Available: https://doi.org/10.1145/3152771.3156149

[114] K. Rogers, J. Funke, J. Frommel, S. Stamm, and M. Weber, "Exploring interaction fidelity in virtual reality: Object manipulation and whole-body movements," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–14. [Online]. Available: https://doi.org/10.1145/3290605.3300644

[115] F. Buttussi and L. Chittaro, "Locomotion in place in virtual reality: A comparative evaluation of joystick, teleport, and leaning," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 1, pp. 125–136, 2021.

[116] X. Li, T. Han, E. Zhang, W. Shao, L. Li, and C. Wu, "Memoride: An exergame combined with working memory training to motivate elderly with mild cognitive impairment to actively participate in rehabilitation," in *Human Aspects of IT for the Aged Population. Supporting Everyday Life Activities*, Q. Gao and J. Zhou, Eds. Cham: Springer International Publishing, 2021, pp. 90–105.

[117] N. Ringgenberg, S. Mildner, M. Hapig, S. Hermann, K. Kruszewski, A. L. Martin-Niedecken, K. Rogers, A. Schättin, F. Behrendt, S. Böckler, S. Schmidlin, R. Jurt, S. Niedecken, C. Brenneis, L. H. Bonati, C. Schuster-Amft, and B. Seebacher, "ExerG: adapting an exergame training solution to the needs of older adults using focus group and expert interviews," *Journal of NeuroEngineering and Rehabilitation*, vol. 19, no. 1, Aug. 2022. [Online]. Available: https://doi.org/10.1186/s12984-022-01063-x

[118] K. M. Gerling, K. K. Dergousoff, and R. L. Mandryk, "Is movement better? comparing sedentary and motion-based game controls for older adults," 2013, p. 133 – 140. [Online]. Available: http://eprints.lincoln.ac.uk/id/eprint/13640/

[119] K. Rogers, S. Karaosmanoglu, D. Wolf, F. Steinicke, and L. E. Nacke, "A best-fit framework and systematic review of asymmetric gameplay in multiplayer virtual reality games," *Frontiers in Virtual Reality*, vol. 2, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frvir.2021.694660

[120] O. Alhadreti, "A comparison of synchronous and asynchronous remote usability testing methods," *International Journal of Human–Computer Interaction*, vol. 38, no. 3, pp. 289–297, Jun. 2021. [Online]. Available: https://doi.org/10.1080/10447318.2021.1938391

[121] B. J. Newton, Z. Rothlingova, R. Gutteridge, K. LeMarchand, and J. H. Raphael, "No room for reflexivity? critical reflections following a systematic review of qualitative research," *Journal of Health Psychology*, vol. 17, no. 6, pp. 866–885, 2012. [Online]. Available: https://doi.org/10.1177/1359105311427615

Evaluating Augmented Locomotion and Range of Reachable Objects for Older Adults in a Virtual Reality Exergame

Sukran Karaosmanoglu, Sebastian Finnern, Frank Steinicke, Katja Rogers

This document consists of supplementary materials of the paper "*Evaluating Augmented Locomotion and Range of Reachable Objects for Older Adults in a Virtual Reality Exergame*".

I. Sample Characteristics

We report further sample characteristics of our user studies.

A. User Study 1: Gaming & VR Habits

Participants' frequency of playing games was varied: for traditional games (e.g., board games, card games): daily (n=1), at least once a week (n=4), at least once a month (n=15), at least once a year (n=6), less frequent (n=1), and never (n=2); for digital games (e.g, mobile games, PC games, VR games): daily (n=7), at least once a week (n=10), at least once a month (n=7), at least once a year (n=3), less frequent (n=1) and never (n=1). Twenty-seven players reported prior VR experience, however, 14 of them used VR at least a once year at most. The remaining reported the following frequency: daily (n=2), at least once a week (n=10), at least once a month (n=1).

B. User Study 1: Physical Exercise Habits

The participants rated their enjoyment of doing physical exercises quite positively (M=5.21, SD=1.40 on a scale from 1/*strongly disagree* to 7/*strongly agree*). Twenty-seven participants reported engaging in weekly physical exercises: 1-2 (n=11), 3-4 (n=12), and 5-7 times (n=4).

C. User Study 2: Gaming & VR Habits

Participants' frequency of playing games was varied: for traditional games: at least once a week (n=7), at least once a month (n=7), at least once a year (n=2), less frequent (n=4), and never (n=4); for digital games: daily (n=6), at least once a week (n=6), at least once a month (n=3), less frequent (n=3), and never (n=6). Eighteen participants had prior VR experience (n=13 using at least once a year and n=5 less frequent), yet almost all indicated that this was due to participation in previous user studies.

Sukran Karaosmanoglu, Sebastian Finnern, and Frank Steinicke are with Human-Computer Interaction group, Universiät Hamburg, Germany. Katja Rogers is with Digital Interactions Lab, University of Amsterdam, the Netherlands.

Corresponding author e-mail: sukran.karaosmanoglu@uni-hamburg.de
Manuscript submitted July 16, 2024.

D. User Study 2: Physical Exercise Habits

The participants rated their enjoyment for doing physical exercises quite positively (M=6, SD=0.93). Almost all participated in weekly physical activities: 1-2 times (n=7), 3-4 times (n=12), 5-7 times (n=4); we note that one data point was missing for this item..

II. Questionnaires

We provide the items of the questionnaires used in our both studies.

A. PXI-Immersion and KIM

The items of the immersion sub-scale of Player Experience Inventory (PXI) can be found in [1], [2]. The Kurzskala intrinsischer Motivation (KIM) questionnaire items are listed in [3].

B. PACES-8

We used the physical activity enjoyment scale (PACES-8) [4] to measure enjoyment of physical activity. The translated German items (based on discussion of three native speakers) are listed in Table I.

C. SSQ

We chose to use SSQ [5] to measure cybersickness symptoms of participants. A German version of the SSQ questionnaire [6] was administered in both studies (see Table II).

D. Exertion Custom Single-item

We used a custom single-item to measure exertion scores of participants (see Table III).

E. User Studies: Younger (User Study 1) and Older Adults (User Study 2)

This semi-interview/open-ended questions guideline is followed during the user studies.

(1) English - Please think back on the version you liked the most. What were the reasons for it?

(1) German - Erinnern Sie sich bitte an die Version zurück, die Ihnen am besten gefallen hat. Was waren die Gründe dafür?

TABLE I
THIS TABLE LISTS THE ENGLISH [4] AND GERMAN PACES-8 ITEMS.

#	ENGLISH	GERMAN
1	I find it pleasurable; I find it unpleasurable	Ich genieße es; Ich genieße es nicht
2	It's no fun at all; It's a lot fun	Es macht überhaupt keinen Spaß; Es macht sehr viel Spaß
3	It's very pleasant; It's very unpleasant	Es ist sehr angenehm; Es ist sehr unangenehm
4	It's very invigorating; It's not at all invigorating	Es ist sehr belebend; Es ist nicht sehr belebend
5	It's very gratifying; It's not at all gratifying	Es gibt mir etwas; Es gibt mir nichts
6	It's very exhilarating; It's not at all exhilarating	Es ist sehr aufregend; Es ist überhaupt nicht aufregend
7	It's not at all stimulating; It's very stimulating	Es ist überhaupt nicht interessant; Es ist sehr interessant
8	It's very refreshing; It's not at all refreshing	Es ist sehr erfrischend; Es ist überhaupt nicht erfrischend

TABLE II
THIS TABLE LISTS THE ENGLISH [5] AND GERMAN [6] SSQ ITEMS. (*) INDICATES ADJUSTMENTS IN THE ITEMS.

#	ENGLISH	GERMAN
1	General discomfort	Allgemeines Unwohlsein
2	Fatigue	Müdigkeit
3	Headache	Kopfschmerzen
4	Eyestrain	Überanstrengung der Augen
5	Difficulty focusing	Probleme beim scharf sehen*
6	Increased salivation	Erhöhter Speichelfluss
7	Sweating	Schwitzen
8	Nausea	Übelkeit
9	Difficulty concentrating	Konzentrationsschwankungen*
10	Fullness of head	Kopfdruck
11	Blurred vision	Verschwommenes Sehen
12	Dizzy (eyes open)	Schwindel bei offenen Augen
13	Dizzy (eyes closed)	Schwindel bei geschlossenen Augen
14	Vertigo	Gelichgewichtsstörungen*
15	Stomach awareness	Magen macht sich bemerkbar
16	Burping	Aufstoßen

TABLE III
THIS TABLE LISTS THE ENGLISH AND GERMAN VERSION OF THE CUSTOM SINGLE-ITEM MEASURE OF EXERTION.

#	ENGLISH	GERMAN
Exertion	I found the game to be physically demanding.	Ich habe das Spiel als körperlich anstrengend empfunden.

(2) English - What influence did the virtual alteration of the movement have on your motivation? Why?
(2) German - Welchen Einfluss hatte die virtuelle Anpassung der Bewegung auf ihre Motivation? Warum?
(3) English - What influence did the virtual alteration of object manipulation have on your motivation? Why?
(3) German - Welchen Einfluss hatte die virtuelle Anpassung der Objektmanipulation auf ihre Motivation? Warum?
(4) English - What aspects of the game would you like to change, and which would you leave as they are now?
(4) German - Welche Aspekte des Spiels würde Sie gerne verändern, welche würden Sie so lassen, wie sie jetzt sind?
(5) English - Do you have any additional comments?
(5) German - Haben Sie zusätzliche Anmerkungen?

F. Semi-structured Interview Guideline: Physiotherapist

This semi-interview guideline is followed during the interview with a physiotherapist.

Thank you for taking part in this interview. In this study, you played a game that explores different degrees of exaggerated interactions on older adults' user experience. In some variations of the game, you experienced exaggerated abilities for locomotion and object manipulation. Now, with this interview, I would like to gather information about how you as a physiotherapist evaluate this game, focusing on motivation, therapeutic value, and safety aspects.

1) Let's start with some questions regarding the effects of exaggerated abilities on the motivations of players.
   1.1 How motivated do you think that older adults would be to play this game?
   1.2 How do you think the exaggerated locomotion feature would affects the motivation of the players? Why?

1.3 How do you think the object manipulation feature would affects the motivation of the players? Why?

2) Let's continue with some questions regarding the therapeutic significance of the game.

2.1 How effective do you think the movements performed in this game were in terms of therapeutic benefits for the older adult players? Why?

2.2 How beneficial do you think the movements would be for the older adults with varying physical abilities (e.g., older adults using wheelchairs)?

2.3 How do you think the degree of exertion was for the players? Why?

3) Let's continue with some questions regarding the assurance of safety for the players.

3.1 Do you have any safety concerns for the older adult players in this game? Why?

3.2 How safe do you think the players' movement sequences are as well as their seated playing experience? Why?

4) Finally, general feedback:

4.1 Do you have additional comments or ideas that could improve the game?

G. Example Code Snippets

Here, we report example R code snippets used to analyze quantitative data of the user studies.

```
library(ARTool)
library(ez)
library(psychReport)
library(rstatix)
library(RVAideMemoire)

# analyzes are performed based on following resources: [1, 2, 3, 4]

# non-parametric ANOVA Results
m = art(dependentvariable ~ MovementManipulation*ObjectManipulation
                    + Error(Participant ID/(MovementManipulation*ObjectManipulation)), data=df)
summary(m)
m.art.anova = anova(m)
print(m.art.anova, verbose=TRUE)
m.art.anova$eta.sq.part = with(m.art.anova, `Sum Sq`/(`Sum Sq` + `Sum Sq.res`))
m.art.anova

#parametric ANOVA Results
model <-ezANOVA(data = df, dv = .(dependentvariable), wid = .
                        (Participant ID), within = .(MovementManipulation, ObjectManipulation),
                        detailed= TRUE, type = 3, return_aov = TRUE)

aovRT <- aovEffectSize(model, effectSize = "pes")
aovDispTable(aovRT)

# pairwise post-hoc tests
pairwise.wilcox.test(df$dependentvariable, df$condition, paired= TRUE, p.adj = "bonferroni")
pairwise.t.test(df$dependentvariable, df$condition, paired= TRUE, p.adj = "bonferroni")

# Friedman tests
friedman.test(dependentvariable ~ condition | Participant ID, data=df)
friedman_effsize(dependentvariable ~ condition | Participant ID, data= df)

# Multinominal tests
multinomial.test(df$bestcondition)
multinomial.multcomp(df$bestcondition, p.method = "bonferroni")
```

References

[1] V. V. Abeele, K. Spiel, L. Nacke, D. Johnson, and K. Gerling, "Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences," *International Journal of Human-Computer Studies*, vol. 135, p. 102370, 2020.

[2] L. Graf, M. Altmeyer, K. Emmerich, M. Herrlich, A. Krekhov, and K. Spiel, "Development and validation of a german version of the player experience inventory (pxi)," ser. MuC '22. New York, NY, USA: Association for Computing Machinery, 2022. [Online]. Available: https://doi.org/10.1145/3543758.3543763

[3] M. Wilde, K. Bätz, A. Kovaleva, and D. Urhahne, "Überprüfung einer kurzskala intrinsischer motivation (kim)," *Zeitschrift für Didaktik der Naturwissenschaften*, vol. 15, pp. 31–45, 2009. [Online]. Available: https://archiv.ipn.uni-kiel.de/zfdn/pdf/15_Wilde.pdf

[4] S. P. Mullen, E. A. Olson, S. M. Phillips, A. N. Szabo, T. R. Wójcicki, E. L. Mailey, N. P. Gothe, J. T. Fanning, A. F. Kramer, and E. McAuley, "Measuring enjoyment of physical activity in older adults: invariance of the physical activity enjoyment scale (paces) across groups and time," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 8, no. 1, pp. 1–9, 2011.

[5] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The International Journal of Aviation Psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[6] A. Hösch, "Simulator sickness in fahrsimulationsumgebungen-drei studien zu human factors," Ph.D. dissertation, Technische Universität Ilmenau, 2018. [Online]. Available: https://www.db-thueringen.de/receive/dbt_mods_00037865

10.4 [Preprint] Exercube vs. Virtual Reality: A Comparative Study of Exergame Technologies for Older Adults

[CorePre3] **Sukran Karaosmanoglu**, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. Exercube vs. virtual reality: A comparative study of exergame technologies for older adults. 2025. DOI 10.25592/uhhfdm.16797. URL https://doi.org/10.25592/uhhfdm.16797.

Supplementary materials for this article are included after the full-text of the paper.

ExerCube vs. Virtual Reality: A Comparative Study of Exergame Technologies for Older Adults

Sukran Karaosmanoglu
Human-Computer Interaction
University of Hamburg
Hamburg, Germany
sukran.karaosmanoglu@uni-hamburg.de

Sebastian Cmentowski
Industrial Design
Eindhoven University of Technology
Eindhoven, Netherlands
s.cmentowski@tue.nl

Lennart E. Nacke
HCI Games Group, Stratford School of Interaction Design and Business, University of Waterloo
Waterloo, Canada
lennart.nacke@acm.org

Frank Steinicke
Human-Computer Interaction
University of Hamburg
Hamburg, Germany
frank.steinicke@uni-hamburg.de

Figure 1: Illustrations of *Winter Wonderland* game variants: (a) a participant approaches a snowflake to catch it in the *VR-HMD* version of *Winter Wonderland*, and (b) a player catches a snowflake in the *ExerCube* version of the game.

Abstract

Insufficient physical activity is a major challenge in our aging society. Although exergames can provide enjoyable exercise opportunities for older adults, it remains unclear which display technology is best suited to reach this goal. This paper compares two popular exergame technologies with different immersion levels: (i) a virtual reality head-mounted display (VR-HMD) and (ii) the ExerCube, a commercial projection-based system. We conducted a within-participants study (N=34) with older adults to evaluate player experience, presence, cybersickness, game performance, and physical exertion. Both display types provided a comparably high player experience and physical exertion that can benefit older adults' physical well-being. The VR-HMD offered superior presence, while the ExerCube led to higher performance and physical activity. Our findings advance the understanding of how different exergame technologies affect older adults' experiences. We present research and design implications to guide the future development of age-appropriate exergames.

CCS Concepts

• **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**; • **Software and its engineering** → **Interactive games**.

Keywords

virtual reality, ExerCube, exergames, movement games, exercise, older adults, display

ACM Reference Format:
Sukran Karaosmanoglu, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. 2025. ExerCube vs. Virtual Reality: A Comparative Study of Exergame Technologies for Older Adults. In . ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Conference'17, Washington, DC, USA

1 Introduction

Regular physical activity is essential for keeping good health: it can help prevent many illnesses like cardiovascular diseases, diabetes, and even cancer and dementia [84]. Exercising is critical for every age group, but can particularly help older adults preserve an independent life [83]. However, participation in regular training can be challenging [55] and most people today do not meet the recommended amounts of daily physical activity [84]. Therefore, the World Health Organization (WHO) strongly advocates for creating "*safe, affordable places and spaces that invite, support and enable people of all ages and abilities to be active in different ways*" [83]. In our increasingly digitalized world, this appeal also calls for digital solutions that meet both peoples' exercise and entertainment needs.

In this context, exergames possess great potential to motivate people to be active by combining exercises with gameplay [56] and are known for their positive effects on well-being [60, 67]. Apart from general users, exergaming is increasingly used by senior homes to keep older adults fit and agile [38, 59]. However, hardly any prior exergame research investigated the technological requirements of older adults, even though it is well known that their use of technology [47, 52], and the resulting experience and performance [15, 16, 45] differs significantly between age groups. Accordingly, we see a need to focus specifically on this user group to better understand their expectations and preferences. Our research targets one specific aspect: the impact of different display systems.

Recently, a lot of attention has been put on how to improve exergaming experiences. Apart from using fully immersive headsets, i.e., virtual reality head-mounted displays (*VR-HMD*s) [37], various commercial solutions, such as the *ExerCube*, have been proposed that trade aspects like immersiveness and visual fidelity with a higher usability. While previous studies showed that display technologies affect players' exergaming experiences [4] and performance [85], there is no empirical knowledge on how general-purpose *VR-HMD*s compare against specialized exergame setups like the *ExerCube*. Given the growing academic and commercial prevalence of both technologies [37, 49, 51, 53, 59], it is crucial to understand their individual strengths and drawbacks towards creating enjoyable and effective workouts. Therefore, our research compares these two highly contrasting display systems: (i) *VR-HMD*s and (ii) the *ExerCube*.

*VR-HMD*s have unique benefits that make them a perfect fit for exergames. They can fully immerse users into the virtual environments, thereby boosting players' motivation [4]. Users can access exercise opportunities without traveling to exercise locations [39] or requiring a large space [38]. Already today, there are plenty of well-selling commercial exergames (e.g., FitXR [20]) and even VR-enhanced gyms [80]. However, the use of VR technology also introduces new challenges, such as cybersickness, usability issues, and discomfort of wearing the HMD. In contrast, the *ExerCube* is a specialized exergame system that is specifically designed as a fitness tool for exercising purposes [50]. Unlike a *VR-HMD*, the *ExerCube* uses three projection walls to display the virtual environment. This system has received attention in academia [42, 50], and is commercially used in multiple training locations. Although the *ExerCube* is colloquially referred to as "immersive" (because it immerses the user into a virtual gameplay), it differs greatly from fully immersive systems like *VR-HMD*s (e.g., no stereoscopy, limited depth perception and field-of-view). Accordingly, it does not meet the minimal requirements of technical immersion [69]. However, research shows that it still provides enjoyable and motivating experiences [50, 51]. While this system allows users to experience the virtual environment without head-worn displays, it also has significantly higher cost and space requirements. Although both systems have specific strengths and drawbacks and have been proven to be suitable for exergames, we do not know how the choice of display system influences the exergame experience, in particular for older adults who have vastly different needs and expectations than other demographic groups. To close this gap, our research addresses the following research question:

RQ How does the technical immersion of an exergame system (i.e., a *VR-HMD* vs. *ExerCube*) influence older adults' exergaming experience, motivation, cybersickness, physical activity enjoyment, workload, game performance, physical exertion (heart rate) and physical activity?

To answer this research question, we first iteratively designed and implemented an exergame (*Winter Wonderland*) tailored to the needs of older adults using a human-centered design approach [21]. The gameplay is exactly identical when playing it with a *VR-HMD* or the *ExerCube*. Using this exergame, we then conducted a within-participants study (*N*=34) with older adults (+65 years) using a mixed-method design [12]. Our quantitative and qualitative findings indicate that both display types yield comparably high levels of enjoyment of gameplay and physical activity. While the *VR-HMD* setup offers greater presence, the *ExerCube* setup leads to higher game performance and physical activity. Overall, both setups contribute to increased physical exertion, which may enhance the physical well-being of older adults. With our work, we challenge the assumption that highly immersive technologies necessarily lead to better performance or increased physical activity. Our analysis of the subtle trade-offs between presence, performance, and physical activity across technologies reveals new opportunities for designing exergaming experiences tailored to individual age groups. We summarize this paper's contributions according to Wobbrock and Kientz [82] categorization in the field of human-computer interaction (HCI):

- By creating an exergame tailored to older adults, we provide an "artifact contribution" [82]. Our detailed design process and rationales offer guidance on how to design exergames targeted at specific user groups.
- Conducting "empirical research" [82], we present the first comparison of *VR-HMD* and *ExerCube* technologies with older adults, which challenges assumptions on the role of technical immersion on player experience.
- We end with a "theoretical" and "opinion" contribution [82]: By linking our findings to prior work and societal impacts, we present implications supporting the research and design of innovative, age-appropriate exergames.

2 Related Work

This section provides an overview of prior work on immersive displays and exergames for older adults.

2.1 Display Technologies for Virtual Environments

When talking about the various output systems used to deliver virtual content, researchers use the term *immersion* to refer to a system's technical quality and capabilities [69], which includes many factors such as stereoscopy, field of view, or resolution. While there is no definitive judgment on what makes a system immersive, the minimal requirement is a head-dependent stereoscopic view of the virtual world that exceeds the players' field of view without requiring more than head-based input [23, 69]. To describe the subjective experience of feeling present in virtual environments, researchers use the term *presence* [68]. While, there are different understandings of these terms (e.g., [46]), we follow the definitions of [23, 68, 69] and use the term immersion to refer to the technical differences between the *ExerCube* and *VR-HMD* systems, while to describe users' experience resulting from these technologies we use the term presence.

VR systems typically consist of two components: the *VR-HMD* offers a stereoscopic view of the virtual world, while the controllers are used for spatial interactions. Today, both components typically feature six-degree-of-freedom positional tracking. These systems are considered fully immersive because of their ability to provide high-quality visuals and audio that transport users into virtual environments [70]. However, *VR-HMD*-based systems also introduce challenges, such as cybersickness [31], usability problems [1], and safety issues [39]. These drawbacks led researchers to develop alternatives, such as CAVEs (i.e., immersive room-based projection-based systems [13]). Although these costly, specialized systems do not feature HMDs, they are still fully immersive since users see a stereoscopic, head-dependent view using shutter glasses and optical tracking systems. Unlike *VR-HMD*s and CAVEs, the *ExerCube* [50] does not fully conform to the technical definition of immersion. While it uses similar projection walls like CAVE installations [13], it does not offer a view-dependent stereoscopic image [50]. Thus, in our paper, we consider this system as non-immersive.

2.2 Experiencing Exergames with Different Technologies

Exergames combine the power of fun with physical activity: users primarily exert physical effort to complete game tasks [56]. Research indicates that these games benefit cognitive, physical, and social well-being [28, 58]. In particular, *VR-HMD* technologies are becoming increasingly popular in exergame research [37] and commercial exergames, both for home use (e.g., Beat Saber [24]) and gym settings (e.g., Black Box VR [80]). Previous studies have shown that *VR-HMD* exergames can yield higher motivation [4] and performance [85] compared to TV-based versions. Contrarily, another study [86] found no significant difference in player experience between these two display types. *VR-HMD* exergames are generally praised for their immersive qualities and motivational power. However, wearing a view-blocking headset during exercising also creates new challenges regarding safety, comfort, and real-world connectivity [39].

As an alternative, the *ExerCube* [50] is a specifically designed exergaming system which is based on the dual-flow model, emphasizing both the psychological flow and the balance between physical intensity and personal fitness [66]. To date, the *ExerCube* system has been established in many locations, such as local gyms and senior living facilities. Users can exercise without the drawbacks of *VR-HMD*s (e.g., sweating, discomfort of the headset's weight) [11, 31] because they do not wear any headset that blocks their real-world connection with their bodies (i.e., limited proprioception). This can potentially increase users' feeling of safety, allowing them to move more freely. Although the *ExerCube* does not fulfill the requirements of technical immersiveness, research shows that *ExerCube*-based exergames still provide engaging and motivating physical exercises [49–51]. However, the *ExerCube* also has disadvantages compared to *VR-HMD* systems, which impact both exergame experience (i.e., delivering a lower-resolution, non-stereoscopic view that is limited to three fixed walls) and financial resources (i.e., costing more and requiring more space and computational power).

Despite increasing research on *ExerCube*-based exergames [42, 49–51, 53], it remains unclear if the promised benefits hold and if it delivers a comparable or even better experience than *VR-HMD*s. To date, the comparison between both systems has received almost no attention. The studies closest to ours were done by Elor et al. [19] and Schmidt et al. [64]; but, both studies compared exergames in *VR-HMD* and CAVEs, not in the *ExerCube*. Elor et al. [19] revealed that the *VR-HMD* outperformed regarding both game performance and engagement. Schmidt et al. [64] reported that players had higher presence and preferred the *VR-HMD*-based rowing game. Even though these studies provide a foundation, their results cannot be readily translated to the comparison between *VR-HMD* and *ExerCube*: the *ExerCube* misses many critical features of CAVEs (e.g., stereoscopy, head-tracking) but offers unique advantages for exergames (e.g., easier calibration, cushioned walls). Yet, no study has compared VR and *ExerCube*. Coupled with a lack of focus on older adults (despite the rapidly increasing use of *VR-HMD* and *ExerCube* systems in senior homes), these differences surface a clear need for an empirical study to inform the design and delivery of effective, enjoyable, and age-appropriate workouts.

2.3 Older Adults & Exergames

According to the activity theory of aging [29, 75], staying engaged in activities can benefit the well-being of older adults. Exergames offer accessible exercise options, one of the major barriers for older adults [55], and support their psychological [60] and physical well-being [67]. Although this shows that older adults could profit greatly from exergames, a recent review surfaced that most studies have been conducted with younger adults [37], showcasing a critical gap in HCI research: we have limited information about exergame experiences of older adults. Yet, it is well-known that older adults have different requirements and abilities than younger adults in the context of exergames [15, 45], for example due to age-related decline in cognitive and physical abilities [25, 34]. Accordingly, Karaosmanoglu et al. [37] caution against the assumption that "*all exergames work for any age group*" to avoid "*harm[ing] people*".

Prior research shows that older adults focus more on health benefits of exergames [71], rate their experiences more positively [45], experience higher presence and less cybersickness in VR [16], and perform worse compared to younger generations [15, 45, 62]. Accordingly, it is challenging to transfer findings from younger to older adults. Instead, designing exergames around older adults' specific preferences and needs, can ensure they enjoy these games [39, 44] and use them as an alternative tool to exercise [38]. But, prior studies also highlight that some older adults feel unsafe when using *VR-HMD*s, potentially experience disconnection from the real world [39], and face usability issues when using typical VR controllers [1, 39]. These issues can prevent older adults from enjoying the experience and raise questions about the suitability of using *VR-HMD*s as exergaming platform.

In contrast, the *ExerCube* setup allows users to see their own body instead of an avatar, like in HMD-based exergames. Due to the age-related decline (e.g., motor skills, reaction time) [25], this advantage may be particularly beneficial for older adults by providing a sense of safety while moving, whereas the immersive nature of *VR-HMD* may be more cognitively stimulating and attention capturing [32]. Unfortunately, research on *ExerCube*-based exergames specifically for older adults is still limited, with a few exceptions [53, 59] focusing on identifying target-user-group requirements. Moreover, all studies comparing CAVE and *VR-HMD* setups focused on younger generations [19, 36, 64]. Given the increasing usage of both *VR-HMD* [37] and *ExerCube* exergames and the aging population [77, 78], we see a clear need to understand how *VR-HMD* and *ExerCube*-based exergames compare in terms of player experience and performance for older adults. Specifically, we evaluate the role of technical immersion in older adults' exergaming experiences, complementing prior research on the influence of technical immersion on player motivation and performance [4, 85].

3 Exergame: Winter Wonderland

To best of our knowledge, there is no commercial game available that can run on both exergame systems and provide comparable experiences across these setups. To answer our research question, we therefore developed a single-player exergame in Unity [74] (v.2022.3.14f1) that users can play using both a *VR-HMD* and the *ExerCube*: *Winter Wonderland* (see also limitations). We used a desktop PC equipped with two NVIDIA GeForce RTX 2080 Ti graphics cards, an Intel Core i9-9900K CPU, and 32 GB of RAM to run our game. Using this game, also allowed us to collect movement and game performance data. To ensure that the virtual mapping was exactly the same in both setups, we carefully configured the alignment of the physical and virtual environments. Consequently, the real-world and virtual-world positioning of game objects was identical, requiring participants to perform exactly the same movements under each condition.

Our game development was informed by a human-centered design approach (i.e., iteratively testing the game with experts and end-users) [21] and previous studies [1, 37–39, 73]. During the development process, we tested our game in each stage with older adults (N=4, 3 women, 1 man, M_{age}= 81, SD_{age}=8.76), a physiotherapist (42 years, man), and a trainer (57 years, woman). Below, we describe the game design rationale and the gameplay of *Winter Wonderland*.

3.1 Game Design Rationale

Following recommendations [37], we describe our exergame based on the goals, people, exercises, design, and technologies taxonomy. Our game's goal is to promote physical activity and provide cognitive and physical training for older adults. Designing exergames targeting physical and cognitive training simultaneously is feasible [38] and beneficial [73]. To design our game, we relied on prior work [38, 45] that targeted older adults and created multiple levels with increasing challenges (see Table 1). Since our main goal was to compare different displays, both conditions are played in the same stationary space, i.e., within the *ExerCube*, and just differ in the type of display type. In the *VR-HMD* version, players see the virtual content through a Valve Index headset, while in the *ExerCube* version, players see the game on the three projection walls. Additionally, players' hands are tracked using two Vive trackers attached to their wrists.

With our game, we target full-body movements. Players play the game in a standing position. As the main interaction, players are tasked with catching colored snowflakes that spawn within one of three vertical planes corresponding to the real *ExerCube*'s walls. Consequently, players see the snowflakes either displayed on the *ExerCube*'s projection surface or in the fully immersive *VR-HMD*. To catch the snowflakes, they walk and perform reaching movements with their arms. Once they touch a snowflake with their hand, it disappears after playing a short animation and sound feedback. We also added similar juicy effects to the arrival of the snowflake; these fly on short, slightly tumbling paths from the distance toward the player and stop at their intended position within the vertical planes. After arriving, their movement changes to a slow self-rotation, and a short sound is played to give players additional feedback that a new snowflake has arrived. Regarding the in-game movements, we received positive feedback from the physiotherapist and trainer. After playing the prototype, the physiotherapist considered the game safe and rich in movements (e.g., catching and walking). He suggested adding a score or leaderboard for a competitive feeling. Consequently, we added a score that reflects how fast and correctly players completed their task. For every successfully caught object, players receive 100 base points and up to 100 bonus points (depending on how fast they reacted, i.e., *bonus points=100/percentage of elapsed time of lifetime of an object*). For every wrongly caught object, 100 points are deducted. No points are subtracted for missed objects.

The game features a daily theme: a winter landscape. We chose this theme to cater to our target population (following [39])—older adults living in a senior living facility in Germany—and timing of running our study during winter. Players experience this environment while standing on a small cliff in the midst of a semi-dense forest featuring pine trees, paths, boulders, and bushes. The scenery is covered in snow and a snowfall effect creates the impression that it is snowing all over the winterly forest, with colored, interactable snowflakes only spawning within the players' proximity. As background music, we use Christmas-themed soundtracks to create audiovisual congruence [1]. The feedback from older adults in the development process was extremely positive. Some participants were moving with the rhythm of the music while completing their tasks: "*I like it, I love it*"-P1 and "*I found that movements, the music and everything fitted nicely together*"-P2. Nevertheless, players provided improvement ideas, such as wishing for a louder sound to indicate a new arriving task. To account for these expectations, we further fine-tuned minor aspects of the game, such audio effects, wording, placement of the tutorial information, and the brightness of the object colors.

To prevent players from leaving the play area, we marked it in the virtual environment with three stone barriers. In the corners, blue crystals emit light rays that indicate the position of the three

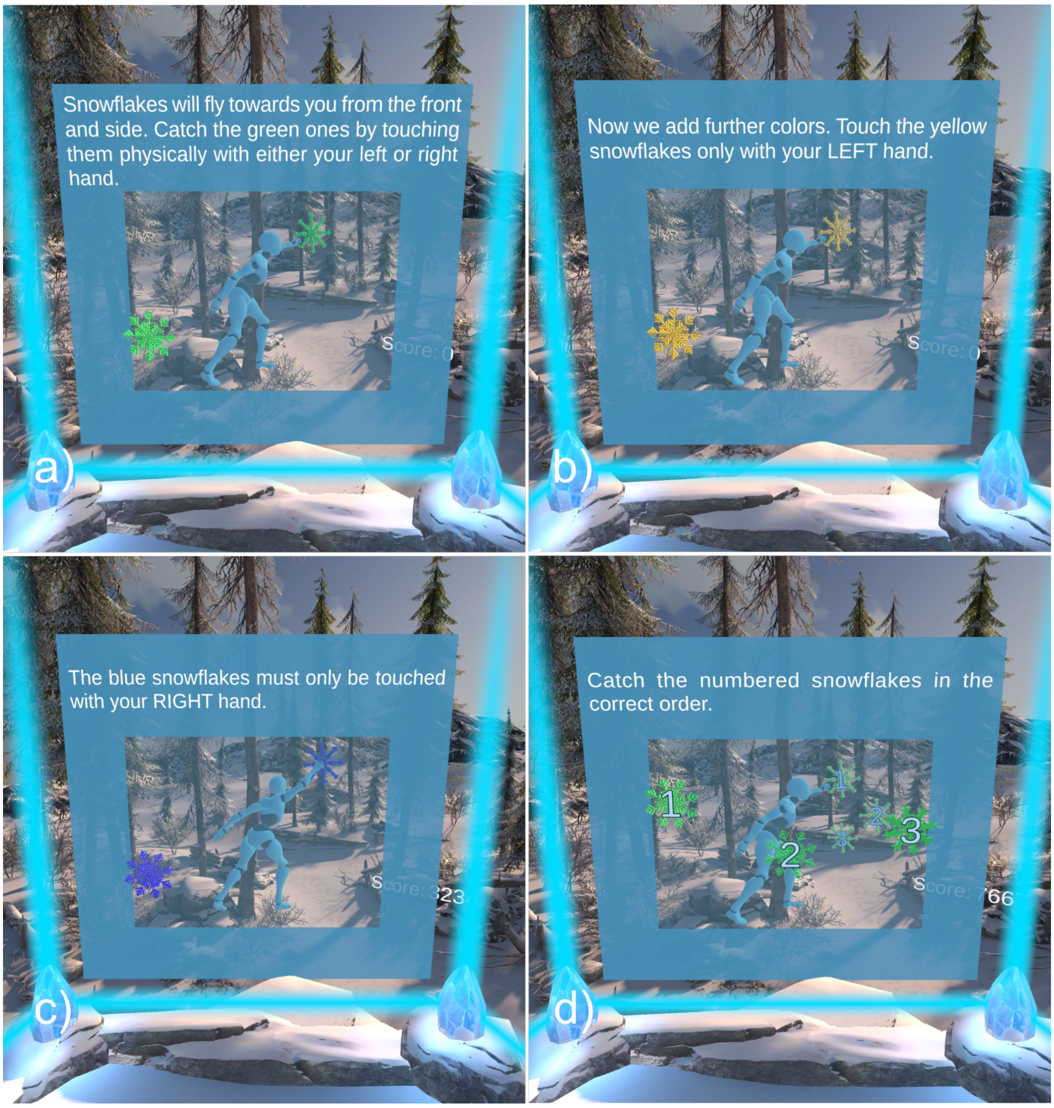

Figure 2: Interactive tutorial of each game level in *Winter Wonderland*: a) players collect green snowflakes with their either hand. b) players collect yellow snowflakes with their left hand. c) players catch blue snowflakes with their right hand. d) in the last level, players catch the number displayed snowflakes in the correct order.

vertical interaction planes corresponding to each *ExerCube* wall. Before the game, players calibrate their proportions to ensure that the game matches the snowflakes' spawning position with their arm reach. Since we targeted older adults, we decided to avoid overly straining movements and limit the vertical snowflake spawn to between 30% and 90% of the player's vertical arm reach.

3.2 Gameplay

In *Winter Wonderland*, players catch colored (green, yellow, blue) snowflakes that spawn on the three vertical interaction planes, inspired by the prior work [38]). To see the impression of the gameplay, please refer to the supplementary video.

The total game duration (excluding tutorials) is five minutes. While the first two levels feature a static number of spawned snowflakes (level 1: 21, level 2: 31), the last level is designed to run until players reach the targeted game time (see Table 1). Depending on how fast they catch each object, players can catch a theoretical maximum of 129 snowflakes across all three levels. These are evenly spread across all three planes, i.e., maximally 43 snowflakes per plane.

The game starts with players calibrating their arm reach by performing a t-pose and an overhead reach pose. Afterward, the game shows a tutorial introducing the in-game task. A simplified 3D avatar catching a green snowflake illustrates the described actions

Table 1: This table lists the game levels in *Winter Wonderland*.

No	Level Names	Game Properties	Description
#1	All Collectable	❄	Catch as many green snowflakes as possible.
#2	Colors for Hands	❄ ❄	Catch yellow snowflakes with left hand and catch blue snowflakes with right hand.
#3	Order Numbers	❄ 1-2-3	Catch the snowflakes in the right order.

(see Figure 2a). To complete the tutorial, players must catch a first tutorial snowflake. Next, they play the first game level. Here, all snowflakes are colored green, indicating that they can be caught with either hand. Each snowflake remains for 15 seconds before disappearing to give players enough time to learn the interactions.

The second level adds hand-dependent colors: yellow snowflakes have to be caught with the left and blue ones with the right hand (designed based on the game of [38] and principles of simultaneous training [73]). As before, players are introduced to these tasks with an interactive tutorial (see Figure 2b and c). The level starts with five yellow snowflakes, followed by five blue snowflakes, before mixing both colors. As before, all spawned objects are spread evenly across the three walls. Also, this level reduces the lifetime of all snowflakes to five seconds to encourage players to react quickly.

The last level reverts to green-colored snowflakes, adding numbers to the snowflakes to indicate in which order they have to be caught. In this level, we provide simultaneous training that involves both cognitive and physical aspects [73]. The trainer particularly liked this level design due to its holistic mix of cognitive and physical tasks. Like before, players are introduced to this task with an interactive tutorial (see Figure 2d). Next, the game begins to spawn sets of five snowflakes. These are numbered beginning from one, with numbers increasing over sets (i.e., first set: 1–5, second set: 6–10). Unlike the prior snowflakes, these do not disappear but stay visible until the player has caught all five in the correct order. After the last flake has been caught, the next set of five flakes is immediately spawned. This process is repeated until the total game time of five minutes (for all three levels combined) has run out. Then, all the remaining flakes disappear, and a congratulations message ends the game.

4 User Study

To compare how different displays affect players' experience, performance, and exertion, we conducted a within-participants study with two conditions: *VR-HMD* and *ExerCube*. We conducted our study with older adults at a senior living facility located in Germany. This senior living facility is home to both people who require institutional care and those living independently. For the study, we only recruited older adults who were at least 65 years old (following the definitions of the United Nations [77, 78]) and were not assigned a care level (i.e., independently living without help). Our study received ethics approval from the local ethics committee of University of Hamburg (#: 004/2023).

4.1 Pre-Game Measures

We used a demographic questionnaire to gather sample characteristics (e.g., age, gender, enjoyment of physical activity, experience with *VR-HMD* and *ExerCube*). To assess the prior cybersickness symptoms, we employed the Simulator Sickness Questionnaire (SSQ) [33, 41]; it has 16 items to be rated on a 4-point scale, resulting in a total SSQ, nausea, oculomotor disturbance, and disorientation score. We used a Polar OH1 armband (1 Hz) to collect heart rate data, which is an established method for measuring objective exertion in exergame research [4, 35, 37]. Before the game, we assessed the heart rate for five minutes and averaged these values to obtain a baseline heart rate (similar to other works [40]).

4.2 In-Game Measures

During the gameplay, we collected game performance metrics: physical activity, score, and collected targets. The physical activity metric was quantified as the total distance players moved both their left and right hands during each level. This traveled distance was calculated by adding the frame-to-frame positional differences of the controllers. The score was calculated based on success rate and reaction speed (for each correctly caught snowflake: $100 + 100 * (1 - t_{catchtime}/t_{lifetime})$). The collected targets indicate the number of snowflakes caught. We also calculated the heart rate during gameplay by collecting the raw data with the same armband and averaging them for each condition.

4.3 Post-Game Measures

We again used the SSQ to capture cybersickness symptoms after each condition. The Igroup presence questionnaire (IPQ) [65] assessed the perceived presence; it includes 14 items rated on a 7-point Likert scale (e.g., *fully disagree/-3 — fully agree/+3*), yielding a general presence score and three subcategories: spatial presence, involvement, and experienced realism. We measured the game interest/enjoyment (i.e., motivation) using the Kurzskala intrinsischer Motivation (KIM) questionnaire [61, 81] (three items on a 5-point Likert scale, *not at all true/0 — very true/4*). To assess overall player experience, we used the Player Experience Inventory (PXI) [2] (German version [26]). The PXI has 30 items (10 constructs) on a 7-point Likert scale (*strongly disagree/-3 — strongly agree/+3*). The ratings for enjoyment of physical activity were collected using the short version of the Physical Activity Enjoyment Scale (PACES-S) [9, 22] (four items on a 5-point Likert scale (*strongly disagree/1 — strongly agree/5*). Lastly, the NASA-Task Load Index (NASA-TLX) [27] was used to measure the perceived workload. The NASA-TLX contains six items (e.g., cognitive workload) on a 21-point Likert scale. The raw values for each item can be calculated from 0 to 100 (e.g.,*very low/0 — very high/100*).

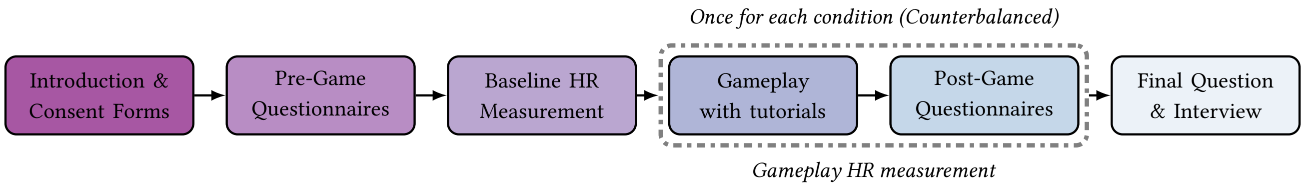

Figure 3: The user study procedure followed in our research. The participants played two game versions in a counterbalanced fashion. We assessed questionnaire, game performance, heart rate, and interview data.

4.4 Final Question & Interviews

After the participants completed all conditions, they answered a single question asking about their preference: "*Which version of the game did you like the best?*". Finally, we conducted audio-recorded semi-structured interviews, asking six questions to compare older adults' experiences with the two systems (see questions in supplementary materials).

4.5 Sample Characteristics

Demographics. Overall, we recruited 36 participants. However, we excluded two participants' data (one due to technical issues and one due to withdrawal), resulting in a total sample size of 34 people. Twenty-five participants were women, while nine participants reported their gender as men. The mean age of participants was 79.41 (*SD*=7.40) years.

Gaming Experience. Nineteen participants reported playing traditional games (e.g., board and card games) never or less than once per year. The remaining play at least once a week (*n*=10) or month (*n*=5). Digital games (e.g., mobile phone and computer games) are played more frequently: daily (*n*=13), weekly (*n*=11), less frequently (*n*=3), never (*n*=7).

Technology Experience. Only 14 participants had *VR-HMD* experience; but all of them used VR only seldomly (at least once a year *n*=2, less frequent *n*=12). Yet, most participants (*n*=29) already had experience with the *ExerCube*, due to its availability in the senior living facility (they can book fitness sessions with the *ExerCube*): *n*=21 played at least once a week, *n*=1 at least once a month, and the rest less frequently (*n*=7). Although we note this familiarity as an unpreventable bias, we ensured that no participant had seen any portion of the testbed game before the study.

Physical Activity Experience. The participants rated their enjoyment of exercising positively (*M*=2.12, *SD*=0.84 on a scale from -3/*strongly disagree* to +3/*strongly agree*). Thirty-two participants reported engaging in weekly physical activity: 1-2 (*n*=12), 3-4 (*n*=11), and 5-7 times (*n*=9).

4.6 Procedure

We announced our study with the help of the senior living facility in Germany where the study was conducted. Although all participants were recruited with the help of the senior living facility, we ensured that everyone participated individually and had not experienced parts of the game before (neither during development or other peoples' participations). At the beginning of our study, participants were given an informed consent form and a safety document informing them about possible negative effects of using *VR-HMD*s (e.g., cybersickness). After signing the form, participants completed the pre-game measures and received a verbal explanation regarding the use of *VR-HMD*s and cybersickness. Next, participants put on the heart rate monitor and wore it on their non-dominant arm. To measure the baseline heart rate, they were instructed to sit still without moving or speaking for five minutes. Afterward, participants played each version of the game (*ExerCube* and *VR-HMD*, including the tutorial and the actual gameplay) in a counterbalanced order. Each game (excluding the tutorials) lasted five minutes. During the gameplay, we recorded the heart rate metrics and all gameplay performance data. After each condition, they completed the post-game questionnaires. During this phase, we closely monitored participants' heart rate (through an app) and waited until it returned to its baseline before playing the other condition. The break between conditions was between 10-15 minutes. After both rounds of gameplay, participants answered the final question and were interviewed (audio-recorded). The entire study took approximately 90 minutes. Participants were compensated with 15 EUR. The full study procedure can be seen in Figure 3.

5 Analysis Methodologies

To answer our research question, we used a mixed-method approach [12]. While our quantitative analysis answers how constructs are affected by the different displays, our qualitative analysis explains why.

Quantitative Methodology. We confirmed the normality of our dependent variables using Shapiro-Wilk tests. When the normality was violated, we conducted Wilcoxon-signed rank tests instead of paired t-tests (one factor: displays *ExerCube* vs *VR-HMD*). For cybersickness scores and heart rate metrics, we conducted one-way ANOVAs or Friedman's ANOVAs (based on normality) and treated the measurement points as time points (one factor) with three levels: pre, *ExerCube*, and *VR-HMD*. Lastly, we performed an exact binomial test to see participants' preferences over the displays. We note that one participant had a single data point missing in the audiovisual appeal, clarity of goals, and ease of control subcategories of PXI. We employed a k-nearest neighbor (kNN) imputation method [43, 76] to determine replacement values for these data points, considering only respective subcategories (k=3).

Qualitative Methodology. To prepare our data, we transcribed it using Dovetail [17], and a native speaker checked for any errors. We

Table 2: The table shows the descriptive and statistical test values of the PXI and SSQ subcategories, and PACES-S.

Condition	KIM-enj. Mean (SD)	PXI-mea. Mean (SD)	PXI-cur. Mean (SD)	PXI-mas. Mean (SD)	PXI-aut. Mean (SD)	PXI-imm. Mean (SD)	PXI-pro. Mean (SD)	PXI-cha. Mean (SD)
ExerCube	3.43 (0.72)	0.86 (1.83)	1.46 (1.62)	1.85 (1.22)	-0.60 (2.27)	1.94 (1.00)	1.24 (1.53)	2.25 (0.84)
VR-HMD	3.44 (0.69)	0.72 (1.81)	1.70 (1.17)	1.90 (0.92)	-0.56 (2.27)	2.07 (0.90)	1.21 (1.44)	2.45 (0.83)
Test	V=80	V=190.5	V=86	V=203.5	V=157.5	V=113	V=124	V=60.5
p	p=0.826	p=0.459	p=0.312	p=0.736	p=0.841	p=0.455	p=0.781	p=0.171
Effect	r=-0.027	r=-0.090	r=-0.122	r=-0.041	r=-0.024	r=-0.091	r=-0.034	r=-0.166

Condition	PXI-aud. Mean (SD)	PXI-eas. Mean (SD)	PXI-cla. Mean (SD)	Total SSQ Mean (SD)	Nausea Mean (SD)	Oculomotor Mean (SD)	Disorientation Mean (SD)	PACES-S Mean (SD)
Pre	–	–	–	9.46 (17.30)	8.42 (21.11)	8.03 (12.09)	8.19 (20.02)	–
ExerCube	2.11 (0.94)	2.59 (0.75)	2.84 (0.31)	1.87 (3.93)	2.24 (4.73)	1.34 (4.36)	1.23 (5.27)	4.75 (0.51)
VR-HMD	2.31 (0.84)	2.62 (0.89)	2.67 (0.57)	3.08 (5.25)	2.53 (6.35)	2.68 (5.87)	2.68 (5.87)	4.5 (0.96)
Test	V=98	V=42.5	V=53	$\chi^2(2)$=11.532	$\chi^2(2)$=3.509	$\chi^2(2)$=13.170	$\chi^2(2)$= 6.75	V=55
p	p=0.228	p=0.330	p=0.082	**p=0.003**	p=0.173	**p=0.001**	**p=0.034**	p=0.055
Effect	r=-0.146	r=-0.118	r=-0.211	W=0.170	W=0.052	W=0.194	W=0.099	r=-0.232

then translated the text into English using DeepL Pro [14]. Again, the same person rechecked the text for any errors. For the analysis of the interviews, we used the reflexive orientation of thematic analysis [6, 7] because it is a well-established method in HCI (for an overview of usage in healthcare HCI, see [5]), provides flexibility, can be used to understand people's behaviors and experiences [10], draws on the researcher' background to analyze patterns in the data, and explicitly emphasizes the role of the researcher and recommends reporting its role [6]. Therefore, this method does not require multiple coders, as it does not aim at generalizability [7]. Overall, our approach followed Braun and Clarke [6, 7]'s steps: one researcher (i) getting familiar with the data through reading it, (ii) coding the data iteratively, (iii) forming initial themes, (iv) reviewing the initial themes, (v) refining final themes, and (vi) reporting. For coding, we used a mix of deductive and inductive codes: First, we used deductive categories and codes (e.g., VR is more motivating (category) –> presence (code)) to capture relevant aspects focusing on our research. However, when additional concepts appeared, we created new codes (i.e., inductive) or refined our existing codes. In the end, we derived themes focused on answering our research question. Following the best practices in qualitative data analysis [7, 57], we provide a positionality statement that contributes to transparency and rigor; the lead researcher who conducted the qualitative analysis has performed multiple studies on the impact of age differences on player experience, performance, and physiological measures. They have experience in developing VR games and conducting VR user studies.

6 Results

This section presents the findings from the questionnaires, objective metrics, and qualitative analysis of the study. For readability, we only report significant results in the text and list all results in Table 2, Table 3, Table 4, and Figure 4.

6.1 Questionnaires

6.1.1 Cybersickness. We found a significant effect of the time points on the Total SSQ ($\chi^2(2)$=11.532, p=0.003, W=0.170), oculomotor disturbance ($\chi^2(2)$= 13.170, p=0.001, W=0.194), and disorientation ratings ($\chi^2(2)$=6.75, p=0.034, W=0.099). The post-hoc tests showed a significant difference between *pre* and post-*ExerCube* Total SSQ (V=241, p=0.005) as well as oculomotor disturbance scores (V=115.5, p=0.004), indicating that participants experienced higher symptoms before the study. However, the post-hoc tests did not reveal any significant differences for the disorientation category.

6.1.2 Motivation, Player Experience, & Physical Activity Enjoyment. We found no significant difference between the *ExerCube* and *VR-HMD* conditions in terms of both game enjoyment, player experience, and enjoyment of physical activity; both setups led to moderate to high scores on these constructs.

6.1.3 Presence. The *VR-HMD* led to significantly higher scores compared to *ExerCube* in all subcategories of the IPQ: general presence (V=41, p=0.017), spatial presence ($t(33)$=-3.805, p=0.001, d=-0.653), involvement (V=78, p=0.001, r=-0.403), and experienced realism ($t(33)$=-3.221, p=0.003, r=-0.552). We illustrate these constructs in Figure 4.

6.1.4 Workload. We found a significant difference for the NASA-TLX's performance subcategory, indicating that the perceived success was higher in the *ExerCube* compared to *VR-HMD* (V=239, p=0.040, r=-0.250). Similarly, participants found it more frustrating to play the exergame in *VR-HMD* compared to playing it in *ExerCube* (V=6.5, p=0.007, r=-0.329).

6.1.5 Preference. Eighteen participants preferred *ExerCube* and 16 preferred *VR-HMD*. A two-sided exact binomial test indicated that these proportions were not statistically significant, p=0.864.

6.2 Objective Metrics

6.2.1 Game Performance Metrics. Participants mostly performed better in the game when they played in the *ExerCube* compared to the *VR-HMD*: correctly collected snowflakes ($t(33)$= 2.279, p=0.029,

Table 3: The table presents the descriptive and statistical values of NASA-TLX and IPQ subcategories.

Condition	NASA-TLX-Men. *Mean (SD)*	NASA-TLX-Phy. *Mean (SD)*	NASA-TLX-Tem. *Mean (SD)*	NASA-TLX-Per. *Mean (SD)*	NASA-TLX-Eff. *Mean (SD)*
ExerCube	28.09 (25.29)	30.59 (27.30)	18.24 (26.02)	81.47 (18.97)	15.29 (22.22)
VR-HMD	35.29 (30.62)	24.71 (26.60)	19.71 (27.08)	74.56 (21.19)	26.91 (27.39)
Test	V=142.5	V=214.5	V=113	V=239	V=75.5
p	p=0.267	p=0.163	p=0.779	**p=0.040**	p=0.059
Effect	r=-0.135	r=-0.169	r=-0.034	r=-0.250	r=-0.229

Condition	NASA-TLX-Fru. *Mean (SD)*	IPQ-Pre. *Mean (SD)*	IPQ-Spa.Pre. *Mean (SD)*	IPQ-Inv. *Mean (SD)*	IPQ-Exp.Rea. *Mean (SD)*
ExerCube	4.12 (9.81)	-0.56 (2.34)	0.02 (1.60)	0.64 (1.39)	-0.14 (1.81)
VR-HMD	9.41 (16.91)	0.41 (2.38)	0.91 (1.21)	1.43 (1.14)	0.70 (1.63)
Test	V=6.5	V=41	t(33)=-3.805	V=78	t(33)=-3.221
p	**p=0.007**	**p=0.017**	**p=0.001**	**p=0.001**	**p=0.003**
Effect	r=-0.329	r=-0.290	d=-0.653	r=-0.403	d=-0.552

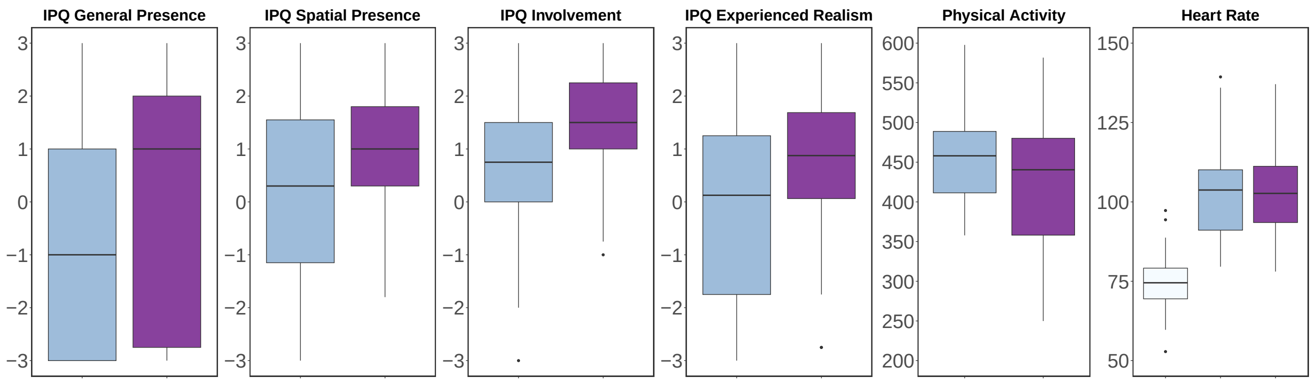

Figure 4: The box plots visualize the IPQ subcategories and physical activity measures between *ExerCube* (blue) and *VR-HMD* (purple); players experienced significantly higher presence in all IPQ subcategories in *VR-HMD*, but moved more in the *ExerCube*. The last plot shows the heart rate at baseline (white) and in the game versions; both setups lead to comparable, yet significantly increased exertion.

Table 4: The table presents the descriptive and statistical values for game performance, physical activity, and heart rate metrics.

Condition	Correct Flakes *Mean (SD)*	Wrong Flakes *Mean (SD)*	Missed Flakes *Mean (SD)*	Score *Mean (SD)*	Physical Activity *Mean (SD)*	Heart Rate *Mean (SD)*
Baseline	–	–	–	–	–	75.45 (9.35)
ExerCube	82.59 (10.00)	1.12 (1.82)	4.91 (3.57)	14729.12 (2438.36)	460.82 (62.91)	102.56 (14.82)
VR-HMD	78.53 (14.69)	1.62 (3.22)	6.59 (6.51)	13670.68 (3387.34)	429.07 (85.45)	104.47 (14.94)
Test	t(33)=2.279	V=33	V=151.5	t(33)=2.625	t(33)=4.713	F(1.28, 42.22)=137.977
p	**p=0.029**	p=0.398	p=0.156	**p=0.013**	**p<0.001**	**p<0.001**
Effect	d=0.391	r=-0.102	r=-0.172	d=0.450	d=0.808	η_P^2 =0.807

d=0.391), score (t(33)=2.625, p=0.013, d=0.450), and physical activity (t(33)=4.713, p<0.001, d=0.808, see Figure 4.

6.2.2 Heart Rate Metrics. We found a significant effect of the time points on heart rate metrics (F(1.28, 42.22)=137.977 (Greenhouse Geisser corrected), p<0.001, η_P^2 =0.807). The post-hoc tests indicated that players had a lower heart rate in the baseline measurement compared to *ExerCube* (t(33)=-11.792, p<0.001) as well as *VR-HMD* (t(33)=-12.704, p<0.001). But, there was no significant difference between *ExerCube* and *VR-HMD* setups (t(33)=-1.959, p=0.176, see Figure 4.

6.3 Qualitative Findings

6.3.1 Theme 1: Both game versions were very well received, but the reasons for preferring either setup differed fundamentally. Regardless of the display option, the majority of participants reacted very positively to the game: "*My only piece of advice is to keep going, that's a fantastic thing*"-P_{31}. Both *ExerCube* and *VR-HMD* were almost equally favored by participants. Half of the participants who chose the *ExerCube* attributed their decision to the ergonomics of the VR headset (e.g., cable, weight, wearing VR headset with glasses). Another notable reason was safety; in the *ExerCube* version "*[they] can see where the walls are*"-P_{16}, leading them to feel safer while moving. Interestingly, one participant expressed concerns about losing control and connection to reality: "*I liked the* ExerCube *better because I didn't really lose the connection, whereas with VR I'm afraid of losing it, a kind of loss of control. And I don't like that [...] It was simply drawing you in-P_3.*" A few noted that their preference stemmed from their familiarity with the *ExerCube*. Similarly, one participant stated that playing the game in the previous condition influenced their decision (*ExerCube*), as they felt more experienced.

In contrast, the novelty of VR technology was the reason for preferring the *VR-HMD* version for only a few participants: "*A completely different experience, a physical experience too*"-P_{14}. Most importantly, however, participants who selected VR connected their decision to technical immersion features, such as presence and spatiality: "*I really felt like I was in this snowy landscape*"-P_{28}. These points were complemented by a few participants commenting on the higher visual fidelity, clear vision, and realism in VR: "*Because you could see everything much more clearly and all the individual figures or trees or something like that were much more pronounced*"-P_1.

6.3.2 Theme 2: Motivation and Immersion can be influenced by familiarity, ergonomics, and perceptual effects. For almost half of the participants, "*the motivation was the same*"-P_2 across the displays. Two players who found the different game variants more motivating emphasized the role of getting familiar with the game: "*I found it a bit higher in the second version, but maybe that had something to do with the fact that you already knew the game a bit*"-P_{10}. For a few, the *ExerCube* was more motivating because they were familiar with the setup: "*I think that has to do with the fact that I do* ExerCube *once a week*"-P_{20}. Similarly, a couple of participants found "*[the* ExerCube*] is easier*"-P_{19} to play. We saw the effect of ergonomics again (e.g., discomfort and safety concerns due to wearing a VR headset), but this time on motivation: "*So I felt a bit more insecure, of course, because i didn't know how the room was confined at all*"-P_{22}. While familiarity was a reason for finding the *ExerCube* more motivating, there were a few participants higlighting the positive impact of the novelty of *VR-HMD* on motivation: "*What motivated me more? [...] Because it was new*"-P_5. Aligning with the preference reasonings, individual participants reported unique motivational features of VR, such as visual fidelity, realism, and presence: "*The graphics were somehow more vivid and you could see it better*"-P_{32}.

A quarter of participants felt that playing on different screens did not affect immersion: "*It [immersion] was actually the same*"-P_8. Again, VR-headset-related issues (e.g., discomfort, safety, and limited field of view) had an impact, but this time on immersion, leading participants to find the *ExerCube* more immersive: "*The glasses prevented me from concentrating even more intensely*"-P_{18}. However, overwhelmingly, for more than half of the participants, the *VR-HMD* was more immersive. For many, this was related to the feeling of being physically there: "*I felt like I was freezing because it was in winter*"-P_{33}. Additionally, a few connected this to visual fidelity and realism: "*it was nicer because the background was clearer*"-P_{30}. Yet, participants found the *VR-HMD* version more challenging, which led them to concentrate more: "*[...] you have to concentrate more, including myself, also in terms of movement, than in the normal environment*"-P_6. A few individuals emphasized the positive role of the novelty of *VR-HMD*s in feeling present. However, there were opposing perspectives on how learning effects contribute to presence; while for two, being familiar with how to play the game had a negative influence on how they felt present in the game, for one participant "*it was an advantage*"-P_{22} since they were already familiar with the controls, they could focus more on being in the experience.

6.3.3 Theme 3: Feeling of safety, connection to the real world, and VR-display-related challenges can impact physical activity. More than half of the participants did not feel a difference in terms of physical activity between the conditions. But, some felt more physically active in connection with feeling safe in the *ExerCube*: "*It was higher in the* ExerCube *because you do feel more secure on your feet*"-P_3. Interestingly, despite the same number of snowflakes and the same distance to reach them in both game setups, two participants felt moving more in the *ExerCube*: "*I have to walk more. I probably walked in the other one too, but somehow I didn't notice it so directly*"-P_{13}. Moreover, a few participants stated that their familiarity with *ExerCube* and the game (i.e., learning effects) contributed to them feeling more active. In contrast, for a few, the *VR-HMD* gave them the feeling of being more active due to the low field of view: "*I had to move my head a lot more to see where the snowflakes were coming from*"-P_{28}. Similarly, a few found the *VR-HMD* version was more strenuous because "*you had to orientate yourself more and see where you were or that you were standing*"-P_{26}. Only one noted that in VR the game felt faster and required concentration, which lead them to move more.

7 Discussion

In this study, we focused on older adults' experiences with two exergame systems that differ in technical immersion. Specifically, we investigated the research question: *How does the technical immersion of an exergame system (i.e., a VR-HMD vs. ExerCube) influence older adults' exergaming experience, motivation, cybersickness, physical activity enjoyment, workload, game performance, physical exertion (heart rate) and physical activity?* Here, we discuss the findings and provide implications for the design and research of future exergames for the older population.

7.1 Safety: Cybersickness & Discomfort

Our quantitative results indicate that participants did not experience severe cybersickness symptoms [63] after playing our game in both setups. Interestingly, we observed a significant reduction in Total SSQ and oculomotor disturbance scores from pre- to post-*ExerCube*. We speculate that this reduction may be related to the high game enjoyment experienced by the participants. It could also be explained by the older adults' attitude of rating their experiences more positively [45] and having positive opinions after testing VR applications compared to pre-use [30].

Our qualitative findings provide nuanced insights into other potential drawbacks of exergame technologies: In line with previous research [31], several participants expressed concerns related to *VR-HMD* ergonomics, such as discomfort from wearing a headset and having a cable attached to the HMD. Since these aspects affect in-game movements, they probably contributed to the significantly higher physical activity in the *ExerCube* condition. To have comparable conditions for our study, we used the same tracking system for both systems resulting in the use of a cable-bound *VR-HMD*. In our case, this configuration did not introduce safety issues, as we monitor the participants at all times during the game. But, we also note that having a person monitor older adults during such exercise routines can increase the workload of senior living facilities [39]. Therefore, when exercising alone, we advocate for standalone headsets to avoid the risk of injury. Although this recommendation is valid for all exergame users, it applies particularly to older adults due to their higher risk of injuries and slower reaction time [25, 34]. In general, while our results emphasize that our game, regardless of the type of display, can be used for physical activity without negatively affecting the well-being of older adults, they also suggest that further research is needed to understand how specific features of exergame technologies (e.g., improved ergonomics) affect player experience and performance.

7.2 Technical Immersion: No impact on player experience

In line with a previous study on CAVE vs. *VR-HMD* exergames [64], our findings reveal that the *VR-HMD* evoked a significantly higher presence than the *ExerCube*. Participants felt like in a winter wonderland and praised the technical quality of the display. However, challenging established assumptions, the higher presence did not translate to significantly higher player enjoyment (i.e., motivation) or player experience. We build on the work by Born et al. [4], who found that younger players felt more motivated when playing immersive (i.e., *VR-HMD*) than non-immersive (i.e., TV-based setups) exergames. Our findings add that, for older adults, we could not find any significant difference in motivation between immersive *VR-HMD* and non-immersive *ExerCube* setups.

Moreover, none of the player experience constructs demonstrated significant differences, which diverges from the work by Elor et al. [19]. We show that their findings of improved player experience with more immersive environments for younger adults are not transferrable to the comparison between *ExerCube* and *VR-HMD* for older adults. While this disparity highlights the nuanced impact of the factor of age on how players experience exergames, our research can only be a first step towards a complete picture. All cited studies were conducted in the Western hemisphere and only focused on a limited subset of available technologies. Consequently, we recommend further research to explore the impact of different demographics and technologies, while also comparing exergames against real-world options such as video-based exercises [45] and personal trainer-guided exercise sessions [51].

7.3 Performance: Higher game performance and physical activity in the ExerCube

While participants enjoyed the physical activity in both conditions comparably, the *ExerCube* outperformed the *VR-HMD* in terms of game performance, physical activity, perceived performance, and frustration. Previous research showed positive results regarding the impact of technical immersion on game performance: Born et al. [4] found that younger adults achieved significantly better game performance in the *VR-HMD* condition compared to the TV-based condition. Elor et al. [19] reported that higher technical immersion (*VR-HMD* vs. CAVE) led to better game performance. Our results diverge from the work of both Born et al. [4] and Elor et al. [19]. Accordingly, we believe that a purpose-oriented design—the *ExerCube* was designed for gamified exercise [50] while the used *VR-HMD* is a general purpose device—is more important than high technical immersion. Speculating about the reasons for this finding, we believe that a major cause lies within the perceived safety: wearing a VR headset decoupled players from their own body (i.e., no proprioception [79]) and their physical surroundings. Therefore, relying on previous work [54], we suspect that they were more cautious about movements and reduced their movements during the *VR-HMD* condition.

Regarding physical exertion, we did not find any significant differences in heart rate metrics between conditions. However, we speculate that increased movements in the *ExerCube* version may translate into greater physical exertion with longer play durations and aim to explore this aspect in future studies. We also observed that both game versions resulted in significantly higher exertion compared to the participants' baseline heart rate. We conclude that the game, regardless of the type of display, promotes physical activity that may improve the physical well-being of older adults.

7.4 Research & Design Implications

We provide implications, guiding future work on researching and developing exergames for older adults.

7.4.1 Novelty may not lead to better experiences in older adults. Many participants emphasized the positive effect of familiarity with the setup, the game, and the learning effects from experiencing the same gameplay in both conditions. While many studies with young adults show that novelty can add to the player experience [40, 72], our study suggests that older adults might approach new technologies cautiously and need some familiarization. Since older adults grew up in a non-digital world [52] and previous work found that they might be hesitant towards new technologies [30], we conclude that novelty might not benefit the experience. Furthermore, older adults typically require more time to adapt to new technology, and their unfamiliarity with the game may have affected their performance, which can make them feel less competent [8]. This can discourage them from using such innovative solutions for exercising. Therefore, we underline the importance of ensuring familiarity with technologies and stimuli to preserve older adults' well-being and reduce their frustration. For example, game designers should provide more extensive trial and acclimatization periods that go beyond the interactive tutorials we provided before the actual game.

7.4.2 Redesign VR headsets to fit older adults' ergonomic needs. Older adults reported having discomfort while wearing a *VR-HMD* and the qualitative data indicates that this discomfort impacted their experience, leading to feelings of insecurity and reduced mobility. Consistent with findings from previous studies that report on the ergonomic discomfort associated with the use of *VR-HMD*s for general purposes [31], our research identified similar issues for exergames. Thus, we emphasize that current commercial VR headsets, especially the model utilized in our study, require improvements in design to accommodate age-related factors better. Issues such as the headset's weight and the comfort of wearing it with glasses are particularly relevant for older users [39]. With increased movement during exergame sessions, these problems and their severity typically worsen. For instance, glasses are prone to fogging. These issues might discourage older adults from using *VR-HMD* setups and create a barrier to the potential benefits of *VR-HMD*-based exergames.

7.4.3 Consider VR-HMDs' perceptually more difficult nature for older adults. In our study, older adults performed worse in the *VR-HMD* condition. This finding contrasts prior work, which showed that younger users perform better in more immersive environments [19, 85]. Looking at our qualitative insights, we see that older adults felt more challenged and had spatial difficulties in *VR-HMD*. We think that some features of *VR-HMD*s make the game inherently more challenging than *ExerCube*: In the *ExerCube*, the virtual world is displayed on three flat panes in the real environment, which enables players to use their natural depth perception. In contrast, the *VR-HMD* replaces the players' view in the real-world completely with a virtual view. It is well known that depth perception in VR is significantly worse than in real life [18]. Applied to the game task of interacting with snowflakes, this technological difference translates to a higher necessary cognitive effort. According to Anders et al. [3], even simple exergames that do not target cognitive exercises feature cognitive processing. Our game includes dedicated cognitive tasks and combines it with the perceptual challenges of VR. Furthermore, the *VR-HMD* only has a limited field of view, whereas in the *ExerCube*, participants retain their natural field of view. While this may not pose a challenge for young adults [19], older adults [25, 34] typically have reduced reaction time and motor skills, which requires them to put more effort into perceiving the entire virtual environment. Altogether, the cognitive task, target user group, and technological differences may further explain differences in game performance between the conditions. This assumption is supported by a tendency in the NASA-TLX mental effort subscale, showing a higher, yet non-significant, effort for the *VR-HMD* condition.

7.4.4 Balance the benefits and drawbacks of different technologies. Our results show that both exergame technologies offer a good player experience and induce physical exertion that can benefit the well-being of older adults. However, we also identified important differences. Instead of having a clearly superior choice, each system offers benefits and disadvantages that should be evaluated carefully based on the individual use case:

The *ExerCube* requires significantly more physical space and financial investment compared to using a *VR-HMD*. Yet, it also encouraged older adults to be more active and feel safer while moving. Considering the vital role of physical activity for older adults [84], we recommend utilizing the *ExerCube* to create safe exercise opportunities for them, provided that financial resources and space are available. Conversely, *VR-HMD*s serve as an appealing and affordable option for older adults interested in engaging with innovative technology without sacrificing player experience and still experiencing exertion. Since *VR-HMD*s provide a higher sense of presence, do not require complex room setups, and are more cost-effective, they may be a more appealing choice for older adults who prioritize these factors. Also, we emphasize the importance of presence for other purposes than exergame enjoyment. For example, it might be used to distract users from painful sensations during the exercise as shown in previous work on pain relief [48]. However, *VR-HMD*s' use also raises safety concerns because older adults may feel insecure due to discomfort and disconnection from the real world. Ultimately, both configurations have unique advantages that cater to specific user preferences.

7.5 Limitations

We conducted our study with the help of the senior living facility that provided access to the *ExerCube*, but of course, it also influenced our sample: many participants had prior experience with the *ExerCube* since it is also used for physical exercise in the facility. However, none of the participants saw our game during any stage of development before our study. Also, all participants were from Germany. Research indicates that socio-cultural factors are important to consider in exergame design [53]. Thus, we recommend conducting further research with a focus on these aspects.

Like any study involving game-based research, our study is subject to limitations related to the design of our game. To ensure a reproducible and insightful study, we put emphasis on the design (e.g., the theme [39], audiovisual congruence [1], having collection tasks that are common in exergames [37]) and precise calibration (i.e., identical mapping in both conditions). However, this rigor reduces the study to a single game, which might limit the generalizability. Hence, while we recommend conducting future work that considers the impact of different game designs/tasks (see [37] for an overview), we note that the evaluation of multiple games could be challenging as the duration of the study may lead to fatigue effects. Moreover, there are inherent features of each setup (e.g., VR's limited field of view, the lower resolution of the *ExerCube*) that we cannot control and may have affected the outcomes of our study.

Finally, the researcher who conducted the qualitative analysis possesses unique strengths for interpreting the data, including a psychological and technical background in designing and implementing VR games (see the positionality statement in the qualitative methodology section), as well as conducting studies, particularly with older adults. While this background facilitates a focused interpretation of the data in these areas, they do not have professional expertise in analyzing qualitative data through the lens of movement science, despite their involvement in exercise and sports.

8 Conclusion

Our study of exergame technologies for older adults reveals that both *VR-HMD* and *ExerCube* systems offer high levels of game and physical activity enjoyment. While *VR-HMD* systems excel in creating a sense of presence, the *ExerCube* system leads to better game performance and more physical activity. Nevertheless, both display technologies provide physical exertion that can potentially benefit older adults’ well-being. Overall, our results challenge the assumption that more immersive technologies always lead to better outcomes in exergaming contexts. We provide four research and design implications that can support efforts to create more inclusive, effective, and enjoyable exergaming experiences that promote physical activity among older adults.

Acknowledgments

We thank the Hospital zum Heiligen Geist Hamburg and our participants for their participation in this research project. This work was funded under the Excellence Strategy of the Federal Government and the Länder. Additionally, this project was supported by the European Union’s Horizon Europe research and innovation program (grant number: 101135025), the grant (number: 16SV8878) from the German Federal Ministry of Education and Research (BMBF), the CIHR AAL EXERGETIC operating grant (number: 02237-000), and the Provost’s Program for Interdisciplinary Postdoctoral Scholars at the University of Waterloo.

References

[1] Vero Vanden Abeele, Brenda Schraepen, Hanne Huygelier, Celine Gillebert, Kathrin Gerling, and Raymond Van Ee. 2021. Immersive Virtual Reality for Older Adults: Empirically Grounded Design Guidelines. *ACM Trans. Access. Comput.* 14, 3, Article 14 (aug 2021), 30 pages. doi:10.1145/3470743 pages #3, #4, #12

[2] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370. doi:10.1016/j.ijhcs.2019.102370 pages #6

[3] Phillipp Anders, Tim Lehmann, Helen Müller, Karoline B. Grønvik, Nina Skjæret-Maroni, Jochen Baumeister, and Beatrix Vereijken. 2018. Exergames Inherently Contain Cognitive Elements as Indicated by Cortical Processing. *Frontiers in Behavioral Neuroscience* 12 (2018). doi:10.3389/fnbeh.2018.00102 pages #12

[4] Felix Born, Sophie Abramowski, and Maic Masuch. 2019. Exergaming in VR: The Impact of Immersive Embodiment on Motivation, Performance, and Perceived Exertion. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, 1–8. doi:10.1109/VS-Games.2019.8864579 pages #2, #3, #4, #6, #11

[5] Robert Bowman, Camille Nadal, Kellie Morrissey, Anja Thieme, and Gavin Doherty. 2023. Using Thematic Analysis in Healthcare HCI at CHI: A Scoping Review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI ’23)*. Association for Computing Machinery, New York, NY, USA, Article 491, 18 pages. doi:10.1145/3544548.3581203 pages #8

[6] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. doi:10.1080/2159676X.2019.1628806 pages #8

[7] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47. doi:10.1002/capr.12360 pages #8

[8] Johnny Salazar Cardona, Jeferson Arango Lopez, Francisco Luis Gutiérrez Vela, and Fernando Moreira. 2023. Meaningful learning: motivations of older adults in serious games. *Universal Access in the Information Society* (2023), 1–16. doi:10.1007/s10209-023-00987-y pages #11

[9] Cheng Chen, Susanne Weyland, Julian Fritsch, Alexander Woll, Claudia Niessner, Alexander Burchartz, Steffen C. E. Schmidt, and Darko Jekauc. 2021. A Short Version of the Physical Activity Enjoyment Scale: Development and Psychometric Properties. *International Journal of Environmental Research and Public Health* 18, 21 (2021), 15 pages. doi:10.3390/ijerph182111035 pages #6

[10] Victoria Clarke and Virginia Braun. 2022. Thematic analysis: A practical guide. *Thematic Analysis* (2022), 1–100. pages #8

[11] Sebastian Cmentowski, Sukran Karaosmanoglu, Lennart E. Nacke, Frank Steinicke, and Jens Harald Krüger. 2023. Never Skip Leg Day Again: Training the Lower Body with Vertical Jumps in a Virtual Reality Exergame. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI ’23)*. Association for Computing Machinery, New York, NY, USA, Article 786, 18 pages. doi:10.1145/3544548.3580973 pages #3

[12] John W Creswell and Vicki L Plano Clark. 2017. *Designing and conducting mixed methods research*. Sage publications, Thousand Oaks, California, United States. pages #2, #7

[13] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. 1993. Surround-screen projection-based virtual reality: the design and implementation of the CAVE. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques* (Anaheim, CA) *(SIGGRAPH ’93)*. Association for Computing Machinery, New York, NY, USA, 135–142. doi:10.1145/166117.166134 pages #3

[14] DeepL SE. 2017. DeepL Translator. DeepL Pro version. Cologne, Germany. https://www.deepl.com/pro?cta=header-pro/. pages #8

[15] Kristen K. Dergousoff, Regan L. Mandryk, and Kathrin Gerling. 2013. Is movement better? Comparing sedentary and motion-based game controls for older adults. (1 2013). https://repository.lincoln.ac.uk/articles/conference_contribution/Is_movement_better_Comparing_sedentary_and_motion-based_game_controls_for_older_adults/25166216 pages #2, #3

[16] Andrew T. Dilanchian, Ronald Andringa, and Walter R. Boot. 2021. A Pilot Study Exploring Age Differences in Presence, Workload, and Cybersickness in the Experience of Immersive Virtual Reality Environments. *Frontiers in Virtual Reality* 2 (2021). doi:10.3389/frvir.2021.736793 pages #2, #3

[17] Dovetail Research PTY. LTD. 2016. Dovetail. Surry Hills, Australia. https://dovetailapp.com/. pages #7

[18] Fatima El Jamiy and Ronald Marsh. 2019. Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality. *IET Image Processing* 13, 5 (2019), 707–712. doi:10.1049/iet-ipr.2018.5920 pages #12

[19] Aviv Elor, Michael Powell, Evanjelin Mahmoodi, Nico Hawthorne, Mircea Teodorescu, and Sri Kurniawan. 2020. On Shooting Stars: Comparing CAVE and HMD Immersive Virtual Reality Exergaming for Adults with Mixed Ability. *ACM Trans. Comput. Healthcare* 1, 4, Article 22 (sep 2020), 22 pages. doi:10.1145/3396249 pages #3, #4, #11, #12

[20] FitXR. 2019. *FitXR - Box. HIIT. Dance. Sculpt. Combat.* Game [VR]. FitXR. London, England. Meta Reviews: https://www.meta.com/experiences/2327205800645550/. pages #2

[21] International Organization for Standardization. 2019. ISO 9241-210:2019 Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. https://www.iso.org/standard/77520.html International Organization for Standardization. Geneva, Switzerland. pages #2, #4

[22] Julian Fritsch, Susanne Weyland, Katharina Feil, Alexander Burchartz, Steffen Schmidt, Alexander Woll, Ulrich Strauch, Benjamin Wienke, and Darko Jekauc. 2022. A Study on the Psychometric Properties of the Short Version of the Physical Activity Enjoyment Scale in an Adult Population. *International Journal of Environmental Research and Public Health* 19, 22 (2022), 11 pages. doi:10.3390/ijerph192215294 pages #6

[23] Borko Furht (Ed.). 2008. *Immersive Virtual Reality*. Springer US, Boston, MA, 345–346. doi:10.1007/978-0-387-78414-4_85 pages #3

[24] Beat Games. 2019. *Beat Saber*. Game [VR]. Beat Games. Los Angeles, CA, USA. https://beatsaber.com/. Steam Reviews: https://store.steampowered.com/app/620980/Beat_Saber/. Accessed: 16.06.2023. pages #3

[25] Kathrin Maria Gerling, Jonas Schild, and Maic Masuch. 2010. Exergame Design for Elderly Users: The Case Study of SilverBalance. In *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology* (Taipei, Taiwan) *(ACE ’10)*. Association for Computing Machinery, New York, NY, USA, 66–69. doi:10.1145/1971630.1971650 pages #3, #4, #11, #12

[26] Linda Graf, Maximilian Altmeyer, Katharina Emmerich, Marc Herrlich, Andrey Krekhov, and Katta Spiel. 2022. Development and Validation of a German Version of the Player Experience Inventory (PXI). In *Proceedings of Mensch Und Computer 2022* (Darmstadt, Germany) *(MuC ’22)*. Association for Computing Machinery, New York, NY, USA, 265–275. doi:10.1145/3543758.3543763 pages #6

[27] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. doi:10.1177/154193120605000909 pages #6

[28] Günnur Hastürk and Müge Akyıldız Munusturlar. 2022. The Effects of Exergames on Physical and Psychological Health in Young Adults. *Games for Health Journal* 11, 6 (2022), 425–434. doi:10.1089/g4h.2022.0093 pages #3

[29] Robert J. Havighurst. 1961. Successful Aging. *The Gerontologist* 1, 1 (03 1961), 8–13. doi:10.1093/geront/1.1.8 pages #3

[30] David Healy, Aisling Flynn, Owen Conlan, Jenny McSharry, and Jane Walsh. 2022. Older Adults’ Experiences and Perceptions of Immersive Virtual Reality: Systematic Review and Thematic Synthesis. *JMIR Serious Games* 10, 4 (6 Dec 2022), e35802. doi:10.2196/35802 pages #11

[31] Teresa Hirzle, Maurice Cordts, Enrico Rukzio, Jan Gugenheimer, and Andreas Bulling. 2021. A Critical Assessment of the Use of SSQ as a Measure of General Discomfort in VR Head-Mounted Displays. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI ’21)*.

Association for Computing Machinery, New York, NY, USA, Article 530, 14 pages. doi:10.1145/3411764.3445361 pages #3, #11, #12

[32] Hunter G. Hoffman, Clayton C. Seibel, Lyla Coron, Laura E. Simons, Sydney Drever, Sylvie Le May, Keira P. Mason, and Herta Flor. 2024. Increasing presence via a more immersive VR system increases virtual reality analgesia and draws more attention into virtual reality in a randomized crossover study. *Frontiers in Virtual Reality* 5 (2024). doi:10.3389/frvir.2024.1452486 pages #4

[33] Anne Hösch. 2018. *Simulator Sickness in Fahrsimulationsumgebungen-drei Studien zu Human Factors.* Ph. D. Dissertation. Technische Universität Ilmenau. https://www.db-thueringen.de/receive/dbt_mods_00037865 pages #6

[34] Wijnand Ijsselsteijn, Henk Herman Nap, Yvonne de Kort, and Karolien Poels. 2007. Digital game design for elderly users. In *Proceedings of the 2007 Conference on Future Play* (Toronto, Canada) *(Future Play '07)*. Association for Computing Machinery, New York, NY, USA, 17–22. doi:10.1145/1328202.1328206 pages #3, #11, #12

[35] Christos Ioannou, Patrick Archard, Eamonn O'Neill, and Christof Lutteroth. 2019. Virtual Performance Augmentation in an Immersive Jump & Run Exergame. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3290605.3300388 pages #6

[36] M. Carmen Juan and David Pérez. 2009. Comparison of the Levels of Presence and Anxiety in an Acrophobic Environment Viewed via HMD or CAVE. *Presence: Teleoperators and Virtual Environments* 18, 3 (06 2009), 232–248. doi:10.1162/pres.18.3.232 arXiv:https://direct.mit.edu/pvar/article-pdf/18/3/232/1624916/pres.18.3.232.pdf pages #4

[37] Sukran Karaosmanoglu, Sebastian Cmentowski, Lennart E. Nacke, and Frank Steinicke. 2024. Born to Run, Programmed to Play: Mapping the Extended Reality Exergames Landscape. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, 28 pages. doi:10.1145/3613904.3642124 pages #2, #3, #4, #6, #12

[38] Sukran Karaosmanoglu, Lucie Kruse, Sebastian Rings, and Frank Steinicke. 2022. Canoe VR: An Immersive Exergame to Support Cognitive and Physical Exercises of Older Adults. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 342, 7 pages. doi:10.1145/3491101.3519736 pages #2, #3, #4, #5, #6

[39] Sukran Karaosmanoglu, Sebastian Rings, Lucie Kruse, Christian Stein, and Frank Steinicke. 2021. Lessons Learned from a Human-Centered Design of an Immersive Exergame for People with Dementia. *Proc. ACM Hum.-Comput. Interact.* 5, CHI PLAY, Article 252 (oct 2021), 27 pages. doi:10.1145/3474679 pages #2, #3, #4, #11, #12

[40] Sukran Karaosmanoglu, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. 2021. Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 443, 15 pages. doi:10.1145/3411764.3445492 pages #6, #11

[41] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology* 3, 3 (1993), 203–220. doi:10.1207/s15327108ijap0303_3 pages #6

[42] Sascha Ketelhut, Lisa Röglin, Anna Lisa Martin-Niedecken, Claudio R. Nigg, and Kerstin Ketelhut. 2022. Integrating Regular Exergaming Sessions in the ExerCube into a School Setting Increases Physical Fitness in Elementary School Children: A Randomized Controlled Trial. *Journal of Clinical Medicine* 11, 6 (2022), 11 pages. doi:10.3390/jcm11061570 pages #2, #3

[43] Alexander Kowarik and Matthias Templ. 2016. Imputation with the R Package VIM. *Journal of Statistical Software* 74, 7 (2016), 1–16. doi:10.18637/jss.v074.i07 pages #7

[44] Lucie Kruse, Sukran Karaosmanoglu, Sebastian Rings, Benedikt Ellinger, and Frank Steinicke. 2021. Enabling Immersive Exercise Activities for Older Adults: A Comparison of Virtual Reality Exergames and Traditional Video Exercises. *Societies* 11, 4 (2021). doi:10.3390/soc11040134 pages #3

[45] Lucie Kruse, Sukran Karaosmanoglu, Sebastian Rings, and Frank Steinicke. 2022. Evaluating Difficulty Adjustments in a VR Exergame for Younger and Older Adults: Transferabilities and Differences. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction* (Online, CA, USA) *(SUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 7, 11 pages. doi:10.1145/3565970.3567684 pages #2, #3, #4, #11

[46] Eugene Kukshinov. 2024. It's (not) Me: Dynamic Nature of Immersive Experiences in Video Game Play. *Human Studies* (2024), 1–22. doi:10.1007/s10746-024-09768-9 pages #3

[47] Christopher Sze Chong Lim. 2010. Designing inclusive ICT products for older users: taking into account the technology generation effect. *Journal of Engineering Design* 21, 2-3 (2010), 189–206. doi:10.1080/09544820903317001 pages #2

[48] Christopher V Maani, Hunter G Hoffman, Michelle Morrow, Alan Maiers, Kathryn Gaylord, Laura L McGhee, and Peter A DeSocio. 2011. Virtual reality pain control during burn wound debridement of combat-related burn injuries using robot-like arm mounted VR goggles. *Journal of Trauma and Acute Care Surgery* 71, 1 (2011), S125–S130. doi:10.1097/TA.0b013e31822192e2 pages #12

[49] Anna Lisa Martin-Niedecken, Andrea Mahrer, Katja Rogers, Eling D. de Bruin, and Alexandra Schättin. 2020. "HIIT" the ExerCube: Comparing the Effectiveness of Functional High-Intensity Interval Training in Conventional vs. Exergame-Based Training. *Frontiers in Computer Science* 2 (2020). doi:10.3389/fcomp.2020.00033 pages #2, #3

[50] Anna Lisa Martin-Niedecken and Elisa D. Mekler. 2018. The ExerCube: Participatory Design of an Immersive Fitness Game Environment. In *Serious Games*, Stefan Göbel, Augusto Garcia-Agundez, Thomas Tregel, Minhua Ma, Jannicke Baalsrud Hauge, Manuel Oliveira, Tim Marsh, and Polona Caserman (Eds.). Springer International Publishing, Cham, 263–275. pages #2, #3, #11

[51] Anna Lisa Martin-Niedecken, Katja Rogers, Laia Turmo Vidal, Elisa D. Mekler, and Elena Márquez Segura. 2019. ExerCube vs. Personal Trainer: Evaluating a Holistic, Immersive, and Adaptive Fitness Game Setup. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3290605.3300318 pages #2, #3, #11

[52] Karyn Moffatt. 2013. Older-adult HCI: why should we care? *Interactions* 20, 4 (July 2013), 72–75. doi:10.1145/2486227.2486242 pages #2, #11

[53] Reza Hadi Mogavi, Juhyung Son, Simin Yang, Derrick M Wang, Lydia Choong, Ahmad Alhilal, Peng Yuan Zhou, Pan Hui, and Lennart E Nacke. 2024. The Jade Gateway to Exergaming: How Socio-Cultural Factors Shape Exergaming Among East Asian Older Adults. *arXiv preprint arXiv:2407.10053* (2024). pages #2, #3, #4, #12

[54] Betty Mohler, Jennifer Campos, Michael Weyel, and Heinrich H Bülthoff. 2007. Gait parameters while walking in a head-mounted display virtual environment and the real world. In *13th Eurographics Symposium on Virtual Environments and 10th Immersive Projection Technology Workshop (IPT-EGVE 2007)*. Eurographics Association, 85–88. https://hdl.handle.net/11858/00-001M-0000-0013-CCBB-0 pages #11

[55] Anna Moschny, Petra Platen, Renate Klaaßen-Mielke, Ulrike Trampisch, and Timo Hinrichs. 2011. Barriers to physical activity in older adults in Germany: a cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity* 8, 1 (2011), 1–10. doi:10.1186/1479-5868-8-121 pages #2, #3

[56] Florian 'Floyd' Mueller, Darren Edge, Frank Vetere, Martin R. Gibbs, Stefan Agamanolis, Bert Bongers, and Jennifer G. Sheridan. 2011. Designing Sports: A Framework for Exertion Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2651–2660. doi:10.1145/1978942.1979330 pages #2, #3

[57] Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand, and Jon Howard Raphael. 2012. No room for reflexivity? Critical reflections following a systematic review of qualitative research. *Journal of Health Psychology* 17, 6 (2012), 866–885. doi:10.1177/1359105311427615 pages #8

[58] Junhyuck Park and JongEun Yim. 2016. A new approach to improve cognition, muscle strength, and postural balance in community-dwelling elderly with a 3-D virtual reality kayak program. *The Tohoku journal of experimental medicine* 238, 1 (2016), 1–8. pages #3

[59] Nathalie Ringgenberg, Sarah Mildner, Marcia Hapig, Sarah Hermann, Katharina Kruszewski, Anna Lisa Martin-Niedecken, Katja Rogers, Alexandra Schättin, Frank Behrendt, Sonja Böckler, et al. 2022. ExerG: adapting an exergame training solution to the needs of older adults using focus group and expert interviews. *Journal of NeuroEngineering and Rehabilitation* 19, 1 (2022), 89. doi:10.1186/s12984-022-01063-x pages #2, #4

[60] Dori Rosenberg, Colin A Depp, Ipsit V Vahia, Jennifer Reichstadt, Barton W Palmer, Jacqueline Kerr, Greg Norman, and Dilip V Jeste. 2010. Exergames for subsyndromal depression in older adults: a pilot study of a novel intervention. *The American Journal of Geriatric Psychiatry* 18, 3 (2010), 221–226. doi:10.1097/JGP.0b013e3181c534b5 pages #2, #3

[61] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion* 30, 4 (2006), 344–360. doi:10.1007/s11031-006-9051-8 The IMI items are available online under this link: http://selfdeterminationtheory.org/intrinsic-motivation-inventory/. Accessed: 02.09.2022. pages #6

[62] Sankaran Nair Sara J. Czaja, Joseph Sharit and Mark Rubert. 1998. Understanding sources of user variability in computer-based data entry performance. *Behaviour & Information Technology* 17, 5 (1998), 282–293. doi:10.1080/014492998119355 pages #3

[63] Dimitrios Saredakis, Ancret Szpak, Brandon Birckhead, Hannah A. D. Keage, Albert Rizzo, and Tobias Loetscher. 2020. Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis. *Frontiers in Human Neuroscience* 14 (2020), 17 pages. doi:10.3389/fnhum.2020.00096 pages #10

[64] Steven Schmidt, Patrick Ehrenbrink, Benjamin Weiss, Jan-Niklas Voigt-Antons, Tanja Kojic, Andrew Johnston, and Sebastian Möller. 2018. Impact of Virtual Environments on Motivation and Engagement During Exergames. In *2018 Tenth*

International Conference on Quality of Multimedia Experience (QoMEX). 1–6. doi:10.1109/QoMEX.2018.8463389 pages #3, #4, #11
[65] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (06 2001), 266–281. doi:10.1162/105474601300343603 pages #6
[66] Jeff Sinclair, Philip Hingston, and Martin Masek. 2009. Exergame development using the dual flow model. In *Proceedings of the Sixth Australasian Conference on Interactive Entertainment* (Sydney, Australia) *(IE '09)*. Association for Computing Machinery, New York, NY, USA, Article 11, 7 pages. doi:10.1145/1746050.1746061 pages #3
[67] Nina Skjæret, Ather Nawaz, Tobias Morat, Daniel Schoene, Jorunn Lægdheim Helbostad, and Beatrix Vereijken. 2016. Exercise and rehabilitation delivered through exergames in older adults: An integrative review of technologies, safety and efficacy. *International Journal of Medical Informatics* 85, 1 (2016), 1–16. doi:10.1016/j.ijmedinf.2015.10.008 pages #2, #3
[68] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557. doi:10.1098/rstb.2009.0138 pages #3
[69] Mel Slater and Maria V. Sanchez-Vives. 2016. Enhancing Our Lives with Immersive Virtual Reality. *Frontiers in Robotics and AI* 3 (2016), 47 pages. doi:10.3389/frobt.2016.00074 pages #2, #3
[70] Frank Steinicke. 2016. *Being really virtual*. Springer, Berlin, Germany. doi:10.1007/978-3-319-43078-2 pages #3
[71] Sruti Subramanian, Yngve Dahl, Nina Skjæret Maroni, Beatrix Vereijken, and Dag Svanæs. 2020. Assessing Motivational Differences Between Young and Older Adults When Playing an Exergame. *Games for Health Journal* 9, 1 (2020), 24–30. doi:10.1089/g4h.2019.0082 pages #3
[72] Philipp Sykownik, Sukran Karaosmanoglu, Katharina Emmerich, Frank Steinicke, and Maic Masuch. 2023. VR Almost There: Simulating Co-located Multiplayer Experiences in Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 789, 19 pages. doi:10.1145/3544548.3581230 pages #11
[73] Jamie L. Tait, Rachel L. Duckham, Catherine M. Milte, Luana C. Main, and Robin M. Daly. 2017. Influence of Sequential vs. Simultaneous Dual-Task Exercise Training on Cognitive Function in Older Adults. *Frontiers in Aging Neuroscience* 9 (2017), 368. doi:10.3389/fnagi.2017.00368 pages #4, #6
[74] Unity Technologies. 2023. Unity Real-Time Development Platform. https://unity.com/ LTS Release 2022.3.14f1. 21 November 2023. https://unity.com/releases/editor/qa/lts-releases?version=2022.3. pages #4
[75] Soraia Teles and Oscar Ribeiro. 2019. *Activity Theory*. Springer International Publishing, Cham, 1–8. doi:10.1007/978-3-319-69892-2_748-1 pages #3
[76] Matthias Templ, Alexander Kowarik, Andreas Alfons, , Gregor de Cillia, Bernd Prantner, and Wolfgang Rannetbauer. 2021. Package 'VIM'. https://cran.r-project.org/web/packages/VIM/VIM.pdf Package 'ARTool' v.6.2.2. pages #7
[77] United Nations, Department of Economic and Social Affairs, Population Division. 2019. World Population Ageing 2019: Highlights (ST/ESA/SER.A/430). doi:10.18356/9df3caed-en pages #4, #6
[78] United Nations, Department of Economic and Social Affairs, Population Division. 2020. World Population Ageing 2019 (ST/ESA/SER.A/444). doi:10.18356/6a8968ef-en pages #4, #6
[79] Irene Valori, Phoebe E. McKenna-Plumley, Rena Bayramova, Claudio Zandonella Callegher, Gianmarco Altoè, and Teresa Farroni. 2020. Proprioceptive accuracy in Immersive Virtual Reality: A developmental perspective. *PLOS ONE* 15, 1 (01 2020), 1–23. doi:10.1371/journal.pone.0222253 pages #11
[80] Black Box VR. 2022. *Black Box VR*. Game [VR]. Black Box VR. Boise, USA. https://www.blackbox-vr.com/. pages #2, #3
[81] Matthias Wilde, Katrin Bätz, Anastassiya Kovaleva, and Detlef Urhahne. 2009. Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften* 15 (2009), 31–45. https://archiv.ipn.uni-kiel.de/zfdn/pdf/15_Wilde.pdf pages #6
[82] Jacob O. Wobbrock and Julie A. Kientz. 2016. Research contributions in human-computer interaction. *Interactions* 23, 3 (April 2016), 38–44. doi:10.1145/2907069 pages #2
[83] World Health Organization. 2022. *Global status report on physical activity 2022*. World Health Organization. https://www.who.int/publications/i/item/9789240059153 (visited 4 September 2023). pages #2
[84] World Health Organization. 2024. Physical Activity. https://www.who.int/news-room/fact-sheets/detail/physical-activity Accessed: 01.09.2024. pages #1, #2, #12
[85] Wenge Xu, Hai-Ning Liang, Kangyou Yu, and Nilufar Baghaei. 2021. Effect of Gameplay Uncertainty, Display Type, and Age on Virtual Reality Exergames. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 439, 14 pages. doi:10.1145/3411764.3445801 pages #2, #3, #4, #12
[86] Wenge Xu, Hai-Ning Liang, Yifan Yu, Diego Monteiro, Khalad Hasan, and Charles Fleming. 2019. Assessing the effects of a full-body motion-based exergame in virtual reality. In *Proceedings of the Seventh International Symposium of Chinese CHI* (Xiamen, China) *(Chinese CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3332169.3333574 pages #3

Supplementary Materials

This document consists of supplementary materials of the paper "*ExerCube vs. Virtual Reality: A Comparative Study of Exergame Technologies for Older Adults*".

Interview Questions

(1a) **German version:** Erinnern Sie sich bitte an die Version zurück, die Ihnen am besten gefallen hat. Was waren die Gründe dafür?

(1b) **English version:** Please think back to the version you liked best. What were the reasons for this?

(2a) **German version:** Haben Sie in irgendeiner Version des Spiels einen Unterschied in Bezug auf die Motivation erlebt? Warum?

(2b) **English version:** Have you experienced a difference in terms of motivation in any version of the game? Why?

(3a) **German version:** Haben Sie in irgendeiner Version des Spiels einen Unterschied in Bezug auf die Immersion erlebt? Warum?

(3b) **English version:** Have you experienced a difference in terms of immersion in any version of the game? Why?

(4a) **German version:** Haben Sie in irgendeiner Version des Spiels einen Unterschied in Bezug auf dievkörperliche Aktivität erlebt? Warum?

(4b) **English version:** Have you experienced a difference in terms of physical activity in any version of the game? Why?

(5a) **German version:** Welche Aspekte des Spiels würde Sie gerne verändern, welche würden Sie so lassen, wie sie jetzt sind?

(5b) **English version:** Which aspects of the game would you like to change, which would you leave as they are now?

(6a) **German version:** Haben Sie zusätzliche Anmerkungen?

(6b) **English version:** Do you have any additional comments?

11 IMMERSIVE VR GAMES FOR SOCIAL INTERACTION

11.1 VR ALMOST THERE: SIMULATING CO-LOCATED MULTIPLAYER EXPERIENCES IN SOCIAL VIRTUAL REALITY

[CorePub6] ✱Philipp Sykownik, **Sukran Karaosmanoglu**, Katharina Emmerich, Frank Steinicke, and Maic Masuch. Vr almost there: Simulating co-located multiplayer experiences in social virtual reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. DOI 10.1145/3544548.3581230. URL https://doi.org/10.1145/3544548.3581230

Supplementary materials for this article are available here: https://doi.org/10.1145/3544548.3581230.

VR Almost There: Simulating Co-located Multiplayer Experiences in Social Virtual Reality

Philipp Sykownik*
Entertainment Computing Group,
University of Duisburg-Essen
Duisburg, Germany
philipp.sykownik@uni-due.de

Sukran Karaosmanoglu*
Human-Computer Interaction,
Universität Hamburg
Hamburg, Germany
sukran.karaosmanoglu@uni-hamburg.de

Katharina Emmerich
High Performance Computing Group,
University of Duisburg-Essen
Duisburg, Germany
katharina.emmerich@uni-due.de

Frank Steinicke
Human-Computer Interaction,
Universität Hamburg
Hamburg, Germany
frank.steinicke@uni-hamburg.de

Maic Masuch
Entertainment Computing Group,
University of Duisburg-Essen
Duisburg, Germany
maic.masuch@uni-due.de

Figure 1: We evaluated a VR application that enables two remote users to play a co-located multiplayer game in VR.

ABSTRACT

Consumer social virtual reality (VR) applications have recently started to enable social interactions at a distance. Yet it is still relatively unknown if and to what extent such applications provide meaningful social experiences in cases where in-person leisure activities are not feasible. To explore this, we developed a custom social VR application and conducted an exploratory lab study with 25 dyads in which we compared an in-person and a virtual version of a co-located multiplayer scenario. Our mixed-methods analysis revealed that both scenarios created a socially rich atmosphere and strengthened the social closeness between players. However, the lack of facial animations, limited body language, and a low field of view led to VR's main social experiential limitations: a reduced mutual awareness and emotional understanding compared to the in-person scenario. We derive implications for social VR design and research as well as game user research.

*Both authors contributed equally to this research.

CHI '23, April 23–28, 2023, Hamburg, Germany
 The definitive Version of Record was published in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*, https://doi.org/10.1145/3544548.3581230.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **Virtual reality**; *Collaborative interaction*; • **Applied computing** → *Computer games*.

KEYWORDS

social virtual reality, multiplayer games, social presence, player experience, social interaction

ACM Reference Format:
Philipp Sykownik, Sukran Karaosmanoglu, Katharina Emmerich, Frank Steinicke, and Maic Masuch. 2023. VR Almost There: Simulating Co-located Multiplayer Experiences in Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3544548.3581230

1 INTRODUCTION

One of the major purposes of computer-mediated interpersonal communication has always been to enable valuable and rich social experiences at a distance, with face-to-face (f2f) interaction often serving as the gold standard for comparison [8, 68]. When considering modern VR technology, simulating this gold standard and offering an alternative when f2f is not feasible seems possible. Compared to established, non-immersive means of computer-based communication (e.g., videoconferencing), multiuser VR holds enormous potential. It enables natural and intuitive interaction with each other in shared virtual environments using real-time tracking of body movement and its mapping to the virtual world. Current consumer multiuser or social VR applications offer various social- and entertainment-related experiences. Recent research indicates that they provide access to valuable social and other experiences that users integrate into their everyday-life [5, 53, 72, 89]. Additionally, there are significant investments and efforts of large enterprises (e.g., *Meta* or *Microsoft*) into a future "Metaverse" [96] that incorporates virtual environments where people meet and interact with each other through immersive technologies like VR [61, 62]. Therefore, we expect more people to gain access to these experiences in the coming years. However, we believe that people will measure these applications' value more by whether they enable a sufficiently meaningful experience when a physical get-together is not possible, and not whether they induce the same experience as a f2f situation.

As one such scenario, we focus on engagement in co-located or local digital multiplayer games. Playing video games with others in person is still a play mode that many players enjoy [21]. However, modern online multiplayer games feature distinctive benefits such as time- and place-independent scheduling of game sessions [95]. Past studies indicate that unique experiential qualities characterize local multiplayer scenarios—i.e., giving a high-five, seeing the others' facial expression [43, 86] — that result in additional sociability, which remote games often cannot offer due to technical limitations. We see the simulation of co-located multiplayer scenarios in social VR as a promising approach to combine the best of both worlds: (i) eliciting the rich sociability associated with co-located gaming while (ii) keeping the independence of place or distance constraints offered by online games.

This paper presents the design and results of a user study comparing the experiential qualities of playing a digital multiplayer game physically co-located with playing the same game within a social VR environment while being physically separated (see Figure 1). Our work is guided by the following two exploratory research questions:

RQ1: How do the player and social experience in VR compare to the experience in the f2f setting?
RQ2: What features enhance or inhibit the player and social experience in VR?

Comparing the experiential qualities of a virtual and a f2f co-located gaming scenario allows assumptions about how well current and future consumer social VR applications function as an alternative site for meaningful popular leisure social activities when f2f meetings are not feasible. Moreover, our work contributes multiuser VR researchers and practitioners, as well as games user researchers, as follows:

C1: assessing the potential of current social VR offerings to enable meaningful social leisure activities over distance,
C2: extending previous findings from similar use cases by applying more nuanced measures of sociality,
C3: continuing and extending previous research comparing co-located and remote multiplayer game scenarios, and
C4: suggesting research, methodological, and design implications for social VR and games user research.

2 MULTIPLAYER EXPERIENCES

Playing digital games is a widely accepted leisure activity [21] that also provides opportunities for rich social experiences when more than one player is engaged in the gaming context [17, 25, 36, 38, 78, 86]. Prior work suggests that multiplayer gaming can provide a higher level of flow [46] and positive mood [39] than playing alone. In addition to entertainment purposes, many players use multiplayer games to socialize, spend time with their families or friends [21], befriend people they meet in-game [97], or connect with others in exceptional or challenging times [15, 34, 70].

Typically, multiplayer games can be divided into those played while co-located with others and those played online while being physically separated from co-players. With the advancement of technology, online multiplayer games became a common leisure activity throughout society [21]. Compared to co-located games, online games have specific advantages. For instance, players can play games with vast groups of co-players, play against or with strangers, and coordinate game sessions independent from their residence [95]. In contrast, co-located multiplayer games are, per definition, played by players who share the same location. They thus offer more coordination effort in terms of planning time and location of game sessions [95]. Though, as usually played with known others, co-located games tend to be associated with higher enjoyment compared to online games with strangers [95].

More specifically, co-located games benefit from game-external social interactions between players [43, 78, 86] that are not offered by current online games that usually connect players via headsets. Players can physically interact with each other (e.g., high-fives), easily see each others' facial expressions and regulate interpersonal distance. In sum, players benefit from an increased mutual awareness and more interaction opportunities [17], which eventually induce the experience of social presence, i.e.:"sense of being with another" [7]. Early studies comparing different social contexts of playing together indicate that co-located games induce higher levels of social presence, increased fun, and competence feelings [24–26].

However, the authors of these early studies note that physical co-presence alone does not guarantee an enhanced experience. Instead, the interrelationship of factors like players' spatial orientation to each other and the required focus on the game or a shared monitor determine how much players can allocate attention to each other to benefit from the social stimuli present in this scenario [17].

3 SOCIAL VR EXPERIENCES

VR applications that allow multiple users to interact with each other in a virtual world have been developed and researched as immersive, collaborative virtual environments (CVEs [6]) for several decades in varying contexts (e.g., [3, 75, 91, 100]). Today, capable yet affordable consumer-grade VR hardware provides leisure time access to multiuser VR applications that enable private individuals to interact with others worldwide. The genre of consumer social VR applications has emerged since 2015 and currently comprises a range of platforms like *Altspace VR, VRChat, Horizon Worlds*, and *RecRoom*[1] that offer a multitude of activities for social experiences [5, 53, 59, 60, 71, 72, 89], i.a.: socializing with known others and strangers, playing and creating games, dancing, and attending community events.

In social VR, users are commonly represented as avatars through which they interact with the virtual environment and each other. Based on sensory immersion, modern VR headsets induce spatial presence, i.e., the sensation of actually being in the virtual world or the place illusion [82, 98]. In addition, by transferring users' real-time tracked body movements to their avatars, a strong illusion of virtual body ownership can emerge, i.e., a sense that one is embodying the avatar and interacting through it within VR [35, 41, 56, 83]. These building blocks can support social presence and enable intuitive interaction with each other in social VR, including (non)-verbal communication.

Recent work indicates that engaging in social VR may indeed satisfy social, but also entertainment and other needs [89], and can have positive psychological outcomes for users [5], e.g., by supporting meaningful relationships [23, 54, 90, 99]. Other recent studies investigating multiuser VR interaction in general, suggest that social VR can have meaningful social outcomes; compared to f2f interaction, multiuser VR can lead to comparable compliance behavior [19], trust [69], conversation patterns [84], and experiences in a get-to-know-you conversation [74]. But, current social VR platforms still have technological limitations. For example, the tracking of facial expressions, which are the second most important social cue to facilitate social presence following gaze information [81], is not yet an established feature. The hardware required for this has only recently become available for the consumer market and developers have yet to integrate these new capabilities into their applications[2,3]. However, users seem to adapt to such limitations when engaging in social VR regularly and increasingly perceive experiences like social presence [28] or presence [40] over time.

Since we are interested in replicating a famous use case of joint media consumption in VR (i.e., experiencing co-located gameplay while being physically separated), we specifically searched for work with a similar goal. In total, we found four studies that investigated the following use cases: video watching [57, 63], photo sharing [48], and another study simulating co-located gaming [88].

One study (N=12) evaluated several custom prototypes for shared video consumption using early consumer-grade VR hardware combined with a video-based telepresence setup (e.g., within a lab-based photo-sphere or a virtual cinema) [57]. Overall, the study found that the VR experience can approximate the co-located scenario in certain social aspects depending on the specific implementation. A more recent study (N=22) also compared different setups for watching videos together [63]: f2f vs. social VR (i.e., *Facebook Spaces*) vs. a custom video-based telepresence system using VR headsets and a *Kinect* sensor. The study found that the custom telepresence variant induced a slightly better quality of interaction and more social meaning than the social VR variant. The participants expressed concerns regarding the limited graphical and behavioral realism of the *Facebook Spaces* avatars, i.e., limited facial expressions triggered by controller input and limited body language. However, the overall experience was similar across all conditions.

Another study (N=52) compared photo-sharing experiences in three conditions: f2f vs. social VR (i.e., *Facebook Spaces*) vs. video-conferencing [48]. The social VR version closely approximated the f2f experience with minor significant differences regarding the perceived quality of interaction but no significant differences in social meaning. Again, some participants criticized that the avatars only supported limited facial expressions; these expressions were triggered by controller input, which limited the spontaneity of the emotional reactions.

A work-in-progress paper that examined the same scenario as we do (N=4) only shares anecdotal insights and appears to have not been continued to date [88]. The authors found in their limited investigation that VR seems similar to the f2f scenario. However, based on the anecdotal nature of the work, we cannot derive reasonable conclusions. Nevertheless, the study provides a blueprint for our take on the scenario.

In summary, previous work indicates that VR scenarios of joint media consumption can approximate its f2f counterpart. However, studies in this context are sparse and the setups used do not reflect the capabilities of today's consumer social VR applications. The earliest study [57] does not reflect modern VR hard- and software and did not use avatars. *Facebook Spaces*, that was used in the other studies [48, 63], has been discontinued and the avatar aesthetics used in its successor, *Horizon Worlds*, also evolved. Similarly, the *Oculus Rift S* used in those studies [48, 63] did not support hand and finger tracking and thus limited gestural communication compared to the *Meta Quest* devices that are now available. Further, the studies employed a photosphere of a physical lab as the virtual background [48, 57, 63] instead of a walkable actual virtual environment as offered in today's consumer social VR applications. Unfortunately, there is also no clear information on the duration of interaction [48], or authors opted for a very brief interaction exposure of 2.5 minutes [63], potentially neglecting adaptation effects to avatar limitations [28, 40]. Moreover, the studies predominantly focused on assessing the social experience in terms of perceived

[1]https://altvr.com/; https://hello.vrchat.com/; https://www.oculus.com/horizon-worlds/; https://recroom.com/
[2]HTC released a face tracker in 2022: https://www.vive.com/us/accessory/facial-tracker/
[3]The Meta Quest Pro features integrated face-tracking and was launched while this paper was under review: https://www.meta.com/de/quest/quest-pro/

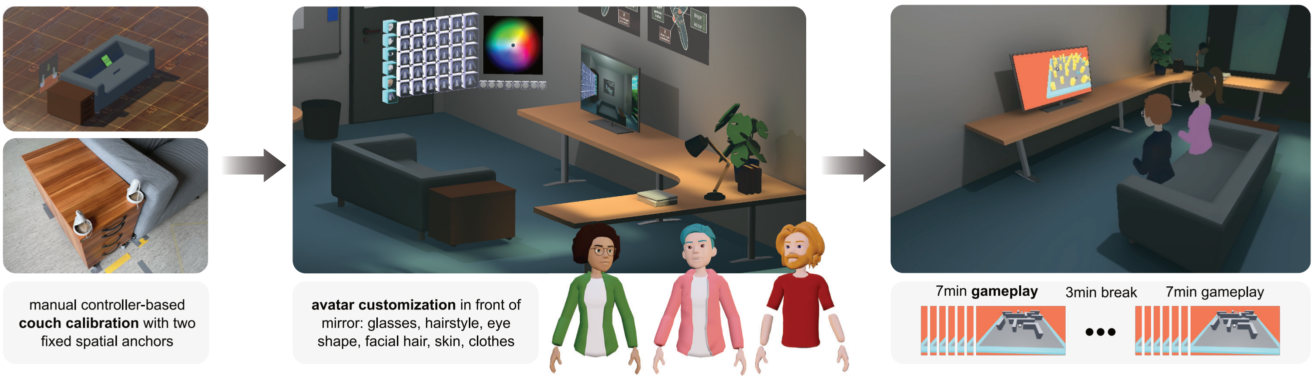

Figure 2: The VR app included three scenes: couch calibration, avatar customization, and the multiplayer scene.

interaction quality, social meaning, and immersion, and mainly report these three aspects [48, 63]. However, other questionnaires, like the Networked Minds Measure of Social Presence (NMMSP) that includes six sub-scales [29], provide a more nuanced assessment of the social experience because it takes more sociality facets into account. Considering those limitations, our work extends the still sparse literature on virtual joint media consumption by providing a timely and nuanced look at our specific use case: simulating co-located multiplayer gaming in social VR.

4 APPLICATION DEVELOPMENT

We want to investigate how well current consumer-grade social VR functions as an alternative site for co-located multiplayer experiences when f2f meetings are not feasible. Thus, our aesthetic goal was not to develop a photorealistic representation of the scenario but a stylized replica whose aesthetic conception resembles current consumer social VR offerings. To compare the VR replica with its f2f counterpart in a controlled lab study, we also had to develop a multiplayer game that we could use equally well in both the VR and the co-located variants with their different input and output modalities. Eventually, we developed two custom applications with the *Unity Engine*[4]: a VR application for the *Meta Quest 2*, a popular, standalone VR headset with integrated hand and finger tracking, and a PC-based multiplayer game to be played on a 55-inch UHD TV screen.

4.1 The Co-located Multiplayer Simulator

We used the following popular social VR applications available on the *Meta Quest 2* as aesthetic references for our VR application: *Rec Room, Altspace VR, BigScreen*, and *Horizon Worlds*. We also included rudimentary avatar customization, a prominent feature of social VR applications. Utilizing the *Meta Quest 2's* integrated camera-based hand tracking, the application allows users to switch between hand and controller tracking. In this regard, the virtual scenario resembles the physically co-located scenario, as users can also pick up and put away controllers there. Additionally, we implemented a calibration system to match the virtual environment with our physical labs so that physically remote users eventually appear to be sitting on the same sofa in VR. The app uses the VR headset's built-in speakers and microphones to provide voice chat with spatial audio. To ensure sufficient performance, our application relies on a server-client architecture with a PC-based server instance that synchronizes the VR clients. The server instance also provides an audio and video live stream of the users within the VR environment.

4.1.1 Application Structure. As illustrated in Figure 2, our app features three consecutive scenes: a calibration scene, an avatar customization scene, and a game scene in which users meet and play together on a virtual monitor. The app is started while holding the *Meta Quest's* controllers in both hands. These are initially used to calibrate the virtual environment so that the virtual and physical sofas align. As a reference point, we used two spatial anchors on which one places the controllers. Users then use their hands to interact with the virtual world. After the calibration, the users enter the customization scene, where the avatar's arm length, body scale, and appearance can be adjusted in front of the virtual TV display that functions as a mirror. Adjusting arms and body size shall ensure a seamless avatar animation using inverse kinematics. Finally, when both remote users sit on the couch in their individual scenes, they enter the game scene together, appearing as sitting next to each other on the virtual couch. Here, they can retake their controllers to play the multiplayer game together. The game is played in two rounds of seven minutes with a three-minute break in-between. We included the break to provide opportunities for open social interaction in the study.

4.1.2 VR Environment. In contrast to related work [48, 63], we designed a 3D virtual environment that users could explore and that is aesthetically consistent with the avatars we provided. Thereby, the virtual environment's layout and interior match the two same-sized physical labs where we conducted the study. In both labs, we placed the same sofa we rebuilt with the exact dimensions in VR.

4.1.3 VR Avatars. In keeping with the aesthetic language of the selected social VR applications, we created humanoid but stylized avatars with reduced anatomical features (i.e., no legs or feet). Avatars in *RecRoom, AltspaceVR,* and *BigScreen* have a torso but lack arms. *Horizon Worlds*' avatars also represent arms. We opted for the middle ground and omitted the representation of the hand joints

[4]Unity 2019 LTS: https://unity.com/releases/2019-lts

and partly the elbow joints (Figure 2), thus preventing visualization of unrealistic joint rotations. Our avatars can be color-customized and individualized with different assets we have prefabricated. All clothing assets are available in a rather masculine or feminine body shape. Assets of different categories can be combined in any way (e.g., a beard with a female body shape and hairstyle). The avatars can represent real-time hand gestures via the hand and finger tracking of the VR headset and thus have increased behavioral realism compared to related work [48, 63]. Further, we implemented simulated blink and pupil movement based on the random fixation of predefined interest points within the avatars' field of view (e.g., nose, eyes, and shoulders of other avatars). Furthermore, mouth animations are triggered by speech input.

4.1.4 Design Rationale for VR Avatars. Despite previously reported concerns regarding avatars with limited graphical realism [48, 63], we opted for stylized avatars. However, in contrast to present them within a real-world photosphere [48, 63], we designed our avatars and the virtual environment aesthetically consistent, preventing potential negative effects of stylistic contrasts within VR. Further, our application provides a longer exposure time compared to the applications in related work. Thus, by allowing our participants to adapt to the stylized but consistent aesthetic of our virtual environment and avatars, we assumed to limit negative reactions to the stylized avatars. However, based on previous findings [48, 63], we decided against including a manual activation of facial animations, which users perceived as burdensome to use [40, 48, 63]. A manual expression system would conflict with the engagement with the multiplayer game. We also rejected an automized approach as reliably triggering different facial expressions based on stimuli like voice input would have required sophisticated algorithms and significantly more nuanced modeling to avoid the uncanny valley and eventually antipathy to our avatars [64, 80]. Moreover, sophisticated simulated facial animations have yet to be an established feature in consumer social VR [93]. Thus, considering that we focus on the overall experience and not the nuances of the avatar system, we did not opt for a facial expression system. Also, past studies indicate that users can acclimate limitations regarding facial emotional expression and recognition in VR if verbal communication is sufficiently supported [40, 65].

4.2 The Co-located Multiplayer Game

In addition to the VR application, we had to develop a multiplayer game that would function equally well in VR on a virtual display and in the co-located scenario on a physical monitor. The game should also be easy to implement as the VR implementation was already extensive. It should run with equal performance in both conditions and be playable with the respective controller peripherals. Eventually, we reviewed game design literature for a game principle that met our requirements to limit conceptual work effort. In particular, we aimed for an easy-to-understand but engaging game mechanic that induces social interaction between players. Thus, we predominantly searched for mechanics to generate player interdependence to motivate player communication [18, 30, 31, 37]. Finally, a recent literature review identified shared control as a mechanic that meets our requirements [73]. It requires multiple players to control the same game object simultaneously and thus generates extreme player interdependence [51, 87]. Since the shared control game described in [87] met all our requirements, we recreated it based on the descriptions in the paper.

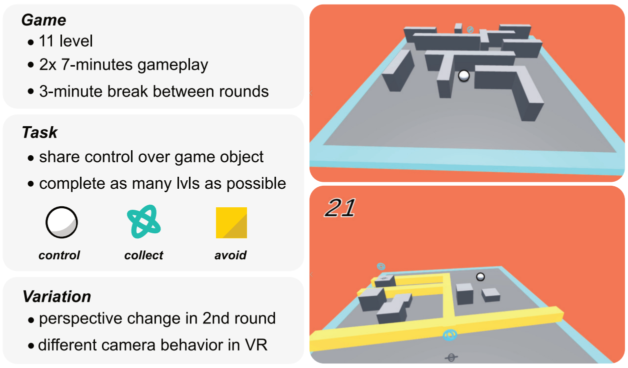

Figure 3: Game design of the shared-control game.

The final game requires two players to navigate a sphere through several obstacle courses. Thereby, the steering input of both players is simultaneously applied to the sphere so that it moves faster if both inputs are similar (see Figure 3). If players do not steer in the same direction, they cannot gain enough speed to overcome certain obstacles in the levels and eventually cannot finish the game. We varied the camera behavior and perspective to introduce more variation for the planned within-participants study design and limit learning effects. For example, in the second round of a condition, the camera perspective changes to the other side of the playing field, so players must now steer in different directions. Consequently, individual levels get implicitly more difficult than before. In addition, the camera behaves differently in the VR version than in the co-located version: camera movement vs. camera panning based on sphere movement.

5 USER STUDY

We compared the co-located multiplayer simulator with its f2f counterpart in a within-participants user study with the two respective conditions *VR* and *Col*. In the *VR* condition, participants played together in VR while being located in separate labs. In the *Col* condition, they were co-located together in the same room and played on a single TV screen (illustrated in Figure 4). Given the sparse literature on joint media consumption in social VR and the identified limitations of prior studies, we opted against a confirmatory approach and did not specify hypotheses. Instead, we conducted exploratory comparisons of the two conditions regarding various facets of the player and social experience (**RQ1**). Further, the study aimed to identify specific aspects of the respective scenarios that promote or limit the social and player experience (**RQ2**).

5.1 Sampling

We applied convenience sampling [4] and advertised the study in lectures of two HCI-related undergraduate and graduate study programs at a university in Germany. The sparse related work on similar study setups did not allow a determination of effect sizes to be expected. We thus did not aim for a confirmatory but an

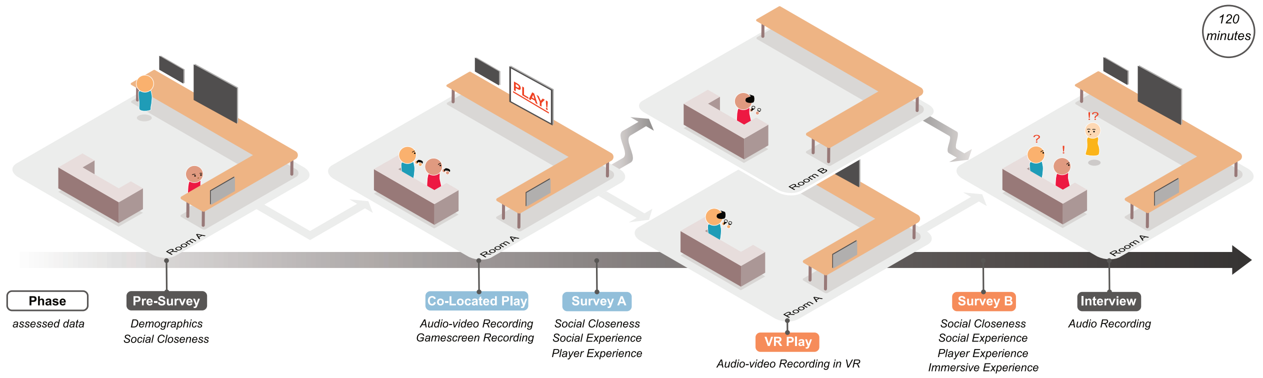

Figure 4: The user study followed a within-participants design with alternated sequence of the co-located and VR play sessions.

exploratory analysis and did not conduct a priori power analysis to calculate a sufficient sample size. Though applying feasibility analysis[5] and aligning with local standards [13] (i.e., comparing to similar studies [20, 30, 48]), we aimed for a sample size of 50.

5.2 Procedure

A study run consisted of four phases and took ca. 120 minutes (Figure 4). At the beginning of the procedure, participants gave their participation consent and agreed with recordings during gameplay and interview. The two conditions were conducted in alternating sequences to counterbalance potential sequence effects.

In the co-located condition, participants played together in the same room, sharing a sofa. After a briefing on the gameplay, a researcher started the screen and camera recording, and left the room. The gameplay lasted 17 minutes, divided into two 7-minute rounds with a 3-minute break. The participants then completed questionnaires in the same room.

Before the VR condition, there was a verbal briefing on the VR headset and application. Afterward, a researcher started the server application and its screen recording in another room. A research assistant then guided one participant into another room. Participants put on their VR headsets in their respective rooms, supported by the researchers. The researchers remotely started the VR applications and guided participants through the first two VR phases: sofa calibration and avatar customization. After ensuring that both participants were in the shared game scene, the researchers left the rooms and met in a third room. Participants again engaged in 14 minutes of gameplay and a 3-minute break. They then completed the PC-based questionnaires alone. During the VR condition, the researchers sat at the server PC that provided a live stream and voice chat of the VR application to intervene in case of technical issues.

Afterward, both participants were interviewed together for 20-30 minutes. Most participants were compensated in the form of certificates of participation if relevant to their study program. There was no other compensation.

5.3 Measures

We used German-translated questionnaires, recordings, and interview questions to assess a broad spectrum of data (Figure 4). Thereby, the questionnaires assessed the player and social experiences in both conditions that we then compared statistically **(RQ1)**. All items of the questionnaires were administered on a 7-point Likert scale (e.g., *Strongly disagree – Strongly agree*)[6]. We used the recordings and interviews to supplement the quantitative data comparison **(RQ1)** and identify explanatory factors for quantitative differences between conditions **(RQ2)**.

5.3.1 Immersive Experience Measures. To evaluate the technical and aesthetic soundness of the VR app, we assessed the immersive experience in the VR condition in terms of the perceived presence, virtual embodiment illusion, and the aesthetic appeal of the virtual environment and avatars. We used the Igroup Presence Questionnaire (IPQ) [79], with its four sub-scales general presence, spatial presence, involvement, and perceived realism. Virtual embodiment illusion was assessed using the ownership and agency scales of the Virtual Embodiment Questionnaire (VEQ) [76]. Finally, we assessed the aesthetic appeal with the audio-visual appeal items of the PXI [1] (regarding environment and avatars).

5.3.2 Player Experience Measures. To evaluate the shared-control game in terms of game design quality and induced player experience **(RQ1)**, we applied the Player Experience Inventory (PXI) and used its subscales perceived mastery, immersion, autonomy, and meaning [1, 27]. For overall game enjoyment, we used three enjoyment items also proposed by the authors of the PXI.

5.3.3 Social Experience Measures. We assessed the social experience **(RQ1)** in terms of three experiential qualities: (i) social presence, i.e., a sense of co-presence and engagement between users [8], (ii) social richness, i.e., the perceived "warmth" of the interaction [50], and (iii) social connectedness as a transformative social outcome. To assess the perceived social presence, we applied the Networked Minds Measure of Social Presence with all its subscales [29]: co-presence, attentional allocation, message and affective understanding, emotional and behavioral interdependence. We assessed social richness using the corresponding sub-scale of

[5]Additional information in the supplementary material.

[6]Anchors varied such as *not at all – very much* depending on the specific instrument.

the Temple Presence Inventory (TPI) [50], which includes seven items of semantic differentials (e.g., unemotional – emotional). Social connectedness was assessed using the Inclusion of Other in Self (IOS) measure [2] consisting of a single pictorial item. Further, participants rated their perceived comfort and feeling of belonging to each other with two items of the friends construct of the Social Connectedness Scale (SCS) [14].

5.3.4 Recordings. To assess participants' social behavior during the gameplay and the break between game rounds, we recorded audio and video of the co-located and VR sessions. These recordings supplement the questionnaire data and provide explanatory information for any identified quantitative experiential differences **(RQ1 & 2)**. In addition, they allow us to check if there were any fundamental technical issues during the sessions, which should be considered accordingly in the analysis. In the co-located condition, we placed a camera next to the TV to record interaction on the sofa. In the VR condition, we only recorded the virtual scenery from several viewpoints within VR.

5.3.5 Interview. We conducted semi-structured interviews with the pairs to gain a more in-depth understanding of the participants' experience and what specific aspects shaped it **(RQ1 & 2)**. The first two authors prepared the interview questions by reviewing, discussing, and refining the questions in detail. The questions were then shared with the rest of the research team to finalize them. Overall, the interview questions were focused on participants' general thoughts on the VR scenario, their perceived sociality during the conditions, their evaluation of the avatars, and their thoughts on using the VR application in their private leisure time.

6 ANALYSIS RATIONALE

In the following, we detail the analysis procedures for the different types of data we assessed in the user study.

6.1 Questionnaire Data

We conducted analyses based on individual participants' scores and initially checked internal consistency for each subscale of the applied instruments. Eventually, we excluded one item from the analysis of the IPQ's spatial presence subscale.

We used frequentist statistics to quantitatively analyze the player and social experience based on questionnaires' subscale data. This analysis comprised subscales' mean and median comparisons between the two conditions. Based on Shapiro-Wilk tests, we checked if the assumption of a normal distribution was met for the participants' score differences. Accordingly, we conducted either paired-sample t-tests or Wilcoxon signed-rank tests. When reporting Wilcoxon signed-rank tests, we report the Hodges-Lehmann estimate, i.e., the median of the individual differences between both conditions ($Mdn_{Pre\text{-}Col}$). Social connectedness scores, which we assessed three times (*Pre*, *Post-VR*, *Post-Col*), were analyzed using a Friedman test and Wilcoxon signed-rank tests for post-hoc pairwise comparisons.

All significance tests were conducted using a .05 α-level (two-tailed). In the case of post-hoc pairwise comparisons, we applied the Bonferroni-Holm adjustment of the α-level [33].

To make our results comparable with those of similar work in the future, we also report effect sizes: Cohen's d for t-tests [16], rank-biserial correlation for Wilcoxon signed-rank tests [94], and Kendall's W for Friedman tests [22]. Additionally, we report 95% confidence intervals for the effect sizes of pairwise comparisons as a measure of their precision.

6.2 Video Data

One of the first authors and the third author conducted the video analysis. A first rough review of the video material revealed that participants' behavior was way more dynamic, complex, and "interesting" during the break compared to the gameplay phase. We thus decided to analyze the gameplay and break phase separately from each other and use different methods for each phase to reduce the analysis effort reasonably. One session was removed entirely from this analysis as participants gave no consent to be recorded during co-located gameplay.

For the gameplay phase, we assessed, with the help of student assistants, instances of socially relevant gaze actions, i.e., one participant looks towards the other. Additionally, we assessed the amount of verbal communication by extracting the accumulated duration of mutually perceivable verbal utterances (e.g., conversation, laughter). These two measures allow us to compare the two conditions in terms of instances in which participants actively perceived each other's presence. One session was removed from verbal communication analysis due to missing audio in the VR condition.

We decided against a quantitative assessment of discrete user behaviors for the break phase analysis. Instead, we opted for an informal qualitative approach, noting and discussing participants' activities and interactions—this process aimed to identify similarities and differences in behavior patterns in the two conditions. Thus, note-taking aimed to provide a focused assessment of social interactions between the participants. The notes were prepared by the third author and then discussed with one of the first authors.

6.3 Interview Data

We transcribed the interviews using automatic transcription[7] followed by manual correction by three native speakers. As we are a multi-lingual team, the transcriptions were then translated to English using translation software[8] and manually corrected for inaccuracies by one of the first authors. Afterward, the two first authors analyzed the interviews, following a hybrid codebook and reflexive method of thematic analysis [11, 12].

In a first phase, we used a codebook-oriented approach to extract some general quantitative insights, e.g., What condition did participants like more? After deciding on initial deductive categories and codes, we started to code the data focusing on participants' general experience evaluation, i.e., were they positive, negative, or neutral about specific aspects? To reduce the time required for this process, each researcher only coded half of the interviews (12 vs. 13). They then reviewed each other's results for disagreements and met to resolve those.

A second phase aimed to reveal how specific aspects of the VR scenario shaped participants' experience. This analysis did not aim

[7]Dovetail: https://dovetailapp.com/
[8]DeepL Pro: https://www.deepl.com/pro?cta=header-pro/

for a final set of codes or agreement on coded paragraphs in the interviews but on inferring coherent themes. We thus applied both aspects of the codebook and reflexive approaches. In particular, we accepted the co-located scenario as the "gold standard". We structured our codes based on this premise, e.g., What aspects inhibited/enhanced the sociality in VR compared to the co-located scenario? How did these affect sociality? Again, we initially constructed deductive categories and codes. We extended these with inductive codes during the analysis. To reduce analysis time, we again decided to split the data set, so the two involved researchers only had to analyze half of the interviews in detail. However, to ensure the validity of the results, both researchers initially synchronized their understanding of initial codes and their perspective on the data by coding three same interviews independently from each other and discussing their results. In this discussion, they did not aim to reach a consensus on the final coding but to extend each others' understanding of the data and use of deductive codes. Only then they did continue independently coding their respective data set (n = 11). Upon completion, they checked each other's non-coded passages in the interviews to prevent information from being missed. They then defined general insights based on their results and presented them to each other. In a final joint reflective discussion, they clustered the insights from the two data sets into coherent themes.

6.4 Reflexivity Statement

The qualitative analysis has been both enriched and potentially biased by the background of the three involved authors [12, 67]. For transparency, we thus specify their backgrounds: One has a background in cognitive systems and psychology and engages in games and VR research. The two others share a computer science and psychology background, with one engaging in multiuser VR research and the other in games user research focusing on social experiences.

7 RESULTS

25 dyads participated in the study (N=50, women=37, men=13, M_{age}=21.74 SD=3.41). All participants grew up in Germany, and almost all were students at the time of the study (n=48). The majority of participants were friends or knew each other well: Fifteen dyads of friends, one of close acquaintances, and one romantic couple. Further, two dyads were loose acquaintances, and four were strangers. Furthermore, in two dyads, participants indicated in-congruent relationship types (in each case, friendship and close acquaintances). Due to this unbalanced distribution of relationship types, we did not consider this factor in further analysis.

Overall, the participants indicated neutral to positive interest (M=4.98, SD=1.80) and passion (M=4.54, SD=1.97) towards digital games and over half (n=35) of them indicated playing at least several hours a month. Prior VR experiences were mostly only made occasionally (n=34) and only one participant had prior experience with social VR applications.

7.1 Immersive Experience

Facets of perceived presence (IPQ), virtual embodiment (VEQ), and audio-visual appeal (PXI) were rated with high scores by at least

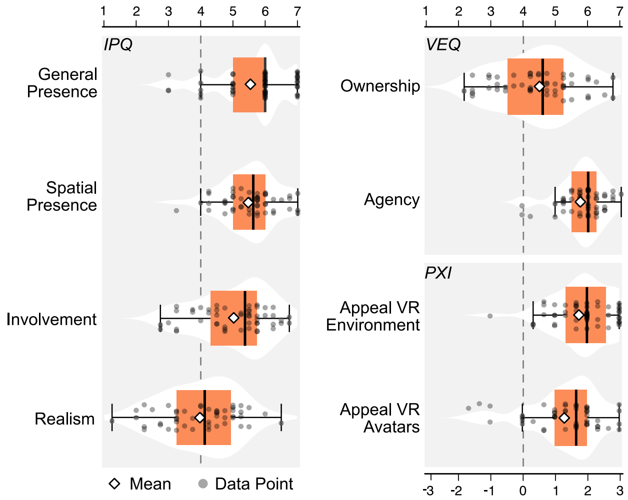

Figure 5: Immersive experience induced by the VR application illustrated by combined box & violin plots with individual data points.

50% and above moderate scores by the majority, as illustrated by median values and interquartile ranges in Figure 5. Participants' scores cover the entire scale for perceived realism and virtual body ownership, and interquartile ranges go below four. But, the median values and violin plots indicate that the bulk of scores is above four. These results indicate a sound technical and aesthetic implementation of the VR application. Thus, we did not consider the immersive experience in the following analysis as a confounding factor.

7.2 Player Experience

All relevant test statistics for the following comparisons are included in Figure 6.

Overall, both conditions seemed to have induced high levels of enjoyment for most participants. No team has finished all game levels; most reached the third-last level. The mastery, autonomy, and meaning scores of the joined gameplay show a central tendency to moderate values across conditions. The perceived immersion seems to have been moderately high in both conditions.

Wilcoxon signed-rank tests revealed significantly higher levels of immersion in VR than in co-located play, but no significant differences concerning enjoyment, mastery, autonomy, and meaning.

7.3 Social Experience

7.3.1 Social Presence. Overall, the participants seemed to have experienced moderate to high degrees of social presence according to central tendencies and distribution of averaged agreement scores of the NMMSP sub-scales, which are illustrated in Figure 7a. The boxplots illustrate larger inter-individual differences regarding the perceived attentional allocation, affective understanding, and emotional interdependency in VR, as scores range from high to rather low values. Further, difference lines, illustrating the participants' individual score differences, indicate inter-individual differences in terms of which condition induced higher levels of social presence.

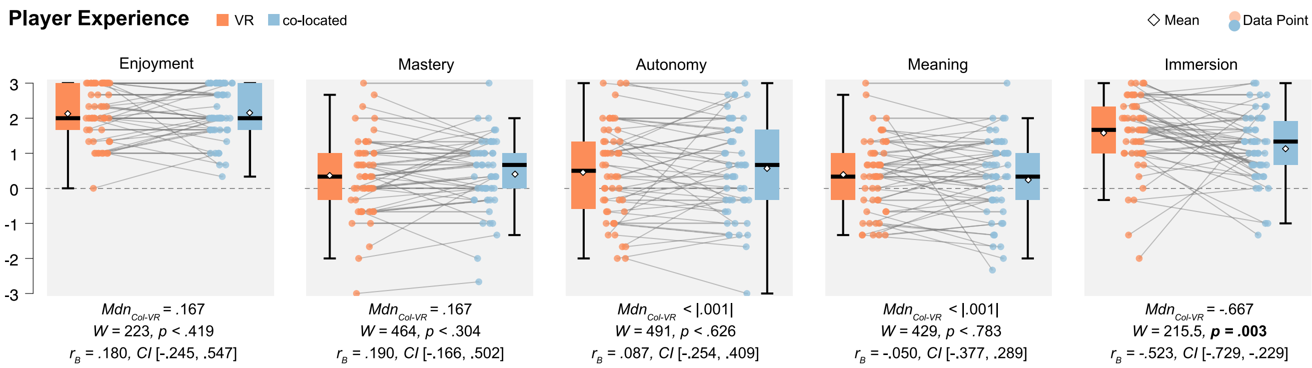

Figure 6: Summary of player experience analysis based on boxplots, individual data points, and test results.

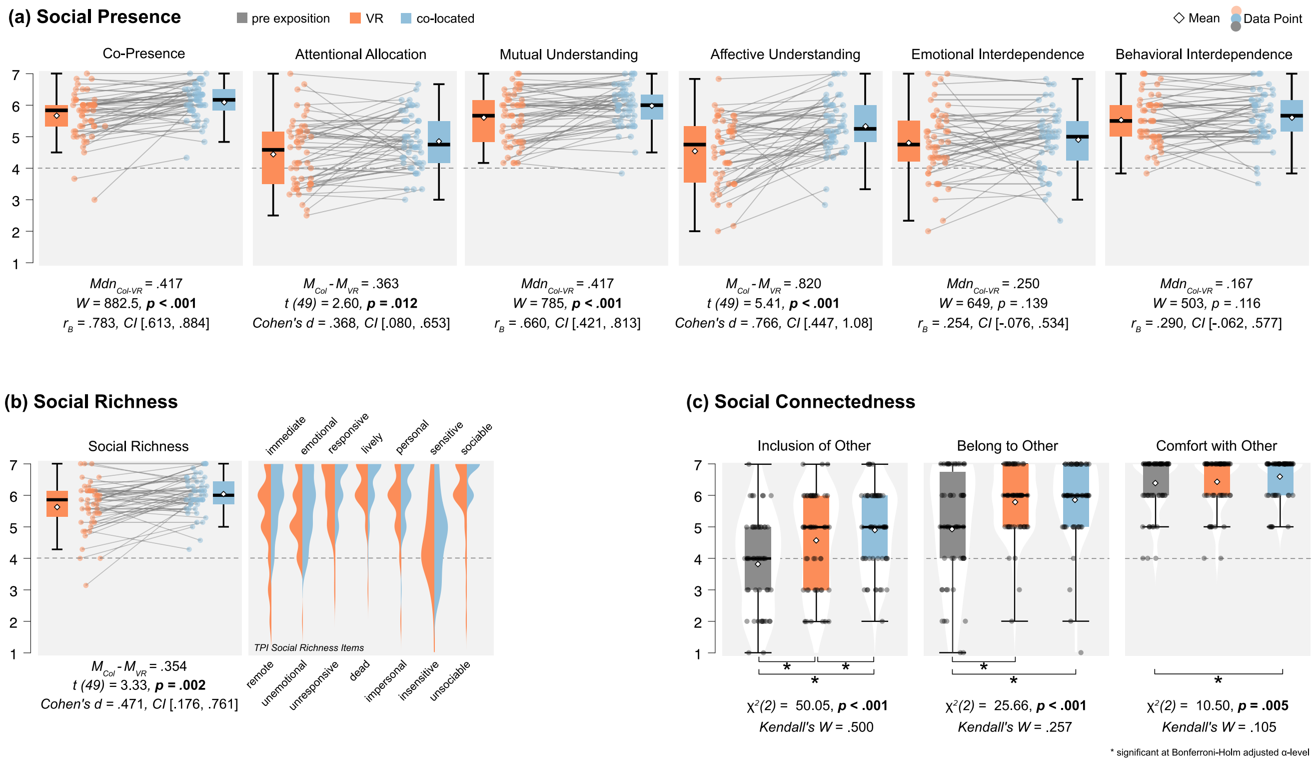

Figure 7: Summary of social experience analysis including individual or combined box- and violin plots, individual data points, and test results.

Wilcoxon signed-rank tests and paired-sample t-tests revealed that perceived co-presence, mutual understanding, attentional allocation, and affective understanding were significantly higher in co-located gameplay. Emotional interdependence and behavioral interdependence were not significantly different in the conditions. All relevant test-statistics are included in Figure 7a.

7.3.2 Social Richness. In both conditions, most participants seem to have perceived high levels of social richness based on how they described the experience using semantic differentials (Figure 7b). Though, social richness was significantly lower in VR than in the co-located condition based on a paired-samples t-test (Figure 7b).

7.3.3 Social Connectedness. The boxplots in Figure 7c indicate that the central tendency of inclusion of other scores changed from a moderate level prior to the gameplay to a slightly increased level after the VR and co-located gameplay. Further, the co-located scores' dispersion seems smaller than the VR scores'. Similarly, the boxplots of belonging scores show a positive change of central tendency from a moderate to a moderately high level and a reduction of score dispersion after gameplay compared to the baseline scores.

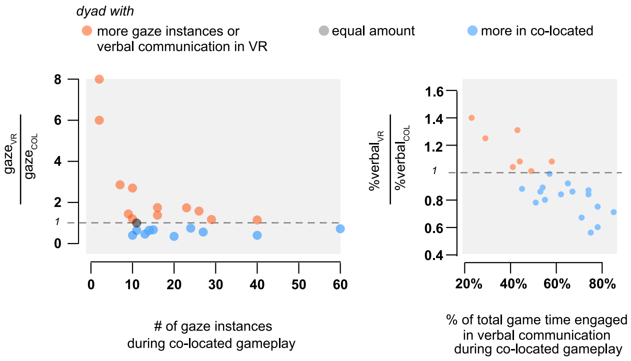

Figure 8: Ratio of social gaze instances and verbal communication amount in VR compared to co-located gameplay.

Perceived comfort of being with the other was high-to-very high at all three measurement times, with some outliers in the moderate level for pre-gameplay and VR measurement.

Three Friedman tests revealed a significant effect of the measurement time (pre vs. VR vs. co-located) on perceived inclusion of other, belonging, and comfort (Figure 7c). Post-hoc analyses with Bonferroni-Holm adjusted α-levels revealed the following significant differences between conditions:

Inclusion of other scored higher after co-located and after VR play compared to study start ($Mdn_{Pre\text{-}Col}$=-1.00, p<0.001, r_B=-1.00; $Mdn_{Pre\text{-}VR}$=-1.00, p<0.001, r_B=-.711). In addition, the scores after the co-located play were higher than after VR play ($Mdn_{Col\text{-}VR}$=1.00, p=0.025, r_B=-.558).

Belonging was perceived higher after co-located and after VR play compared to study start ($Mdn_{Pre\text{-}Col}$=-1.50, p<0.001, r_B=-.746; $Mdn_{Pre\text{-}VR}$=-1.50, p<0.001, r_B=-.764).

The perceived comfort of being with the other was higher after the co-located gameplay than at baseline ($Mdn_{Pre\text{-}Col}$=-1.00, p=0.001, r_B=-.875)

7.4 Verbal & Gaze Behavior during Gameplay

The scatter plots in Figure 8 map the gaze instances and amount of verbal communication within the VR condition's dyads' behavior in the co-located condition. The plots reveal that for the gaze behavior, there was no clear tendency that VR was generally more or less looking at the partner. In about the same number of sessions there were more glances towards the partner in VR as there were sessions in which there were more glances towards the partner in co-located. However, in four dyads with very few attempts to look at the other during co-located play, participants gave each other significantly more partner-directed glances in VR.

Overall, participants tended to engage in more verbal communication in the co-located condition, as the plot in Figure 8 shows more data points representing corresponding dyads. In contrast to gaze behavior, a pronounced influence of the fundamental interaction dynamics in the dyads is evident. The more communicative a pair of participants was in the co-located condition, the more likely they were to be less communicative in VR than in-person.

7.5 Behavior during Co-located Breaks

Most co-located breaks started with a brief game-related conversation. They then drifted either into an off-topic conversation or relatively silent phases of waiting for the break to end. The off-topic conversations, which occurred in half of all co-located breaks, were related to "gaming" in general or to mutual experiences or plans of both participants, e.g., regarding courses at university. Besides, participants who played the co-located game version after the VR version often referred to the VR session, i.e.: they talked about their VR experience and the VR headsets. They discussed both positive and negative aspects of the VR setting (e.g., uncomfortable headsets, VR room feeling comfier). The silent waiting phases occurred in nine sessions. Participants waited seemingly impatiently for the time to pass and felt visibly uncomfortable because they did not know what to say or do. In some of those cases, participants expressed feeling unpleasantly observed by the camera.

During all co-located breaks, participants remained sitting on the couch, looked at each other occasionally while talking, and looked around the room to some extent. Across all co-located breaks, we observed only one instance of social touch. One team performed a high-five at the beginning of the break to celebrate the level they had just won. We noticed no other instances of social touch nor any remarkable gesture-based interaction.

7.6 Behavior during VR Breaks

During the VR breaks, we observed very different behavior. However, the fundamental social dynamics between teammates hardly changed in most cases, as dyads talking a lot in co-located breaks also interacted more in the VR break compared to quieter groups. Predominantly, participants started to explore the interaction possibilities in the VR environment. Except for only eight dyads, the participants tried to touch each other and observed with interest what happened when their avatars collided. Often they started with poking the other's body or face, then trying different forms of physical interaction, including clap games, fist bumps, handshakes, head patting, and boxing. The dominant emotions during such interactions were amusement and fascination, as participants laughed a lot and were curious to try different things.

Most dyads also took a closer look at their avatars during the break. About half of the teams commented on the visual appearances of their avatars, sometimes laughing at eccentric looks (e.g., green hair) or discussing their design choices during avatar customization. Moreover, several participants noted their "missing" legs and wrists, discussing it as strange or irritating. In addition to focusing on their teammates, many participants also explored the VR environment. They moved through the room, looked around, and tried to touch some virtual items and furniture.

7.7 Insights from Interviews

Our first analysis phase revealed the insights illustrated in Figure 9. Around 75% of the participants stated that they liked one of the two conditions better, with almost 50% indicating the in-person variant. Still, about 25% were in favor of the VR version. About 20% of the participants were undecided and could not state a clear preference. The following subsections present results from our second analysis phase and provide detailed insights into the scenario.

Figure 9: Stacked bar charts with percentage distribution of responses to selected interview questions.

7.7.1 ***Theme 1: Most players liked the shared-control game for its simplicity regarding design, controls, and mechanics, which created a motivational pull to proceed.*** Overall, the game was very well received by the players (Figure 9). The shared-control mechanic was well received as the source of coordination and challenge. Though, it was sometimes perceived as annoying or not convincing as a mechanic at all: "*I found it quite funny that both can control everything, but it was equally frustrating.* ". In particular, referring to the later levels of the game, players reported how frustrating it was to fail because of a disproportionally increase in difficulty. A few specifically expressed dissatisfaction with the not-adjustable camera perspective.

7.7.2 ***Theme 2: Avatars and voice chat were the building blocks of the social experience. However, some were unsatisfied with avatars, not only due to technical issues.*** The majority of participants described how they perceived a social presence in the VR condition and how they felt as they were actually sitting next to each other: "*So, it really felt like we were with each other. [...] When I took the glasses off, I was confused for a moment because you weren't in the room with me [...]* "

This was largely related to simply being able to see each other in VR. Many participants emphasized how the avatars fundamentally allow for and signal mutual attention. The mere perception of the other's avatar, its movement through the virtual environment, and its orientation partly induced intense feelings of social presence. Some participants described how they avoided running into each other when taking their controllers. Moreover, during the game, it just felt natural to look to the side and see someone sitting there: "*So the avatar also helped a lot. So that when you looked to the side, you saw someone sitting there.* "

Together, mutual visual perception and the possibility of verbal communication formed the basis for the perceived social presence, naturalness, and emotionality of the interaction in VR. A recurring argument was that one could see and talk to each other just like in the co-located scenario: "*Well, the conditions were relatively the same, because we could see each other somewhere in virtual reality and we could also hear each other.* "

However, some players did not perceive interacting through avatars in VR as natural. They described detrimental effects on sociality in terms of perceived naturalness, social presence, closeness, and emotionality: "*Yeah, well, I think "natural" is hard to say. [...] because in VR it's just an avatar sitting next to you.* "

In individual instances, software-based issues with the avatars were associated with negative effects on the perceived naturalness or realism during the interaction: "*It was sometimes just a little bit with the calibration of the hands, but so that it looked weird for a second, maybe to the other person, but otherwise nothing.* "

7.7.3 ***Theme 3: Avoiding Uncanny Valley and providing customization opportunities can create a pleasant and personal experience, but upper-body-only avatars may annoy some users.*** The graphic style of the avatars was not only predominantly positively received (Figure 9) but was also partially described in the interviews as having a positive effect on the social experience, i.a., by generally contributing to a pleasant atmosphere. Specifically referring to the Uncanny Valley, it was noted by some participants that the style combined a pleasant level of realism with appropriate abstraction so that it did not annoy when looking at each other: "*I think it was pleasant. One didn't get scared now, because you look too realistic [...]* "

In several interviews, participants described how the customization may enhance the personal closeness of the interaction, as customization choices reveal something about themselves: "*[...] it says a lot about the person.* " Further, customization prevents avatars from all looking the same and ensures that they express individualism, which was associated with increased realism. In addition, the design choices induced conversations and triggered mutual laughter in cases of deliberately eye-catching appearances.

For some participants, a match between their interaction partner's physical and avatar appearance can have a positive impact on sociality in terms of the personal closeness, naturalness, or realism of the interaction: "*And of course, they don't look like you, but you can bring them a little bit closer to your appearance, and so you can create a different kind of closeness in VR.* "

In several instances, the anatomical properties of the avatars were critically discussed. In particular, the lack of legs and, in some cases, the lack of wrists harmed the perceived naturalness or realism of the interaction: "*Because it would complete it. It was one of the first things I noticed when I saw him. He has no legs.* " Further, the anatomic limitations of avatars impair opportunities for expression: "*Because half of the body was missing, so to speak. And half of the feelings were also missing, I thought.* " However, a few participants reported that legs were not necessary and they were not annoyed by it: "*It was funny that we were just standing there, just with the upper body, on this couch [...] So it wasn't absolutely necessary.* "

7.7.4 ***Theme 4: Behavioral realism in head movement, body orientation, and gesturing enhanced sociality. However, participants missed mimics and nuanced avatar postures. Social touch in VR emphasized physical separation.*** Most users indicated that sociality was heavily supported by the ability to look at each other and to perceive each others' gaze direction in VR. This, in particular, was supportive during verbal communication: "*[...] it's actually somehow more polite if I then talk and look at her. She sees that, too. And then I thought that was actually relatively natural, yes.*" In particular, seeing that the partner turns his head and body in one's direction establishes an understanding of mutual awareness. Many participants described this as an essential contributor to their social experience. Despite being limited, the body movement indicated where and in which orientation users sat on the sofa. This way, they could derive rough impressions of each others' attentional state: "*[...] I remember very well this one part where I turned to you and your body was turned to me.[...] I was like: Okay, that's so open and somehow so "turned towards me" and I found that positive.* "

Besides the voice and the avatars' visibility, the possibility to gesticulate with the hands positively influenced the perceived fun, social presence, liveliness, naturalness, and emotionality during social interaction. Gesturing accompanied verbal communication but also allowed gesture-based communication, such as a high-five or a pointing gesture to the virtual monitor: "*But I think the most important thing that I communicated with, apart from my voice, was gestures with my hands.* "

Although limited, the existing facial stimuli of the avatar system were occasionally described as enhancing sociality. Whereas the simulated gaze behavior was not actively recalled, the simulated blinking and the mouth animation during voice input, in particular, were described as increasing the realism and social presence: "*I liked that the mouth moved when you spoke. That definitely supported that the other person was saying something. I also liked it when the avatars blinked, I think they blinked, that it felt a little more natural.* "

In many interviews, the participants described the lack of facial expressions inhibited overall sociality. Some of them specifically referred to the perceived emotionality, social presence, personal closeness, and naturalness: "*[...] but you couldn't see the facial expressions. When someone strains and looks strained. You just didn't notice that.* " In particular, players emphasized how mimics could have enhanced the experience of shared laughter: "*So you also laughed a lot via the headset, but when you see it on the spot, it's something else again, so you feel it even more. Then this sharing of feelings is even more intense.* "

Another inhibitor of perceived emotionality and realism was the lack of nuanced body language, such as detailed body poses. The participants, for example, complained that the avatars could not express the tension during gameplay: "*But you could really see [in co-located] whether the person [...] was suddenly really tense and controlled, that is, was also very concentrated and focused, or somehow played a bit more "laid-back" and chilled out. That has been one of the great aspects, in comparison.* "

Social touch was a social interaction that participants occasionally described as an inhibitor of social presence. In particular, they reported that when a virtual touch happens, but no physical touch is perceived, one becomes aware of the physical distance again. This was further supported by having no visual collisions happening in VR, as avatar models "reached through" each other: " *Except just this obstacle of not feeling each other, even though you still have the visual.* "

However, the participants also relativized the importance of the limitations above. In particular, during gameplay, one tends to focus on the game. Then, limitations such as the lack of facial expressions of the game partner's avatar played a lesser role: "*I don't really think it has an impact because I wasn't looking over at her the whole time [...], I was just looking at the screen and [...] it felt like she was sitting next to me. It doesn't matter what she looks like.* " For some participants, it was also sufficient to hear each others' voices, and thus not a significant issue, that avatars suffered from limited social stimuli: "*I didn't even find the facial expressions absolutely necessary. [...] Because you could take a lot out of the voice, the pitch of the voice, the color of the voice. So how the emotional situation is, how the mood is.* "

7.7.5 ***Theme 5: The majority of participants transferred their relationship dynamics to VR.*** Most participants indicated that knowing the other person was necessary for their overall experience in the VR scenario. For example, knowing their partner's voice and how they look in reality relativized certain avatar limitations. Moreover, knowing each other resulted in isolated benefits for the social and player experience, e.g., a blind mutual understanding while controlling the game was just as possible in VR as in co-located. Further, friends who play a lot together online via voice chat are used to communicate mainly via voice and are adapted to understand each other emotionally in this situation. Furthermore, the majority of participants paraphrased how the usual social dynamics between them transferred to VR: " *Yeah, exactly. So, we were just being ourselves. [...] Like we always are, just not in the same room.* " In individual cases, the participants noted how their behavior and emotions slightly changed in VR, e.g., hitting each other for fun but emphasizing that they would not do this in-person. In another case, the perceived social distance affected the emotional reaction to in-game failure: "*[...] it was just somehow so distant, that it has already stressed me a bit, if it did not work as I wanted in the moment. I didn't have that before [in presence], but then in VR, it made me a little bit angry: Why isn't the ball jumping now?* "

7.7.6 ***Theme 6: Novelty effect of VR, technological aspects, and study design affect sociality.*** The interviews revealed that being in VR enhanced sociality in terms of increased liveliness and enjoyment. This was predominantly attributed to the novelty of VR for many participants. As being one of only few prior experiences or even the first VR encounter at all, the participants reported that this lack of experience stimulated their curiosity and tempted them to interact with each other: "*[...] when we were just on the couch during the break, we weren't really interacting with each other much, [...] But in the VR world, we were there playing around a little bit more, teasing and seeing what you could do [...].* "

However, in some cases, the VR environment distracted participants to the point that they, e.g., engaged less in cooperation or personal conversation: "*Not really, except that in presence, in pause, we talked about something personal. And in VR, we rather tested the limits of the game.* "

The participants saw the 3-minute break as an opportunity for intense social interaction in VR. A recurring comparison was that they did not really interact with each other at all during the break in the co-located scenario, neither did they stand up: "*Yes, yes, with the avatars you also tried to shake hands or something like that. [...] And if you're just sitting on the couch, then, I don't know, only the everyday conversations come up, I would say.* "

One of the prevailing notions of how VR negatively influenced sociality was related to the VR headsets' technical specifications. In particular, due to the restricted field of view, one could not perceive the partner in VR's peripheral field of view. Though, the other person would still be visible in the exact spatial constellation in the co-located situation. This impaired, especially, the perceived social presence during gameplay: "*But when you're looking straight ahead, you can't see the avatar next to you. But when you're sitting next to each other, you still feel the physical presence of the other person.*"

The condition sequence was occasionally discussed as a sociality-enhancing factor, i.a.: getting used to each other in the first condition and then being more relaxed in the following VR condition, or a decrease of excitement or social focus when transitioning from VR to co-located: "*The first time was just like: cool VR, this is all fun, we're in a game, so let's play. The second time was more: Come on, let's go through the levels now.*"

The study design and context were also occasionally described as limiting certain aspects of the sociality. For example, some participant pairs that did not know each other well noted less social activity in VR, in case it was the first condition because they had yet to warm up to each other. In another instance, a participant had to get used to the study context and the audio-video recording. "*In the first one (VR) you were still rather biased because of....you knew that you were being recorded and that you were being listened to. And so on. And that's the kind of thing you forget at some point [...]*" Further, because of the study context, some participants consciously dispensed crazy avatar designs in the customization, as this would have been inappropriate in the study context.

7.7.7 *Theme 7: Simulation co-located gaming in VR was accepted by users and found superior to existing options. However, VR technology presents some challenges that require attention*.

Most players could imagine using Couch-Coop VR in their private leisure time (Figure 9), in particular, as it allows them to overcome physical distance and to feel close to friends and family members in specific use cases. In some cases, it was also noted as a convenient alternative to in-person meetings, e.g., when you do not want to drive late in the evening. A few explicitly emphasized that they would have liked to use such an application during the COVID-19's social distancing and quarantine measures.

Most participants evaluated the scenario as superior to current alternatives of playing online games or communicating over a distance. They attributed this mainly to the avatars and the intensified perception of social presence within a spatial context, which induces a more personal experience. Concerning the game experience, some users described how being immersed in VR helped them to focus more on the actual game. Though, a recurring request was that such a scenario must offer a variety of games on the virtual screen. In addition, a few participants hinted that they would like to play a "typical" VR game together and deviate from playing couch-coop games in front of a virtual monitor. Besides that, some participants wished for scenarios or features that would provoke more direct social interactions. Relating to the avatars, the participants would add mimics and legs, as well as more diverse customization features, i.a.: hairs, head-coverings like hijabs, and body shapes.

The interviews also revealed specific challenges and general concerns about using such a scenario or VR technology in the future. Participants mostly noticed challenges due to current technical limitations: blurry view due to limited resolution, a narrow field of view, and bad headset ergonomics regarding weight and fit. They also found it too expensive and worried that it is or will not be widely owned. Further, we identified two general concerns towards VR that were occasionally described: the increased probability of pathological escapism and reduced face-to-face meetings.

8 DISCUSSION

The main goal of our research was to explore how well social VR may function as an alternative to f2f when leisure activities in person are not feasible. We thus determined, in an exploratory user study, how the experience of a co-located multiplayer game scenario replicated in social VR compares to its f2f counterpart. Answering our research questions, we can summarize the following findings.

RQ1: Comparison of Player and Social Experience. In our study, players had a similar player experience in both conditions, with immersion even a bit more intense in VR. In terms of sociality, however, the VR version induced a slightly less rich experience. Furthermore, the VR version generated less co-presence, mutual attention and understanding, and affective understanding. However, we consider the identified differences marginal since both scenarios induced high social richness and social presence overall. Additionally, both conditions increased social connectedness after playing with each other. The qualitative findings show that the participants could transfer their usual relationship dynamics into VR and that most would use the VR scenario privately.

RQ2: Enhancing and Inhibiting Factors. In principle, the shared control game generated social interaction in both conditions. However, in VR, a high level of interaction was particularly evident during the three-minute break. The voice chat and the avatars' presence were essential for the social experience in VR. In addition, avatar customization allowed the introduction of a personal touch, which induced conversation. Moreover, VR generated a high motivation for interaction for novice users. However, we identified specific limitations of the avatar system as the main inhibitors to the social experience: missing lower body and a lack of facial expressions. Those limitations are mainly due to the current tracking capabilities of the *Meta Quest 2*. Similarly, mutual awareness during gameplay in VR suffered from limited peripheral vision compared to human's natural field of view.

8.1 Joint Media Consumption in Social VR

Our results align with and extend previous work that compared f2f and VR interactions in the context of joint media consumption [48, 57, 63, 88].

In the scenarios of joint video watching and photo sharing, VR performed slightly worse in terms of interaction quality than the f2f variant [48, 57, 63]. Perceived interaction quality includes aspects such as mutual attention and perceived emotions [48]. As we applied the NMMSP, which assesses these aspects separately, we can specify previous insights: both aspects are individually less pronounced in VR. However, emotional interdependence did not differ statistically in our setting; even if there were differences in mutual perception between VR and f2f, the users could perceive and reveal so much about each other that they were emotionally interdependent in both scenarios to a similar degree. By looking at our qualitative results, we mainly attribute this to the voice chat that enabled co-players to assess each other's emotions.

The contexts of video watching and photo sharing yielded ambivalent results regarding perceived social meaning in VR compared to f2f. While for video watching, it was lower in VR [63], it was not significantly lower in VR for photo sharing [48]. Social meaning stems from feelings of togetherness, enjoyment, and bonding processes [48]. These components reflect our measures of enjoyment, the inclusion of and belonging to the other, and co-presence. While co-presence and inclusion of the other were marginally lower in our VR condition than in f2f, we found no significant differences in perceived belonging. Thus, our results align with both previous works but give a more nuanced insight. As we assessed social closeness at the study start, we can extend the photosharing results by finding positive social outcomes from collaborative interaction in VR. The differences between ours and others' findings are presumably related to the different spatial orientations of users to each other: facing each other during photosharing vs. predominantly looking at a screen during gameplay and video watching.

8.2 Activity Context Counterbalances Inhibitors

The themes we extracted from interview data align well with previously derived themes, i.a.: avatars with limited behavioral realism due to a lack of or poor implementation of mimics and body language annoyed users, and novelty effect excites users [48, 63, 65]. Though, we extend these by also reflecting on aspects like the role of player relationships, avatar customization, and virtual social touch.

The identified inhibitors in our scenario had limited practical relevance, as we found relatively marginal experiential differences. Co-presence and affective understanding differed most between the scenarios, and we primarily relate these differences to the limited field of view and the lack of facial expressions. As our qualitative findings point out, the participants could not perceive each other within their peripheral view during gameplay, which inhibited their perceived co-presence and attentional allocation compared to the f2f setting. At the same time, specific missing social stimuli did not stand out so much during gameplay as during the break. Though, when actively looking at the partner in VR during gameplay, the lack of mimics and nuanced body postures presumably led to the significant difference in terms of affective understanding and social richness we found. Our qualitative results indicate that these limitations were primarily perceivable during the break, where participants actively interacted with and focused on each other in VR. However, the interaction during the break was again conducive to sociality. In short, the avatars were both facilitators and inhibitors of overall sociality in VR. However, the inhibiting aspects can become less salient depending on the activity context, i.e., playing and focusing on a game together. Additionally, avatar limitations become even less critical if the scenario provides a well-functioning voice chat as the primary source for mutual emotional understanding [40, 48, 65]. We assume that providing a different yet socially stimulating focus point than the avatars is also the main reason previous work in similar contexts found VR to be quite similar to f2f despite using less sophisticated social VR environments [48, 57, 63]. Thus, we infer that joint media consumption is a suitable use case to be replicated in social VR despite its current technical limitations. Moreover, our qualitative analyses highlight that next to the activity context, VR-exclusive features, like avatar customization, create opportunities for social interaction that does not occur in the f2f scenario. In this regard, we consider the break as a crucial design decision supporting our VR scenario's overall sociality.

8.3 Novelty Effects

In line with other studies [19, 48], we also observed novelty effects based on recorded behavior and participant statements in the interviews. Moreover, in our case, they certainly compensated for sociality-inhibiting effects by triggering more social interaction during the game break. However, whether the experience in VR would be significantly worse than in f2f in the long run due to habituation and the omission of novelty effects cannot be conclusively assessed. Recent studies indicate adaptation processes of the users, which may even lead to an increase of individual aspects of the experience over time [28, 40]. Furthermore, as partly desired in the interviews and suggested by related work [48], we emphasize that VR designers should utilize such novelty effects and intentionally deviate from reality [58, 75, 92].

8.4 An Alternative for Socially Meaningful Game Experiences

From a broader perspective, our findings regarding positive social outcomes, i.e., increased social connectedness, align with literature describing the benefits that users already derive from consumer social VR platforms [5, 53, 69, 89]. Our qualitative findings further align with previous observations indicating that existing social dynamics naturally transfer into VR [65]. Consequently, our results also extend the literature comparing different types of multiplayer gaming [24–26] by proposing a new form of online gaming that theoretically allows the interaction aspects particular to local multiplayer scenarios [43, 78, 86] to be experienced in an online context. Thus, our results blur the boundaries between local and online multiplayer, despite certain stimuli, such as mimics, haptic feedback, and body postures, are still limited or missing in our specific VR scenario. However, to avoid undermining the openness to VR technology, we emphasize the importance of adequate software features and interaction designs that compensate for confounding effects that technical limitations may have.

Eventually, we interpret the identified experiential differences as being of low practical relevance in real world use cases, where a f2f meeting is not feasible. Users may experience the social VR variant as slightly different in direct comparison, but we assume it can yet be a sufficient source for social experiences over distance.

8.5 Limitations

Our results are subject to the following limitations. First, they stem from a WEIRD (Western, Educated, Industrialized, Rich, and Democratic) convenience sample [49] and we did not specify a certain population that it may represent. Thus, our results cannot be statistically generalized to a specific population, e.g., all potential future social VR users. However, our exploratory research perspective and findings based on quantitative as well as qualitative data informs more specific follow-up research for extended theory-building [4].

Secondly, the studied scenario particularly thrives on playing it at home in a cozy environment and possibly with more than two players and varying game modes. Since some participants also mentioned the influence of the study setting as inhibiting, we admit that a lab-based study context cannot easily recreate an equivalent atmosphere. However, given our within-participant design, we assume that the lab setting did affect both conditions similarly. Ultimately, longitudinal in situ research [65] is necessary to assess the real-world value of our VR scenario in different social and game contexts (e.g., compete with remote groups of friends or strangers while being at home) and the extent of assumed habituation and adaptation effects [28, 40].

Finally, as the recently launched *Meta Quest Pro* features integrated face-tracking, follow-up studies may reveal even more similar experiences between VR and f2f. However, costing three times as much as the *Quest 2* and primarily addressing business customers, face tracking may now be an available but also expensive and not yet widely integrated feature. Additionally, the *Quest 2* is currently by far the most popular consumer VR headset[9]. Therefore, we assume our prototype continues to reflect the social VR venues and experiential qualities that most VR users will have access to in the foreseeable future. Thus, our results retain external validity until technological advancements become more economically affordable, widely implemented, and functionally optimized. Also, tracking technologies were not the focus of our research but only one of many factors shaping the overall player and social experience in a specific social VR use case. However, despite technical limitations, our VR prototype already produced an experience very similar to the f2f situation. Accordingly, future studies have to reveal if and to what extent more advanced VR hardware like the *Quest Pro* will enhance the overall experience and how users may reevaluate the lack of features like face-tracking once they have experienced them. Thus, ours and others’ findings [48, 63] make a lasting contribution to the field, by demonstrating what current widespread consumer hardware is capable of and by providing a benchmark for evaluating the experiential benefits and user adoption of technological advancements in future studies.

8.6 Social VR Design Implications

We provide the following implications for developers and researchers to design and create compelling social VR experiences based on our results.

As long as VR hardware limitations exist (e.g., the restricted field of view and the lack of face tracking), software solutions should be designed to compensate for inhibiting effects. For example, another user’s presence or attention and affective state could be indicated by indicators in the field of view [10, 77].

If relevant in the context, avatars should provide mimics, one of the most important stimuli to contribute to sociality [68, 81]. Since manually triggered facial expressions or emojis can be burdensome [40, 48], approaches that generate facial expressions from intuitive-to-use stimuli like voice or hand gestures seem promising (as is done in *Rec Room* or *Horizon Worlds*. Alternatively, different forms of sharing affective states, e.g., based on biosignal visualization, open up exciting design spaces [47]. However, privacy-related concerns are discussed in recent social VR literature [55, 66, 90]. We thus expect that certain users are not willing to use face or other tracking technologies in individual contexts. For such instances, alternative approaches of manual and automated sharing of affective states should be explored and reconciled with user privacy claims.

We question whether photorealistic avatars based on users’ physical appearances are something future users ultimately want or expect in leisure activity contexts. Our qualitative results indicate, in line with other findings, openness for deviations from graphical realism [32, 52, 101]. As existing social VR applications like *VRChat* allow users even to use non-anthropomorphic avatars, ongoing research on the individual contributions of graphical and behavioral realism to a compelling social experience [9, 44, 45, 68, 74, 101] become more practically relevant. In particular, future work should consider how beyond-human avatars that deviate from anthropomorphic traits limit or advance social experiences in social VR encounters. For example, a prior work reported that embodying virtual animals (in a single-user context) induces enjoyment [42], and we believe that social encounters generally could benefit from the mere entertainment values of such approaches. However, we consider the following aspects as factors that might shape social experiences with beyond-human avatars: (i) user preferences and openness to embody and interact with such avatars, (ii) relevance of presented or missing social stimuli (e.g., eye contact) in specific contexts, (iii) presence of intelligible social stimuli, even when presented on graphically non-anthropomorphic avatars (e.g., an animal with human mimics), and (iv) user habituation effects to limited or unfamiliar social stimuli (e.g., exclusively offered stimuli in VR for social interaction like particle effects for high-five interaction [58, 75, 92]). Consequently, avatar design choices should consider the aforementioned factors and how they might stimulate sociality in a specific use case independent from the overall graphical or behavioral realism.

Finally, we are aware that in a real-world use case, users most likely do not have the same room layouts and interior designs as they had in our lab setup. Thus, developers and researchers should continue exploring previous approaches that enable convenient locomotion and interaction within a shared virtual environment that does not inhibit a natural social interaction [85].

8.7 Games User Research Implications

Based on our results, we also define the following implications for games user research.

Shared control demonstrated as a simple yet scalable game mechanic to create high player interdependence and communication. Thus, we recommend it as a mechanic to investigate and promote social dynamics in multiplayer game contexts.

Using VR to simulate co-located multiplayer gaming, where players traditionally sit together, deviates from the paradigm of utilizing VR’s spatiality to engage users in walking around. Thus, we see it as a worthwhile scenario for players who do not want or cannot engage in such activity but nevertheless want to play with others over distance in a socially rich manner. Further, this scenario introduces the possibility to experience a previously “friend-exclusive" game situation with strangers and is therefore possibly an access to social dynamics that other game media do not offer so far. Thus, as

[9]Global Headsets Market Share 2022: https://www.counterpointresearch.com/global-xr-ar-vr-headsets-market-share/

a prior work suggests [88], research should further explore this scenario as a novel online gaming context with unique opportunities to not only create innovative experiences but also offer meaningful experiences for players with diverse and individual requirements.

Similar to suggestions from related work, simulating co-located multiplayer in VR, provides a tool for remote testing of co-located games in a standardized environment [88] that inherently provides various opportunities for behavioral observation during gameplay.

9 CONCLUSION

Social VR is a comparatively new communication technology that has yet to prove its capability to provide meaningful social experiences in contexts people value and where in-person encounters may not be feasible. Therefore, we developed a custom social VR application for such a context and compared it in detail with its f2f counterpart: co-located multiplayer gaming. We found that the VR application matched the player experience and closely approximated the social experience of the conventional in-person scenario. However, the lack of facial animations, a limited body language, and a low field of view inhibited facets of the social experience. Our exploratory findings and implications inform follow-up research for extended theory-building regarding the interplay between social VR features, user characteristics, and experiential qualities. In a real-world use case where people may not have the option to decide between VR and f2f, we consider the identified experiential differences as being of low practical relevance and social VR as a source for sufficient social experiences. Consequently, if consumer VR technology advances and becomes more affordable, we anticipate an increasing everyday value of virtual social leisure activities as a means to connect over distance.

ACKNOWLEDGMENTS

We would like to express our gratitude to all those who supported the realization of this work: Erique Schnorbus, Anastasia Chernetskaya, Dana Schulte, Aileen Kossira, Katja Rogers, and Lennart E. Nacke. This work was supported by the German Federal Ministry of Education and Research (BMBF).

REFERENCES

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies* 135 (2020), 102370. https://doi.org/10.1016/j.ijhcs.2019.102370 pages #6

[2] Arthur Aron, Elaine N Aron, and Danny Smollan. 1992. Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology* 63, 4 (1992), 596. https://doi.org/10.1037/0022-3514.63.4.596 pages #7

[3] Jeremy N Bailenson, Jim Blascovich, Andrew C Beall, and Jack M Loomis. 2001. Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments* 10, 6 (2001), 583–598. https://doi.org/10.1162/105474601753272844 pages #3

[4] Sebastian Baltes and Paul Ralph. 2022. Sampling in Software Engineering Research: A Critical Review and Guidelines. *Empirical Softw. Engg.* 27, 4 (jul 2022), 31 pages. https://doi.org/10.1007/s10664-021-10072-8 pages #5, #14

[5] Miguel Barreda-Ángeles and Tilo Hartmann. 2022. Psychological benefits of using social virtual reality platforms during the covid-19 pandemic: The role of social and spatial presence. *Computers in Human Behavior* 127 (2022), 107047. https://doi.org/10.1016/j.chb.2021.107047 pages #2, #3, #14

[6] Steve Benford, John Bowers, Lennart E Fahlén, Chris Greenhalgh, and Dave Snowdon. 1995. User embodiment in collaborative virtual environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (Denver, Colorado, USA). ACM Press/Addison-Wesley Publishing Co., USA, 242–249. https://doi.org/10.1145/223904.223935 pages #3

[7] Frank Biocca, Chad Harms, and Judee K. Burgoon. 2003. Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoperators and Virtual Environments* 12, 5 (2003), 456–480. https://doi.org/10.1162/105474603322761270 pages #2

[8] Frank Biocca, Chad Harms, and Jenn Gregg. 2001. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. *4th annual international workshop on presence, Philadelphia, PA* (2001), 1–9. pages #2, #6

[9] Jim Blascovich. 2002. A Theoretical Model of Social Influence for Increasing the Utility of Collaborative Virtual Environments. In *Proceedings of the 4th International Conference on Collaborative Virtual Environments* (Bonn, Germany) *(CVE '02)*. Association for Computing Machinery, New York, NY, USA, 25–30. https://doi.org/10.1145/571878.571883 pages #15

[10] Riccardo Bovo, Daniele Giunchi, Muna Alebri, Anthony Steed, Enrico Costanza, and Thomas Heinis. 2022. Cone of Vision as a Behavioural Cue for VR Collaboration. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 502 (nov 2022), 27 pages. https://doi.org/10.1145/3555615 pages #15

[11] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806 pages #7

[12] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47. https://doi.org/10.1002/capr.12360 pages #7, #8

[13] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498 pages #6

[14] Annemaree Carroll, Julie M. Bower, and Sandy Muspratt. 2017. The conceptualization and construction of the Self in a Social Context—Social Connectedness Scale: A multidimensional scale for high school students. *International Journal of Educational Research* 81 (2017), 97–107. https://doi.org/10.1016/j.ijer.2016.12.001 pages #7

[15] Sebastian Cmentowski and Jens Krüger. 2020. *Playing With Friends - The Importance of Social Play During the COVID-19 Pandemic.* Association for Computing Machinery, New York, NY, USA, 209–212. https://doi.org/10.1145/3383668.3419911 pages #2

[16] J Cohen. 1988. Statistical power analysis for the behavioral sciences (Revised ed.). Hillsdale, NJ: Lawrence Earlbaum Associates. pages #7

[17] Yvonne A. W. de Kort and Wijnand A. Ijsselsteijn. 2008. People, places, and play: player experience in a socio-spatial context. *Computers in Entertainment* 6, 2 (2008), 1–11. https://doi.org/10.1145/1371216.1371221 pages #2, #3

[18] Ansgar E Depping and Regan L Mandryk. 2017. Cooperation and interdependence: How multiplayer games increase social closeness. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play.* Association for Computing Machinery, New York, NY, USA, 449–461. https://doi.org/10.1145/3116595.3116639 pages #5

[19] Elena Dzardanova, Vlasios Kasapakis, Damianos Gavalas, and Stella Sylaiou. 2022. Virtual reality as a communication medium: a comparative study of forced compliance in virtual reality versus physical world. *Virtual Reality* 26, 2 (2022), 737–757. https://doi.org/10.1007/s10055-021-00564-9 pages #3, #14

[20] Katharina Emmerich and Maic Masuch. 2017. The Impact of Game Patterns on Player Experience and Social Interaction in Co-Located Multiplayer Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17)*. Association for Computing Machinery, New York, NY, USA, 411–422. https://doi.org/10.1145/3116595.3116606 pages #6

[21] Entertainment Software Association (ESA). 2021. 2021 Essential Facts about the Video Game Industry. https://www.theesa.com/resource/2021-essential-facts-about-the-video-game-industry/ Washington, D.C., USA. pages #2

[22] Andy P Field. 2005. Kendall's coefficient of concordance. *Encyclopedia of Statistics in Behavioral Science* 2 (2005), 1010–11. pages #7

[23] Guo Freeman and Dane Acena. 2021. Hugging from A Distance: Building Interpersonal Relationships in Social Virtual Reality. In *ACM International Conference on Interactive Media Experiences* (Virtual Event, USA) *(IMX '21)*. Association for Computing Machinery, New York, NY, USA, 84–95. https://doi.org/10.1145/3452918.3458805 pages #3

[24] Brian Gajadhar, Yvonne de Kort, and Wijnand IJsselsteijn. 2008. Influence of Social Setting on Player Experience of Digital Games. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (Florence, Italy) *(CHI EA '08)*. Association for Computing Machinery, New York, NY, USA, 3099–3104. https://doi.org/10.1145/1358628.1358814 pages #2, #14

[25] B. J. Gajadhar, Y. A.W. de Kort, and W. A. Ijsselsteijn. 2009. Rules of Engagement: Influence of Co-Player Presence on Player Involvement in Digital Games. *International Journal of Gaming and Computer-Mediated Simulations* 1, 3 (2009), 14–27. https://doi.org/10.4018/jgcms.2009070102 pages #2

[26] Brian J. Gajadhar, Yvonne A. W. de Kort, and Wijnand A. Ijsselsteijn. 2008. Shared Fun Is Doubled Fun: Player Enjoyment as a Function of Social Setting. In *Fun and Games*, Panos Markopoulos, Boris de Ruyter, Wijnand IJsselsteijn, and Duncan Rowland (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 106–117. https://doi.org/10.1007/978-3-540-88322-7_11 pages #2, #14

[27] Linda Graf, Maximilian Altmeyer, Katharina Emmerich, Marc Herrlich, Andrey Krekhov, and Katta Spiel. 2022. Development and Validation of a German Version of the Player Experience Inventory (PXI). In *Mensch und Computer 2022 - Tagungsband*, Bastian Pfleging, Kathrin Gerling, and Sven Mayer (Eds.). ACM, New York, 265–265. https://doi.org/10.1145/3543758.3543763 pages #6

[28] Eugy Han, Mark Roman Miller, Nilam Ram, Kristine L Nowak, and Jeremy N Bailenson. 2022. Understanding Group Behavior in Virtual Reality: A Large-Scale, Longitudinal Study in the Metaverse. *72nd Annual International Communication Association Conference, Paris, France* (2022). pages #3, #14, #15

[29] Chad Harms and Frank Biocca. 2004. Internal Consistency and Reliability of the Networked Minds Social Presence Measure. In *Seventh Annual International Workshop: Presence 2004* (Valencia: Universidad Politecnica de Valencia), In M. Alcaniz & B. Rey (Ed.). pages #4, #6

[30] John Harris and Mark Hancock. 2019. To Asymmetry and Beyond! Improving Social Connectedness by Increasing Designed Interdependence in Cooperative Play. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300239 pages #5, #6

[31] John Harris, Mark Hancock, and Stacey D Scott. 2016. Leveraging asymmetries in multiplayer games: Investigating design elements of interdependent play. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 350–361. https://doi.org/10.1145/2967934.2968113 pages #5

[32] Daniel Hepperle, Christian Felix Purps, Jonas Deuchler, and Matthias Wölfel. 2022. Aspects of visual avatar appearance: self-representation, display type, and uncanny valley. *The Visual Computer* 38, 4 (2022), 1227–1244. pages #15

[33] Sture Holm. 1979. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* 6, 2 (1979), 65–70. http://www.jstor.org/stable/4615733 pages #7

[34] Ioanna Iacovides and Elisa D. Mekler. 2019. The Role of Gaming During Difficult Life Experiences. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300453 pages #2

[35] Wijnand A IJsselsteijn, Yvonne A W de Kort, and Antal Haans. 2006. Is this my hand I see before me? The rubber hand illusion in reality, virtual reality, and mixed reality. *Presence: Teleoperators and Virtual Environments* 15, 4 (2006), 455–464. https://doi.org/10.1162/pres.15.4.455 pages #3

[36] Katherine Isbister. 2010. Enabling Social Play: A Framework for Design and Evaluation. In *Evaluating User Experience in Games*, Reg-ina Bernhaupt (Ed.). Springer London, London, 11–22. https://doi.org/10.1007/978-1-84882-963-3_2 pages #2

[37] Sukran Karaosmanoglu, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. 2021. Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 443, 15 pages. https://doi.org/10.1145/3411764.3445492 pages #5

[38] Linda K. Kaye and Jo Bryce. 2012. Putting The "Fun Factor" Into Gaming: The Influence of Social Contexts on Experiences of Playing Video-games. *International Journal of Internet Science* 7, 1 (2012), 23–36. pages #2

[39] Linda K Kaye and Jo Bryce. 2014. Go with the flow: The experience and affective outcomes of solo versus social gameplay. *Journal of Gaming & Virtual Worlds* 6, 1 (2014), 49–60. https://doi.org/10.1386/jgvw.6.1.49_1 pages #2

[40] Negar Khojasteh and Andrea Stevenson Won. 2021. Working together on diverse tasks: A longitudinal study on individual workload, presence and emotional recognition in collaborative virtual environments. *Frontiers in Virtual Reality* 2 (2021), 643331. https://doi.org/10.3389/frvir.2021.643331 pages #3, #5, #14, #15

[41] Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* 21, 4 (2012), 373–387. https://doi.org/10.1162/PRES_a_00124 pages #3

[42] Andrey Krekhov, Sebastian Cmentowski, Katharina Emmerich, and Jens Krüger. 2019. Beyond Human: Animals as an Escape from Stereotype Avatars in Virtual Reality Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 439–451. https://doi.org/10.1145/3311350.3347172 pages #15

[43] Sonja Kröger and Thorsten Quandt. 2014. Get Together: Console Playing as a Group Experience. In *Multiplayer*, Thorsten Quandt and Sonja Kröger (Eds.). Routledge, London, 147–161. pages #2, #14

[44] Christos Kyrlitsias and Despina Michael-Grigoriou. 2022. Social Interaction With Agents and Avatars in Immersive Virtual Environments: A Survey. *Frontiers in Virtual Reality* 2 (2022). https://doi.org/10.3389/frvir.2021.786665 pages #15

[45] Marc Erich Latoschik, Daniel Roth, Dominik Gall, Jascha Achenbach, Thomas Waltemate, and Mario Botsch. 2017. The Effect of Avatar Realism in Immersive Social Virtual Realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg, Sweden) *(VRST '17)*. Association for Computing Machinery, New York, NY, USA, Article 39, 10 pages. https://doi.org/10.1145/3139131.3139156 pages #15

[46] Christopher Lee, Peta Wyeth, Daniel Johnson, and Joshua Hall. 2015. Flow during Individual and Co-operative Gameplay. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '15*, Anna L. Cox, Paul Cairns, Regina Bernhaupt, and Lennart Nacke (Eds.). ACM Press, New York, New York, USA, 103–107. https://doi.org/10.1145/2793107.2793142 pages #2

[47] Sueyoon Lee, Abdallah El Ali, Maarten Wijntjes, and Pablo Cesar. 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* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 425, 15 pages. https://doi.org/10.1145/3491102.3517451 pages #15

[48] Jie Li, Yiping Kong, Thomas Röggla, Francesca De Simone, Swamy Ananthanarayan, Huib de Ridder, Abdallah El Ali, and Pablo Cesar. 2019. Measuring and Understanding Photo Sharing Experiences in Social Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300897 pages #3, #4, #5, #6, #13, #14, #15

[49] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How WEIRD is CHI?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 143, 14 pages. https://doi.org/10.1145/3411764.3445488 pages #14

[50] Matthew Lombard, Theresa B Ditton, and Lisa Weinstein. 2009. Measuring presence: the temple presence inventory. In *Proceedings of the 12th annual international workshop on presence*. 1–15. pages #6, #7

[51] Anna Loparev, Walter S. Lasecki, Kyle I. Murray, and Jeffrey P. Bigham. 2014. Introducing Shared Character Control to Existing Video Games. https://doi.org/10.1184/R1/6470180.v1 pages #5

[52] Fang Ma and Xueni Pan. 2022. Visual Fidelity Effects on Expressive Self-avatar in Virtual Reality: First Impressions Matter. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Christchurch, New Zealand, 57–65. https://doi.org/10.1109/VR51125.2022.00023 pages #15

[53] Divine Maloney and Guo Freeman. 2020. Falling Asleep Together: What Makes Activities in Social Virtual Reality Meaningful to Users. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Canada) *(CHI PLAY '20)*. Association for Computing Machinery, New York, NY, USA, 510–521. https://doi.org/10.1145/3410404.3414266 pages #2, #3, #14

[54] Divine Maloney, Guo Freeman, and Andrew Robb. 2021. Stay Connected in An Immersive World: Why Teenagers Engage in Social Virtual Reality. In *Interaction Design and Children* (Athens, Greece) *(IDC '21)*. Association for Computing Machinery, New York, NY, USA, 69–79. https://doi.org/10.1145/3459990.3460703 pages #3

[55] Divine Maloney, Samaneh Zamanifard, and Guo Freeman. 2020. Anonymity vs. Familiarity: Self-Disclosure and Privacy in Social Virtual Reality. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) *(VRST '20)*. Association for Computing Machinery, New York, NY, USA, Article 25, 9 pages. https://doi.org/10.1145/3385956.3418967 pages #15

[56] Antonella Maselli and Mel Slater. 2013. The building blocks of the full body ownership illusion. *Frontiers in human neuroscience* 7 (2013), 83. https://doi.org/10.3389/fnhum.2013.00083 pages #3

[57] Mark McGill, John H Williamson, and Stephen Brewster. 2016. Examining the role of smart TVs and VR HMDs in synchronous at-a-distance media consumption. *ACM Transactions on Computer-Human Interaction (TOCHI)* 23, 5 (2016), 1–57. https://doi.org/10.1145/2983530 pages #3, #13, #14

[58] Joshua McVeigh-Schultz and Katherine Isbister. 2021. The Case for "Weird Social" in VR/XR: A Vision of Social Superpowers Beyond Meatspace. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, Article 17, 10 pages. https://doi.org/10.1145/3411763.3450377 pages #14, #15

[59] Joshua McVeigh-Schultz, Anya Kolesnichenko, and Katherine Isbister. 2019. Shaping Pro-Social Interaction in VR: An Emerging Design Framework. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300794 pages #3

[60] Joshua McVeigh-Schultz, Elena Márquez Segura, Nick Merrill, and Katherine Isbister. 2018. What's It Mean to "Be Social" in VR? Mapping the Social VR Design Ecology. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems* (Hong Kong, China) *(DIS '18 Companion)*. Association for Computing Machinery, New York, NY, USA, 289–294. https://doi.org/10.1145/3197391.3205451 pages #3

[61] Meta. 2020. Facebook Horizon Welcomes Its First Virtual Explorers. Retrieved August 09, 2022 from https://about.fb.com/news/2020/08/facebook-horizon-welcomes-first-virtual-explorers/ pages #2

[62] Microsoft. 2022. Altspace VR Beginner's Guide. Retrieved August 09, 2022 from https://altvr.com/altspacevr-beginners-guide/ pages #2

[63] Mario Montagud, Jie Li, Gianluca Cernigliaro, Abdallah El Ali, Sergi Fernández, and Pablo Cesar. 2022. Towards socialVR: evaluating a novel technology for watching videos together. *Virtual Reality* 26 (2022), 1–21. https://doi.org/10.1007/s10055-022-00651-5 pages #3, #4, #5, #13, #14, #15

[64] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine* 19, 2 (2012), 98–100. https://doi.org/10.1109/MRA.2012.2192811 pages #5

[65] Fares Moustafa and Anthony Steed. 2018. A Longitudinal Study of Small Group Interaction in Social Virtual Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 22, 10 pages. https://doi.org/10.1145/3281505.3281527 pages #5, #14, #15

[66] Vivek Nair, Gonzalo Munilla Garrido, and Dawn Song. 2022. Exploring the Unprecedented Privacy Risks of the Metaverse. https://doi.org/10.48550/ARXIV.2207.13176 pages #15

[67] Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand, and Jon Howard Raphael. 2012. No room for reflexivity? Critical reflections following a systematic review of qualitative research. *Journal of Health Psychology* 17, 6 (2012), 866–885. https://doi.org/10.1177/1359105311427615 pages #8

[68] Catherine S. Oh, Jeremy N. Bailenson, and Gregory F. Welch. 2018. A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI* 5 (2018). https://doi.org/10.3389/frobt.2018.00114 pages #2, #15

[69] Ye Pan and Anthony Steed. 2017. The impact of self-avatars on trust and collaboration in shared virtual environments. *PloS one* 12, 12 (2017), e0189078. https://doi.org/10.1371/journal.pone.0189078 pages #3, #14

[70] Katy E Pearce, Jason C Yip, Jin Ha Lee, Jesse J Martinez, Travis W Windleharth, Arpita Bhattacharya, and Qisheng Li. 2022. Families playing animal crossing together: coping with video games during the COVID-19 pandemic. *Games and Culture* 17, 5 (2022), 773–794. pages #2

[71] T. S. Perry. 2016. Virtual reality goes social. *IEEE Spectrum* 53, 1 (2016), 56–57. https://doi.org/10.1109/MSPEC.2016.7367470 pages #3

[72] Roosa Piitulainen, Perttu Hämäläinen, and Elisa D Mekler. 2022. Vibing Together: Dance Experiences in Social Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 188, 18 pages. https://doi.org/10.1145/3491102.3501828 pages #2, #3

[73] Katja Rogers, Sukran Karaosmanoglu, Dennis Wolf, Frank Steinicke, and Lennart E. Nacke. 2021. A Best-Fit Framework and Systematic Review of Asymmetric Gameplay in Multiplayer Virtual Reality Games. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.694660 pages #5

[74] Shane L. Rogers, Rebecca Broadbent, Jemma Brown, Alan Fraser, and Craig P. Speelman. 2022. Realistic Motion Avatars are the Future for Social Interaction in Virtual Reality. *Frontiers in Virtual Reality* 2 (2022). https://doi.org/10.3389/frvir.2021.750729 pages #3, #15

[75] Daniel Roth, Gary Bente, Peter Kullmann, David Mal, Chris Felix Purps, Kai Vogeley, and Marc Erich Latoschik. 2019. Technologies for Social Augmentations in User-Embodied Virtual Reality. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 5, 12 pages. https://doi.org/10.1145/3359996.3364269 pages #3, #14, #15

[76] Daniel Roth and Marc Erich Latoschik. 2020. Construction of the Virtual Embodiment Questionnaire (VEQ). *IEEE Transactions on Visualization and Computer Graphics* 26, 12 (2020), 3546–3556. https://doi.org/10.1109/TVCG.2020.3023603 pages #6

[77] Sylvia Rothe, Alexander Schmidt, Mario Montagud, Daniel Buschek, and Heinrich Hußmann. 2021. Social viewing in cinematic virtual reality: a design space for social movie applications. *Virtual Reality* 25, 3 (2021), 613–630. https://doi.org/10.1007/s10055-020-00472-4 pages #15

[78] Katie Salen Tekinbaş and Eric Zimmerman. 2004. *Rules of play: Game design fundamentals*. MIT press. pages #2, #14

[79] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments* 10, 3 (06 2001), 266–281. https://doi.org/10.1162/105474601300343603 arXiv:https://direct.mit.edu/pvar/article-pdf/10/3/266/1623697/105474601300343603.pdf pages #6

[80] Valentin Schwind, Katrin Wolf, and Niels Henze. 2018. Avoiding the Uncanny Valley in Virtual Character Design. *Interactions* 25, 5 (aug 2018), 45–49. https://doi.org/10.1145/3236673 pages #5

[81] Navya N. Sharan, Alexander Toet, Tina Mioch, Omar Niamut, and Jan B. F. van Erp. 2022. The Relative Importance of Social Cues in Immersive Mediated Communication. In *Human Interaction, Emerging Technologies and Future Systems V*, Tareq Ahram and Redha Taiar (Eds.). Springer International Publishing, Cham, 491–498. pages #3, #15

[82] Mel Slater. 2009. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1535 (2009), 3549–3557. https://doi.org/10.1098/rstb.2009.0138 pages #3

[83] Mel Slater, Daniel Pérez Marcos, Henrik Ehrsson, and Maria V Sanchez-Vives. 2008. Towards a digital body: the virtual arm illusion. *Frontiers in human neuroscience* 2 (2008), 6. https://doi.org/10.3389/neuro.09.006.2008 pages #3

[84] Harrison Jesse Smith and Michael Neff. 2018. Communication Behavior in Embodied Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173863 pages #3

[85] Misha Sra, Aske Mottelson, and Pattie Maes. 2018. Your Place and Mine: Designing a Shared VR Experience for Remotely Located Users. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 85–97. https://doi.org/10.1145/3196709.3196788 pages #15

[86] Jaakko Stenros, Janne Paavilainen, and Frans Mayra. 2011. Social interaction in games. *International Journal of Arts and Technology* 4, 3 (2011), 342. https://doi.org/10.1504/IJART.2011.041486 pages #2, #14

[87] Philipp Sykownik, Katharina Emmerich, and Maic Masuch. 2018. Exploring Patterns of Shared Control in Digital Multiplayer Games. In *Advances in Computer Entertainment Technology*, Adrian David Cheok, Masahiko Inami, and Teresa Romão (Eds.). Springer International Publishing, Cham, 847–867. pages #5

[88] Philipp Sykownik, Katharina Emmerich, and Maic Masuch. 2020. *Like in the Good Old Times, but Virtual - A Case for Simulating Co-Located Multiplayer Games in VR*. Association for Computing Machinery, New York, NY, USA, 379–383. https://doi.org/10.1145/3383668.3419885 pages #3, #13, #16

[89] Philipp Sykownik, Linda Graf, Christoph Zils, and Maic Masuch. 2021. The Most Social Platform Ever? A Survey about Activities & Motives of Social VR Users. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Lisboa, Portugal, 546–554. https://doi.org/10.1109/VR50410.2021.00079 pages #2, #3, #14

[90] Philipp Sykownik, Divine Maloney, Guo Freeman, and Maic Masuch. 2022. Something Personal from the Metaverse: Goals, Topics, and Contextual Factors of Self-Disclosure in Commercial Social VR. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 632, 17 pages. https://doi.org/10.1145/3491102.3502008 pages #3, #15

[91] Philipp Sykownik and Maic Masuch. 2020. The Experience of Social Touch in Multi-User Virtual Reality. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology* (Virtual Event, Canada) *(VRST '20)*. Association for Computing Machinery, New York, NY, USA, Article 30, 11 pages. https://doi.org/10.1145/3385956.3418944 pages #3

[92] Philipp Sykownik and Maic Masuch. 2021. Developing a Framework for Designing Augmentations of Social Touch in Multiuser Virtual Reality. *Mensch und Computer 2021-Workshopband* (2021). https://doi.org/10.18420/muc2021-mci-ws16-241 pages #14, #15

[93] Theresa Jean Tanenbaum, Nazely Hartoonian, and Jeffrey Bryan. 2020. "How Do I Make This Thing Smile?": An Inventory of Expressive Nonverbal Communication in Commercial Social Virtual Reality Platforms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376606 pages #5

[94] Maciej Tomczak and Ewa Tomczak. 2014. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in sport sciences* 1, 21 (2014), 19–25. pages #7

[95] Kellie Vella, Madison Klarkowski, Daniel Johnson, Leanne Hides, and Peta Wyeth. 2016. The Social Context of Video Game Play: Challenges and Strategies. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems - DIS '16*, Marcus Foth, Wendy Ju, Ronald Schroeter, and Stephen Viller (Eds.). ACM Press, New York, New York, USA, 761–772. https://doi.org/10.1145/2901790.2901823 pages #2

[96] Kurt Wagner. 2021. Zuckerberg Says Facebook's Future Lies in Virtual 'Metaverse'. Retrieved August 09, 2022 from https://www.bloomberg.com/news/articles/2021-07-29/mark-zuckerberg-explains-metaverse-vision-to-facebook-fb-investors-analysts pages #2

[97] Dmitri Williams, Nicolas Ducheneaut, Li Xiong, Yuanyuan Zhang, Nick Yee, and Eric Nickell. 2006. From Tree House to Barracks: The Social Life of Guilds in World of Warcraft. *Games and Culture* 1, 4 (2006), 338–361. https://doi.org/10.

1177/1555412006292616 arXiv:https://doi.org/10.1177/1555412006292616 pages #2

[98] Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240. https://doi.org/10.1162/105474698565686 pages #3

[99] Samaneh Zamanifard and Guo Freeman. 2019. "The Togetherness That We Crave": Experiencing Social VR in Long Distance Relationships. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing* (Austin, TX, USA) *(CSCW '19)*. Association for Computing Machinery, New York, NY, USA, 438–442. https://doi.org/10.1145/3311957.3359453 pages #3

[100] Xiaolong Zhang and George W. Furnas. 2002. Social Interactions in Multiscale CVEs. In *Proceedings of the 4th International Conference on Collaborative Virtual Environments* (Bonn, Germany) *(CVE '02)*. Association for Computing Machinery, New York, NY, USA, 31–38. https://doi.org/10.1145/571878.571884 pages #3

[101] Katja Zibrek, Sean Martin, and Rachel McDonnell. 2019. Is photorealism important for perception of expressive virtual humans in virtual reality? *ACM Transactions on Applied Perception (TAP)* 16, 3 (2019), 1–19. https://doi.org/10.1145/3349609 pages #15

11.2 A best-fit framework and systematic review of asymmetric gameplay in multiplayer virtual reality games

[CorePub7] Katja Rogers, **Sukran Karaosmanoglu**, Dennis Wolf, Frank Steinicke, and Lennart E. Nacke. A best-fit framework and systematic review of asymmetric gameplay in multiplayer virtual reality games. *Frontiers in Virtual Reality*, 2, 2021. ISSN 2673-4192. DOI10.3389/frvir.2021.694660. URL https://doi.org/10.3389/frvir.2021.694660

Supplementary materials for this article are available here: https://doi.org/10.3389/frvir.2021.694660.

frontiers in Virtual Reality

SYSTEMATIC REVIEW
published: 27 July 2021
doi: 10.3389/frvir.2021.694660

A Best-Fit Framework and Systematic Review of Asymmetric Gameplay in Multiplayer Virtual Reality Games

Katja Rogers[1,2,3], Sukran Karaosmanoglu[4], Dennis Wolf[5], Frank Steinicke[4] and Lennart E. Nacke[1,2,3]*

[1]HCI Games Group, The Games Institute, University of Waterloo, Waterloo, ON, Canada, [2]Department of Communication Arts, University of Waterloo, Waterloo, ON, Canada, [3]Stratford School of Interaction Design and Business, University of Waterloo, Waterloo, ON, Canada, [4]Human-Computer Interaction, Universität Hamburg, Hamburg, Germany, [5]Institute of Media Informatics, Ulm University, Ulm, Germany

OPEN ACCESS

Edited by:
Bernhard E. Riecke,
Simon Fraser University, Canada

Reviewed by:
Ernst Kruijff,
Hochschule Bonn-Rhein-Sieg (H-BRS), Germany
Sebastian Oberdörfer,
Julius Maximilian University of Würzburg, Germany

***Correspondence:**
Katja Rogers
katja.rogers@acm.org

Specialty section:
This article was submitted to Technologies for VR, a section of the journal Frontiers in Virtual Reality

Received: 13 April 2021
Accepted: 15 June 2021
Published: 27 July 2021

Citation:
Rogers K, Karaosmanoglu S, Wolf D, Steinicke F and Nacke LE (2021) A Best-Fit Framework and Systematic Review of Asymmetric Gameplay in Multiplayer Virtual Reality Games. Front. Virtual Real. 2:694660. doi: 10.3389/frvir.2021.694660

Increasingly, virtual reality (VR) design and research leverages gameplay asymmetries, flattening discrepancies of interface, abilities, information or other aspects between players. A common goal is to induce social interactions that draw players without head-mounted displays into a shared game world. Exploring these asymmetries resulted in many artifacts, creating an innovative yet disparate research landscape that showcases points for improvement in coverage of the field and theoretical underpinnings. In this article, we present a literature review of asymmetry in multiplayer VR games, using a framework synthesis method to assess the field through a lens of existing literature on asymmetries in gameplay. We provide an overview of this emerging subfield and identify gaps and opportunities for future research. Moreover, we discuss how research artifacts address prior theoretical work and present a "best fit" framework of asymmetric multiplayer VR games for the community to build upon.

Keywords: VR, virtual reality, asymmetry, games, systematic review, asymmetric games

1 INTRODUCTION

Virtual reality (VR) technology has received immense interest from researchers, developers, designers, and the games industry alike in recent years. Proponents praise VR's ability to facilitate novel and immersive experiences. Nevertheless, the immersive quality of VR technologies, in particular head-mounted displays (HMDs), has also received criticism for its potentially isolating characteristics, both technologically and socially (Boland and McGill, 2015). This isolation risk has inspired game designs that enhance the VR experience by including bystanders as co-players. These game designs can leverage the benefits of multiplayer game experiences (e.g., supporting social connectedness; Woods, 2009; Vella et al., 2019) without requiring additional HMDs. Examples include giving bystanders insight into the virtual environments (VEs) inhabited by the HMD player (e.g., via a monitor displaying their camera perspective; Jeong et al., 2020), or an active role in the game (e.g., giving them essential information that the HMD player requires, thus enforcing communication; Liszio and Masuch, 2016; Smilovitch and Lachman, 2019).

These interactions of HMD players with non-HMD players can be considered a type of asymmetry: *a difference in the interface with which users interact with the VE*. However, while asymmetry of player interfaces is increasingly common in VR experiences (rooted in cross-reality paradigms), there are many other asymmetry types that can be designed (Harris et al., 2016). Asymmetric games explicitly incorporate and design for differences between players in how they interact with the game, and for the abilities and information the players possess within the game world. With the influx of research in this context, asymmetric multiplayer games are emerging as a subfield of VR games research.

The VR artifacts[1] being produced—and related findings on how asymmetric game designs affect player experience (PX)—are novel and innovative because this is an emerging research field. Nevertheless, the artifacts are also disparate in their focus. In particular, few designers or researchers deeply engage with prior theories on designing for and catering to gameplay asymmetries in general (outside of VR). We argue that a systematic approach and close integration of theory-driven games research is essential to leverage the full potential of asymmetric multiplayer VR games. As a first step, we aim to gain an overview of how asymmetries are being incorporated in multiplayer VR games, investigating what asymmetries multiplayer VR games contain and how multiplayer VR asymmetries affect player experience.

To map out this research area, we conducted a literature review on asymmetric multiplayer VR games. Beginning with 481 identified records from *The ACM Guide to Computing Literature* and *Scopus* covering the majority of the human-computer interaction (HCI) literature, we followed a systematic screening and snowballing approach (via PRISMA; Moher et al., 2009) to arrive at a corpus of 25 relevant papers on asymmetric multiplayer VR games.

Given the novelty of asymmetric multiplayer VR, we employed a method based on framework synthesis (Carroll et al., 2011; Dixon-Woods, 2011) to guide our analysis of these papers against the backdrop of prior literature. Papers were synthesized within a "best fit" framework based on theoretical work on asymmetry in games, gameplay aesthetics, social aspects of PX, and shared control. This methodology works well for research fields in which there is no clear "best fit" theory: it applies multiple theories or models to reduce limitations of a single one. This allows researchers to "engage with theory but not be constrained by it" (Carroll et al., 2013). Further, as it is open to improvements on the *a priori* framework, it turns the review into an opportunity to reflect on and refine existing theory.

The goal of the review was to explore the following questions: 1) what kinds of asymmetry are being explored in research on multiplayer VR games, and 2) how asymmetry in these games affects PX. For this purpose, we sought out papers that touched on the design and/or evaluation (for any participant group, with any intervention, comparison, or outcome) of any kind of asymmetric gameplay within a multiplayer game experience featuring at least one HMD-VR player.

The contributions of this work are twofold: 1) Our work provides an overview of this emerging subfield of VR games research, shows how it engages with prior theoretical research on asymmetry, and identifies gaps in the literature to guide future research. Specifically, we point to the following as opportunities for future work: multiplayer games with more than two players, alternative interfaces to monitor variants, mirrored and unidirectional interdependence between players, remote play, shared control within the game world are rare in the design of asymmetric multiplayer VR games. Additionally, effects of player skill or familiarity with interface and partner are rarely considered. Explicitly shared physical spaces, embodied physical interaction, and the presence of a human co-player emerged as the key drivers of positive PX in asymmetric multiplayer VR. These key findings are summarized in **Table 3**. 2) Furthermore, we present, apply, and refine the "best fit" framework for the asymmetric VR games field to employ and build upon (see **Figures 2**, **3**). While time will bear out the use of the refined framework, we suggest that it has generative, structural, and analytical potential: it could inform the design of asymmetric multiplayer VR games to systematically explore the overall design space (generative), scaffold reporting and description of such games (structural), and guide both future evaluations and literature reviews of such games (analytical).

2 BACKGROUND

Progress in VR has long been entangled with "other" forms of realities resulting in mixed realities (MR; Milgram and Kishino, 1994), which enrich the real world with virtual content as in augmented reality (AR) or augmented virtuality (AV), which "*mutually reflect, influence, and merge into each other*" (Lifton, 2007). Such mutual interactions between realities is often described as *cross-reality* (XR; Want, 2009) or *blended reality* (Schmidt et al., 2019), and increasingly refers to any exchange of information between realities "*to [. . .] a meaningful and discernable effect*" (Coleman, 2009). A close and bidirectional connection between VR and the real world is necessary for collaboration in scenarios that see one user located in VR and one in the real world (Coleman, 2009; Reilly et al., 2010). The general increase in VR usage in recent years is also driving a trend toward XR designs and research (Efstratios et al., 2018; George et al., 2019; Nakagawa and Sonobe, 2019; Kiourt et al., 2020; Simeone et al., 2020; Wang et al., 2020). In particular, the social aspects of cross-reality VR research are becoming popular (Yassien et al., 2020). For example, VR researchers are exploring the questions of understanding how different environments (e.g., public, private) and familiarity among users can affect XR experiences (O'Hagan et al., 2020), how they can provide VR users with an awareness of the presence of those around them (McGill et al., 2015; O'Hagan and Williamson, 2020), and whether bystanders can understand the experiences of VR users by observing them (George et al., 2019).

VR is increasingly available to consumers. However, the usage of multiple HMDs at the same time remains rare because of the cost, required space, and potential collisions between users. VR's most appealing feature—its immersiveness—can also engender isolation (Boland and McGill, 2015; Mütterlein and Hess, 2017; Rogers et al., 2019). Thus, cross-reality consisting of the inclusion of non-VR users into the VR experience (asymmetry of interface) has been explored as a way to increase social interaction for both the HMD user and bystanders (e.g., Gugenheimer et al., 2017a; Zhou et al., 2019; Lee et al., 2020), with promising results for

[1]We use the term artifact to refer to games, prototypes, or technical systems.

enjoyment and social connection. Furthermore, it may offer an alternative to users suffering from VR sickness (Jerald, 2015; Peck et al., 2020).

Commercial VR games have also begun employing asymmetry of interface between an HMD player and one or multiple non-HMD players, such as *Keep Talking And Nobody Explodes* (Steel Crate Games, 2015), *Acron: Attack of the Squirrels* (Resolution Games, 2019), *Carly and the Reaperman—Escape from the Underworld* (Odd Raven Studios, 2019) and *Panoptic* (Team Panoptes, 2019). Yet games offer a much broader design space for asymmetries, including aspects *beyond* differences in interface. Multiplayer games enable social environments wherein players can engage, interact, and develop trust toward each other (Depping et al., 2016). They can foster relatedness between players, which supports well-being (Ryan and Deci, 2000; Deci and Ryan, 2011). However, designing for multiplayer engagement holds additional challenges, as multiple players may require the game to accommodate different abilities, preferences, or technical equipment. By integrating asymmetries in their mechanics, dynamics, and aesthetics (Harris et al., 2016), games can engage many different players.

Previous works (Manninen and Korva, 2005; Zagal et al., 2006; Beznosyk et al., 2012) have addressed points of asymmetric game design (such as distribution of information, goals, and varying levels of responsibility) in the design of collaborative games; similar aspects are reflected in game balancing to address differences in player skill (Cechanowicz et al., 2014; Vicencio-Moreira et al., 2014; Abuhamdeh et al., 2015). Harris et al. (2016) introduced a first conceptual framework of asymmetric games, using the widely adopted (Mechanics-Dynamics-Aesthetics) MDA framework (Hunicke et al., 2004) as an analytical lens for asymmetric games. In their framework, they formulated themes specific to that context (i.e., mechanics of asymmetry, dynamics of asymmetry, and aesthetics of asymmetry), which we will discuss in detail below. Further, they showed that asymmetric gameplay can increase social presence, connectedness, and immersion compared to a symmetric game (Harris and Hancock, 2019). Their findings highlight the resulting increased interdependence between players—their degree of reliance on each other (Beznosyk et al., 2012; Harris and Hancock, 2019)—as fostering positive effects on PX (e.g., higher social presence and connectedness). A recent example in VR found similar effects (Hansen et al., 2020).

In summary, we observe asymmetry in VR (particularly—but not only—of interface) as an emerging trend in XR. In the context of games, the inclusion of non-VR players in VR games has the potential to combat isolation, enhance engagement, and cater to social interaction motivations, thus prompting this review. We address further theoretical related work in more detail below as part of our systematic review of this research field.

3 METHODS

To gain a deeper understanding of this emerging research field, we conducted a systematic review of the literature. We build on the research gap and motivation outlined in the previous section: the combination of VR and non-VR users is a promising way to increase immersion and reduce potential isolation. This led us to articulate our research questions and inclusion criteria for this systematic review to focus on artifacts with at least one VR user. Our two research questions were: RQ_1:"*What kinds of asymmetry are being designed in multiplayer VR games?*" and RQ_2: "*How does VR asymmetry in multiplayer VR games affect PX?*"

3.1 Approach

We followed the PRISMA protocol for systematic reviews (Moher et al., 2009) to identify and screen records; our PRISMA protocol (Shamseer et al., 2015) is provided in the **Supplementary Materials**. We employed a technique based on framework synthesis (Carroll et al., 2011; Dixon-Woods, 2011) to thematically analyze relevant results through a theoretical lens of existing literature on asymmetric player experience.

Figure 1's PRISMA flowchart (Moher et al., 2009) details the stages (and the corresponding number of records) prior to the synthesis of this review: initial identification (N = 481 records), removal of duplicates (→N = 399), rigorous screening based on inclusion criteria (→N = 21), a snowball approach to identify additional papers (n = 6) based on the initial screening selection (→N = 27), and full-text screening for final inclusion (→N = 25). This final corpus of N = 25 records (of which n = 11 were included in quality assessment) underwent the "best fit" framework synthesis analysis.

3.1.1 Identification: Data Sources and Queries

We used two online databases to identify potential papers: The *ACM Guide to Computing Literature* published by the Association for Computing Machinery (ACM) and *Scopus*[2], Elsevier's abstract and citation database. With ACM focusing on computing sciences and Scopus providing a larger but more general coverage, these two databases together offer comprehensive access to research on VR, games, and human-computer interaction (HCI). While this necessitates two different search strings, the use of more than one database is a mark of quality and comprehensive coverage in reviews.

Starting with our research questions, we phrased and tested search queries to target asymmetry in VR games via the keywords asymmetr*, VR/virtual reality, and game*. Additionally, in a recent study exploring inter-dependencies between players within a custom asymmetric VR game, agency emerged as an important aspect in the asymmetric gameplay experience (Karaosmanoglu et al., 2021). In this study, we followed the definition put forth by Murray (2017) as "the satisfying power to take meaningful action and see the results of our decisions and choices." This PX component is particularly relevant in asymmetric VR games: creating interdependence between players can easily cause one player to have a more active or dominant role, which manifests as a stronger ability to take meaningful action. This is potentially amplified in VR when users are separated from each other by

[2]For detailed information on Scopus's indexed resources (e.g., Institute of Electrical and Electronics Engineers; IEEE), please refer to the website: https://www.scopus.com/sources and the Scopus Content Coverage Guide: https://www.elsevier.com/?a=69451.

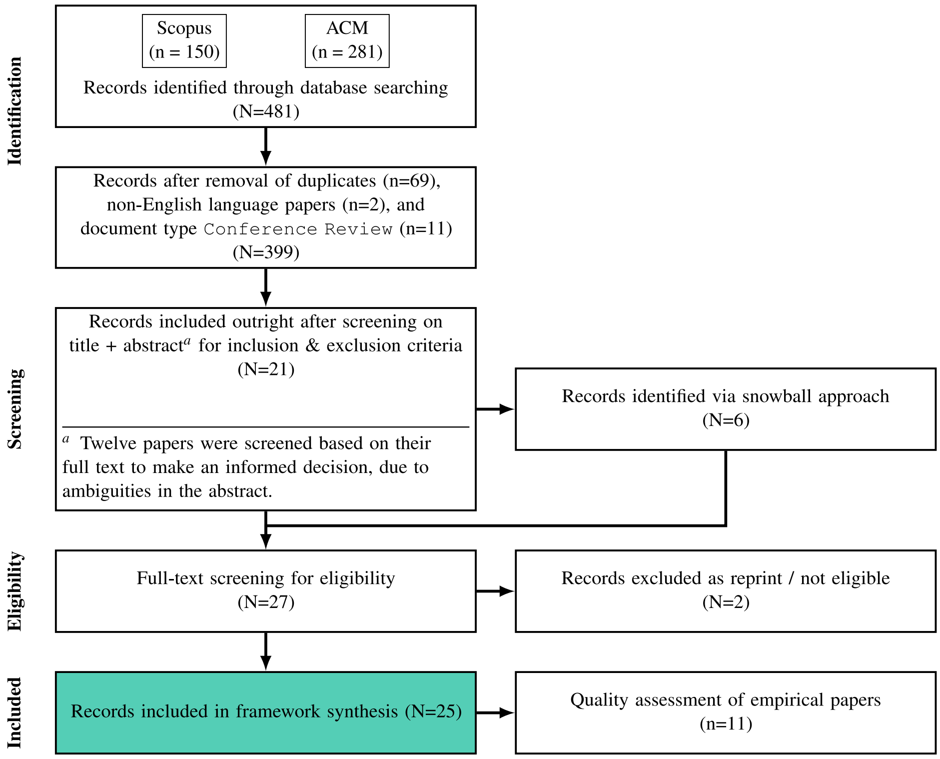

FIGURE 1 | PRISMA flowchart detailing the identification, screening, and eligibility stages prior to the framework synthesis.

TABLE 1 | Data sources, queries, dates, numbers and publication dates of results.

Data source	Query/Description of process	Results	Publication dates
Scopus *Date: May 13, 2020*	(ABS (VR OR "virtual reality") AND ABS (asymmetr* OR agency) AND ALL (game*))	150	All times
ACM *Date: May 13, 2020*	"Query": {Abstract: (VR OR "virtual reality") AND Abstract: (asymmetr* OR agency) AND AllField: (game*)} "Filter": {article type: research article}	281	All times
Merged dataset		481	

the VR headset. Karaosmanoglu et al. (2021) have shown that low/high agency can be felt keenly by players of an asymmetric VR game (and also that low agency is not inherently something to be avoided, nor is high agency always something to be sought). Because of this, and in the expectation some research in this context might refer to "asymmetry" in roles, mechanics, or interfaces as a difference in agency, we included the term agency as well. As this was added via an "OR" operator the additional search term does not exclude any relevant papers. We applied a filter to all keywords except game* to check for occurrence in the abstract, to exclude papers that do not focus on VR and asymmetric interaction. Thus, we aimed for papers that mention the term "game" at least somewhere in the paper (as a minimum requirement), but that are focused closely enough on VR and asymmetry that they mention both in the abstract.

Table 1 shows the exact queries, and the number of papers identified thereby in the ACM and Scopus databases (prior to duplicate removal).

Beyond requiring the involvement of (any kind of) asymmetry in a multiplayer game with at least one VR player, we did not exclude papers based on study characteristics (e.g., participant groups). We otherwise considered any English-language, peer-reviewed original research article. Merging of the database exports and removal of duplicate[3], non-English, and ineligible-type[4] papers resulted in an initial pool of 399 papers to enter the screening stage.

3.1.2 Initial Screening and Inclusion Criteria

We performed our screening of the 399 records identified after duplicate removal based on their Abstract and Title, using the following inclusion criteria:

[3]Based on title regardless of capitalization and year, as databases were not consistent in their use of DOI, ID, and URL fields.

[4]Document type Conference Review. Not based on DOI because databases are not consistent in their use of DOI, ID, and URL fields.

1. Does it address an HMD-based VR gameful experience?
2. Does it make a statement about any type(s) of asymmetric gameplay in VR games?
3. Does it make a statement about how players experience any type(s) of asymmetric gameplay in VR games?

To be included, papers had to fulfill item 1, as well as either item 2 and/or item 3. Rating based on these inclusion criteria was conducted by three of the authors, with each coder rating two-thirds of the papers (n = 266). Each paper received inclusion/exclusion votes by two authors. In case of disagreement, the third coder was asked to break the tie[5]. As a result, a total of 21 papers were identified in this stage as included for further steps; 378 papers were excluded. Of these, n = 42 papers were excluded because they did not involve a VR game experience; n = 366 papers involved no asymmetric VR game experience. Several excluded papers involved asymmetric collaborative scenarios (VR or otherwise), but were not games (e.g., Piumsomboon et al., 2018). While we note that such papers may certainly be interesting to reflect on collaboration in asymmetric scenarios, for the scope of this paper we wanted to focus on games or gameful experiences in particular, as we expect these kinds of experiences to leverage higher intrinsic motivation among users/players. How the findings in asymmetric games can be generalized to collaborative work scenarios (or vice versa) is a question for future work. Additionally, we note that the ACM database is loose in matching keywords, and so a portion of our excluded papers were ones that did not actually mention asymmetr* or agency, but rather provided results for subsets of the keywords, or did match the keywords but only in metadata outside the abstract field (despite the ABS parameter[6]). Overall, however, the queries may have caught more than strictly relevant, and were subsequently filtered in the screening process via human coders. We considered this a safer approach compared to more restrictive queries that might miss relevant papers, and it is common in systematic literature reviews.

We calculated inter-rater reliability (IRR) between the three coders for the paper screening (included/excluded) via Cohen's Kappa test (Gamer et al., 2012). Results indicated high IRR between the first and second (97.7%; κ = 0.901, z = 10.4, $p < 0.0001$), between second and third (100%; κ = 1, z = 11.5, $p < 0.0001$), and between the first and the third (100%; κ = 1, z = 11.5, $p < 0.0001$) authors.

3.1.3 Snowball Approach

To avoid being too dependent on the specifics of our search terms, and to catch papers that describe asymmetric gameplay while not using that terminology, we applied a backwards snowball approach (Wohlin, 2014) to the 21 initially identified papers. This snowball approach consisted of the following steps: 1) For each initially included paper, we extracted references that appeared relevant based on title. We expanded this list based on reading each of the initially included papers' introduction and related work/background sections. 2) Potential additions were checked for not being duplicates. We then applied our review's inclusion criteria to their abstract, and decided based on consensus from two authors. Resulting from this procedure, six more articles were identified for our review (see **Table 2**), yielding a total of 27 papers for the next steps of the review.

3.1.4 Full-Text Eligibility

We checked the full text of all identified papers for eligibility. During this process, two articles were removed from the pool of included papers entirely[7]. Removing these two papers resulted in the final selection of 25 papers (the final corpus; see **Table 2**).

3.1.5 Quality Assessment

As recommended for literature reviews, we conducted a quality assessment (QA) stage (Aromataris and Munn, 2020). Given the early stage of the research field, however, we consider even informal playtesting to hold potentially relevant information, and so we did not aim to exclude papers with only informal empirical findings based on the QA procedure. Instead, we use this stage to gain insight into the applied methodology and the quality of empirical work in this field. Further, some papers did not contain studies at all, instead consisting of theoretical work or system descriptions. While these are still useful for the framework synthesis stage, there is—to our knowledge—no checklist for assessing quality of that kind of work. The QA scoring was thus only applied to papers with a clearly identifiable quantitative and/or qualitative user study; papers without this were excluded from the QA process (see **Table 2** for an overview).

The first nine papers (36%) were assessed based on their full text by the first two authors together (double data extraction) to ensure consensus in rating, while the remaining papers were split to be assessed separately. The JBI Manual for Evidence Synthesis (Aromataris and Munn, 2020) suggests that two-reviewer quality QA meets best practices for systematic reviews. Further, we conducted double data extraction for 36% of papers, which further reduces error (Buscemi et al., 2006), and ensures reliability of coding for the QA process (McDonald et al., 2019).

In total, 11 papers were found to fulfill the prerequisites for the QA stage. To these, we applied Kmet et al. (2004)'s QA criteria checklist for quantitative (n = 14 checklist items) and qualitative (n = 10 items) studies. For example, these verify that papers describe their research questions and study designs appropriately, but equally focus on reporting of results, methodology, data collection methods, and reliability and validity of the measurements. It is rated on a scale of *Yes* = 2, *Partial* = 1,

[5]Twelve papers were initially flagged as undecided based on only the Abstract and Title. For these papers, we accessed the papers' full text to make an informed decision, yielding their exclusion.

[6]We have been in touch with the ACM digital library team about such irregularities but have so far not received a definitive answer.

[7]Streck et al. (2019b) was identified as the reprinted version of a paper already in our list Streck et al. (2019a) and Woodworth et al. (2017) was identified as not game related. We further note that #24 was almost excluded because although it describes a game, it does not seem to consider any other involved people except the HMD playeras "players." Nevertheless, the non-HMD "actuators" are incentivized to help the HMD player, they have a goal, a physical challenge, and they reportedly enjoyed it; we thus kept this paper included as a gameful multiplayer experience.

TABLE 2 | Quality assessment scores of the identified papers in our literature review. The range of overall quantitative and qualitative scores were 0–1; 1 represents the highest possible point (Kmet et al., 2004). For appropriateness scores (Connolly et al., 2012), 6 constitutes the highest possible score while 2 equates to the lowest score. In the synthesis, we refer to papers by ID.

ID	Papers	Notes	Included in QA		Appropriate-for-review scores
			Quantitative	Qualitative	
#1	Jeong et al. (2020)	Experimental results and analysis	×	—	5
#2	Bortolaso et al. (2019)	Early evaluation	—	×	4
#3	Kerure and Freeman (2018)	Prototype	—	—	5
#4	Lee et al. (2020)	Experimental results and analysis	×	—	5
#5	Li et al. (2017b)	Almost-duplicate of #11	—	—	5
#6	Serubugo et al. (2018b)	Experiment 1	×	×	4
#7	Serubugo et al. (2018a)	Experiment design	×	—	5
#8	Sra et al. (2016)	Preliminary deployment	—	—	5
#9	Gugenheimer et al. (2017a)	User study	×	—	6
#10	Zhou et al. (2019)	Study	×	×	6
#11	Li et al. (2017a)	Almost-duplicate of #5	—	—	5
#12	Sra et al. (2017)	Pilot study	—	—	5
#13	Gugenheimer et al. (2018a)	Demo paper of #9	—	—	4
#14	Streck et al. (2019a)	—	—	—	3
#15	Smilovitch and Lachman (2019)	Early tests	—	—	5
#16	Liszio et al. (2017)	Study on social entities and VR PX	×	—	6
#17	Furukawa et al. (2019)	Preliminary evaluation with authors/exhibition	—	—	4
#18	Gugenheimer et al. (2018b)	Evaluation	×	×	6
#19	Serubugo et al. (2017)	Public library testing	—	—	4
Added via snowball approach:					
#20	Sajjadi et al. (2014)	Evaluation	×	×	6
#21	Gugenheimer et al. (2017b)	Extended abstract to #18	—	—	4
#22	Liszio and Masuch (2016)	Case study	—	—	4
#23	Schmitz et al. (2015)	Exhibition	—	—	5
#24	Cheng et al. (2014)	Lab study	×	—	5
#25	Knierim et al. (2016)	Preliminary study	—	—	5
Mean and standard deviation - *M* (*SD*) - of quality assessment scores			0.42 (0.19)	0.58 (0.24)	4.84 (0.8)
Adjusted quality assessment scores			0.63 (0.21)	—	—

No = 0, *Not applicable (n/a)* = -2. Based on Kmet et al. (2004)'s instructions, the papers' QA scores were calculated by dividing the sum score by the total possible score (quantitative: 28; qualitative: 20). This scoring is punitive to more informal studies (as *n/a* ratings lead to minus points). Additionally, all empirical studies comparing an HMD experience to a non-HMD experience cannot realistically blind participants and investigators to the study conditions[8]. Low scores for quantitative studies are the result. We thus excluded two questions to also calculate an adjusted QA score for the quantitative papers.

Additionally, we used two items based on Connolly et al. (2012)'s checklist to determine the relevance of each paper to our review[9]. We applied this rating based on a scale of *high* = 3, *medium* = 2, or *low* = 1 appropriateness for this review; the overall appropriateness score was calculated by summing up the two item ratings.

[8]Items six and seven in Kmet et al. (2004)'s quantitative checklist.

[9]1) "*How relevant is the particular focus of the study (including conceptual focus, context, sample and measures) [, paper, or system description] for addressing the question or sub-questions of this review?*" 2) "*To what extent can the study findings [or paper contributions] be trusted in answering the study question(s) [or paper objective(s)]?*"

Table 2 illustrates which papers were included in the QA, the mean QA score for empirical studies (qualitative, quantitative, and adjusted quantitative), and each paper's appropriateness scores to be included for review.

3.1.6 Framework Synthesis

For the synthesis stage of our literature review, we drew on framework synthesis as a methodology, which is based on framework analysis (Pope, 2000; Ritchie and Spencer, 2002; Dixon-Woods, 2011), a qualitative method for data analysis. This method consists of five stages (Pope, 2000): 1. familiarization (becoming acquainted with the raw data), 2. identifying a thematic framework (development of our *a priori* framework), 3. indexing (applying the framework to the data), 4. charting (extracting and summarizing data into a framework overview), and 5. mapping and interpretation (synthesizing the themes and findings from the framework overview). When used for a literature review, this is called framework synthesis (Carroll et al., 2011; Dixon-Woods, 2011).

In particular, we used a variant of framework synthesis that employs a "best fit" strategy (Carroll et al., 2011; Dixon-Woods, 2011; Carroll et al., 2013): Researchers use a theoretical framework as a guiding lens through which they explore relevant literature. With the "best fit" strategy, the lens can be an existing framework, or alternatively an extension or combination of (an) existing framework(s). The benefit is that the lens can provide a frame through which the literature review can be explored and synthesized systematically, without first creating a comprehensive framework through time-consuming inductive methods. For this reason, it has found uptake in health science for policy-making (Dixon-Woods, 2011); it is considered a reliable synthesis method and allows "*testing [of] existing potentially generalisable theories and models within a specific context*" (Carroll et al., 2011).

Given the novelty of the field, we consider it a useful approach here as well, because there is no seminal, confirmed model of asymmetric game experiences. Applying framework synthesis enabled us to explore both how the current papers in this emerging field engage with prior theoretical work, and how well suited the *a priori* framework (based on prior theoretical work) is to classify asymmetric VR games. We consider now to be a crucial time to conduct this systematic review, because there are a few candidates for theoretical frameworks to draw from (see next section), and enough relevant papers to review to generate a spectrum of the emerging field in our synthesis. Yet it is also "early" enough in the subfield's emergence for our findings to yield a thorough foundation for future work.

After developing our "best fit" *a priori* framework for asymmetric multiplayer VR games (see next section), we conducted an informal thematic analysis (Braun and Clarke, 2006; Nowell et al., 2017; Braun and Clarke, 2020), extracting text relating to and describing how each column or facet of the framework applied to each paper. We consider our approach to be a hybrid orientation of both codebook and reflexive thematic analysis (Braun and Clarke, 2020): we employed the framework as our codebook, and eschewed inter-rater reliability, but we nevertheless valued consensus between coders as a means to drive discussion surrounding the coding of papers in the framework, and the validity of the framework itself. We followed a consensus coding approach between two coders for nine papers (36%), and then coded the remaining papers separately. Disagreements and arising uncertainties in coding were discussed and resolved (thus making a second inter-rater reliability calculation moot) in recurring meetings throughout the synthesis stage[10].

3.2 Developing the *A Priori* Framework

The "best fit" *a priori* framework (see **Figure 2**) began with the work by Harris et al. (2016). Based on several game design iterations and a user study (Harris et al., 2014; Harris et al., 2016; Harris and Hancock, 2019), their work has resulted in a conceptual design framework for asymmetric games. In turn, their framework builds upon the mechanics, dynamics, and aesthetics (MDA) framework (Hunicke et al., 2004), which describes games through the designed rules and logic (*mechanics*), the resulting gameplay based on players' input (*dynamics*), as well as players' emotional and immersive experience while interacting with the game (*aesthetics*). Harris et al. (2016) provide an overview of dimensions of asymmetry across the MDA categories. In game mechanics, they describe asymmetry of ability, challenge, interface, information, investment and goal/responsibility. For dynamics, they classify the type/direction of interdependence between players (mirrored, unidirectional, or bidirectional dependence) and temporal aspects thereof (asynchronous, sequential, expectant, concurrent, coincident). We included these categories to systematize our coding and synthesis of the described VR games.

A thematic analysis of PX in an asymmetric game reported in the same paper also generated "salient themes most relevant to the design of asymmetric games" (Harris et al., 2016). These themes were posited based on their analysis of asymmetric games and their own design work (creating and evaluating an asymmetric game using mixed methods) — under application of the MDA framework (Hunicke et al., 2004) (including aesthetics) as an analytical lens. We thus use the Harris et al.'s aesthetics of asymmetry in our paper to build on this prior work and its specific application within asymmetric games. These categories of aesthetics of asymmetry consist of: leadership and primacy, effects of player skill, familiarity with interface or partner, interdependence and necessity, and coordination. Because the aesthetics relate to players' actual experience, we applied these categories to reports of empirical user studies or in formal playtesting, as well as speculation about expected PX.

The **MDA framework** on which Harris et al. (2016)'s conceptual work is based further discusses several dimensions of players' aesthetic experience: sensation, fantasy, narrative, challenge, fellowship, discovery, expression, and submission (Hunicke et al., 2004). Because the aesthetics categories described by Harris et al. (2016) are specific to the game stimuli they tested in their study, we included the MDA aesthetics in our *a priori* "best fit" framework, to enable us to capture a more extensive understanding of PX in asymmetric VR games.

Further, we added two more exploratory framework parts, in the spirit of the "best fit" nature of this methodological approach: one designed to explore **social asymmetry** such as differences in shared space, age, and abilities (custom categories) and factors put forth by Kaye (2016) regarding teamwork and communication in games. The dimension of shared space emerged from the initial full-text eligibility check of the selected papers when we noticed that some papers allow for or address remote play, meaning that there can be a difference in location or space (These options of remote vs. co-located were later extended during the review by the observation that differences also occur within co-located setups in how the space and proximity been players are framed and utilised.). We included the categories of age and abilities with the expectation that differences between players could frame players of different ages (e.g., young children and parents or grandparents), or different cognitive or physical abilities (e.g., for rehabilitative training games in clinical settings). Further, we include work by Kaye (2016) which has suggested factors that facilitate an experience of flow within social play in digital games: communication, teamwork, and knowledge of others (e.g., in terms of skills relevant to the task). While not relating specifically to asymmetry (e.g., of communication), we saw this as an

[10]The final framework coding is presented in **Supplementary Materials**.

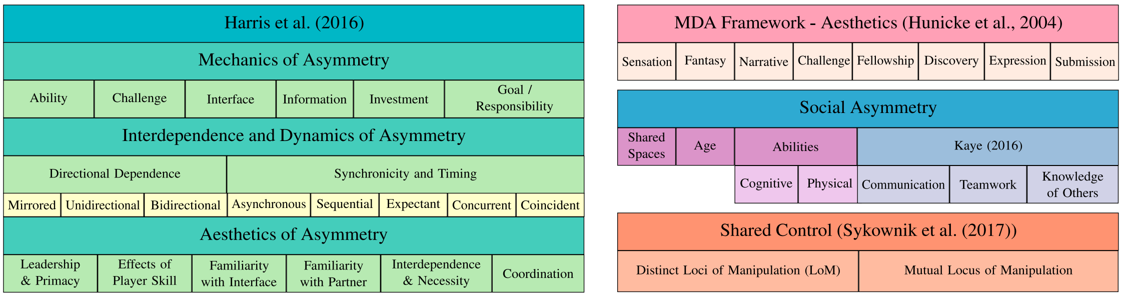

FIGURE 2 | Our *a priori* "best fit" framework for asymmetric VR games.

opportunity to both look out for such factors, and capture relevant factors for social multiplayer gaming experiences in general.

The final aspect of the framework consisted of patterns of **shared control** suggested by Sykownik et al. (2017). This distinguishes between different ways that players have control over points or entities in a digital game, which are termed loci of manipulation. They distinguish between distinct loci of manipulation (players control different points/entities in the game), and mutual locus of manipulation (players share control of the same points/entities in the game). We added this because shared control is described as another type of interdependence between players.

The full *a priori* "best fit" framework is illustrated in **Figure 2**. Its categories make up the variables for which data was collected by looking at each paper's full text.

4 RESULTS

This section presents the synthesis for the 25 papers in our corpus, derived from the PRISMA procedure described above and illustrated in **Figure 1**. We begin with an overview of the empirical papers (a detailed overview of study characteristics is provided in the **Supplementary Materials**), then discuss all papers through the lens of our *a priori* framework. We refer to papers by their index in the corpus (i.e., #*x* for the paper with ID *x*, see **Table 2**). When referring to numbers of papers, we count #11 and #5 as one paper given their very close similarity, as well as #13 and #9 due to the former being a demonstration of the latter. #21is an earlier iteration of #18, yet these were counted separately as they do vary in content.

A summarized overview of our key takeaways is also provided in **Table 3**. Further, based on our synthesis process and findings, we propose changes to our "best fit" framework to develop a *posthoc* framework (addressed in detail within the discussion of this paper). How our findings fit within the *posthoc* framework is shown in **Tables 4, 5**.

4.1 Summary of Empirical Research

A total of 11 papers were identified as containing user studies of some kind (n = 4 mixed-design, n = 6 quantitative, n = 1 qualitative). All featured one or multiple custom game(s); commercial asymmetric VR games were only used in one of the papers (#10): *AudioShield* (Fitterer, 2016) and *Keep Talking Then Nobody Explodes* (Steel Crate Games, 2015), in addition to a custom game.

Across all studies, the corpus reports on the experience of a total of 289 participants [an average of 26.27 (*SD* = 23.35) participants per paper]. Gender was not always clearly reported for all study participants. A few papers reported no gender information at all (#4, #6), or only reported it partially (e.g., reporting only the number of female participants, but not specifying the remainder of participants' gender as in #18 or #24). Additionally, it was sometimes unclear how gender was assessed (e.g., which answer options were presented). We summarized gender based on reported distributions: of n = 221 participants with reported gender, n = 81 were female (37%), while n = 140 male (63%); non-binary participants are not reported. This pattern of unclear reporting repeated for the age of participants; while some papers reported both average age/age range and standard deviation (#9, #16, #18), others only presented the age range or no age at all. Overall, PX was tested with fairly young participants (average = 22.43). The context of recruitment was often also unclear but can be assumed to have been a university/lab setting in many papers. A notable exception are the public museum settings in Serubugo et al.'s work (#6, #7), where detailed demographics are omitted in favor of in-the-wild empirical playtesting. Almost all papers focused on asymmetry of interface as the main independent variable in their study: namely an HMD player, a (or multiple) non-HMD player(s), and in some cases additional spectators. A few studies instead explored varying viewpoints of the non-HMD player (#1, #4) as their independent variable: first-person point, third-person point of view, or seeing both (#4)/switching between the two (#1). One paper used the type of social entities in VR (human vs. agent) and their interaction with the HMD player (interactive vs. non-interactive) as independent variables (#16).

For quantitative studies, we observed similarities in how studies assessed PX of participants; the majority of these papers utilized questionnaires. For general PX, many employed the Game Experience Questionnaire (GEQ; IJsselsteijn et al., 2013; core and post-game modules, or specifically the positive experience modules therein); a few only used custom items (#24, #6). Emotions were assessed via the self-assessment manikin (SAM; Bradley and Lang, 1994) (#9, #10, #18). Approximately half of the empirical quantitative

TABLE 3 | Summarized overview of takeaways and research gaps from the review synthesis, as well as meta takeaways that led to the adjustment of the *a priori* framework to develop the *posthoc* framework.

Takeaways & Research Gaps	Meta Takeaways
Harris et al. (2016)	
Mechanics: Asymmetry of interface, ability/challenge, information, and responsibility most common. More than two players and alternative interfaces underexplored.	Ability & challenge are so interconnected that they can be merged. Goal & responsibility should be addressed separately: same goal but different responsibilities is common
Interdependence/dynamics: Bidirectional interdependence most common.	Synchronicity & timing don't make sense for games as a whole (although they may well be applicable for examining separate game actions or segments)
(Designed) Aesthetics: Effects of player skill, familiarity with interface and partner underexplored. Leadership & primacy, interdependence & necessity, and coordination most commonly addressed aspects of designed PX. Communication largely verbal - other alternatives (gestures, physical social contact) rare but promising. Trust and power level / dominance as important factors; mixed views on targeting asymmetry vs. balance.	
Reporting of asymmetric game design often unclear	Some categories may only apply to collaborative games (e.g., goal, interdependence)
MDA framework - Aesthetics (Hunicke et al., 2004)	
Key drivers of enjoyment: explicitly shared physical space, embodied physical interaction, and human co-player (see social asymmetry). Immersion and role-playing (fantasy category) barely assessed at all. PX aesthetics surprisingly often not reported, even engagement/enjoyment/fun.	Mostly just sensation and fellowship categories applied. Should replace/extend with more nuanced alternatives for sensation (PX) and fellowship (social PX), respectively.
Social Asymmetry	
Shared Spaces: Most games co-located, with explicitly closely-shared physical space as a rarer subtype and promising research direction. Remote scenarios underexplored; shared space scenario often not clearly reported.	Shared spaces as emerging distinction.
Asymmetry of Age & Abilities: No papers addressed addressed asymmetries in age or abilities but could be interesting for accessibility purposes.	
Kaye's flow experience model in cooperative gaming: Communication, teamwork, and knowledge of others all rarely addressed.	Despite rare occurrence in the corpus, these aspects could significantly enhance social experiences of players. Alternative models of players' social experience could be explored in future iterations of the framework.
More rigour needed in describing participant samples: gender and age often only roughly described. (Especially considering potential correlations of sex on simulator sickness in VR.)	
Shared Control (Sykownik et al., 2017)	
Most commonly (essentially always) distinct loci of manipulation. Entity vs. environment as common subtype of this asymmetry; no loci as another.	Difficult to code when one player lacking a distinct in-game entity/avatar. Subcategories for distinct loci emerged: all players control game world elements directly vs. some have only indirect control (e.g., no loci, verbal assistance only); all players control entities vs. some players control the environment.

papers measured presence. This was assessed with one of three different questionnaires: Witmer et al. (2005)'s presence questionnaire (#1, #4); the Igroup Presence Questionnaire (IPQ; Schubert et al., 2001) (#16); and Slater-Usoh-Steed (SUS; Slater et al., 1994) (#9, #18) questionnaires. #16 employed the immersion subscale of the GEQ, yet otherwise none used dedicated immersion questionnaires. Only one paper used a measure for simulator sickness: the Simulator Sickness

TABLE 4 | An overview of how the final corpus of surveyed papers fits into the (Harris et al. part of the) *posthoc* framework. Emptier columns indicate existing research gaps in the field.

		Mechanics						Interdependence			Aesthetics					
ID	Papers	Ability/Challenge	Interface	Information	Investment	Goal	Responsibility	Mirrored	Unidirectional	Bidirectional	Leadership & Primacy	Effects of Player Skill	Familiarity with Interface	Familiarity with Partner	Interdependence & Necessity	Coordination
#1	Jeong et al. (2020)	●	●	●		●	●	●		●	●				●	●
#2	Bortolaso et al. (2019)	●	●							●	●				●	●
#3	Kerure and Freeman (2018)	●	●	●		●					●					
#4	Lee et al. (2020)	●	●	●			●			●	●					●
#5	Li et al. (2017b)	●	●	●		●	●		●							
#6	Serubugo et al. (2018b)	●	●	●			●			●	●				●	●
#7	Serubugo et al. (2018a)	●	●	●			●			●	●		●		●	●
#8	Sra et al. (2016)	●	●	●		●										
#9	Gugenheimer et al. (2017a)	●	●	●			●			●	●				●	●
#10	Zhou et al. (2019)	●	●	●			●			●	●			●	●	●
#11	Li et al. (2017a)	●	●	●		●	●		●							
#12	Sra et al. (2017)	●	●	●			●			●	●				●	●
#13	Gugenheimer et al. (2018a)	●	●	●			●			●	●				●	●
#14	Streck et al. (2019a)	●	●	●					●		●				●	
#15	Smilovitch and Lachman (2019)	●	●	●			●	●							●	●
#16	Liszio et al. (2017)	●	●				●			●				●		●
#17	Furukawa et al. (2019)	●	●				●			●						
#18	Gugenheimer et al. (2018b)	●	●	●			●			●	●	●		●	●	
#19	Serubugo et al. (2017)	●	●			●	●		●						●	●
#20	Sajjadi et al. (2014)	●	●	●			●			●	●	●	●		●	●
#21	Gugenheimer et al. (2017b)	●	●				●	●			●					
#22	Liszio and Masuch (2016)	●	●	●			●			●					●	
#23	Schmitz et al. (2015)	●	●	●			●			●					●	●
#24	Cheng et al. (2014)	●	●				●		●			●		●		●
#25	Knierim et al. (2016)		●	●												

Questionnaire (SSQ; Kennedy et al., 1993) (#16). The same paper (#16) was also the only one to measure individual differences in participants' inherent traits, such as their immersive tendencies (Immersive Tendencies Questionnaire, ITQ; Witmer and Singer, 1998), and their attitude toward loneliness (Preference for Solitude Scale; Nestler et al., 2011).

Surprisingly few papers used questionnaires to measure social factors of PX. While a variety of social experience questionnaires were in use, the majority of these were employed in a single paper (#16). The questionnaires used for this aspect of PX were: the GEQ's social presence submodule (#1, #4, #18; and the submodule's behavioral involvement subscale in #9), the Social Presence in Gaming Questionnaire (SPGQ; De Kort et al., 2007) (#16, #20) and the Cooperative Social Presence scale (CSP; Hudson and Cairns, 2014) (#16), the relatedness score within the Player Experience of Needs Satisfaction (PENS; Rigby and Ryan, 2007) questionnaire (#16), an adapted version of UCLA Loneliness Scale (Russell et al., 1978) (#16), and the Inclusion of Other in the Self Scale (Aron et al., 1992) (#10).

Qualitative studies largely employed interviews and video analysis to assess PX, although one paper used cooperative performance metrics (Seif El-Nasr et al., 2010) to analyze their observations (#20). Two papers employed some form of thematic analysis for the interview and video analysis (#10, #18). The qualitative papers largely did not clarify the type of thematic analysis used (e.g., codebook vs. reflexive; Braun and Clarke, 2020) nor did they provide a detailed description of the process they followed.

Overall, much of the empirical research reported in the reviewed papers was of an informal nature; only few consisted of formal experiments. Fourteen papers contained either user testing too informal to be included in the QA process, or technical descriptions of their system/implementation. This highlights the potential for more comprehensive empirical work in this subfield,

TABLE 5 | An overview of how the final corpus of surveyed papers fits into the (MDA, social asymmetry and shared control parts of the) *posthoc* framework. Emptier columns indicate existing research gaps in the field. Codings are subjective and relate to the paper as a whole (i.e., covering multiple applications).

		MDA								Social Asymmetry									Shared Control			
										Shared Spaces									Distinct			
ID	Papers	Sensation	Fantasy	Narrative	Challenge	Fellowship	Discovery	Expression	Submission	Non-Colocated	Co-located	Explicitly Shared	Age	Cognitive Abilities	Physical Abilities	Communication	Teamwork	Knowledge of Others	Entities	Environment	Mutual Locus	Indirect/None
#1	Jeong et al. (2020)	●			●	●					●					●			●			●
#2	Bortolaso et al. (2019)					●					●					●			●	●		
#3	Kerure and Freeman (2018)	●									●								●	●		
#4	Lee et al. (2020)	●				●					●								●			●
#5	Li et al. (2017b)									●	●								●	●		
#6	Serubugo et al. (2018b)					●					●								●			●
#7	Serubugo et al. (2018a)	●				●					●								●			●
#8	Sra et al. (2016)	●								●	●								●			●
#9	Gugenheimer et al. (2017a)	●				●					●	●							●	●		
#10	Zhou et al. (2019)	●				●					●	●					●		●			●
#11	Li et al. (2017a)									●	●								●	●		
#12	Sra et al. (2017)					●					●					●			●			●
#13	Gugenheimer et al. (2018a)	●				●					●	●							●			
#14	Streck et al. (2019a)										●								●	●		
#15	Smilovitch and Lachman (2019)	●									●								●	●		
#16	Liszio et al. (2017)	●	●		●	●					●								●			
#17	Furukawa et al. (2019)										●								●	●		
#18	Gugenheimer et al. (2018b)	●				●					●	●							●	●		
#19	Serubugo et al. (2017)	●				●					●								●			●
#20	Sajjadi et al. (2014)	●			●	●					●					●	●		●			●
#21	Gugenheimer et al. (2017b)										●	●							●	●		
#22	Liszio and Masuch (2016)	●				●					●								●	●		
#23	Schmitz et al. (2015)	●								●	●								●			●
#24	Cheng et al. (2014)	●			●	●					●	●							●	●		
#25	Knierim et al. (2016)	●				●					●									●		

and we look forward to future work that builds upon these early explorations.

4.2 Mechanics, Dynamics, and Aesthetics—Harris et al. (2016)

4.2.1 Mechanics of Asymmetry in Game Design

All papers explored an asymmetry of more than one mechanic specified by Harris et al. (2016); most commonly, this consisted of asymmetry of the interface (all papers), as well as ability/challenge, information, and responsibility. In terms of **interface**, the majority of papers (19 total) described games designed for two players, one of whom wore an HMD. The non-HMD players then usually interacted with the game via an interactive display medium (PC, smartphone, tablet device, projection, or Sifteo Cubes). In several cases, non-HMD player involvement was markedly subtle: they only viewed the virtual world (mostly via a monitor) without interaction, assisting the HMD player through verbal communication. Two papers employed additional props or devices: one game featured an upright visual screen projection plus a tangible robot avatar of the HMD player (#17). Both game variants in #9 offered the non-HMD playera monitor-attached hand-held controller, once with a lightsaber prop. In one game variant in #8, the non-HMD player(s) had no visual cues at all, and instead reacted only to auditory cues or based on the HMD player's movement in the physical world (In their other game variant, all players had an HMD, but one of the player roles additionally held a nerf gun controller.). Further, in one notable example across two papers, the non-HMD player still interacts with the HMD while it is worn by the HMD player, by pressing on attached touchscreens (#18, #21).

Four papers explored game designs with more than two players. The breadth of variation here shows a wide array of options to explore: two HMD players while the non-HMD player views a projector; one HMD player while two non-HMD player use tablet devices; one HMD player, one co-located player in an AR environment, and a remote player with a tablet device. The gameful experience in #24 employed people "actuators"; they were instructed on how to carry and move the HMD player via smartphones attached to the HMD player.

Almost all papers featured games with differing **ability** assigned to specific players, making it a common type of asymmetry. For example, the HMD player is often tasked with navigating and interacting with the virtual environment directly, while the non-HMD player may be presented with a top-down view and perform a guiding or assisting task. In one artifact these roles were reversed: #23's games put the HMD player in the guiding/assisting role, while the non-HMD players interacted with the game world (CAVE system) more directly. It was difficult to distinguish ability from **challenge**, as abilities are generally directly bound to specific challenges. We included this in our framework only when an ability was explicitly discussed in terms of challenge in the paper. Only two papers addressed challenge in their game design at all (#16, #24); none discussed challenge in terms of a difference between players.

In many papers, player abilities also involved the asymmetric distribution of **information**. It was not always clear which player had what kind of information. Commonly, however, the non-HMD player would have information that the HMD player relied upon to either complete their goal at all or to complete it faster. Generally, this was designed to give the non-HMD player something to do or to encourage communication. Only #15 featured an informational imbalance in both directions: "*To ensure interdependence, players receive the instructions for the other player*" (Smilovitch and Lachman, 2019).

Asymmetry was never discussed in terms of time **investment** required. However, in discussing the papers, we noted that this category could have been useful to clarify and classify degrees of involvement. For example, several papers featured game designs in which the HMD player is not entirely reliant on non-HMD player but can simply achieve their goal faster or easier with minor involvement or assistance from the non-HMD player.

In discussing the papers in terms of **goal/responsibility**, we found that the category may need to be refined in two ways. First, for the majority of collaborative games, players commonly had the same goal, but different responsibilities. For games with lower degrees of investment on behalf of the non-HMD player, the goal was difficult to define: is their goal merely to support the HMD player (different goals and different responsibilities), or is their shared goal to win the game (same goal and different responsibilities)? The category may therefore need to be split into two separate ones (goal vs. responsibility). Second, the category is difficult to apply to competitive games, as here players may be considered to have the same goal (i.e., winning) but in opposition to each other.

4.2.2 Dynamics of Asymmetry in Game Design

We report the dynamics designed or speculated to arise from asymmetry for **directional dependence** categories (mirrored, unidirectional, and bidirectional dependence), and **synchronicity and timing** between player actions (Harris et al., 2016). For **mirrored** dependence, the nature of each player's reliance on the other(s) is the same (Harris et al., 2016). While the nature of reliance was sometimes difficult to discern, we observed this kind of dependence in only three papers (#1, #15, #21). For example, both non-HMDand HMD player need to work together by performing slicing actions to cut fruits to win (Gugenheimer et al., 2017b)'s collaborative game variant (#21). For #1, only the first-person view point variant of the game provides mirrored dependence. #15 constitutes a stronger mirrored interdependence: here instructions for each player are conveyed to the co-player; they rely on each other to pass on the information. For **unidirectional dependence**, we identified four papers (#5, #14, #19, #24). In general, these contained games that could be completed without the help of other players. #19 featured a maze game: the HMD player completed it faster with the non-HMD player's guidance, however, the outcome of the game was achievable without this help.

The majority of identified papers (n = 15) featured **bi-directional dependence**; players relied on (the) other player(s) but in different ways. The degree of this type of dependence between players varied, placing one player in a more active role or less reliant role. However, we note that in many cases it was not clear what the non-HMD player's goal was, making it difficult to determine the nature of reliance between the players: for example, in a maze game navigated by the HMD player, but guided by the non-HMD player, does the non-HMD player rely on the HMD player to physically navigate the virtual world in their stead? This was mainly due to the role definition of the non-HMD player; we often had difficulty determining whether this player was a facilitator or an essential element of the game.

Four papers (#3, #4, #9, #25) featured competitive games, which proved difficult to code as a category of directional dependence, as here players generally do not rely on each other as in the collaborative games—except in the sense that a positive gameplay experience requires a certain degree of motivation and effort from the other players.

Synchronicity and timing was also difficult to apply, as we were discussing and coding entire games, as opposed to specific game mechanics or segments. We observed multiple categories of synchronicity and timing for different player actions in the games. Further, these player actions were often not clearly described in terms of temporal occurrence or dependence, exacerbating the coding process. For instance, in one game (#6), one player had the role of navigator while the other held the role of the explorer (Serubugo et al., 2018b). Yet the game's description did not clarify whether players performed their own respective actions regardless of the others' (**asynchronous timing**) or whether the explorer requires information from the navigator to start to perform their action (**sequential timing**).

4.2.3 Emergent/Expected Aesthetics of Asymmetry in Game Design

Here we coded empirically reported PX (or speculations thereon), based on the aesthetics components proposed by Harris et al. (2016) (based on analytical work employing, but not identical to MDA aesthetics). Leadership and primacy, interdependence and necessity, and coordination were addressed most commonly by the papers. For **leadership and primacy**, this was often only a very implicit coding of the theme (e.g., based on a player role being described as "assistant" to another), but nevertheless some papers point toward how this kind of asymmetry was induced, e.g., through informational asymmetry (#2). A few papers explicitly highlighted trust as an important factor in asymmetric VR gameplay (#12, #21). Gugenheimer et al.'s works (#9, #18) especially touch on power and skill imbalances, as does Kerure and Freeman (2018) (#3) and Zhou et al. (2019) (#10). Gugenheimer et al. report that the difference in "*power level*" (e.g., through asymmetry in information, ability, and interface) drives enjoyment and need not necessarily be equally balanced, as players restrained themselves (although they also knew they would switch roles). Generally their players were aware of the non-HMD player's "*more dominant role,*" which therefore required trust and a more "*more responsible*" playing style, given the potential for abuse of this power. Conversely, when an early iteration of #3's game found an imbalance in favor of the non-HMD player due to an easier game mechanic (voice commands as input); they adjusted this to increase difficulty (voice commands based on pitch as input) and reduce this asymmetry. Zhou et al. (2019) also highlight a balance in responsibilities as a positive factor increasing enjoyment, with an appealing "*give and take' interaction style.*"

The categories **interdependence and necessity** and **coordination** were closely related: many games fostered **interdependence** through game mechanics that required communication and coordination (e.g., one player waits for instructions from another to continue). #22 emphasized that mutual dependencies "*force players either to collaborate or compete*" and result in complex and unique PX. They also speculate player choice over their role increases involvement. **Coordination** was often only verbal in the reviewed games, or not explicitly described. #15 is a notable example for interdependence explicitly enforced through a communication constraint: at some point during the game, audio chat functionality is removed, and players have to develop their own communication system through gestures to succeed (Smilovitch and Lachman, 2019). #10is another interesting example: here, interdependence had a strong physical nature; "*physical social contact*" became a key coordination strategy to win the game (Zhou et al., 2019).

The other aspects in the framework were rarely addressed. #18 and #20 mention **effects of player skill** as contributing to blame attribution and frustration, respectively. The latter also mention this as a potential adaptation or customization factor in #20. **Familiarity with the interface** was barely discussed, even though most studies stem from a time when many participants likely were experiencing VR via HMDs for the first time. Similarly, although the most common motivation for asymmetric VR was including non-HMD players in the HMD player's social experience, only few papers discussed **familiarity with the partner** as a factor in PX. #16 enforced gameplay between strangers by placing one of the examiners as the co-player, but did not report how this affected PX (Liszio et al., 2017). #24's participants expressed greater acceptance to play with friends or family as opposed to the public. #18 report that couples displayed higher degrees of intimacy in interaction; it also affected players' acceptance of physical contact (Gugenheimer et al., 2018b). Zhou et al. (2019) discuss this aspect in the most detail, describing the social connection between players as a strong factor in high social and affective experience ratings. It was reported to enhance relationships and support communication and synchronization.

4.3 MDA Aesthetics in Gameplay

We applied the MDA framework of aesthetics (Hunicke et al., 2004) to obtain more comprehensive information about PX in asymmetric VR, therefore coding study results for PX. **Sensation** and **fellowship** were addressed by the majority of papers, but PX relating to **narrative**, **discovery**, **expression**, and **submission** was not reported by any papers.

The **sensation** aspect of aesthetics was strongly present in the data, consisting of various PX factors (e.g., enjoyment, positive/negative affect, and engagement). Surprisingly, several papers did not collect any measure of fun or enjoyment (#12, #14, #17, #2, #21), but most papers did report some measure of multiplayer engagement for their participants. Several papers discuss which elements of the game provided or induced these positive experiences. A few of them linked players' sense of game enjoyment to player abilities featured by game roles (#1, #4). For example, comparable levels of presence and social interaction were attributed to different view capabilities (#1, #4). The specifics of game design likely strongly influence PX results; for example, Gugenheimer et al. (2017a) reported higher valence scores for their *ShareVR* (#9) system, which fosters explicitly shared physical space with a projection space and controller-mounted props compared to a gamepad-and-TV condition. Similarly, Zhou et al. (2019) (#10) attributed the high enjoyment for their HMD player and non-HMD player to the richness of the highly embodied, physical interaction. Interestingly, they point out some enjoyment was derived from the peripheral interaction with spectators, who also reported high enjoyment and engagement ("*It was really funny [. . .] I was screaming for you to get the one on the right*"; Zhou et al. (2019)). Further, Liszio et al. (2017) indicated PX improved due to playing with a human co-player instead of an agent/bot co-player (#16).

Unsurprisingly, **fellowship** was a commonly coded aesthetic of asymmetric game play; here we coded aspects relating to social interaction, relatedness, or perceived loneliness. In terms of questionnaires, interviews, and behavioral observations, the papers here indicate that asymmetry successfully involved non-HMD players and created interesting and positive social experiences between the players (e.g., "*by providing various roles to the non-HMD users and accordingly expanding their interactions, the social relationship with the HMD users increased*"; Jeong et al., 2020). Cheng et al. (2014) suggest that existing social bonds could be leveraged to improve player satisfaction: playing with family and friends as opposed to strangers (#24). Yet many of the reviewed papers do not discuss whether or how well participants knew each other.

In #9, this positive social experience was shown to rate "*significantly more socially engaging*" (Gugenheimer et al., 2017a) compared to a less physically enmeshed asymmetric setup. A few papers report that the social interaction reported between players was enabled through physical social proximity (#10, #9, #13, #18, #21). In particular, Zhou et al. (2019) highlight social touch, increasing interpersonal relationships, and enabled social affordances between the HMD player, the non-HMD player, and spectators as key contributions to PX (#10). Similarly, Gugenheimer et al. (2018b) report emergent positive social interaction between both players via enabling close proximity (#18) and players' self-regulation of imbalances in perceived "*power level.*" Their collaborative game led to higher empathy and less negative feelings compared to their competitive one; perhaps due to a reliance on the higher-power player (non-HMD player) to self-regulate and avoid abusing their power, which could negatively affect PX. The paper comparing an asymmetric VR game with a human co-player to a version with a virtual agent and a single-player variant (#16) is particularly interesting for the fellowship lens. Here, players with an interactive human co-player reported significantly less loneliness than players whose human co-player did not interact with them, or than single-players.

As **fantasy**, we coded aspects like enjoyment of role-playing and immersion. Surprisingly, this was only reported once; Liszio et al. (2017) (#16) reported higher immersion for a single player game compared to game play with an agent and interactive co-player. **Challenge** (as aesthetics/during gameplay, as opposed to the games' mechanics/dynamics) was addressed by four papers (#1, #16, #20, #25). For instance, two papers described both the HMD player and non-HMD player reporting similar degrees of challenge as a design goal; implicitly (#1) or explicitly (#20).

4.4 Social Asymmetry

The category of **shared spaces** emerged as a useful distinction. The large majority of games were designed as co-located, meaning players are located in the same room (A few do not address this directly, but can be inferred to be co-located from context or images.). Only very few address a remote game design (i.e., asymmetry of space): #5 has two co-located players, but the tablet player is remote. The games in #23 are not explicitly remote, but also not quite co-located; players are within hearing distance but "*separated by distance and projection curtains*" (Schmitz et al., 2015). While #8's evaluation is co-located, it is described as capable of remote play—likely, most of the games are capable of or easily adaptable to remote play.

A further interesting distinction within co-located games emerged: some games encourage or enforce *closely* co-located games, inviting direct physical contact between players (#18, #21, some game variants in #9), #24, and #10). Herein, players were in close proximity to each; they touched the other player's HMD, "hit" them with an inflatable prop, physically carried or held another player's body, or danced together. This gameplay almost makes conventional co-located settings appear to demonstrate asymmetry of space, and resulted in engaging and dynamic social experiences. We term this co-located and explicitly shared physical space.

Differences in **age** or **abilities** (cognitive/physical) were not explored by any of the papers in this review. Nevertheless, we note that this is a gap within the field worth exploring, as we address in more detail in the discussion.

4.5 Shared Control

This part of the framework was based on Sykownik et al. (2017)'s patterns of shared control: **distinct loci of manipulation** vs. **mutual locus of manipulation**. Discerning the VR games's locus of manipulation was often difficult when no game entity is controlled, complicating coding in more specific sub-categories.

Distinct loci of manipulation—both players control different entities in the game—was the most commonly observed category of shared control between players. The games generally featured a specific game element or entity that the players were responsible to control; for example, in #9's *Be My Light*, the HMD player controls a sword to damage monsters, while the non-HMD player controls a light to locate them. We note that Sykownik et al. (2017) also mention a sub-type of distinct shared control wherein the loci of manipulation "*establish a coherent entity*" (e.g., one player controls a game entity's legs for locomotion, while the other controls the same entity's hands during object manipulation). This did not occur in the reviewed games.

Instead, we noted other patterns of distinct shared control not described by Sykownik et al. (2017): we noted several cases in which one player controls a game entity while the other player controls the game *environment*. We consider this a subtype of distinct loci of manipulation, but one that may hold greater asymmetry than when players control distinct game entities. For example, in Smilovitch and Lachman (2019)'s *BirdQuestVR*, the HMD player physically interacts as an entity in the spaceship environment, while one of the non-HMD players only controls parts of the ship's systems through an interface on their tablet device, thereby affecting the overall game environment without controlling an in-game entity. Interestingly, there was only a single paper that did not feature in-game entity control for any players (#25).

In a new pattern of shared control that emerged in our review, only one of the players has a distinct loci of manipulation. While the other player may have additional information that the first does not, they do not directly control any game entity: they have no loci of manipulation. Instead, they generally provide only verbal guidance to the player capable of affecting the game world. In most papers that displayed this pattern of *indirect control/none*, the non-HMD player engaged in verbal communication to guide the HMD player (e.g., #6, #7, #19). Twice this was reversed: Schmitz et al. (2015)'s (#23) *Coral Rift* featured a non-HMD player within a CAVE system and an HMD player (VR). The HMD player observes the sea and warns the non-HMD players about obstacles, but only the non-HMD player actively balanced and navigated the ship. Similarly, Sajjadi et al. (2014)'s HMD player provided guidance to the non-HMD player who manipulated the actual game pieces. Alternatively, sometimes the non-HMD player's control within the virtual world was indirect instead of none: for Sra et al. (2017), the non-HMD player could trigger a galvanic vestibular response to impact the HMD player's navigation.

Mutual locus of manipulation (through control alternation or an input processing function) did not occur in the reviewed games. We note that it could be argued that a game like's *FaceDisplay* might represent this pattern. In this game's cooperative variant, both players use HMD-attached touch screens (front vs. side) to "slice" fruit projectiles. If one considers the locus of manipulation to be the virtual screen on which players slice to interact with the world, then this could be considered a mutual locus of manipulation pattern. However we argue that the players have separate loci of control wherever they slice on the virtual screen, thus creating distinct loci (even though they can at times overlap).

5 DISCUSSION

We separate our discussion of the synthesis into implications of the systematic review for the field of asymmetric multiplayer VR games, and reflections on the framework and how it can benefit the VR research community. Takeaways from both aspects of the discussion are summarized in our overview in **Table 3**.

5.1 Implications and Opportunities for Asymmetric VR Games

The review showcases a vibrant emerging research field, yet also traces several gaps in the literature that represent opportunities for future research.

5.1.1 Commonalities

The classification based on Harris et al. (2016)'s mechanics and dynamics dimensions indicates that most of the existing asymmetric VR games consist of two players, one of which is the HMD player, and provide information to an non-HMD player through a secondary screen of some sort. In addition to asymmetry of interface, asymmetries of ability/challenge, information, and responsibility are most common. Player roles are frequently designed to contain bidirectional interdependence and generally expected to affect leadership, interdependence and coordination between players (see **Table 4**).

5.1.2 Research Gaps

Alternative interfaces [e.g., physiological input (Karaosmanoglu et al., 2021)] are rare, as are games that accommodate more than two players (although the few artifacts in this space are promising). Gaps in the literature are visible in terms of effects of player skill, familiarity with the interface, and familiarity with the co-player (s) (see **Table 4**). A lot of valuable research in games and HCI could inform these factors in design considerations, expectations for PX, and study design implications, for example, attribution theory ("*how people assign causes to effects*"; Depping and Mandryk, 2017), novelty effects (Wells et al., 2010), and pre-existing relationships in social gameplay (Eklund, 2015; Perry et al., 2018). Differences in age and abilities also represent a large gap in the literature, even though asymmetries of age are of course already being explored in non-VR games: there is a long history of inter-generational gameplay (Othlinghaus et al., 2011; Voida and Greenberg, 2012; Osmanovic and Pecchioni, 2016). Increasingly, asymmetries in terms of cognitive or physical ability are also being explored in games research (Cairns et al., 2019). For example, Gerling et al. (2014) investigated the use of game balancing approaches to accommodate players with different physical abilities (e.g., players with or without a wheelchair), to avoid reducing self-esteem and relatedness. A more recent example by Graf et al. (2019) featured a projection-based AR exergame, which enabled children with and without a wheelchair to play together. We emphasize that the context of VR may hold additional challenges for asymmetries in player age or ability (e.g., to design suitable inter-dependencies between players with dementia and their relatives), or make certain constellations entirely inappropriate (e.g., the HMD may constitute a safety hazard for people with high injury risk). However, perhaps the immersive and socially engaging qualities of asymmetric VR can be transferred to games with such asymmetries while still addressing these challenges.

We also point out that the focus on immersive and social qualities of asymmetric VR that is prevalent in the corpus (and thus also this review) may in itself represent a research gap. Player behavior and performance were rarely explored in detail in the papers we surveyed. Performance is likely disregarded in favor of focusing on aspects more closely aligned with common goals of asymmetric mechanics: increased communication/coordination and social connection. However, at the line of distinction between competitive and collaborative games, performance may become a more relevant factor. Player behavior was largely discussed within the context of socially motivated behavior in our corpus, yet other aspects of player behavior may also be interesting (and also showcase differences between competitive and collaborative game designs). For example, the use of F formations (Ciolek, 1983; Kendon, 2010) may serve to inform explorations of player behavior in terms of orientation [see Marquardt et al. (2012) for an exploration of how proxemics apply to people and devices in cross-device contexts].

Furthermore, we note that physicality and explicitly shared physical spaces are a particularly interesting dynamic in asymmetric VR games. This was particularly noticeable in *FaceDisplay* (#21; Gugenheimer et al., 2017b) and *AstaireVR* (#10; Zhou et al., 2019). We assume that this is a valuable factor for future research; not only do players have a need for social interaction and relatedness (Ryan et al., 2006), there is also strong evidence for an appreciation of or need for embodied interaction and tangible as well as kinaesthetic experiences (Hall, 1966; Dourish, 2004; Hornecker, 2005; Kim and Schneider, 2020). Prior work has discussed phenomena such as gestural excess (Simon, 2009; Harper and Mentis, 2013) and embarrassing or uncomfortable interactions (Benford et al., 2012; Deterding et al., 2015) as a valuable tool and affordance within (body-focused) game design. This is likely also part of the appeal of asymmetric VR games, yet VR users are also inherently (technologically and immersively) isolated from the real world to a degree (Boland and McGill, 2015; Mütterlein and Hess, 2017). Benford et al. (2012) in particular described isolation and giving up control to other people as examples for uncomfortable interactions in HCI—without negative connotation, necessarily. We argue that their research is (implicitly) built upon in asymmetric VR games research. While discomfort is often a negative aspect of user experience, it can be less so in gameful contexts (Bopp et al., 2016). Yet we also point out that Benford et al. (2012) have discussed this kind of design in terms of ethical considerations, suggesting that issues of justification, informed consent, risk management, as well as rights to withdrawal, privacy, and anonymity require special focus in this context. Given the rather tangential and cursory engagement with prior theoretical work displayed by many of the reviewed papers (which we address in more detail below), we find it particularly important to highlight this connection.

5.1.3 Methodology

Conducting this review revealed that a lot of the games and systems were only roughly described (e.g., missing information, information scattered throughout the sections). We suggest that our post-hoc framework, and in particular the dimensions based on Harris et al. (2016)'s mechanics and dynamics, can scaffold game design reporting in a way that would make it easier to understand the design and resulting PX. For example, explicitly addressing players' different goals would make classification and comparison of games a lot easier. It would also increase our understanding of the results in terms of potential effects of the framing with which each role was presented (Further, it may also be worth assessing how players *actually* understand their role, as this could differ.). This could also prove useful partially for the reporting of non-game mixed-reality applications.

With regards to measurement, our review shows that there were surprisingly few quantitative psychometrics for enjoyment: this consisted largely of the GEQ (or a submodule thereof) and SAM, as well as some interview questions. The GEQ's factor structure has been called into question in recent years (Johnson et al., 2018; Law et al., 2018), making this a potentially problematic choice depending on the submodule used (we suggest using alternative questionnaires such as the Player Experience Inventory by Abeele et al., 2020 or The Player Experience of Need Satisfaction by Ryan et al., 2006). We further propose that a wider range of participants is worth exploring: in age and abilities but also in terms of pre-existing social relationships and gender (Burtscher and Spiel, 2020). Especially when considering interdependency relating to power imbalances, isolation, and embarrassing interactions, a broader range of demographics would be highly beneficial. It is also surprising how few papers reported measures for players' perception of their social relationship with each other.

Existing questionnaires may not be well suited to investigate many aspects of PX that are prevalent in asymmetric VR games (e.g., experienced leadership). In light of this, we find it surprising that only five papers in our review employed qualitative methods, which may be more flexible in this regard. When qualititative methods were applied, their reporting could also be improved (e.g., what specific *kind* of thematic analysis was used; Braun and Clarke, 2020) and would increase methodological rigor in this research field.

5.1.4 Connection to Theory

Our review indicates that a lot of research on asymmetric VR games does not deeply engage with existing theoretical work on asymmetry in games. For example, although several papers cite Harris et al. (2016)'s work, very few actively incorporate aspects thereof in their design. Reeves et al. (2005)'s taxonomy for the design of spectator experiences of public interfaces is similarly relevant but was addressed by only a single paper. The taxonomy classifies whether users' manipulations and/or the effects of their manipulation are visible to or hidden from spectators (ranging from "*secretive*" to "*expressive*"; Reeves et al., 2005). Other recent work on users' experience of performing interaction in front of other people (Martínez-Ruiz et al., 2019) and a general spectrum of the degree of interactivity provided to an audience (Striner et al., 2019) may be useful for future research in this area as well. This tenuous theoretical foundation is likely a side effect of the laudably innovative nature of the designed VR artifacts and the youth of the field. Yet a closer connection to the theory in terms of asymmetric game design and social factors in digital games—especially at this early stage—would allow this field to grow more systematically and gain better coverage across relevant dimensions in game design and aesthetic PX.

5.2 Reflections on the Framework

Overall, the "best fit" *a priori* framework was an extensively useful guiding lens for our synthesis of asymmetric VR games, giving us both a high-level overview of the field's thematic focus points, and dimensions along which to more deeply consider results and implementations explored so far. However it also showcased ways in which the *a priori* framework can be improved: we detail these aspects here, and provide the suggested *posthoc* framework in **Figure 3**.

Dimensions of **Harris et al. (2016)'s conceptual framework** for the design of asymmetric games were largely applicable to the reviewed asymmetric VR games, allowing systematic descriptions and categorisations of the different kinds of asymmetry in the reviewed games. However, as we pointed out in the results, some categories proved difficult to apply. In mechanics of asymmetry, we found it difficult to distinguish between ability and challenge; generally challenge was not explicitly addressed at all, and in our discussions we saw this as so closely tied to differences in ability that a distinction may not be necessary. However we note that asymmetry of interface can sometimes result in asymmetry of challenge—even in cases when abilities remain the same (e.g., targeting abilities may be easier when controlled by head tracking than via mouse, even though the ability). In those cases, asymmetry of ability and challenge might have to be separate again.

The category of goal/responsibility yielded many interesting observations. First, it was often difficult to determine players' goals. This difficulty partially resulted from game descriptions—which were often unclear or across several sections—but was especially prominent for competitive games. In competitive games, it is more difficult to determine individual players' goals, because these can be subject to interpretation. In a collaborative game, players seek to achieve a shared goal. In competitive games, opponents' goals can be framed as the same or shared as well (e.g., "win X points"). But it could also be seen as either asymmetric, or symmetric but opposed (player A: "win X points more than B"; player B: "win X points more than A"). Second, as described in more detail above, we found that in many cases players had the same goal, but different responsibilities. We thus suggest that this category of goal/responsibility may need to be split up in two. Further, its use should be carefully considered for competitive asymmetric games, because it may not be directly applicable. An additional way in which competitive games were different relates to dependence: in competing against one another, opponents generally do not depend on one another. Nevertheless, there may be cases of players embracing a voluntary or emergent dependence that could be considered a kind of mirrored dependence: For example, in a competitive racing game, players might rely on other players' unspoken adherence to not purposefully causing collisions [what Salen and Zimmerman (2004) refer to as the implicit rules of the game that will cause a game to function beyond its operational and constituative rules]. However such voluntary dependence is easily circumvented by accident (e.g., due to lack of skill) or wilful noncompliance (e.g., purposeful crashing).

For dynamics, directional dependence was suitable for defining interdependencies between players, but clarifications of the dependence degree can avoid confusion between unidirectional and bidirectional dependence. While some papers displayed a mirrored dependence (#1, #15, #21), it was hard to say that "*the nature of each player's reliance on each other is identical*" (Harris et al., 2016) because of their interface asymmetry. Further, synchronicity and timing was challenging to apply to games as a whole; even determining which types of synchronicity occur over the course of the game was often arduous because of unclear descriptions in the reviewed game designs. Overall, we thus suggest to merge ability and challenge, to split goal and responsibility, and to eliminate the synchronicity category in future iterations of this framework, as displayed in **Figure 3**.

Regarding dependence, we note that this classification is the subject to the researchers' understanding of dependence. This was also shaped by many sessions of discussions, and re-coding upon assessing new items in the corpus. This understanding might have to be adjusted further in the future. A more formalized definition of dependence (and subtypes thereof) would be useful for future work. In empirical work, assessment tools like the perceived behavioral and affective interdependence subscales in Networked Minds Social Scale (Harms and Biocca, 2004) might be of use. We further also refer to prior work that has categorized degrees of interdependence in collaborative settings [e.g., tightly vs. loosely coupled, as also touched upon by Harris and Hancock (2019)]. For example, Sigitov et al. (2019) have observed different user roles for tightly coupled collaborative work when interacting with a shared large display using mobile phones for interaction purposes. Similarly, different types of interdependence have been addressed for shared PC monitor (Tse et al., 2004), and co-located digital tabletops (Scott et al., 2003; Morris et al., 2004). How well such degrees of interdependence (designed or observed) can be applied to asymmetric VR settings remains to be explored in the future, but we note the long history of research on trade-offs between mechanics for individual system use vs. for collaborative system use (Gutwin and Greenberg, 1998).

The **MDA aesthetics** part of the framework worked well for sensation, fantasy, and fellowship; however, the other aspects like narrative, challenge, or discovery were not addressed by the papers. As we do believe these have potential application in asymmetric VR, we suggest to keep this in the framework for future use. However, we also suggest that more nuanced alternatives to address different or more specific kinds of engagement, immersion, or social experience may be a useful addition or replacement in the future.

For **social asymmetry**, shared spaces were the most addressed category. Based on our review, we suggest that this category would be particularly useful if further distinguished as *co-located*, *remote*, and *co-located and explicitly shared physical space*. These different shared space setups fostered different kinds of highly engaging asymmetries of social and physical interaction, pointing toward new potential categories for a future iteration of the framework. Further, while age or abilities were not explored as asymmetry by any of the reviewed games, we acknowledge that these may exist as papers that do not use the keywords we employed in our review (see limitations, below). As discussed, asymmetries in these categories could constitute a valuable design space to engage underexplored (combinations of) demographics, such as young children and their parents, or caregivers/physical therapists and older adults with cognitive impairments, or users with varying visual abilities (Gonçalves et al., 2021).

Kaye (2016)'s theoretical factors for social group play were surprisingly under-represented in our review, barring communication—which itself was reported, but often not clearly. We do believe that—with clearer descriptions of expected and observed player communication, this could be a useful part of the framework for future reviews, as well as research and design work. However to fully embrace the social factors involved in asymmetric and interdependent VR games, this part of the framework may need expanding or partially replacing by more detailed theories of social communication in games. We note that while there is research on how players communicate in games (Klimmt and Hartmann, 2008; Walther et al., 2015; Leavitt et al., 2016), much less is known about communication between players in VR games (Rubio-Tamayo et al., 2017). Yet asymmetry and resulting interdependency are likely to heighten experiences of group or team flow (Borderie and Michinov, 2017) in VR, as well.

The **shared control patterns** were sometimes challenging to apply in our identified papers. We found it difficult to determine what constituted a loci of manipulation, especially when there was no distinct game entity through which the player acted (e.g., one player controls the game environment, or when there is no visible game entity representation). One pattern that emerged saw some players without control over the game world: neither distinct loci, nor mutual loci, but a player with *no* or*indirect* loci of control (e.g., only providing verbal support, or impacting the HMD player's virtual-world navigation through real-world actuation). Comparing this kind of non-HMD player experience to a more active one may be an interesting starting point for future work. The other pattern that emerged, giving (usually) the non-HMD player control over the game environment, is one way to create a more active role. We thus suggest a distinction between players having direct control over distinct loci of control (either entities or the environment), sharing locus of control, or having only indirect or no control (**Figure 3**).

In current asymmetric games, distinct loci of manipulation are more prevalent, yet mutual locus of manipulation could reveal interesting social dynamics through highly interdependent games.

Based on our findings resulting from the application of the *a priori* framework, we believe that it is largely well suited for this field, albeit with the adjustments described in this section. We therefore suggest a *post-hoc* "best fit" framework in **Figure 3** that incorporates these changes. We believe that future systematic reviews will be able to re-use and build upon this framework as the field of research and designed artifacts grows. Further, we highlight those categories that are strong potential factors for PX in asymmetric VR games yet are thus far underexplored (indicated via dashed lines and greater transparency in **Figure 3**).

In **Tables 4, 5**, we illustrate which papers in our corpus apply the framework's specific categories—based on the *post-hoc* framework—to also highlight existing gaps in the research. The research gaps evident in **Tables 4, 5** complement the takeaways of this work described in **Table 3**. We note that our framework claims no completeness with regards to research gaps: Other aspects beyond those uncovered by our framework and review may need to be added in the future. For example, the papers in our corpus rarely explored performance as a metric. We assume that this is because performance is rarely the goal for including asymmetries in game design, although that does not mean it could not be used for it.

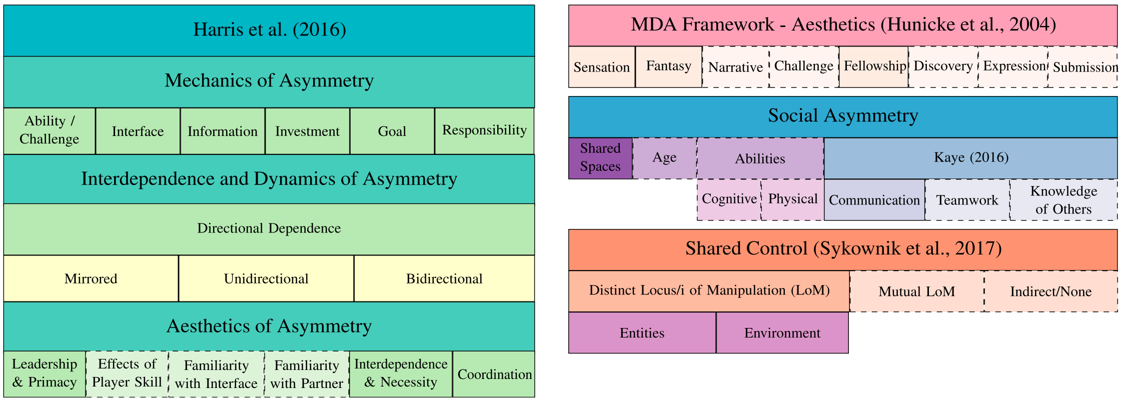

FIGURE 3 | Suggested post-hoc "best fit" framework, with dashed lines and more transparent color for categories that are under-represented in existing asymmetric multiplayer VR games research so far.

Regardless of the focus of future work designed for this field, the *post-hoc* framework provides a lens through which to more clearly describe and design future asymmetric VR games, more clearly delineate desired, expected, and observed results for PX, and offers dimensions for comparison. We further refer readers to related work by Márquez Segura et al. (2021) for their suggested lenses of designed, expected, and observed to describe and analyze player behavior; this could be an interesting addition to our framework in terms of analytical application and methodology.

Finally, we note that adjacent research areas of co-located collaborative systems may hold interesting relevant findings for (collaborative) asymmetric VR games, and thus for future extensions to the framework. We here refer the reader to Brudy et al. (2019)'s taxonomy of cross-device designs and Olin et al. (2020)'s design considerations for cross-device collaboration that include VR. Pinelle et al. (2003)'s mechanics of collaboration framework may be another option for extending our framework. Further, Ouverson and Gilbert (2021) very recently published similar research to our own, positing a framework for asymmetric VR—which they define more narrowly as only asymmetry of interface, opposed to the broader asymmetric VR design space covered in our framework. We believe that our framework can stand concurrently to their five dimensions of asymmetry in VR (spatial co-presence, transportation, informational richness, team interdependence, and balance of power) in future practice. Moreover, given that our chosen methodology has a strong theoretical grounding in multiple theories, our own *posthoc* framework and our findings with regards to existing research gaps (illustrated in **Tables 3–5**, as well as **Figure 3**) may be able to augment their framework. While our work is derived from the context of games, some aspects may also transfer well to the non-game space, and thus enrich the application of both frameworks in parallel. Future work will have to explore this in more detail.

5.2.1 Limitations

As all methodologies, the "best fit" strategy in framework synthesis has limitations. A framework carefully created through comprehensive inductive methods would be preferable to a patchwork-style framework. However, there is no definitive theoretical framework for asymmetric multiplayer VR yet. Perhaps this work can provide a stepping stone toward this. The youth of the field is precisely why this systematic review—using "best fit" framework synthesis—is important at this time.

Further limitations must be acknowledged in terms of review execution. Relevant work that uses different vocabulary could have been missed. As the field (hopefully) moves toward more consistent terminology, this factor should be reduced in future reviews. Yet our added snowball approach mitigates this limitation. Further, although two authors conducted the synthesis together in close communication and over many extensive discussions, bias is still likely. Following recommendations for qualitative methods, we provide a statement on reflexivity (Newton et al., 2012; Berger, 2015): the authors conducting the synthesis have a Computer Science and Cognitive Systems background, respectively, and both have previously conducted VR and VR games research.

We also note that our search query no longer results in the same number of documents as it did at the time of initial data collection. The ACM digital library presents irregularities in terms of its database query results (with lower numbers for our query over time). While we are in correspondence with ACM and the company that built their search engine, at this stage it is unclear why this is the case. However, the initial search resulting in the corpus of this review yielded the highest number of results in our (re-)sampling over the past year. This suggests that our analysis simply screened a larger pool of publications. Yet we point out that even if a few papers were missed due to using different terms, this would not limit the validity of the framework and the review's synthesis.

Our review is based on 25 papers, relating 30 VR artifacts (plus design variants) and 17 studies of sort (plus reports of informal playtesting). This of course only provides a glimpse into the potential overall design space, so while we did speculate to a degree based on non-VR research (e.g., suggesting the design of asymmetry in age or abilities), we cannot claim completeness, only a first step toward it. In future work, we will explore commercial asymmetric VR games as well. Additionally, limitations of our thematic findings are bound to the limitations of the reviewed papers—for example, as also found in general VR research (Peck et al., 2020; MacArthur et al., 2021), this subfield also has predominantly investigated PX for male participants, which may have introduced bias into results.

6 CONCLUSION

This paper presents the first literature review on how asymmetric game design is thus far being leveraged in multiplayer VR. Based on "best fit" framework synthesis, we draw on existing theoretical games research (within and outside of VR research) to gain insight into the state of the field, identify opportunities for more complete coverage of the design space, and point out where the field can improve in methodological rigor. Our final suggestions for a *post-hoc* framework can also be used by future papers to describe asymmetric VR games, as well extended in future systematic reviews as the field grows. The results showcase a novel field with great promise in including non-HMD players, facilitating multiplayer engagement, minimizing VR isolation, and providing room to research a complex range of social dynamics. We hope this review can both spark discussion and orient the field in the journey toward achieving these aims.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

KR and SK contributed to the development of the protocol, including search strategy, selection criteria, risk of bias assessment strategy and data extraction criteria. LN provided feedback on this process. KR conducted the search and removal of duplicates; KR and SK conducted the screening phases; DW acted as third reviewer (tie-breaker). KR and SK developed the *a priori* framework and applied the framework synthesis methodology. KR led the drafting of the manuscript, in collaboration with SK; LN and FS assisted with phrasing and framing. All authors read, provided feedback and approved the final manuscript.

ACKNOWLEDGMENTS

We thank our reviewers, and are especially glad for the thoughtful examples of asymmetry that improved this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2021.694660/full#supplementary-material

REFERENCES

Abeele, V. V., Spiel, K., Nacke, L., Johnson, D., and Gerling, K. (2020). Development and Validation of the Player Experience Inventory: A Scale to Measure Player Experiences at the Level of Functional and Psychosocial Consequences. *Int. J. Human-Computer Stud.* 135, 102370. doi:10.1016/j.ijhcs.2019.102370

Abuhamdeh, S., Csikszentmihalyi, M., and Jalal, B. (2015). Enjoying the Possibility of Defeat: Outcome Uncertainty, Suspense, and Intrinsic Motivation. *Motiv. Emot.* 39, 1–10. doi:10.1007/s11031-014-9425-2

Aromataris, E., and Munn, Z. (2020). JBI Manual for Evidence Synthesis. Available at: https://synthesismanual.jbi.global.doi:10.46658/JBIMES-20-01

Aron, A., Aron, E. N., and Smollan, D. (1992). Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness. *J. Personal. Soc. Psychol.* 63, 596–612. doi:10.1037/0022-3514.63.4.596

Benford, S., Greenhalgh, C., Giannachi, G., Walker, B., Marshall, J., and Rodden, T. (2012). "Uncomfortable Interactions," in CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, May 5–10, 2012 (New York, NY, USA: Association for Computing Machinery), 2005–2014. doi:10.1145/2207676.2208347

Berger, R. (2015). Now I See it, Now I Don't: Researcher's Position and Reflexivity in Qualitative Research. *Qual. Res.* 15, 219–234. doi:10.1177/1468794112468475

Beznosyk, A., Quax, P., Lamotte, W., and Coninx, K. (2012). "The Effect of Closely-Coupled Interaction on Player Experience in Casual Games," in 11th International Confernece on Entertainment Computing (ICEC), Bremen, Germany, September 2012 (Berlin/Heidelberg, Germany: Springer), 243–255. doi:10.1007/978-3-642-33542-6_21

Boland, D., and McGill, M. (2015). Lost in the Rift. *Xrds* 22, 40–45. doi:10.1145/2810046

Bopp, J. A., Mekler, E. D., and Opwis, K. (2016). "Negative Emotion, Positive Experience?," in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, May 7–12, 2016 (New York, NY, USA Association for Computing Machinery), 2996–3006. doi:10.1145/2858036.2858227

Borderie, J., and Michinov, N. (2017). *Identifying Social Forms of Flow in Multiuser Video Games*. New York: Routledge, 32–64. doi:10.4324/9781315629308-3

Bortolaso, C., Bourdiol, J., and Graham, T. C. N. (2019). "Enhancing Communication and Awareness in Asymmetric Games," in Joint International Conference on Entertainment Computing and Serious Games, Arequipa, Peru, November 11–15, 2019 (Cham, Switzerland: Springer), 250–262. doi:10.1007/978-3-030-34644-7_20

Bradley, M. M., and Lang, P. J. (1994). Measuring Emotion: the Self-Assessment Manikin and the Semantic Differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi:10.1016/0005-7916(94)90063-9

Braun, V., and Clarke, V. (2020). One Size Fits All? what Counts as Quality Practice in (Reflexive) Thematic Analysis? *Qual. Res. Psychol.* 1, 1–25. doi:10.1080/14780887.2020.1769238

Braun, V., and Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* 3, 77–101. doi:10.1191/1478088706qp063oa

Brudy, F., Holz, C., Rädle, R., Wu, C.-J., Houben, S., Klokmose, C. N., et al. (2019). *Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning across Multiple Devices*. New York, NY, USA: Association for Computing Machinery, 1–28.

Burtscher, S., and Spiel, K. (2020). "But where Would I Even Start?," in MuC '20: Mensch und Computer 2020 — Tagungsband (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3404983.3405510

Buscemi, N., Hartling, L., Vandermeer, B., Tjosvold, L., and Klassen, T. P. (2006). Single Data Extraction Generated More Errors Than Double Data Extraction in Systematic Reviews. *J. Clin. Epidemiol.* 59, 697–703. doi:10.1016/j.jclinepi.2005.11.010

Cairns, P., Power, C., Barlet, M., Haynes, G., Kaufman, C., and Beeston, J. (2019). Enabled Players: The Value of Accessible Digital Games. *Games Cult.* 16, 262–282. doi:10.1177/1555412019893877

Carroll, C., Booth, A., and Cooper, K. (2011). A Worked Example of "best Fit" Framework Synthesis: A Systematic Review of Views Concerning the Taking of Some Potential Chemopreventive Agents. *BMC Med. Res. Methodol.* 11, 1–9. doi:10.1186/1471-2288-11-29

Carroll, C., Booth, A., Leaviss, J., and Rick, J. (2013). "best Fit" Framework Synthesis: Refining the Method. *BMC Med. Res. Methodol.* 13, 37. doi:10.1186/1471-2288-13-37

Cechanowicz, J. E., Gutwin, C., Bateman, S., Mandryk, R., and Stavness, I. (2014). "Improving Player Balancing in Racing Games," in CHI PLAY '14: Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play, October 19–21, 2014 (New York, NY, USA: Association for Computing Machinery), 47–56. doi:10.1145/2658537.2658701

Cheng, L.-P., Lühne, P., Lopes, P., Sterz, C., and Baudisch, P. (2014). "Haptic Turk," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 3463–3472. doi:10.1145/2556288.2557101

Ciolek, T. M. (1983). The Proxemics Lexicon: A First Approximation. *J. Nonverbal Behav.* 8, 55–79. doi:10.1007/bf00986330

Coleman, B. (2009). Using Sensor Inputs to Affect Virtual and Real Environments. *IEEE Pervasive Comput.* 8, 16–23. doi:10.1109/MPRV.2009.60

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., and Boyle, J. M. (2012). A Systematic Literature Review of Empirical Evidence on Computer Games and Serious Games. *Comput. Educ.* 59, 661–686. doi:10.1016/j.compedu.2012.03.004

De Kort, Y. A., IJsselsteijn, W. A., and Poels, K. (2007). "Digital Games as Social Presence Technology: Development of the Social Presence in Gaming Questionnaire (SPGQ)," in PRESENCE 2007. Vol. 195203 (Stockholm, Sweden: Starlab), 195–203.

Deci, E. L., and Ryan, R. M. (2011). Self-determination Theory. *Handbook theories Soc. Psychol.* 1, 416–437. doi:10.4135/9781446249215.n21

Depping, A. E., Mandryk, R. L., Johanson, C., Bowey, J. T., and Thomson, S. C. (2016). "Trust Me," in CHI PLAY '16: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play, October 16–19, 2016 (New York, NY, USA: Association for Computing Machinery), 116–129. doi:10.1145/2967934.2968097

Depping, A. E., and Mandryk, R. L. (2017). "Why Is This Happening to Me?," in CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, May 6–11, 2017 (New York, NY, USA: Association for Computing Machinery), 1040–1052. doi:10.1145/3025453.3025648

Deterding, S., Lucero, A., Holopainen, J., Min, C., Cheok, A., Waern, A., and Walz, S. (2015). "Embarrassing Interactions," in CHI EA '15: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, April 18–23, 2015 (New York, NY, USA: Association for Computing Machinery), 2365–2368. doi:10.1145/2702613.2702647

Dixon-Woods, M. (2011). Using Framework-Based Synthesis for Conducting Reviews of Qualitative Studies. *BMC Med.* 9, 39. doi:10.1186/1741-7015-9-39

Dourish, P. (2004). *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, MA, USA: MIT press.

Efstratios, G., Michael, T., Stephanie, B., Athanasios, L., Paul, Z., and George, P. (2018). "New Cross/augmented Reality Experiences for the Virtual Museums of the Future," in Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection, Nicosia, Cyprus, October 29–November 3, 2018 (Cham, Switzerland: Springer International Publishing), 518–527. doi:10.1007/978-3-030-01762-0_45

Eklund, L. (2015). Playing Video Games Together with Others: Differences in Gaming with Family, Friends and Strangers. *journal gaming virtual worlds* 7, 259–277. doi:10.1386/jgvw.7.3.259_1

Fitterer, D. (2016). *Audioshield*. Vashon, WA, United States: Game [VR, PC]. Audiosurf LLC. https://store.steampowered.com/app/412740/Audioshield/.

Furukawa, T., Yamamoto, D., Sugawa, M., Peiris, R., and Minamizawa, K. (2019). "TeleSight." in SIGGRAPH '19: ACM SIGGRAPH 2019 Emerging Technologies, July 28–August 1, 2019 (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3305367.3335040

Gamer, M., Lemon, J., Gamer, M. M., Robinson, A., and Kendall's, W. (2012). *Package 'irr'*. Various Coefficients of Interrater Reliability and Agreement. https://cran.r-project.org/web/packages/irr/irr.pdf.

George, C., Janssen, P., Heuss, D., and Alt, F. (2019). "Should I Interrupt or Not?," in DIS '19: Proceedings of the 2019 on Designing Interactive Systems Conference, June 23–28, 2019 (New York, NY, USA: Association for Computing Machinery), 497–510. doi:10.1145/3322276.3322363

Gerling, K. M., Miller, M., Mandryk, R. L., Birk, M. V., and Smeddinck, J. D. (2014). *Effects of Balancing for Physical Abilities on Player Performance, Experience and Self-Esteem in Exergames*. New York, NY, USA: Association for Computing Machinery, 2201–2210. doi:10.1145/2556288.2556963

Gonçalves, D., Rodrigues, A., Richardson, M. L., de Sousa, A. A., Proulx, M. J., and Guerreiro, T. (2021). "Exploring Asymmetric Roles in Mixed-Ability Gaming," in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (New York, NY: Association for Computing Machinery). doi:10.1145/3411764.3445494

Graf, R., Benawri, P., Whitesall, A. E., Carichner, D., Li, Z., Nebeling, M., and Kim, H. S. (2019). "iGYM," in CHI PLAY '19: Proceedings of the Annual Symposium on Computer-Human Interaction in Play, October 22–25, 2019 (New York, NY, USA: Association for Computing Machinery), 31–43. doi:10.1145/3311350.3347161

Gugenheimer, J., Stemasov, E., Frommel, J., and Rukzio, E. (2018a). "A Demonstration of Sharevr: Co-located Experiences for Virtual Reality between Hmd and Non-hmd Users," in 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Reutlingen, Germany, March 18–22, 2018 (Piscataway, NJ: IEEE), 755–756. doi:10.1109/VR.2018.8446551

Gugenheimer, J., Stemasov, E., Frommel, J., and Rukzio, E. (2017a). "ShareVR," in CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, May 6–11, 2017 (New York, NY, USA: Association for Computing Machinery), 4021–4033. doi:10.1145/3025453.3025683

Gugenheimer, J., Stemasov, E., Sareen, H., and Rukzio, E. (2017b). "FaceDisplay," in CHI EA '17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, May 6–11, 2017 (New York, NY, USA: Association for Computing Machinery), 369–372. doi:10.1145/3027063.3052962

Gugenheimer, J., Stemasov, E., Sareen, H., and Rukzio, E. (2018b). "FaceDisplay," in CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, April 21–26, 2018 (New York, NY, USA: Association for Computing Machinery), 1–13. doi:10.1145/3173574.3173628

Gutwin, C., and Greenberg, S. (1998). "Design for Individuals, Design for Groups," in CSCW '98: Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work, November 14–18, 1998 (New York, NY, USA: Association for Computing Machinery), 207–216. doi:10.1145/289444.289495

Hall, E. T. (1966). *The Hidden Dimension. 1966*. New York, NY, USA: New York: Doubleday.

Hansen, A., Larsen, K. B., Nielsen, H. H., Sokolov, M. K., and Kraus, M. (2020). "Asymmetrical Multiplayer versus Single Player: Effects on Game Experience in a Virtual Reality Edutainment Game," in Augmented Reality, Virtual Reality and Computer Graphics (AVR 2020), Lecce, Italy, September 7–10, 2020 (Cham, Switzerland: Springer International Publishing), 22–33. doi:10.1007/978-3-030-58465-8_2

Harms, C., and Biocca, F. (2004). "Internal Consistency and Reliability of the Networked Minds Measure of Social Presence," in *Seventh Annual International Workshop: Presence 2004*. Editors M. Alcaniz and B. Rey (Valencia: Universidad Politecnica de Valencia).

Harper, R., and Mentis, H. (2013). "The Mocking Gaze," in CSCW '13: Proceedings of the 2013 Conference on Computer Supported Cooperative Work (New York, NY, USA: ACM), 167–180. doi:10.1145/2441776.2441797

Harris, J., Hancock, M., and Scott, S. (2014). ""beam Me 'round, Scotty!"," in CHI PLAY '14: Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play (New York, NY, USA: Association for Computing Machinery), 417–418. doi:10.1145/2658537.2661311

Harris, J., Hancock, M., and Scott, S. D. (2016). "Leveraging Asymmetries in Multiplayer Games," in CHI PLAY '16: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play (New York, NY, USA: Association for Computing Machinery), 350–361. doi:10.1145/2967934.2968113

Harris, J., and Hancock, M. (2019). "To Asymmetry and beyond!," in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: ACM), 9. doi:10.1145/3290605.3300239

Hornecker, E. (2005). "A Design Theme for Tangible Interaction: Embodied Facilitation," in ECSCW 2005, Paris, France, September 18–22, 2005 (Dordrecht, Switzerland: Springer), 23–43. doi:10.1007/1-4020-4023-7_2

Hudson, M., and Cairns, P. (2014). "6 Measuring Social Presence in Team-Based Digital Games," in Interacting with Presence: HCI and the Sense of Presence in Computer-mediated Environments, 83. doi:10.2478/9783110409697.6

Hunicke, R., LeBlanc, M., and Zubek, R. (2004). "MDA: A Formal Approach to Game Design and Game Research," in Proceedings of the AAAI Workshop on Challenges in Game AI. Vol. 4, 1722.

IJsselsteijn, W. A., de Kort, Y. A., and Poels, K. (2013). *The Game Experience Questionnaire*. Eindhoven: Technische Universiteit Eindhoven, 3–9.

Jeong, K., Kim, J., Kim, M., Lee, J., and Kim, C. (2020). Asymmetric Interface: User Interface of Asymmetric Virtual Reality for New Presence and Experience. *Symmetry* 12, 53. doi:10.3390/sym12010053

Jerald, J. (2015). *The VR Book: Human-Centered Design for Virtual Reality*. San Rafael, CA: Association for Computing Machinery and Morgan & Claypool.

Johnson, D., Gardner, M. J., and Perry, R. (2018). Validation of Two Game Experience Scales: The Player Experience of Need Satisfaction (PENS) and Game Experience Questionnaire (GEQ). *Int. J. Human-Computer Stud.* 118, 38–46. doi:10.1016/j.ijhcs.2018.05.003

Karaosmanoglu, S., Rogers, K., Wolf, D., Rukzio, E., Steinicke, F., and Nacke, L. E. (2021). "Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer Vr Games," in CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3411764.3445492

Kaye, L. K. (2016). Exploring Flow Experiences in Cooperative Digital Gaming Contexts. *Comput. Hum. Behav.* 55, 286–291. doi:10.1016/j.chb.2015.09.023

Kendon, A. (2010). "Spacing and Orientation in Co-present Interaction," in Development of Multimodal Interfaces: Active Listening and Synchrony (Berlin/Heidelberg, Germany: Springer), 1–15. doi:10.1007/978-3-642-12397-9_1

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *Int. J. aviation Psychol.* 3, 203–220. doi:10.1207/s15327108ijap0303_3

Kerure, A. A., and Freeman, J. (2018). "Audio Source Localization as an Input to Virtual Reality Environments," in Audio Engineering Society Convention, Milan, Italy, May 23–26, 2018 (New York, NY: Audio Engineering Society), 144.

Kim, E., and Schneider, O. (2020). "Defining Haptic Experience: Foundations for Understanding, Communicating, and Evaluating Hx," in CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, April 25–30, 2020 (New York, NY, USA: Association for Computing Machinery), 1–13. doi:10.1145/3313831.3376280

Kiourt, C., Theodoropoulou, H. G., Koutsoudis, A., Ioannakis, J. A., Pavlidis, G., and Kalles, D. (2020). "Exploiting Cross-Reality Technologies for Cultural Heritage Dissemination," in Applying Innovative Technologies in Heritage Science (Hershey, PA: IGI Global), 85–108. doi:10.4018/978-1-7998-2871-6.ch005

Klimmt, C., and Hartmann, T. (2008). Mediated Interpersonal Communication in Multiplayer Video Games: Implications for Entertainment and Relationship Management. *Mediated interpersonal Commun.* 309, 323–344. doi:10.4324/9780203926864-26

Kmet, L. M., Cook, L. S., and Lee, R. C. (2004). *Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of fields*. Canada: Institute of Health Economies, IHE, University of Alberta. doi:10.7939/R37M04F16

Knierim, P., Funk, M., Kosch, T., Fedosov, A., Müller, T., Schopf, B., Weise, M., and Schmidt, A. (2016). "UbiBeam++," in NordiCHI '16: Proceedings of the 9th Nordic Conference on Human-Computer Interaction, October 23–27, 2016 (New York, NY, USA: Association for Computing Machinery). doi:10.1145/2971485.2996747

Law, E. L.-C., Brühlmann, F., and Mekler, E. D. (2018). in CHI PLAY '18: Systematic review and validation of the game experience questionnaire (geq) - implications for citation and reporting practice (New York, NY, USA: Association for Computing Machinery), 257–270. doi:10.1145/3242671.3242683

Leavitt, A., Keegan, B. C., and Clark, J. (2016). "Ping to Win?," in CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 4337–4350. doi:10.1145/2858036.2858132

Lee, J., Kim, M., and Kim, J. (2020). Rolevr: Multi-Experience in Immersive Virtual Reality between Co-located Hmd and Non-hmd Users. *Multimed Tools Appl.* 79, 979–1005. doi:10.1007/s11042-019-08220-w

Li, J., Deng, H., and Michalatos, P. (2017a). "CatEscape," in CHI PLAY '17: Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play (New York, NY, USA: Association for Computing Machinery), 585–590. doi:10.1145/3130859.3130860

Li, J., Deng, H., and Michalatos, P. (2017b). "MagicTorch," in CHI PLAY '17: Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play (New York, NY, USA: Association for Computing Machinery), 431–436. doi:10.1145/3130859.3131341

Lifton, J. H. (2007). *Dual Reality: An Emerging Medium*. PhD thesis. (Massachusetts, MA: Massachusetts Institute of Technology). http://hdl.handle.net/1721.1/41707.

Liszio, S., Emmerich, K., and Masuch, M. (2017). "The Influence of Social Entities in Virtual Reality Games on Player Experience and Immersion," in FDG '17: Proceedings of the 12th International Conference on the Foundations of Digital Games (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3102071.3102086

Liszio, S., and Masuch, M. (2016). "Designing Shared Virtual Reality Gaming Experiences in Local Multi-Platform Games," in International Conference on Entertainment Computing (Springer), 235–240. doi:10.1007/978-3-319-46100-7_23

MacArthur, C., Grinberg, A., Harley, D., and Hancock, M. (2021). "You're Making Me Sick," in CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3411764.3445701

Manninen, T., and Korva, T. (2005). "Designing Puzzles for Collaborative Gaming Experience - Case: Escape," in DiGRA Conference: Selected Papers Proceedings of Digital Games Research Association's Second International Conference, 233–247.

Marquardt, N., Hinckley, K., and Greenberg, S. (2012). "Cross-device Interaction via Micro-mobility and F-Formations," in UIST '12: Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (New York, NY, USA: Association for Computing Machinery), 13–22. doi:10.1145/2380116.2380121

Márquez Segura, E., Rogers, K., Martin-Niedecken, A. L., Niedecken, S., and Vidal, L. T. (2021). "Exploring the Design Space of Immersive Social Fitness Games: The Imsofit Games Model," in CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3411764.3445592

Martínez-Ruiz, F. J., Rauh, S. F., and Meixner, G. (2019). "Understanding Peripheral Audiences: From Subtle to Full Body Gestures," in Advances in Intelligent Systems and Computing (Springer International Publishing), 489–495. doi:10.1007/978-3-030-25629-6_76

McDonald, N., Schoenebeck, S., and Forte, A. (2019). Reliability and Inter-rater Reliability in Qualitative Research. *Proc. ACM Hum.-Comput. Interact.* 3, 1–23. doi:10.1145/3359174

McGill, M., Boland, D., Murray-Smith, R., and Brewster, S. (2015). "A Dose of Reality," in CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 2143–2152. doi:10.1145/2702123.2702382

Milgram, P., and Kishino, F. (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Inf. Syst.* E77-D (12), 1321–1329.

Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. Prisma Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the Prisma Statement. *PLoS Med.* 6 (7), e1000097. doi:10.1371/journal.pmed.1000097

Morris, M. R., Ryall, K., Shen, C., Forlines, C., and Vernier, F. (2004). "Beyond "social Protocols"," in CSCW '04: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (New York, NY, USA: Association for Computing Machinery), 262–265. doi:10.1145/1031607.1031648

Murray, J. H. (2017). *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. Cambridge, MA, USA: MIT press.

Mütterlein, J., and Hess, T. (2017). "Immersion, Presence, Interactivity: towards a Joint Understanding of Factors Influencing Virtual Reality Acceptance and Use," in Twenty-third Americas Conference on Information Systems (AMCIS) AMCIS2017 Proceedings: Adoption and Diffusion of Information Technology (SIGADIT).

Nakagawa, R., and Sonobe, K. (2019). "Encounters," in SA '19: SIGGRAPH Asia 2019 XR (New York, NY, USA: Association for Computing Machinery), 6–8. doi:10.1145/3355355.3361886

Nestler, S., Back, M. D., and Egloff, B. (2011). Psychometrische Eigenschaften zweier Skalen zur Erfassung interindividueller Unterschiede in der Präferenz zum Alleinsein. *Diagnostica* 57, 57–67. doi:10.1026/0012-1924/a000032

Newton, B. J., Rothlingova, Z., Gutteridge, R., LeMarchand, K., and Raphael, J. H. (2012). No Room for Reflexivity? Critical Reflections Following a Systematic

Review of Qualitative Research. *J. Health Psychol.* 17, 866–885. doi:10.1177/1359105311427615

Nowell, L. S., Norris, J. M., White, D. E., and Moules, N. J. (2017). Thematic Analysis. *Int. J. Qual. Methods* 16, 160940691773384. doi:10.1177/1609406917733847

O'Hagan, J., Williamson, J. R., and Khamis, M. (2020). "Bystander Interruption of Vr Users," in PerDis '20: Proceedings of the 9TH ACM International Symposium on Pervasive Displays (New York, NY, USA: Association for Computing Machinery), 19–27. doi:10.1145/3393712.3395339

O'Hagan, J., and Williamson, J. R. (2020). "Reality Aware Vr Headsets," in PerDis '20: Proceedings of the 9TH ACM International Symposium on Pervasive Displays (New York, NY, USA: Association for Computing Machinery), 9–17. doi:10.1145/3393712.3395334

Odd Raven Studios (2019). *Carly and the Reaperman—Escape from the Underworld*. Stockholm, Sweden: Game [HTC Vive, Windows]. Odd Raven Studios.

Olin, P. A., Issa, A. M., Feuchtner, T., and Grønbæk, K. (2020). "Designing for Heterogeneous Cross-Device Collaboration and Social Interaction in Virtual Reality," in OzCHI '20: 32nd Australian Conference on Human-Computer Interaction (New York, NY, USA: Association for Computing Machinery), 112–127. doi:10.1145/3441000.3441070

Osmanovic, S., and Pecchioni, L. (2016). Beyond Entertainment. *Games Cult.* 11, 130–149. doi:10.1177/1555412015602819

Othlinghaus, J., Gerling, K. M., and Masuch, M. (2011). "Intergenerational Play: Exploring the Needs of Children and Elderly," in Workshop-Proceedings of Mensch & Computer 2011 (Chemnitz, Germany: Universitätsverlag der Technischen Universität Chemnitz).

Ouverson, K. M., and Gilbert, S. B. (2021). A Composite Framework of Co-located Asymmetric Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 5, 1–20. doi:10.1145/3449079

Peck, T. C., Sockol, L. E., and Hancock, S. M. (2020). Mind the gap: The Underrepresentation of Female Participants and Authors in Virtual Reality Research. *IEEE Trans. Vis. Comput. Graphics* 26, 1945–1954. doi:10.1109/TVCG.2020.2973498

Perry, R., Drachen, A., Kearney, A., Kriglstein, S., Nacke, L. E., Sifa, R., et al. (2018). Online-only Friends, Real-Life Friends or Strangers? Differential Associations with Passion and Social Capital in Video Game Play. *Comput. Hum. Behav.* 79, 202–210. doi:10.1016/j.chb.2017.10.032

Pinelle, D., Gutwin, C., and Greenberg, S. (2003). Task Analysis for Groupware Usability Evaluation. *ACM Trans. Comput.-Hum. Interact.* 10, 281–311. doi:10.1145/966930.966932

Piumsomboon, T., Lee, G. A., Hart, J. D., Ens, B., Lindeman, R. W., Thomas, B. H., and Billinghurst, M. (2018). "Mini-Me," in CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 1–13. doi:10.1145/3173574.3173620

Pope, C. (2000). Qualitative Research in Health Care: Analysing Qualitative Data. *BMJ* 320, 114–116. doi:10.1136/bmj.320.7227.114

Reeves, S., Benford, S., O'Malley, C., and Fraser, M. (2005). "Designing the Spectator Experience," in CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 741–750. doi:10.1145/1054972.1055074

Reilly, D. F., Rouzati, H., Wu, A., Hwang, J. Y., Brudvik, J., and Edwards, W. K. (2010). "TwinSpace," in UIST '10: Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology (New York, NY, USA: Association for Computing Machinery), 119–128. doi:10.1145/1866029.1866050

Resolution Games (2019). *Acron: Attack of the Squirrels*. Stockholm, Sweden: Game [Oculus Quest, iOS/Android]. Resolution Games. https://www.resolutiongames.com/acron.

Rigby, S., and Ryan, R. (2007). *The Player Experience of Need Satisfaction (Pens): An Applied Model and Methodology for Understanding Key Components of the Player Experience*.

Ritchie, J., and Spencer, L. (2002). Qualitative Data Analysis for Applied Policy Research. *The Qual. researcher's companion* 573, 305–329. doi:10.4135/9781412986274.n12

Rogers, K., Funke, J., Frommel, J., Stamm, S., and Weber, M. (2019). "Exploring Interaction Fidelity in Virtual Reality," in CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 1–14. doi:10.1145/3290605.3300644

Rubio-Tamayo, J., Gertrudix Barrio, M., and García García, F. (2017). Immersive Environments and Virtual Reality: Systematic Review and Advances in Communication, Interaction and Simulation. *Multimodal Technol. Interact.* 1, 21. doi:10.3390/mti1040021

Russell, D., Peplau, L. A., and Ferguson, M. L. (1978). Developing a Measure of Loneliness. *J. Personal. Assess.* 42, 290–294. doi:10.1207/s15327752jpa4203_11

Ryan, R. M., and Deci, E. L. (2000). Self-determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *Am. Psychol.* 55, 68–78. doi:10.1037/0003-066x.55.1.68

Ryan, R. M., Rigby, C. S., and Przybylski, A. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motiv. Emot.* 30, 344–360. doi:10.1007/s11031-006-9051-8

Sajjadi, P., Cebolledo Gutierrez, E. O., Trullemans, S., and De Troyer, O. (2014). "Maze Commander," in CHI PLAY '14: Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play (New York, NY, USA: Association for Computing Machinery), 227–236. doi:10.1145/2658537.2658690

Salen, K., and Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. Cambridge, MA, USA: MIT press.

Schmidt, S., Nunez, O. J. A., and Steinicke, F. (2019). "Blended Agents: Manipulation of Physical Objects within Mixed Reality Environments and beyond," in Symposium on Spatial User Interaction, SUI 2019, New Orleans, LA, October 19-20, 2019. Editors C. W. Borst, A. K. Kulshreshth, G. Bruder, S. Serafin, C. Sandor, K. Johnsen, et al. (New York, NY: Association for Computing Machinery), 6:1–6:10. doi:10.1145/3357251.3357591

Schmitz, M., Akbal, M., and Zehle, S. (2015). "SpielRaum," in CHI PLAY '15: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play (New York, NY, USA: Association for Computing Machinery), 427–432. doi:10.1145/2793107.2810306

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators & Virtual Environments* 10, 266–281. doi:10.1162/105474601300343603

Scott, S. D., Grant, K. D., and Mandryk, R. L. (2003). "System Guidelines for Co-located, Collaborative Work on a Tabletop Display," in Proceedings of the Eighth Conference on European Conference on Computer Supported Cooperative Work - ECSCW'03 (Springer), 159–178. doi:10.1007/978-94-010-0068-0_9

Seif El-Nasr, M., Aghabeigi, B., Milam, D., Erfani, M., Lameman, B., Maygoli, H., and Mah, S. (2010). "Understanding and Evaluating Cooperative Games," in CHI '10: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 253–262. doi:10.1145/1753326.1753363

Serubugo, S., Skantarova, D., Evers, N., and Kraus, M. (2018a). "Facilitating Asymmetric Collaborative Navigation in Room-Scale Virtual Reality for Public Spaces," in Interactivity, Game Creation, Design, Learning, and Innovation, Heraklion, Greece, October 30–31, 2018 (Cham, Switzerland: Springer), 64–73. doi:10.1007/978-3-319-76908-0_7

Serubugo, S., Skantarova, D., Evers, N., and Kraus, M. (2018b). "Self-overlapping Maze and Map Design for Asymmetric Collaboration in Room-Scale Virtual Reality for Public Spaces," in Interactivity, Game Creation, Design, Learning, and Innovation, Heraklion, Greece, October 30–31, 2018 (Cham, Switzerland: Springer), 194–203. doi:10.1007/978-3-319-76908-0_19

Serubugo, S., Skantarova, D., Evers, N., and Kraus, M. (2017). "Walkable Self-Overlapping Virtual Reality Maze and Map Visualization Demo," in VRST '17: Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3139131.3141774

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., et al. (2015). Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015: Elaboration and Explanation. *BMJ* 349, g7647. doi:10.1136/bmj.g7647

Sigitov, A., Hinkenjann, A., Kruijff, E., and Staadt, O. (2019). Task Dependent Group Coupling and Territorial Behavior on Large Tiled Displays. *Front. Robot. AI* 6, 128. doi:10.3389/frobt.2019.00128

Simeone, A., Khamis, M., Kljun, M., Isokoski, P., Esteves, A., Daiber, F., et al. (2020). "International Workshop on Cross-Reality (XR) Interaction," in

Companion Proceedings of the 2020 Conference on Interactive Surfaces and Spaces, Virtual Event, Portugal (New York, NY: Association for Computing Machinery), 111–114. doi:10.1145/3380867.3424551

Simon, B. (2009). Wii Are Out of Control: Bodies, Game Screens and the Production of Gestural Excess. *SSRN J.* doi:10.2139/ssrn.1354043

Slater, M., Usoh, M., and Steed, A. (1994). Depth of Presence in Virtual Environments. *Presence: Teleoperators & Virtual Environments* 3, 130–144. doi:10.1162/pres.1994.3.2.130

Smilovitch, M., and Lachman, R. (2019). "Birdquestvr: A Cross-Platform Asymmetric Communication Game," in CHI PLAY '19: Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (New York, NY, USA: Association for Computing Machinery), 307–313. doi:10.1145/3341215.3358246

Sra, M., Jain, D., Caetano, A. P., Calvo, A., Hilton, E., and Schmandt, C. (2016). "Resolving Spatial Variation and Allowing Spectator Participation in Multiplayer Vr," in UIST '16 Adjunct: Proceedings of the 29th Annual Symposium on User Interface Software and Technology (New York, NY, USA: Association for Computing Machinery), 221–222. doi:10.1145/2984751.2984779

Sra, M., Xu, X., and Maes, P. (2017). "GalVR," in VRST '17: Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3139131.3141219

Steel Crate Games (2015). *Keep Talking and Nobody Explodes*. Ottawa, Canada: Game [VR, PC]. Steel Crate Games. https://keeptalkinggame.com/.

Streck, A., Stepnicka, P., Klaubert, J., and Wolbers, T. (2019b). "Neomento - towards Building a Universal Solution for Virtual Reality Exposure Psychotherapy," in 2019 IEEE Conference on Games (CoG), London, United Kingdom, August 20–23, 2019 (Piscataway, NJ: IEEE), 1–2. doi:10.1109/cig.2019.8848110

Streck, A., Stepnicka, P., Klaubert, J., and Wolbers, T. (2019a). "Neomento SAD - VR Treatment for Social Anxiety," in 2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), San Diego, CA, December 9–11, 2019 (Piscataway, NJ: IEEE). doi:10.1109/aivr46125.2019.00054

Striner, A., Azad, S., and Martens, C. (2019). "A Spectrum of Audience Interactivity for Entertainment Domains," in Interactive Storytelling, Salt Lake City, UT, November 19–22, 2019 (Cham, Switzerland: Springer International Publishing), 214–232. doi:10.1007/978-3-030-33894-7_23

Sykownik, P., Emmerich, K., and Masuch, M. (2017). "Exploring Patterns of Shared Control in Digital Multiplayer Games," in International Conference on Advances in Computer Entertainment, London, United Kingdom, December 14–16, 2017 (Cham, Switzerland: Springer), 847–867. doi:10.1007/978-3-319-76270-8_57

Team Panoptes (2019). *Panoptic*. Brussels, Belgium: Game [VR, PC]. Team Panoptes. https://panopticgame.com/.

Tse, E., Histon, J., Scott, S. D., and Greenberg, S. (2004). "Avoiding Interference," in CSCW '04: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work (New York, NY, USA: Association for Computing Machinery), 252–261. doi:10.1145/1031607.1031647

Vella, K., Klarkowski, M., Turkay, S., and Johnson, D. (2019). Making Friends in Online Games: Gender Differences and Designing for Greater Social Connectedness. *Behav. Inf. Technol.* 39, 917–934. doi:10.1080/0144929x.2019.1625442

Vicencio-Moreira, R., Mandryk, R. L., Gutwin, C., and Bateman, S. (2014). "The Effectiveness (Or Lack Thereof) of Aim-Assist Techniques in First-Person Shooter Games," in CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA: Association for Computing Machinery), 937–946. doi:10.1145/2556288.2557308

Voida, A., and Greenberg, S. (2012). Console Gaming across Generations: Exploring Intergenerational Interactions in Collocated Console Gaming. *Univ. Access Inf. Soc.* 11, 45–56. doi:10.1007/s10209-011-0232-1

Walther, J. B., Van Der Heide, B., Ramirez, A., Jr., Burgoon, J. K., and Peña, J. (2015). *Interpersonal and Hyperpersonal Dimensions of Computer-Mediated Communication*. (West Sussex, United Kingdom: John Wiley & Sons), 1–22. doi:10.1002/9781118426456.ch1

Wang, M., Ryoo, J., and Winkelmann, K. (2020). Preface to the Special Issue on Cross Reality (XR) and Immersive Learning Environments (ILE) in Education. *Interactive Learn. Environments* 28, 539–542. doi:10.1080/10494820.2019.1696845

Want, R. (2009). Through Tinted Eyeglasses. *IEEE Pervasive Comput.* 8, 2–4. doi:10.1109/mprv.2009.58

Wells, J. D., Campbell, D. E., Valacich, J. S., and Featherman, M. (2010). The Effect of Perceived novelty on the Adoption of Information Technology Innovations: A Risk/reward Perspective. *Decis. Sci.* 41, 813–843. doi:10.1111/j.1540-5915.2010.00292.x

Witmer, B. G., Jerome, C. J., and Singer, M. J. (2005). The Factor Structure of the Presence Questionnaire. *Presence: Teleoperators & Virtual Environments* 14, 298–312. doi:10.1162/105474605323384654

Witmer, B. G., and Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence* 7, 225–240. doi:10.1162/105474698565686

Wohlin, C. (2014). "Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering," in Proceedings of the 18th international conference on evaluation and assessment in software engineering, 1–10. doi:10.1145/2601248.2601268

Woods, S. J. (2009). (play) Ground Rules: The Social Contract and the Magic circle. *Observatorio (OBS*)* 3, 204–222. doi:10.15847/obsOBS312009243

Woodworth, J. W., Ekong, S., and Borst, C. W. (2017). "Virtual Field Trips with Networked Depth-Camera-Based Teacher, Heterogeneous Displays, and Example Energy center Application," in 2017 IEEE Virtual Reality (VR) (IEEE), 471–472. doi:10.1109/VR.2017.7892384

Yassien, A., ElAgroudy, P., Makled, E., and Abdennadher, S. (2020). "A Design Space for Social Presence in Vr," in NordiCHI '20: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (New York, NY, USA: Association for Computing Machinery). doi:10.1145/3419249.3420112

Zagal, J. P., Rick, J., and Hsi, I. (2006). Collaborative Games: Lessons Learned from Board Games. *Simulation & Gaming* 37, 24–40. doi:10.1177/1046878105282279

Zhou, Z., Márquez Segura, E., Duval, J., John, M., and Isbister, K. (2019). "Astaire." in CHI PLAY '19: Proceedings of the Annual Symposium on Computer-Human Interaction in Play (New York, NY, USA: Association for Computing Machinery), 5–18. doi:10.1145/3311350.3347152

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

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

11.3 Feels Like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games

[CorePub8] **Sukran Karaosmanoglu**, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. Feels like team spirit: Biometric and strategic interdependence in asymmetric multiplayer vr games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, (CHI '21), New York, NY, USA, 2021. ISSN 2673-4192. Association for Computing Machinery. ISBN 9781450380966. DOI 10.1145/3411764.3445492. URL https://doi.org/10.1145/3411764.3445492

Supplementary materials for this article are available here: https://doi.org/10.1145/3411764.3445492.

Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games

Sukran Karaosmanoglu
Human-Computer Interaction, Universität Hamburg
karaosmanoglu@informatik.uni-hamburg.de

Katja Rogers
HCI Games Group, Stratford School of Interaction Design and Business, University of Waterloo
katja.rogers@acm.org

Dennis Wolf
Institute of Media Informatics, Ulm University
dennis.wolf@uni-ulm.de

Enrico Rukzio
Institute of Media Informatics, Ulm University
enrico.rukzio@uni-ulm.de

Frank Steinicke
Human-Computer Interaction, Universität Hamburg
frank.steinicke@uni-hamburg.de

Lennart E. Nacke
HCI Games Group, Stratford School of Interaction Design and Business, University of Waterloo
lennart.nacke@acm.org

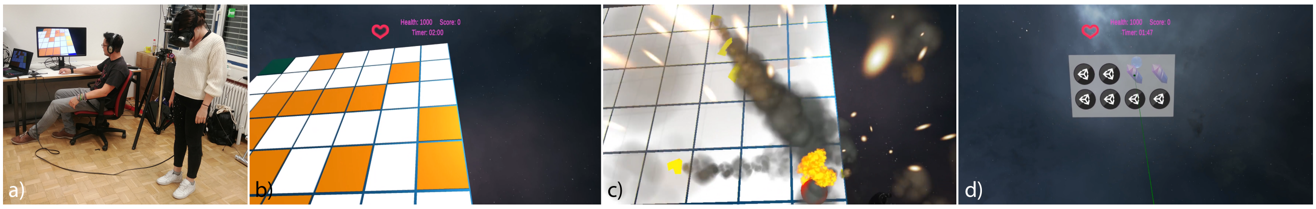

Figure 1: Illustration of our asymmetric VR game concept: (a) both players share the in-game view (controlled by the HMD player), but only the non-HMD player can see additional information such as traps (b, orange tiles) to be avoided by the HMD player (c) and puzzle pieces that need to be paired in a memory game (d).

ABSTRACT

Virtual reality (VR) multiplayer games increasingly use asymmetry (e.g., differences in a person's capability or the user interface) and resulting interdependence between players to create engagement even when one player has no access to a head-mounted display (HMD). Previous work shows this enhances player experience (PX). Until now, it remains unclear whether and how an asymmetric game design with interdependences creates comparably enjoyable PX for both an HMD and a non-HMD player. In this work, we designed and implemented an asymmetric VR game (different in its user interface) with two types of interdependence: *strategic* (difference in game information/player capability) and *biometric* (difference in player's biometric influence). Our mixed-methods user study (N=30) shows that asymmetries positively impact PX for both player roles, that interdependence strongly affects players' perception of agency, and that biometric feedback—while subjective—is a valuable game mechanic.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; • **Software and its engineering → Interactive games**.

CHI '21, May 8–13, 2021, Yokohama, Japan

KEYWORDS

VR; virtual reality; multiplayer; asymmetry; interdependence; strategic; biometric

ACM Reference Format:
Sukran Karaosmanoglu, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. 2021. Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3411764.3445492

1 INTRODUCTION

Current virtual reality (VR) systems for home entertainment such as the HTC Vive [41] support high audiovisual sensory immersion for the player who wears the head-mounted display (HMD) [6, 87]. For this setup, mirroring this player's view on a TV screen/monitor is a common way to provide external access to the VR world to users outside of the VR environment. This setup limits bystanders' interaction to passive consumption rather than active participation because the external display form factor does not provide the same kind of sensory immersion. In some cases, this limited or passive participation is the preferred option for bystanders because they can learn from and enjoy watching others [92] and some people may be more hesitant to try out new games or systems [45]; this may also apply in VR [96]. However, increasingly, VR systems and games that include bystanders more actively have reported positive effects on player experience (PX). For example, Gugenheimer et al.

[34] previously introduced an exploratory HMD prototype that allows both passive and active participation in the same physical space for an HMD and a non-HMD player. Their results indicate that integrating a non-HMD player in a VR experience is feasible, yielding improved enjoyment and presence for both players. However, in their study, PX remained more immersive and enjoyable for the HMD player. This raises the question of whether it is possible to design a VR experience that induces similar engagement for both users regardless of how the virtual world allows them to interact (either with or without an HMD).

Similar approaches to merging the real world with the virtual world have been explored by other researchers [16, 29, 83]. Still, the empirical work often focuses on the novelty and innovation of the HMD implementation, while comprehensive evaluations of PX in asymmetric VR game design remain sparse. Nevertheless, asymmetric games have an enormous potential to increase social interaction between—and thereby wellbeing of—people with different abilities or hardware access. By applying asymmetric VR game design, researchers and developers can connect people despite a potentially single-owned HMD and counteract potential isolation stemming from its use.

To leverage the positive effects of asymmetric VR games, we draw on the conceptual framework by Harris et al. that describes ways to design asymmetric games and create interdependence between players [36–38]. While Harris et al. explored the effects of different *degrees* of interdependence, we focus on different *types* of interdependence. Therefore, we designed a VR game with an asymmetric interface (one player wearing the HMD; the other viewing the virtual world on a monitor). We created two types of asymmetry: *strategic* (different information and interaction opportunities) and *physiological* (different biometric influence).

Multiplayer games have often featured asymmetry as a difference in interaction opportunities and information provided to players (e.g., *Keep Talking and Nobody Explodes* [32], *Panoptic* [69], *Black Hat Cooperative* [88], or *Battlefield 1942* [30]). A notable example is *Savage: The Battle for Newerth* [31], which placed one user in a commander's role. Other users played a first-person shooter (with third-person perspective for melee) role, following the commander's orders. In our design, the non-HMD player is provided with additional information about the VR world that needs to be communicated to the HMD player to navigate and win the game. We expect that this asymmetry will result in interdependence that we term: **strategic interdependence (SI)**, which requires one player to rely on another for information and capability.

We further expand our game design's asymmetry to include a novel, physiological aspect: a difference in biometric influence, resulting in **biometric interdependence (BI)**; it requires one player to rely on another one's physiological responses. The non-HMD player's heart rate (HR) is linked to the game difficulty, increasing it when it passes a predefined threshold (determined in a pilot study; n=10). The game industry has explored this integration of biometric feedback into games in commercial VR games (e.g., *Left4Dead 2* [1, 90], *Alien Swarm* [1, 2, 91], *Nevermind* [26], and *Bring to Light* [73]) and academic research has reported a positive impact of biometric feedback on PX and user experience in games [39, 40, 59, 67]. However, variants that support multiple users or players remain largely underexplored [22], and its usage in asymmetric games has not been investigated yet. We expect BI between players to intensify the players' feelings of responsibility for the other and enhance social PX (similar to previous research exploring physiological linkage and increased social presence among participants [24]). Further, we argue that VR is a fascinating context for this feature because using the HMD isolates the users from the real world while improving embodiment and immersion into the VR world [7, 75].

To evaluate the levels of enjoyment, presence, affective state, and immersion this game design could create for *both* players, regardless of the player role (between-participants: HMD or non-HMD), we conducted a mixed-design user study (N=30). Our findings show how game designers and researchers can use asymmetry (and resulting interdependences) to create enjoyable and engaging experiences between users inside and outside of VR. We further explored PX with or without biometric influence. Moreover, we investigated the effects of different multimodal indicators of this influence (within-subjects: visual, auditory, and audiovisual biometric indicators) to determine how best to provide this kind of feedback for emotion regulation without distracting from gameplay.

Our results indicate that enjoyment, presence, affective state, and immersion scores were comparably high for both player roles, showing that interdependence can help integrate players across different displays of the same medium. While the physiological influence did not yield significant effects, we can nevertheless draw implications for biometric integration in VR game design; for example, a need for familiarization and less subtle impact. Furthermore, our participant interviews emphasize the importance and variety of communication between players despite the interface's asymmetry, players' understanding of agency in the different roles, and the multitude of preferences for biometric indicators. Our key contributions are as follows:

- we demonstrate and discuss how a non-HMD player can be integrated into a VR game while enhancing PX for both the HMD player and the non-HMD player,
- discuss qualitative design considerations for this design goal,
- and report a first exploration of biometric influence as an interdependence type in asymmetric VR.

2 RELATED WORK

There have been attempts to include non-VR users in collaborative VR settings, and there is prior work showing asymmetries in games can enhance PX; however, there are few comprehensive explorations of asymmetric VR games. Further—while biometric feedback has been explored in games—there is little work exploring its integration in VR.

2.1 Collaborative VR Experiences that Include Non-VR Users

Asymmetric VR setups have been explored outside of games in many contexts. Stafford et al. [85, 86] propose the addition of a top-down view for a non-HMD user, so that they can provide more effective instructions for the HMD user. Their results indicate that purely auditory instructions, which are often used for guidance [3, 53], are less efficient than visual cues for the HMD user, but they did not explore effects on presence and enjoyment of either user.

To share the VR experience of an HMD player with the outside world—beyond simple mirroring of the point-of-view [42]—previous work has proposed approaches for the active participation of non-HMD players. Yang et al. [95] addressed the challenge of non-HMD users crossing the tracking area of an HMD player by visualizing them as "shields" within the virtual environment [95]. Their focus, however, was on reducing interference with the VR experience rather than integrating non-HMD users.

Physical interfaces have been explored as mediums for interaction in VR in multi-user scenarios. For example, Mai et al. proposed enabling the collaboration of HMD and non-HMD users through a physical surface that serves as an interface to the virtual world via two-dimensional, bi-directional input and output [53]. Their results suggest that communication between HMD and non-HMD users improved task performance and presence of the HMD user. However, the non-HMD user's presence was not evaluated, which hinders comparison between the roles.

In general, the focus of these systems is on the user experience of the HMD-wearing user and ignores the experience of the non-HMD user. In a game setting, however, integrating non-HMD players into a VR game makes them players; this creates an additional crucial design goal in creating an engaging experience for *both* players.

2.2 Asymmetry in Games & VR Games

With increasingly reliable technology, multiplayer gaming environments have become virtual meeting places where people can socialize. They can also contain a multitude of differing abilities, interfaces, and preferences of players, and it can be challenging to foresee how these differences interact with or affect PX. Yet these differences, or *asymmetries*, can also significantly enhance PX and social connectedness, by inducing interdependences between players, for example, making one player rely on another. Many game designers and developers are integrating and catering to asymmetries that foster interdependence between players to facilitate multiplayer engagement.

Harris et al. [38] have introduced a first conceptual framework of asymmetric game design, in which they address how video game elements can cater to differences between players: asymmetry of mechanics, dynamics, and aesthetics. Moreover, they observed higher social connectedness and social presence values for asymmetric gameplay where the players have asymmetry of ability, information, and interface, when compared to a symmetric one [36].

VR is particularly suited to explore asymmetry in game design. It is highly immersive, yet research indicates that immersive VR experiences can also be isolating [7, 62, 75] because people feel self-conscious in front of potential onlookers. Several commercial VR games already include asymmetric displays of the VR world to provide a less isolating experience, involving both an HMD and one or multiple non-HMD players (e.g.,*Ruckus Ridge VR Party* [27], *Acron: Attack of the Squirrels* [74], or *Carly and the Reaperman—Escape from the Underworld* [68], all praised for being engaging).

In academic research, Sajjadi et al. [78] showed comparable PX for both the HMD player and the non-HMD player using Sifteo cubes as an interface, showing equal levels of satisfaction for different interaction modes. Some papers have examined the contribution of sharing both first-person and third-person points-of-view to the non-HMD players' PX, yielding comparable levels of PX for both players in terms of presence and social interaction [44, 50]. A recent prototype, *SilhouetteVR* [47], presents the VR world to non-HMD users through a dynamic view frustum displayed on a one-way mirror/screen which reflects the HMD player's embodiment within the VR world. While they do not yet integrate the non-HMD users as players, their initial results for the enjoyment of the non-HMD user are promising. Furthermore, VR prototypes *FaceDisplay* and *FrontFace* allowed non-HMD players to participate in the virtual world through touch-screens that were attached to the HMD, enabling co-located interaction techniques for both users [14, 35]. In an evaluation of *FaceDisplay*, both roles reported enjoyment, however presence and arousal were significantly higher for the HMD player than for the non-HMD player. The concept of integrating non-HMD players was implemented more comprehensively with *ShareVR*, which extends room-scale VR with whole-body interaction for non-HMD players via top-down projection and a hand-held monitor; an example of an augmented-reality approach [34]. The researchers found that both the HMD and non-HMD player experienced higher presence and enjoyment than in the baseline (wherein the non-HMD player used a game pad and a TV set). However, the non-HMD player still reported significantly lower presence and enjoyment than the HMD player for both conditions. The cause of this imbalance remained undiscussed but might be explained by the specific design of interdependence between roles.

These findings motivated us to further explore asymmetry of information and ability (strategic dependence) with different interfaces (HMD-VR and monitor-display): we aimed to design a comparably enjoyable and engaging experience for *both* players, and explore which factors in game mechanics and dynamics contribute to highly engaging PX.

2.3 Biometric Feedback

Integration of biometric feedback has been shown to increase engagement and immersion in single-player games [39, 67]. A common approach is that an increase in player arousal (measured with physiological metrics) leads to a more difficult game, classified as "challenge me" gameplay by Gilleade et al. [33]. Further, these games often require players to self-regulate [52] which can improve stress-management skills [9]. Nacke et al. propose that indirect physiological input such as HR and galvanic skin response should be mapped to features of the game world rather than direct actions [63]. This finding motivated us to link the current game difficulty to the non-HMD player's HR signal.

Kuikkaniemi et al. found that explicit biometric feedback allows players to self-regulate and increases immersion [48]. Sinclair et al. [80, 81] showed that use of HR metrics effective in controlling an exergame to meet the level of exercise desired by players. To aid self-regulation, previous work explored different cueing mechanisms that represent the current biometric state [57, 79]. In *Life Tree*, Patibanda et al. [70] provide biometric feedback about the player's breathing through changes in the VR environment (e.g., a tree object that expands and contracts to match their breathing pattern). Sra et al. [84] have used breathing as a physiological input mechanism in VR, and suggest that this kind of physiological factor can increase presence and players' connection to the physical world. Chen et al. found that players enjoy audio feedback of their HR, while visual cues were described as distracting [15]. Similarly, Dey et al. explored the influence of an artificially accelerated and decelerated auditory HR cue on players' physiological signal [21]. Their results indicate that auditory cues could affect player emotions but did not affect their HR. These findings prompted us to explore the impact of biometric cues on PX in our user experiment.

In multiplayer games, sharing each other's HR was found to have no significant influence on players' emotional state [22] but did improve engagement with an exertion activity [93]. While previous work has explored biometric feedback for symmetrical multiplayer games [22], it has not yet been explored in an asymmetric VR game setup, wherein we argue that it could represent a novel type of interdependence. Therefore, we expand the asymmetry of our game design's SI to the physiological aspect by using HR metrics of the non-HMD player[1] to influence the game world and difficulty. We expect this BI to increase the players' experience of relatedness, as well as the non-HMD player's feeling of responsibility for the HMD player, thereby intensifying the experience of *both* players.

3 RESEARCH QUESTIONS

Our primary research question concerns an asymmetric collaborative VR game, in which only one player acts in the VR world via HMD, yet both players feel engaged and immersed. This holds some game design challenges, because when only one player is actively immersed in VR, the HMD player could easily feel isolation or self-consciousness because of the immersive nature of VR. We thus aimed to design asymmetry of information to include a non-HMD player in the VR game. Thus, the non-HMD player has information required by the HMD player, creating player interdependence.

Hence, the first research question **(RQ1)** asks:

- *Can imbalanced asymmetric information between players lead to one player feeling less in control (i.e., like they are following the other player's instructions without their own agency)?* Agency is an important factor in PX [61], making this a game design challenge for our development team.

For the second research question, we were interested in the effect of biometric feedback in this asymmetric VR experience. **RQ2**:

- *Does biometric influence over game difficulty dynamically affect PX in VR (compared to static difficulty)?* Further, does it make a difference *a)* whether players are provided an in-game indicator of whether their own biometric state is currently increasing game difficulty, or *b)* in *what modality* this indicator is represented in-game.

As a final overarching research theme, we were interested in comprehensively exploring players' experience of an asymmetric VR game, to gain insight into design factors that shape balanced PX between both players.

[1]In early pilot tests, we explored the feasibility of gathering biometric data of both player roles, but only the more stationary non-HMD player's data resulted in reliable measurements.

4 ASYMMETRIC VR GAME IMPLEMENTATION

To explore our research questions in an empirical study, we designed and implemented a VR game as a stimulus. The collaborative VR game for two players featured a design, in which only one person is wearing an HMD (see Figure 1a). The game mechanics are distributed asymmetrically: the player wearing the HMD plays a more active role but has to rely on information provided by the player outside of VR (non-HMD player). The non-HMD player has more information about the game world (which they view via a PC monitor), and their physiological state—indicated by their HR—affects the game parts described below.

Technical Setup. The game was implemented using C# in the Unity game engine (version 2018.3.8f1) [89], with the addition of Valve's SteamVR v1.2.3 [17] and the Virtual Reality Toolkit (VRTK) v3.3 plugin [28]. The experiment was conducted using an i7-7700HQ CPU and a GeForce GTX1060 graphics card.

The VR setup consisted of the HTC Vive [41] with the usual two base stations and one controller for the HMD player. For the non-HMD player, the setup consisted of a PC with a 27-inch monitor (2560x1440 resolution) and headphones. The HR of the non-HMD player was acquired using an Empatica E4 wristband [25]; the corresponding software and a Windows Bluetooth low-energy (BLE) server were used to stream the HR data to the game. The wristband has a 64 Hz sampling rate for blood volume pulse measurement, yielding HR values (sampling rate 1 Hz) and inter-beat interval.

4.1 Game Design

The game was designed as a custom collaborative game that features different roles and asymmetrical information for two players. We chose to focus on collaborative gameplay as the literature highlights improved PX and increased interdependence for HMD and non-HMD users in such scenarios [35, 96]. Following that, the game design choices (e.g., sharing the same camera view, employing memory puzzle) were made to promote communication and increase interdependence between players.

The game consists of three levels. In each level, players are given limited resources (i.e., two minutes time and three lives) to fulfil the tasks: navigate a virtual grid and complete a memory puzzle. The grid is populated with traps; their layout varies with each level. The game world also spawns lasers (at the current height of the HMD player's headset) with a default frequency of 15 seconds.

Asymmetric Roles. The player in VR (*HMD player*) physically navigates the game space grid (Figure 1b). This includes avoiding lasers (Figure 2), traps (Figure 1c), and activating buttons in the memory puzzle (i.e., uncovering images, see Figure 1d). However, they are reliant on their co-player's knowledge of the game world: they cannot see the traps on the grid, nor approaching lasers.

The non-VR player (*non-HMD player*) watches the HMD player's view of the game world via a PC monitor, but has additional information—that is not visible in VR—overlaid on top of this on the monitor display. This player's view includes traps (i.e., the orange-coloured grid cells in Figure 1b), the remaining time (see Figure 1c,d), and they are able to see and hear incoming lasers. They are also able to see uncovered images in the memory puzzle (see Figure 1d).

Figure 2: Lasers (only visible to the non-HMD player) spawned at roughly HMD height, at a distance of 50m from the grid (distance marked by Earth object as a reference point). The HMD player ducked based on the non-HMD player's instructions on timing.

Biometric Influence: Lasers. The HR of the non-HMD player dynamically affects the difficulty of the game, by affecting how frequently lasers appear (every 15 seconds by default). Prior to the game, the non-HMD player's HR is collected as a baseline over a time period of 5 minutes. The non-HMD player wears the wristband on their non-dominant hands, and they were instructed not to talk or move to avoid noise in the data (during the game, they were allowed to talk, but were still instructed not to move). For the rest of the game, a time interval of two seconds is used as a sliding window to acquire continuous HR data from the non-HMD player. Their current HR is then compared to the baseline measurement to calculate a multiplier[2]. This multiplier was used to dynamically affect the frequency of the lasers. *CurrentHeartRate* here refers to the variable that is averaged over a two-second time interval of continuous HR data using the sliding window approach, while *BaselineHeartRate* is the average of the 5-min baseline measurement prior to gameplay. In this way, a multiplier is <1 when the non-HMD player's HR is higher than the baseline, and >1 when it is lower. Values higher than the baseline HR[3] thus triggered more frequent lasers.

Lasers were then spawned at a height of 1–3 cm below the player's current headset height, 50 meters away from the current playing field and reaching the main play area after ~7–8 seconds. They were accompanied by a laser-style sound effect, to give the non-HMD player an understanding of its active position. When the laser was at a distance of 10 meters to the HMD player, an additional sound effect was played for the non-HMD player (3-beat proximity alert).

Furthermore, a secondary measure was applied to ensure participants could not continuously crouch once lasers had spawned. We applied a logarithmic function between the laser and the HMD positions on the playing field so that the former moved toward the latter, adjusting for vertical changes. As a result, the co-players had to communicate with the HMD player to avoid lasers in a timely fashion. The non-HMD player could observe the visual approach of the laser (when the HMD player looked up), and use the laser sound effects as a more continuous auditory warning.

Indicators of Biometric Influence. In some variants of the game, indicators were used to represent the participants' excitement level, and acted as a warning signal for the higher frequency of lasers that accompanied increased excitement. The game variants featured different types of indicators (and combinations thereof) which were then compared in the study: auditory cueing, visual cueing, and combined audio-visual cueing.

Auditory Cueing: The sound of heart beats was used to represent the increased excitement level of the players. When the non-HMD player's excitement exceeded the baseline measure, heart beat sounds were played via the headphones. The sound was played at a constant frequency irrespective of the player's actual HR: the time between beats was 0.65 seconds; each beat consisted of two amplitude peaks separated by 0.25 seconds[4].

Visual Cueing: A red frame flashing around the in-game view—visible only for the non-HMD player (see Figure 3)—was used as a visual cue to represent the increased excitement level of the non-HMD player, and thus conveyed the increased likelihood of lasers. The frequency of the visual cueing was identical to the auditory variant: The red frame always flashed twice separated by 0.25 seconds; the time between paired flashes was 0.65 seconds.

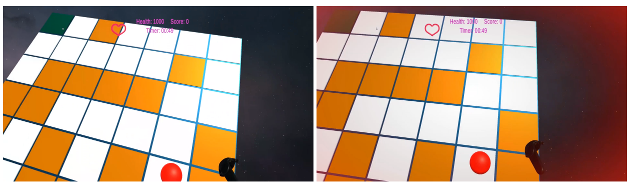

Figure 3: With visual cueing, the non-HMD player's view (left) is augmented with a flashing red overlay frame (right) to indicate an increased HR (and higher frequency of lasers).

[2]Following the equation by Dekker et al. [20]): *Multiplier = 1 / (CurrentHeartRate / BaselineHeartRate)*

[3]In a pilot study (N=10) we found that higher thresholds set the laser frequency too low, inducing less enjoyment.

[4]These values were determined via another pilot test with three pairs of participants in the early tests.

5 USER STUDY

The goal of the study was to explore whether the game design would elicit a positive experience for both player roles, and gain some insight into the factors that resulted in this experience. Further, we wanted to test the influence of the asymmetrical biometric feedback loop on PX. To inform the design of in-game representations of biometric influence, we compared biometric indicators of different modalities (auditory, visual, audiovisual) with regards to their effect on both player roles' experience and the non-HMD player's ability to control their physiological arousal.

Methodology. A mixed design was used for the study, with *player role* and *type of biometric influence* as independent variables. Player roles (HMD and non-HMD) were randomly assigned at the beginning of the study and not switched (between-participants), while the type of biometric influence (including its in-game representation) varied as a within-participants variable over five playthroughs.

We chose a mixed-methods approach followed a triangulation-convergence model [18], which places the quantitative and qualitative measures of our study at equal importance. With the quantitative aspects of our study, we aimed to measure and compare PX across player roles and conditions. In complement to this, there is no questionnaire for measuring underlying factors to PX in asymmetric VR games, so we used qualitative methods (semi-structured interviews) to more deeply and flexibly investigate players' understanding of the roles and the designed asymmetry.

5.1 Conditions

The participants played the game with five different conditions in the study:

(1) *Baseline Game—No Biometric Interdependence (NBI).* In this variant, the non-HMD player's HR had no influence on the frequency of lasers (i.e., set at the default frequency of 15 seconds). To keep conditions comparable (and for later analysis), the non-HMD player still wore the Empatica wristband and their biometric data was recorded.
(2) *Biometric Interdependence (BI).* In this variant, the collected biometric data affected the frequency of lasers in the form described above. However, the non-HMD player was not informed of their current physiological state through any in-game representation (no auditory or visual cueing).
(3) *Biometric Interdependence with Auditory Indicators (BI-A).* This variant was identical to *BI*, however the non-HMD player was informed of their heightened excitement via auditory cueing as described above.
(4) *Biometric Interdependence with Visual Indicators (BI-V).* Here the game employed visual cueing to indicate heightened arousal on part of the non-HMD player.
(5) *Biometric Interdependence with Audio-Visual Indicators (BI-AV).* This variant employed both auditory and visual cueing to signal excitement levels above the baseline.

An overview of these conditions (when biometric influence was present, and how it was indicated to the non-HMD player in their monitor-display overlay) is presented in Table 1.

Attributes	Conditions of the Experiment: *NBI*	*BI*	*BI-A*	*BI-V*	*BI-AV*
Biometric Manipulation	—	x	x	x	x
Indicator Usage	—	—	x	x	x
Auditory Indicator	—	—	x	—	x
Visual Indicator	—	—	—	x	x

Table 1: An overview of the conditions of the experiment.

5.2 Participants

The study was conducted with 30 participants (13 female, 16 male, 1 non-binary) with an average age of 26.03 years (*SD*=3.18). 18 participants (10 in the HMD role) had prior VR experience while 12 of the participants (5 in the HMD role) did not. Gender was roughly similarly distributed within the assigned player roles (HMD player: 7 female, 8 male; non-HMD player: 6 female, 8 male, 1 non-binary). All participant dyads reported that they knew each other.

5.3 Measures

We used physiological and psychometric measures for all participants. For a subset of 14—7 male (4 HMD player, 3 non-HMD player); 6 female (3 HMD player, 3 non-HMD player); 1 non-binary (non-HMD player)—the study was concluded with an optional interview asking in more detail about their experience of the different game variants.

Physiological Measures. We calculated two HR metrics for our analysis, *average HR* per condition (regardless of baseline), and *variance in HR difference to baseline*. The average HR was calculated by averaging the data points collected for each condition, while the variance in HR differences subtracted the individual baseline measurement from each data point, prior to calculating their standard deviation. This type of HR measure has been evaluated and found to represent player arousal in various game research studies [23, 54–56, 58].

In-Game Metrics. We logged descriptive values for each condition/playthrough: how often the player teams died in each condition, and how many levels they completed, the playthrough duration, and experienced number of lasers and trigger events. While we do not report these in the results for scope, we provide a table with average values per condition in the supplementary materials.

Post-Game Questionnaires. After each playthrough, we assessed participants' affective state, immersion, and presence via questionnaires. Affective state was measured as arousal, valence, and dominance with the three 7-point pictorial scales of the self-assessment manikin (SAM) [10]: 1=*calm*/*unhappy*/*controlled*; 7=*excited*/*happy*/*dominant*. For immersion, we employed the Immersive Experience Questionnaire (IEQ) [43], which consists of the subfactors real-world dissociation, challenge, control, emotional and cognitive involvement (7-point Likert scales: 1=*not at all*; 7=*a lot*)[5] as well as

[5]The labels here are presented as examples, as they differ depending on item phrasing for the IEQ as well as the SUS (e.g., *very often* or *definitely no*), but higher numbers represent positive scores.

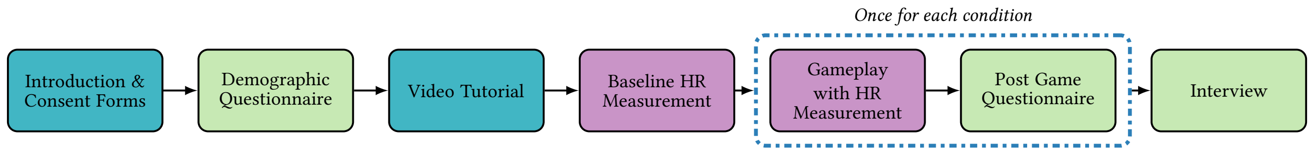

Figure 4: The study procedure followed a mixed-design and assessed psychometric, physiological, and interview data. For a subset of 14 participants, we conducted an interview on their experience of the different game variants.

a single-item measure of immersion (10-point scale of same direction). A presence rating was acquired via the Slater-Usoh-Steed Questionnaire (SUS) [82]; the questionnaire consists of 6 items on a 7-point Likert scale (1=*not at all*; 7=*very much so*). We also employed a custom single-item measure of enjoyment ("*I enjoyed the experience in this condition*") on a 7-point Likert scale (1=*not at all*; 7=*very much so*).

Interview. An interview with a total of 32 pre-determined questions (14 questions posed to the HMD player; 18 posed to the non-HMD player) was conducted to explore the participants' experiences towards collaborative game attitudes, their experience of their player role, the different game variants, and the asymmetrical feedback. While we followed the interview guideline, we also diverged from it for follow-up questions or clarifications based on the participants' responses (semi-structured interview). The full list of questions is provided in the supplementary materials.

5.4 Procedure

We announced our experiment in many different digital distribution forms: mailing lists, messaging groups and the university boards. Subsequently, the participants voluntarily participated in the experiment by contacting the experimenter or registering for an empty time slot using an online portal.

At the beginning of the study, participants were asked to sit in front of two separate monitors, and introduced to the study purpose via handout. After signing the consent form and filling out a demographic questionnaire, participants read a form describing the two different player roles, and then watched the corresponding video tutorial (ca. 2 minutes) for their (randomly) assigned role. Following that, the non-HMD player was asked to sit in front of the designated monitor, where a 5-minute baseline measurement was taken of their resting HR. The participant was instructed not to move or talk during the measurement. At the same time, the HMD player was instructed on how to put on the Vive HMD, and given time to become accustomed to the controllers.

Once the baseline measurement was complete, the experiment was run for the five conditions in counterbalanced order (*5x5 Latin square*). Each condition consisted of gameplay (with the corresponding game variant) and subsequent post-game questionnaires. Further, participants were asked if they wanted to participate in the optional interview to inform us of their experience of the different game variants. To ensure the comfort of participants, we offered an opt-out in case participants preferred not to be interviewed afterwards (e.g., due to time constraints or hesitation in speaking English, which was the second language for most participants). As a result, the interview was conducted (separately) with 7 HMD and 7 non-HMD participants (1 non-binary, 6 female, and 7 male).

The study procedure is illustrated in Figure 4. Participants' remuneration for study (75-90 minutes duration) consisted of 10 EUR.

6 ANALYSIS AND RESULTS

The analysis of the psychometric and physiological data was conducted with parametric tests where data were normally distributed. For not normally distributed cases, we employed non-parametric tests as suggested by Wobbrock and Kay [94]. This is described in more detail below.

For the interviews (n=14: n=7 HMD and n=7 non-HMD player; 2:02 hours of audio recordings total), we applied a thematic analysis methodology using an approach that uses elements from both the *reflexive* and *codebook* orientations of thematic analysis [11, 13]. Our approach consisted of the following: an *a priori* deductive categorization of codes, a reflexive perspective on inductive code and theme generation, and two coders for both consensus and nuanced, collaborative construction of codes. We defined three overarching deductive categories at the beginning of our analysis: *collaboration in VR*, *asymmetry of player roles*, and *biometric asymmetry and indicators*. The interview data were then inductively coded by the first two authors independently, with codes placed into the three *a priori* categories. In four meetings (after coding four interviews in each batch, except the last one), the authors discussed all applied codes, and resolved different readings by adding codes, removing them, or merging/splitting them. We note that discrepancies were thus not necessarily seen as a conflict to be resolved, but instead could be reflected through additional and alternative codes. Themes were then developed from the codes by re-reading and synthesizing the coded quotes and discussed between the first two authors.

6.1 Player Experience

For the mixed-design analysis of PX measures, we used ANOVA-type statistics (ATS) for non-parametric mixed designs from the `nparLD` R package [66], which is reported with adjusted degrees of freedom[6]. None of the factors (for immersion, presence, enjoyment or affective state) showed a significant effect of role or condition, nor an interaction effect. Descriptive statistics separated by role are listed in Table 2; results by condition in Table 3.

Effects of VR Experience. Related work has suggested that novelty may be an initial factor in VR experiences [64, 72, 76], we

[6] "*The adjusted degrees of freedom [(DoF)] used for the approximation of the distribution of ATS may appear to be quite different from the conventional [DoF] employed in the traditional repeated measures ANOVA. However, such an adjustment [...] can be viewed as a generalization of the conventional [DoF] in the heteroscedastic case.*" [66]

Role	Arousal		Valence		Dominance		Enjoyment		Presence		IEQ Single Item	
	Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
HMD	6	4–6	6	5–7	5	3–6	7	6–7	5	4.17–5.5	9	6.5–10
non-HMD	5	4–6	6	5–7	5	4–6	6	5–7	4.67	3.92–5.17	8	7–9
	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p
Main effect	0.04	0.85	0.00	0.99	0.42	0.52	1.40	0.24	0.66	0.42	0.03	0.86
	Control		Challenge		Cogn. Inv.		Emot. Inv.		Real-World Diss.		IEQ Sum	
	Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
HMD	5	4.4–5.8	5	4.5–5.5	6.22	5.56–6.44	5.67	4.67–6.33	5	4–5.86	174	150.5–182.5
non-HMD	5	4.4–5.6	5.25	4.5–5.5	6	5.28–6.62	5.5	4.92–6.17	4.86	4.14–5.64	165	149.5–184
	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p	$F_{ATS}(1)$	p
Main effect	0.00	0.99	0.65	0.42	0.00	0.95	0.00	1	0.01	0.93	0.00	0.97

Table 2: PX results were positive and did not differ significantly between player roles.

Condition	Arousal		Valence		Dominance		Enjoyment		Presence		IEQ Single Item	
	Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
NBI	5	4.25–6	6	5–7	5	4–6	7	6–7	4.67	3.92–5.67	8	7–9.75
BI	5	4–6	6	6–7	5	4–5	6	6–7	5	4.17–5.33	8	7–10
BI-A	6	4–6	6	6–7	5	3.25–5.75	6	5.25–7	4.67	4.21–5.13	8	7–9
BI-V	6	4–6	6	5–7	5	4–6	6	6–7	4.83	4.08–5.33	9	7.25–10
BI-AV	6	4–6	6.5	5.25–7	5	3–6	6.5	6–7	4.83	4.08–5.33	8	7.25–10
	$F_{ATS}(3.30)$	p	$F_{ATS}(2.86)$	p	$F_{ATS}(3.10)$	p	$F_{ATS}(3.34)$	p	$F_{ATS}(3.35)$	p	$F_{ATS}(2.80)$	p
Main effect	0.62	0.61	0.88	0.44	1.31	0.27	1.14	0.33	0.42	0.76	0.75	0.51
	Control		Challenge		Cogn. Inv.		Emot. Inv.		Real-World Diss.		IEQ Sum	
	Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
NBI	5	4.6–5.4	5.25	4.75–5.5	6.17	5.61–6.5	5.58	4.67–6.29	4.71	4.04–5.57	170.5	150–183
BI	5	4.25–5.4	5	4.5–5.5	6.06	5.36–6.64	5.75	4.83–6.29	5.07	4–5.82	169.5	148.8–180
BI-A	5.1	4.45–5.6	5.25	4.56–5.75	6.17	5.44–6.44	5.58	5–6.29	4.86	4–5.43	169.5	151–179
BI-V	5	4.4–5.75	5	4.56–5.25	5.89	5.33–6.33	5.58	4.88–6.13	5.14	4.04–5.86	168	147.2–182.5
BI-AV	5	4.4–5.8	5.13	4.5–5.5	6.22	5.58–6.56	5.5	4.88–6.25	4.71	4.14–5.71	169.5	152.2–184
	$F_{ATS}(3.59)$	p	$F_{ATS}(3.61)$	p	$F_{ATS}(3.51)$	p	$F_{ATS}(3.01)$	p	$F_{ATS}(3.82)$	p	$F_{ATS}(3.06)$	p
Main effect	0.71	0.56	2.07	0.09	0.65	0.61	0.32	0.81	1.48	0.21	0.25	0.86

Table 3: PX results also similarly positive across conditions.

VR Experience	Valence		Emotional Inv.		Challenge		Real World Diss.		IEQ Single Item		IEQ Sum Total	
	Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
with VR Experience	6	5—7	5.33	4.33—6	5	4.5—5.5	4.5	3.86—5.71	7	6—9	162.5	147—180.5
without VR Experience	6.5	6—7	6	5.33—6.54	5.25	4.75—5.5	5.14	4.54—5.89	8	7—9	175.5	163—183.25

Table 4: Prior VR experience had effect on some factors of PX.

Measures	Baseline (resting)		*NBI*		*BI*		*BI-A*		*BI-V*		*BI-AV*	
	Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*
Average HR	76.80	72.94–78.87	79.42	72.51–84.22	78.60	75.9–83.95	76.48	73.5–84.41	77.15	74.1–83.78	77.38	75.37–84.32
Variance in HR difference to Baseline	–	–	10.48	8.9–11.93	9.6	8.64–10.71	8.33	7.74–12.47	9.84	7.62–12.3	10.22	7.8–12.32

Table 5: HR measurements per condition (bpm, non-HMD player only).

therefore tested to see if participants' prior VR experience had an effect on their PX. For this between-participants comparison (with vs. without VR experience), we conducted Mann-Whitney U tests. The VR experience was analyzed for all participants (both roles, not separately) because of the subsample size. Data points were treated as independent as players within dyads experienced distinct gameplay (i.e., task, medium, physical engagement). The results indicated that having VR experience significantly affected the SAM valence scale, U=2185, p=0.04, d=-0.30, resulting in higher scores for players without such experience with the technology. Furthermore, there was a similar significant effect of VR experience on five subfactors of immersion: emotional involvement (U=1779, p<0.001, d=-0.69), challenge (U=2179, p=0.044, d=-0.34), real-world dissociation (U=1886.5, p=0.001, d=-0.54), the single-item immersion (U=2020.5, p=0.008, d=-0.54), and IEQ Sum Total (U=1995.5, p=0.007, d=-0.55). The descriptive statistics can be found in Table 4.

6.2 Heart Rate Analysis

This section reports analyses carried out only for the non-HMD player role. The descriptive data (including the baseline measurements prior to gameplay) are listed in Table 5; this includes both the average HR, and the variance in participants' difference to their baseline HR.

We conducted a one-way repeated measures ANOVA on the averaged HR data. There was no significant effect of condition on average HR. We then calculated participants' variance metric, meaning the difference of their HR measurement per condition respective to their own baseline measurement. Following guidelines by Wobbrock and Kay [94] for non-parametric test assumptions, we conducted a Friedman's ANOVA across the gameplay conditions; there was no significant effect of condition.

6.3 Learning Effects Across Playthroughs

To check for learning effects in the data, we conducted the following tests to determine how PX was affected across the five different playthroughs based on the order in which players experienced the game. For scope, we only report the main and post-hoc tests; the descriptive values and visualizations of the significant differences are presented in the supplementary materials.

Heart Rate Metrics. We conducted one-way repeated measures ANOVA with the within-subjects factor of playthrough order to explore learning effects on participant HR. There was a significant main order effect on average HR, $F(4,56)$=11.93; p=0.001, η_p^2=0.11. Bonferroni post-hoc tests showed that average HR was significantly higher in the first condition when compared to third, fourth, and fifth playthroughs, as well as the second in comparison to the last.

Player Experience Questionnaires. There was a significant order effect on presence, $F_{ATS}(2.74)$=9.72, p<0.001. Post-hoc, the first playthrough differed from all other playthroughs; later playthroughs displayed an increase in presence. Furthermore, a significant order effect on enjoyment was observed ($F_{ATS}(2.88)$=8.21, p<0.001). Here, the first playthrough also differed from all subsequent playthroughs, displaying a similar increase over time. Moreover, there was a significant order effect on the SAM valence scale, $F_{ATS}(3.22)$=8.43, p<0.001. Post-hoc comparisons revealed the same pattern: the first playthrough was significantly lower than all subsequent playthroughs. The SAM arousal scale also showed a significant order effect, $F_{ATS}(2.72)$=4.05, p=0.008. Here, though, the first condition again scored lower compared to the second and fourth playthroughs. The SAM dominance scale was significantly lower for the first condition compared to the last condition, $F_{ATS}(2.17)$=4.81, p=0.007.

There was a significant order effect on several immersion subfactors, displaying similar increases in scores over time. Cognitive involvement was significantly lower for the first playthrough than all others, as was the second for subsequent ones, $F_{ATS}(3.04)$=11.51, p<0.001. Emotional involvement was significantly lower for the first playthrough than when compared to the fourth or fifth playthrough, $F_{ATS}(2.67)$=6.36, p<0.001. There was a significant order effect for challenge, $F_{ATS}(3.21)$=3.99, p=0.006, which post-hoc comparisons indicated lay between the first playthrough compared to the third and fourth playthroughs. The order effect was also observed on control, $F_{ATS}(3.06)$= 4.82, p=0.002; the first playthrough was again significantly lower than all other playthroughs. For IEQ Sum Total, the first playthrough was again significantly lower than all subsequent playthroughs, $F_{ATS}(3.33)$= 7.83, p<0.001. For real-world dissociation there was also a significant order effect, $F_{ATS}(3.80)$=3.40, p<0.01 (disappeared after post-hoc), as well as an interaction order effect with player role, $F_{ATS}(3.80)$=5.31, p<0.001. The interaction effect shows that real-world dissociation tended to increase for the non-HMD player (first playthrough: Mdn=4.14, IQR=3.93–5.07; last: Mdn=5, IQR=4–5.57), while there was a slight decrease for the HMD player (first: Mdn=5.29, IQR=4–5.71; last: Mdn=4.43, IQR=4.14–5.79).

No significant effect was found for order effect and role on the single-item immersion score.

6.4 Summary of Quantitative Findings

Our quantitative results show that PX metrics (i.e., affective state, enjoyment, presence, and immersion) were comparably high for both players, with no significant difference between player roles or conditions and their interaction effects. Further, our physiological measures did not yield a significant difference across conditions.

We also observed novelty effects of VR for some subcategories of quantitative metrics (see Table 4), indicating higher values for the players without prior VR experience. Finally, learning effects results show how PX—for most of the factors—improved over five playthroughs (see supplementary materials).

6.5 Interview Findings

We report quotes based on the session number of the participant pair and add their role (HMD as h; non-HMD as nh) as subscript (e.g., P1_h for the HMD player of the first pair). Our thematic analysis reflects the quantitative results in that both roles were considered "*enjoyable*"-P5_{nh}, and "*engaging*"-P4_h and confirms the increase in positive PX over playthroughs. The results also support positive impacts of multiplayer interaction ("*I like the social aspect a lot [...] I have that feeling of achieving something together [...] so it was really cool to have to experience together and achieve that together.*"-P2_h). Additionally, we constructed four themes through an analysis of our inductive coding.

6.5.1 Theme 1: *Collaborative asymmetric VR experiences require time for adaptation because of the challenge of communication and coordination, yet this is also a key factor in their appeal.*

Players found communication essential to their enjoyment and to succeed in the game: "*Without [it], I think the game wouldn't be as much fun*"-P7_{nh} and "*it was necessary at all to play the game. And for me it made it more fun*"-P2_{nh}. Developing strategies for effective communication, however, took some time: "*I would* *also give some adjustment trial, but not to the virtual reality, but to the team players, [...] getting to know each other and how well they can perform how well they can play together, because I think this is requires time as well*"-P5$_h$. This also explains the increase in enjoyment over time: "*the engagement kind of rose with, with playing time [...] So I got more engaged because I knew what I was doing at some point*"-P7$_h$.

Participants attributed several factors to their inter-communication (and the attached learning curve). They described these as part of the experience's appeal: first, the dynamic characteristics of each gameplay depending on their co-player ("*it's another degree of input or another degree of output, which is pretty interesting, pretty dynamic because I'm pretty sure when I would be playing with someone else [...] the experience would be pretty different*"-P4$_h$); second, the reliance on another person ("*you have a team mate, who you have to have very good communication with, and that could also be a weakness because your team mate is entirely dependent. I mean, the person who's wearing the VR is dependent on the person who's on the screen [...] it works as both a strength and a weakness*"-P5$_{nh}$). Third, some HMD players attributed enjoyment to having a human co-player: "*if we imagine that the game itself is giving me these orders like consistently and without variation that would be I think less interesting than a real human giving me these orders and varying a strategy, varying the wording [...] If all these commands were coming from an onscreen user interface or like audio recordings, there were not dynamic or anything, it would be probably way more boring*"-P4$_h$. However, for some HMD players, their co-player as a voice entity reduced their immersion in VR: one mentioned it was "*a little bit difficult to focus on the voice*"-P5$_h$, while two mentioned it as a potential break in immersion, and a reminder of the artificiality of the game: "*this communication is something that is to me, not part of the game, but more part of the physical environment I'm in*"-P4$_h$ and "*I hear him [non-HMD role] through the room rather than through the HMD. So I was still connected to the real world so the immersion suffered a little bit, but enjoyment was increased because I had social interactions while playing*"-P3$_h$.

6.5.2 Theme 2: ***Asymmetry in interface and information can affect players' perception of agency, dominance, and control in varied ways. However, this is not necessarily a bad thing, and it induced feelings of interdependence.*** Our players demonstrated different opinions of which player role had more prominent agency, dominance, or control in the game. There was, however, a greater tendency towards the perception that non-HMD player had more control. Players that attributed greater agency/control to the HMD player based this on the explicit interaction with the game world: "*[As the non-HMD player] I never really felt like I was actively engaging in the game cause I had no control in terms of the immediate environment because the HMD player is like the intermediary. So because of that, I felt like I was maybe not as immersed as he was*"-P1$_{nh}$. However, a larger number of HMD players felt that agency/control had been transferred to the non-HMD player through the asymmetry of information: "*First, I thought I would have the more active role [as the HMD player] because I have to do the actions but in the end, I was more like the actor and he was commanding me. I basically just followed the instructions*"-P2$_h$, and in contrast "*[as the non-HMD player] I felt like Houston on the mission control, telling him what to do [laughs]*"-P3$_{nh}$. Moreover, one player made a distinction between dominance and importance: "*I don't think my role [as HMD player] was the dominant one but it was like the key role so the game can continue. I think this was my success if we die or not*"-P5$_h$.

Regardless, players' perception of this asymmetry in agency or control did not necessarily translate to inadequate PX. Some HMD players enjoyed that giving up control was accompanied by a more physically active role ("*I had a lot of fun playing that, like moving through the maze, watching out, talking in the right moment. I felt really like adventurous and cool [...] so I really liked that I couldn't see anything and like move through the maze*"-P7$_h$). Inversely, the greater control of the non-HMD player role was also perceived as "*engaging, but very stressful*"-P1$_{nh}$, "*more challenging and more unsettling*"-P7$_{nh}$. Some of the non-HMD players felt a strong sense of responsibility due to their game role, which not all of them appreciated: "*I would love to be HMD player, because [my non-HMD role] is so much responsibility*"-P6$_{nh}$.

Further, it led to interdependence, as expected based on our design and on theoretical work; both player roles reported this: "*I couldn't have moved through the maze without my partner telling me what to do [...] I needed the guidance*"-P7$_h$, "*I felt kind of bad sometimes because for most of the feedback variants, I was pretty excited and thus, we had to deal with a lot of the lasers [...] but I think we manage quite good*"-P7$_{nh}$, and "*It adds to the fun that you [as non-HMD player] are not the person who is controlling everything. you have to make sure that the person you're a good partner to, the communication has to be good enough for the game to be successful [...] everything is not in your control and it's fun*"-P5$_{nh}$. This perception of the interdependence of roles reflects on the different but complementary way that the game incorporated their strengths: "*adding skills to each other [...] you're playing kind of different games*"-P7$_h$.

6.5.3 Theme 3: ***There is high subjectivity in whether players notice biometric influence, especially HMD players. But when present and noticed, biometric influence affected PX (enjoyment, immersion, stress) and performance.*** Whether players noticed the biometric influence when it was present was highly subjective, which may explain the lack of effects found in the quantitative results. Particularly the HMD players, lacking direct visualization of the biometric feedback, reported that while their engagement was high, they did not notice an effect of the biometric influence condition: "*for me, there were no really differences in the conditions because I couldn't see lasers or anything*"-P3$_h$. However, the interviews provide insight into the effects of biometric influence when it *was* present and noticed.

Some players reported an effect—albeit indirect—on performance: "*when she was tense, there were more lasers coming and I was like dying all the time [...] And like when she was tension, for instance, you can say like left instead of right. Like she can mix the directions*"-P6$_h$. We also observed some players trying to calm down with breathing exercises during the game. More commonly, players (particularly non-HMD players) mentioned an effect on PX, largely immersion: "*[biometric feedback] definitely helped with the immersion just because it felt more real [...] you were aware that you as a being even though you're not in the virtual environment yourself, I'm just looking at the screen but you have—your body has an impact on the game*

[... I] was constantly aware that I am triggering the laser basically, which was exciting for me"-P7$_{nh}$. This biometric influence was often perceived as both immersive *and* stressful: "*Game engagement was very high. The experience was incredibly stressful [laughs]*"-P1$_{nh}$. For some players, effects on their co-player's emotions then, in turn, affected their own: "*when she gets more excited, I also felt more excited and I think it was more enjoyable both for me*"-P5$_{h}$.

6.5.4 Theme 4: ***Preferences for the modality of indicators of biometric feedback for the non-HMD player are highly subjective, whereas some HMD players surprisingly extracted biometric feedback information from the non-HMD player's voice.*** There was no consensus among non-HMD players about the best kind of modality for biometric feedback. Some found the biometric feedback useful for self-regulation ("*help[ed] me to focus on keeping my breath low and trying to be not that excited*"-P2$_{nh}$). Many of the others, however, either did not notice the feedback or found it had an opposite effect on them: "*they made it worse because I was aware that I'm being excited, which made me more excited [...] it was not helpful, but it was interesting*"-P7$_{nh}$. For each of the different biofeedback modalities in our study conditions, a participant preferred it or disliked it.

However, a noteworthy aspect of the interviews is that some of the *HMD players* also remarked on biofeedback they perceived. They interpreted their non-HMD player co-player's emotional state from their voice, and reacted accordingly: "*when my coplayer panicking, I can hear it from his sound. So it will affect me at a time. It's like, okay, I have to go, go to this path faster*"-P1$_{h}$. They attributed this specifically to the tone of voice: "*I could sense her excitement from the voice and the voice of tone basically*"-P5$_{h}$.

7 DISCUSSION

In the following, we discuss the key takeaways about integrating a non-HMD player into a VR game, the impact of interdependence on player agency (including that neither high nor low agency is necessarily good or bad in our scenario), and how biometric interdependence can affect players' immersion and experienced stress.

Integrating a Non-HMD Player in a VR Game. Our VR game included the non-HMD player through some design choices with which we gave the non-HMD player a way to impact the VR world, and perceive results thereof. This impact was largely implemented by facilitating SI between the HMD and non-HMD player (i.e., through difference in information and difference in player capability) so that they have to communicate and strategize together. Further, we introduced and explored a novel type of interdependence—biometric—between the players. Hereby, non-HMD player's HR affected game difficulty (triggering lasers that the HMD player had to avoid). Our psychometric results suggest that this overall design worked well, successfully extending the findings of previous work [34, 35] to integrate a non-HMD player into a VR experience so that *both* players achieve high enjoyment and presence without a significant difference between player roles. All psychometric PX factors were not significantly different between player roles: presence, enjoyment, immersion, and affective states were highly rated for both players. While biometric interdependence did not impact quantitative PX (possibly because of a subtle implementation and high subjectivity in players' perception thereof), the thematic analysis shows it has promise as a game mechanic to impact immersion and stress (as we will discuss below). Thus, despite the difference in display (HMD vs. non-HMD) and task (executing actions vs. guiding them), both players shared a strong sense of engagement and presence. Further, the interviews highlighted the communication and interdependence between players as contributing to PX regardless of players' experienced high or low agency.

Further, our results indicate positive social interaction between both player roles, supporting our psychometric findings. We argue that this stems from both interdependences featured in our game. The difference in information in particular forced players to communicate and strategize. In future work, it would be interesting to explore effects of the game on players' experience of loneliness (as explored by Liszio et al. [51]), and their relationship between each other (as reported in an asymmetric game by Zhou et al. [96]).

We further observed novelty effects of VR, which should be considered in asymmetric VR games but also VR research in general: whether participants had prior VR experience affected their immersion and affective state. With low to medium effects sizes, a lack of VR experience elicited higher immersion and valence (arousal, dominance, presence, and enjoyment were unaffected). We speculate that these effects may occur because novices can be more distracted and potentially overwhelmed by the first exposure to VR [49]. This may be of interest to VR researchers, and motivates future research into exploring how to design interaction in VR to remain immersive and engaging over time, and even for players already familiar with the medium. Nevertheless, we note that the exploration of this dynamic in VR research is rather limited; it largely reported only as demographic information of participants and not investigated further. One of the few works that explore longitudinal VR usage found that while novelty wore off, immersion did not [72]; our results reflect and build on this finding for a shorter timeframe.

Interdependence Affects Agency—But Perceptions Thereof Vary. The SI in our game carried the risk of either player experiencing a lack of agency because the power between their roles is imbalanced. We had suspected that the HMD player might experience lower agency: synchronizing with the non-HMD player's instructions required concentration, timing, and physical interaction, yet they largely virtually enacted the non-HMD player's instructions, and had access to less information about the game world (i.e., the approach of lasers or trap layout). Nevertheless, most HMD players reported a positive PX and many saw their role as key to the game.

Inversely, while many non-HMD players described their experience in terms of high agency or control, this was not always an inherently positive factor: high agency was also experienced as strong responsibility towards their co-player and the outcome of the game, sometimes resulting in stress. This stress was mentioned in the context of both types of interdependence, but particularly often in the context of BI (which we discuss in more detail below). Overall, we found that players' perception of agency was strongly affected by the asymmetric design and resulting interdependence—however, "low" vs. "high" agency is not inherently bad or good, respectively. Therefore, our work contrasts with the identified design factors of Gugenheimer et al. [34], which suggest to create an equal "power distribution" between the players, especially for collaborative gameplay. Based on our findings, we argue that uneven power distributions can also create enjoyable gameplay experiences, and in fact, individual playing motivations may determine whether players prefer what they perceive as high or low agency roles. This could in part be linked to which role players feel more competent at and comfortable with, relating to players' need for competence [19, 77]; alternatively it could be linked to players' familiarity with either interface [78]. Finally, we note that our findings reflect on work by Benford et al. [4] on the value of uncomfortable interactions in human-computer interaction. Games constitute suitable scenarios for such interactions, as they often provide an environment for negative or unusual emotions [8] on purpose, and so power imbalances in games make sense as a mechanic to explore for the creation of engaging experiences.

Biometric Interdependence as a Game Mechanic. While BI did not impact quantitative measures, the interviews indicate high variance in whether participants perceived it within the game. There were no significant effects of biometric feedback (or the modality of its representation) on PX. We attribute this to a too subtle implementation of BI, and perhaps also a matter of requiring prolonged exposure to the stimuli (see limitations below). The results are perhaps unsurprising for the HMD players who had no biometric feedback; a less subtle implementation (e.g., increasing lasers and giving HMD players information about their co-player's biometric status) could yield different results. For players who did perceive the BI, however, interviews and gameplay observations revealed several interesting ways in which it did impact PX. Players that felt more immersed because their physiological state had an impact on the game world can be interpreted as a clear example of agency [61]: agency increased as they had a way to noticeably affect the game world, thus positively impacting their immersion. Further, in many cases, players reported a mirroring effect due to BI, as HMD players noticed the non-HMD player's increased excitement and/or stress. Interestingly, however, in addition to the designed biofeedback provided to the non-HMD player, some HMD players *also* perceived biofeedback about their co-player based on their tone of voice. Voice communication (specifically, hearing the voice of a bystander) has been reported as a factor of potential comfort for VR users [75], and has been used as a VR input mechanism [46] in prior work, but otherwise is largely unexplored in VR research.

Finally, while our results show promise for employing BI as a game mechanic in VR, we note that learning effects for HR must be considered. In our study, average HR dropped significantly after two playthroughs, indicating that participants were calmer after having played the game twice. This finding informs future research: it motivates letting players experience VR games twice prior to testing an experiment (if possible), as early measurements may be biased by an increased HR. Further, we conclude that more longitudinal data is necessary to measure and explore VR experiences (of which there are few examples [60, 71, 72]). This may be particularly relevant for game mechanics that include physiological measures and BI.

7.1 Limitations

We note that physiological measures can introduce some limitations. As stated above, there was a learning effect for HR measurements after two playthroughs. Future experiments should adjust their baseline dynamically or re-sample it after longer exposure. We also cannot rule out that some participants may have already been excited during the baseline measurement, introducing higher arousal as bias. Additionally, minor changes to physiological data could have been missed due to the sampling rate of the Empatica E4 device. Largely, however, we assume that the design of the biometric feedback was simply too subtle to make a strong impact—especially for this time frame of stimuli exposure. We will have to re-iterate on the design to refine this.

As roles were not switched within pairs, we cannot prove that PX would have stayed the same between roles in a within-participants design. While our study works as a proof-of-concept that a VR game can be designed to create comparable PX even for a non-HMD player, future work will have to explore the differences in these kinds of player roles in more detail (for example, when players take turns with the roles, with different degrees of directional dependence, or across different game genres).

The distribution of traps did not change with the different playthroughs (i.e., it differed only by level). As such, players could learn the path through each level by heart and then expend less effort in navigating the grid (or instructing its navigation). This could have partially induced the learning effects. Moreover, our participants pairs knew each other, which could also introduce bias. However, this adds to external validity as this kind of local game setup would very likely be played by players that are familiar with each other.

In terms of methodology, we employed a triangulation-convergence approach to leverage advantages of both quantitative (e.g., generalisation) and qualitative findings (e.g., deep details) [18]. However, combining the different types of data is difficult and can introduce bias. Further, we conducted the interviews as an extension of the study, and then conducted the analysis afterwards. While no new codes emerged in the final analysis session, it is possible that new codes would have occurred for a larger sample (cf. [12] on saturation in thematic analysis). Finally, we note as a statement on reflexivity [5, 65] that the two authors who conducted the thematic analysis have a cognitive systems and computer science background, respectively. As both have prior experience of varying duration with VR and VR games, this may have introduced bias into the theme development phase.

8 SUMMARIZING CONCLUSION

This work introduced an asymmetric game design that integrates a non-HMD player into a VR experience. Our goal was to increase PX for *both* the HMD and non-HMD player via SI and BI. In a user study (N=30), we explored effects of resulting interdependences on PX and found that SI induces a comparably high amount of enjoyment, presence, immersion, and affective state for both player roles. BI and audiovisual indicators thereof were subject to learning effects and need more longitudinal data for a comprehensive analysis of its impact on PX. However, the qualitative findings point towards effects on both players' experience in terms of immersion and stress. Moreover, our interviews show that interdependences resulting from asymmetric game design affect player agency—yet also, neither high nor low agency is inherently perceived as good or bad in our game prototype and study.

Our results have shown that it is feasible to integrate a non-HMD player into a VR experience and achieve comparable PX levels to the HMD player. Further, we discussed design implications for future asymmetric game designs, by showcasing the potential of imbalanced strategic asymmetry in games, emphasizing the importance of agency for interdependence, and demonstrating the use of biometric asymmetry as a game mechanic. Our work can thus inform future VR developers as they create immersive asymmetric VR games and experiences, to create multiplayer engagement and shared social environments across interfaces.

ACKNOWLEDGMENTS

The authors of this paper have partly been funded by the German Federal Ministry of Education and Research (BMBF) as well as the German Research Foundation (DFG), the European Union's Horizon 2020 research and innovation program and the Federal Ministry for Economic Affairs and Energy (BMWi). This work has been supported by the DFG project "Empirical Assessment of Presence and Immersion in Augmented and Virtual Realities (RU 1605/4-1)". This work was also supported by Lennart Nacke's NSERC Discovery Grant 2018-06576, the Canada Foundation for Innovation John R. Evans Leaders Fund 35819 "SURGE—The Stratford User Research and Gameful Experiences Lab," the NSERC CREATE SWaGUR grant, Mitacs Accelerate, and CIHR.

REFERENCES

[1] Mike Ambinder. 2011. Biofeedback in gameplay: How valve measures physiology to enhance gaming experience. In *Game Developers Conference*, Vol. 22. Game Developers Conference, San Francisco,USA, 27–33.

[2] Ambinder, Mike. 2020. Brain-Computer Interfaces: One Possible Future for How We Play. https://www.youtube.com/watch?v=Qhj3C1H5JWo Game Developers Conference, San Francisco, USA.

[3] Martin Bauer, Timo Heiber, Gerd Kortuem, and Zary Segall. 1998. A Collaborative Wearable System with Remote Sensing. In *Proceedings of the 2Nd IEEE International Symposium on Wearable Computers (ISWC '98)*. IEEE Computer Society, Washington, DC, USA, 10–. https://doi.org/10.1109/ISWC.1998.729524

[4] Steve Benford, Chris Greenhalgh, Gabriella Giannachi, Brendan Walker, Joe Marshall, and Tom Rodden. 2012. Uncomfortable Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 2005–2014. https://doi.org/10.1145/2207676.2208347

[5] Roni Berger. 2015. Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative research* 15, 2 (2015), 219–234. https://doi.org/10.1177/1468794112468475

[6] Frank Biocca and Ben Delaney. 1995. *Immersive Virtual Reality Technology*. L. Erlbaum Associates Inc., USA, 57–124.

[7] Daniel Boland and Mark McGill. 2015. Lost in the rift: engaging with mixed reality. *XRDS: Crossroads, The ACM Magazine for Students* 22, 1 (2015), 40–45. https://doi.org/10.1145/2810046

[8] Julia Ayumi Bopp, Elisa D. Mekler, and Klaus Opwis. 2016. Negative Emotion, Positive Experience? Emotionally Moving Moments in Digital Games. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2996–3006. https://doi.org/10.1145/2858036.2858227

[9] Stéphane Bouchard, François Bernier, Éric Boivin, Brian Morin, and Geneviève Robillard. 2012. Using biofeedback while immersed in a stressful videogame increases the effectiveness of stress management skills in soldiers. *PloS one* 7, 4 (2012), e36169. https://doi.org/10.1371/journal.pone.0036169

[10] Margaret M Bradley and Peter J Lang. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry* 25, 1 (1994), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

[11] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[12] Virginia Braun and Victoria Clarke. 2019. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health* 0, 0 (Dec. 2019), 1–16. https://doi.org/10.1080/2159676x.2019.1704846

[13] Virginia Braun and Victoria Clarke. 2020. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology* 0, 0 (2020), 1–25. https://doi.org/10.1080/14780887.2020.1769238

[14] Liwei Chan and Kouta Minamizawa. 2017. FrontFace: Facilitating Communication Between HMD Users and Outsiders Using Front-facing-screen HMDs. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vienna, Austria) *(MobileHCI '17)*. ACM, New York, NY, USA, Article 22, 5 pages. https://doi.org/10.1145/3098279.3098548

[15] Hao Chen, Arindam Dey, Mark Billinghurst, and Robert W. Lindeman. 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* (Brisbane, Queensland, Australia) *(OZCHI '17)*. ACM, New York, NY, USA, 108–116. https://doi.org/10.1145/3152771.3152783

[16] Lung-Pan Cheng, Patrick Lühne, Pedro Lopes, Christoph Sterz, and Patrick Baudisch. 2014. Haptic Turk: A Motion Platform Based on People. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3463–3472. https://doi.org/10.1145/2556288.2557101

[17] Valve Corporation. 2018. *SteamVR v1.2.3*. Microsoft Windows. https://github.com/ValveSoftware/steamvr_unity_plugin/releases/download/1.2.3/SteamVR.Plugin.unitypackage

[18] John W Creswell and Vicki L Plano Clark. 2017. *Designing and conducting mixed methods research*. Sage publications, Thousand Oaks, California, United States.

[19] Edward L Deci and Richard M Ryan. 2011. Self-determination theory. *Handbook of theories of social psychology* 1 (2011), 416–433.

[20] Andrew Dekker and Erik Champion. 2007. Please Biofeed the Zombies: Enhancing the Gameplay and Display of a Horror Game Using Biofeedback. In *DiGRA'07-Proceedings of the 2007 DiGRA International Conference: Situated Play*, Vol. 39. Springer, Tokyo, Japan, 43–52. https://doi.org/10.25917/5d1443e8af4a0

[21] Arindam Dey, Hao Chen, Mark Billinghurst, and Robert W. Lindeman. 2018. Effects of Manipulating Physiological Feedback in Immersive Virtual Environments. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play* (Melbourne, VIC, Australia) *(CHI PLAY '18)*. ACM, New York, NY, USA, 101–111. https://doi.org/10.1145/3242671.3242676

[22] Arindam Dey, Thammathip Piumsomboon, Youngho Lee, and Mark Billinghurst. 2017. Effects of Sharing Physiological States of Players in a Collaborative Virtual Reality Gameplay. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 4045–4056. https://doi.org/10.1145/3025453.3026028

[23] Anders Drachen, Lennart E. Nacke, Georgios Yannakakis, and Anja Lee Pedersen. 2010. Correlation between Heart Rate, Electrodermal Activity and Player Experience in First-Person Shooter Games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games* (Los Angeles, California) *(Sandbox '10)*. Association for Computing Machinery, New York, NY, USA, 49–54. https://doi.org/10.1145/1836135.1836143

[24] Inger Ekman, Guillaume Chanel, Simo Järvelä, J.Matias Kivikangas, Mikko Salminen, and Niklas Ravaja. 2012. Social Interaction in Games: Measuring Physiological Linkage and Social Presence. *Simulation & Gaming* 43, 3 (2012), 321–338. https://doi.org/10.1177/1046878111422121 arXiv:https://doi.org/10.1177/1046878111422121

[25] Empatica. 2016. Empatica E4 wristband. https://www.empatica.com/en-int/research/e4/

[26] Flying Mollusk. 2015. *Nevermind*. Game [Windows]. https://store.steampowered.com/app/342260/Nevermind/ Flying Mollusk, Los Angeles, CA, United States.

[27] ForeignVR. 2016. *Ruckus Ridge VR Party*. Game [VR]. https://store.steampowered.com/app/443800/Ruckus_Ridge_VR_Party/ ForeignVR, Palo Alto, USA.

[28] Theston E. Fox. 2019. VRTK - Virtual Reality Toolkit. https://www.vrtk.io/

[29] Taichi Furukawa, Daisuke Yamamoto, Moe Sugawa, Roshan Peiris, and Kouta Minamizawa. 2019. TeleSight: Enabling Asymmetric Collaboration in VR between HMD User and Non-HMD Users. In *ACM SIGGRAPH 2019 Emerging Technologies*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3305367.3335040

[30] EA Games. 2002. *BattleField 1942*. Game [PC]. https://www.ea.com/games/battlefield/battlefield-1942

[31] S2 Games. 2003. *Savage: The Battle for Newerth*. Game [PC]. https://www.newerth.com/

[32] Steel Crate Games. 2015. *Keep Talking And Nobody Explodes*. Game [VR, PC]. Steel Crate Games, Ottawa, Canada. https://keeptalkinggame.com/.

[33] Kiel Gilleade, Alan Dix, and Jen Allanson. 2005. Affective videogames and modes of affective gaming: assist me, challenge me, emote me. *DiGRA 2005: Changing Views–Worlds in Play*. 8 (2005), 547–554.

[34] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683

[35] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. 2018. FaceDisplay: Towards Asymmetric Multi-User Interaction for Nomadic Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 54, 13 pages. https://doi.org/10.1145/3173574.3173628

[36] John Harris and Mark Hancock. 2019. To Asymmetry and Beyond! Improving Social Connectedness by Increasing Designed Interdependence in Cooperative Play. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300239

[37] John Harris, Mark Hancock, and Stacey Scott. 2014. "beam Me 'round, Scotty!": Exploring the Effect of Interdependence in Asymmetric Cooperative Games. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play* (Toronto, Ontario, Canada) *(CHI PLAY '14)*. Association for Computing Machinery, New York, NY, USA, 417–418. https://doi.org/10.1145/2658537.2661311

[38] John Harris, Mark Hancock, and Stacey D. Scott. 2016. Leveraging Asymmetries in Multiplayer Games: Investigating Design Elements of Interdependent Play. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 350–361. https://doi.org/10.1145/2967934.2968113

[39] Samory Houzangbe, Olivier Christmann, Geoffrey Gorisse, and Simon Richir. 2018. Fear as a Biofeedback Game Mechanic in Virtual Reality: Effects on Engagement and Perceived Usability. In *Proceedings of the 13th International Conference on the Foundations of Digital Games* (Malmö, Sweden) *(FDG '18)*. Association for Computing Machinery, New York, NY, USA, Article 12, 6 pages. https://doi.org/10.1145/3235765.3235787

[40] Samory Houzangbe, Olivier Christmann, Geoffrey Gorisse, and Simon Richir. 2020. Effects of voluntary heart rate control on user engagement and agency in a virtual reality game. *Virtual Reality* 24 (2020), 665–681. https://doi.org/10.1007/s10055-020-00429-7

[41] HTC and Valve Corporation. 2016. HTC Vive. https://www.htcvive.com

[42] Akira Ishii, Masaya Tsuruta, Ippei Suzuki, Shuta Nakamae, Junichi Suzuki, and Yoichi Ochiai. 2019. Let Your World Open: CAVE-based Visualization Methods of Public Virtual Reality Towards a Shareable VR Experience. In *Proceedings of the 10th Augmented Human International Conference 2019* (Reims, France) *(AH2019)*. ACM, New York, NY, USA, Article 33, 8 pages. https://doi.org/10.1145/3311823.3311860

[43] Charlene Jennett, Anna L Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. 2008. Measuring and defining the experience of immersion in games. *International journal of human-computer studies* 66, 9 (2008), 641–661. https://doi.org/10.1016/j.ijhcs.2008.04.004

[44] Kisung Jeong, Jinmo Kim, Mingyu Kim, Jiwon Lee, and Chanhun Kim. 2020. Asymmetric interface: user interface of asymmetric virtual reality for new presence and experience. *Symmetry* 12, 1 (2020), 53. https://doi.org/10.3390/sym12010053

[45] Calestous Juma. 2016. *Innovation and its enemies: Why people resist new technologies.* Oxford University Press, Oxford, England.

[46] Agneya A Kerure and Jason Freeman. 2018. Audio Source Localization as an Input to Virtual Reality Environments. In *Audio Engineering Society Convention 144*. Audio Engineering Society, Audio Engineering Society, New York, NY, USA. http://www.aes.org/e-lib/browse.cfm?elib=19415

[47] Andrey Krekhov, Daniel Preuß, Sebastian Cmentowski, and Jens Krüger. 2020. Silhouette Games: An Interactive One-Way Mirror Approach to Watching Players in VR. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Canada) *(CHI PLAY '20)*. Association for Computing Machinery, New York, NY, USA, 561–571. https://doi.org/10.1145/3410404.3414247

[48] Kai Kuikkaniemi, Toni Laitinen, Marko Turpeinen, Timo Saari, Ilkka Kosunen, and Niklas Ravaja. 2010. The Influence of Implicit and Explicit Biofeedback in First-Person Shooter Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 859–868. https://doi.org/10.1145/1753326.1753453

[49] Chen-Lin C. Kulik and James A. Kulik. 1991. Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior* 7, 1 (1991), 75 – 94. https://doi.org/10.1016/0747-5632(91)90030-5

[50] Jiwon Lee, Mingyu Kim, and Jinmo Kim. 2020. RoleVR: Multi-experience in immersive virtual reality between co-located HMD and non-HMD users. *Multimedia Tools and Applications* 79, 1-2 (2020), 979–1005. https://doi.org/10.1007/s11042-019-08220-w

[51] Stefan Liszio, Katharina Emmerich, and Maic Masuch. 2017. The Influence of Social Entities in Virtual Reality Games on Player Experience and Immersion. In *Proceedings of the 12th International Conference on the Foundations of Digital Games* (Hyannis, Massachusetts) *(FDG '17)*. Association for Computing Machinery, New York, NY, USA, Article 35, 10 pages. https://doi.org/10.1145/3102071.3102086

[52] Adam Lobel, Marientina Gotsis, Erin Reynolds, Michael Annetta, Rutger C.M.E. Engels, and Isabela Granic. 2016. Designing and Utilizing Biofeedback Games for Emotion Regulation: The Case of Nevermind. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. ACM, New York, NY, USA, 1945–1951. https://doi.org/10.1145/2851581.2892521

[53] Christian Mai, Sarah Aragon Bartsch, and Lea Rieger. 2018. Evaluating Shared Surfaces for Co-Located Mixed-Presence Collaboration. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia* (Cairo, Egypt) *(MUM 2018)*. ACM, New York, NY, USA, 1–5. https://doi.org/10.1145/3282894.3282910

[54] Regan Mandryk, Margaret Atkins, and Kori Inkpen. 2006. A continuous and objective evaluation of emotional experience with interactive play environments, In Proceedings of the SIGCHI conference on Human Factors in computing systems. *Conference on Human Factors in Computing Systems - Proceedings* 2, 1027–1036. https://doi.org/10.1145/1124772.1124926

[55] Regan L. Mandryk. 2008. *Physiological measures for game evaluation.* Isbister K. Schaffer N., (Eds). Morgan Kaufmann, San Francisco, USA, Chapter 14, 207–235.

[56] Regan L. Mandryk and Stella M. Atkins. 2007. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International journal of human-computer studies* 65, 4 (2007), 329–347. https://doi.org/10.1016/j.ijhcs.2006.11.011

[57] Regan L. Mandryk, Shane Dielschneider, Michael R. Kalyn, Christopher P. Bertram, Michael Gaetz, Andre Doucette, Brett A. Taylor, Alison Pritchard Orr, and Kathy Keiver. 2013. Games as Neurofeedback Training for Children with FASD. In *Proceedings of the 12th International Conference on Interaction Design and Children* (New York, New York, USA) *(IDC '13)*. Association for Computing Machinery, New York, NY, USA, 165–172. https://doi.org/10.1145/2485760.2485762

[58] Regan L. Mandryk, Kori M. Inkpen, and Thomas W. Calvert. 2006. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & information technology* 25, 2 (2006), 141–158. https://doi.org/10.1080/01449290500331156

[59] Regan L. Mandryk and Lennart E. Nacke. 2016. *Biometrics in a Data Driven World: Trends, Technologies, and Challenges* (1st ed.). Taylor & Francis/CRC Press, USA, Chapter Biometrics in Gaming and Entertainment Technologies, 191–224. https://doi.org/10.1201/9781315317083-7

[60] Fares Moustafa and Anthony Steed. 2018. A Longitudinal Study of Small Group Interaction in Social Virtual Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 22, 10 pages. https://doi.org/10.1145/3281505.3281527

[61] Janet Horowitz Murray. 2017. *Hamlet on the holodeck: The future of narrative in cyberspace.* MIT press, London, England.

[62] Joschka Mütterlein and Thomas Hess. 2017. Immersion, presence, interactivity: towards a joint understanding of factors influencing virtual reality acceptance and use. *Twenty-third Americas Conference on Information Systems* 1 (2017), 117–127.

[63] Lennart E. Nacke, Michael Kalyn, Calvin Lough, and Regan L. Mandryk. 2011. Biofeedback Game Design: Using Direct and Indirect Physiological Control to Enhance Game Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 103–112. https://doi.org/10.1145/1978942.1978958

[64] Jeff Nagy and Fred Turner. 2019. The Selling of Virtual Reality: Novelty and Continuity in the Cultural Integration of Technology. *Communication, Culture and Critique* 12, 4 (Nov. 2019), 535–552. https://doi.org/10.1093/ccc/tcz038

[65] Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand, and Jon Howard Raphael. 2012. No room for reflexivity? Critical reflections following a systematic review of qualitative research. *Journal of Health Psychology* 17, 6 (2012), 866–885. https://doi.org/10.1177/1359105311427615

[66] Kimihiro Noguchi, Yulia R Gel, Edgar Brunner, and Frank Konietschke. 2012. nparLD: an R software package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal of Statistical Software* 50, 12 (2012). https://doi.org/10.18637/jss.v050.i12

[67] Pedro A Nogueira, Vasco Torres, Rui Rodrigues, Eugénio Oliveira, and Lennart E Nacke. 2016. Vanishing scares: biofeedback modulation of affective player experiences in a procedural horror game. *Journal on Multimodal User Interfaces* 10, 1 (2016), 31–62. https://doi.org/10.1007/s12193-015-0208-1

[68] Odd Raven Studios. 2019. *Carly and the Reaperman—Escape from the Underworld.* Game [HTC Vive, Windows]. Odd Raven Studios, Stockholm, Sweden.

[69] Team Panoptes. 2019. *Panoptic.* Game [VR, PC]. https://panopticgame.com/ Team Panoptes, Brussels, Belgium.

[70] Rakesh Patibanda, Florian 'Floyd' Mueller, Matevz Leskovsek, and Jonathan Duckworth. 2017. Life Tree: Understanding the Design of Breathing Exercise Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17)*. Association for Computing Machinery, New York, NY, USA, 19–31. https://doi.org/10.1145/3116595.3116621

[71] J. Porter, K. Kohm, and A. Robb. 2020. Lingering Effects Associated with Virtual Reality: An Analysis Based on Consumer Discussions Over Time. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, Atlanta, GA, USA, 806–807. https://doi.org/10.1109/VRW50115.2020.00254

[72] John Porter and Andrew Robb. 2019. An Analysis of Longitudinal Trends in Consumer Thoughts on Presence and Simulator Sickness in VR Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 277–285. https://doi.org/10.1145/3311350.3347159

[73] Red Meat Games. 2018. *Bring to Light*. Game [VR],[PC]. https://store.steampowered.com/app/636720/Bring_to_Light/ Red Meat Games, Halifax, NS, Canada.

[74] Resolution Games. 2019. *Acron: Attack of the Squirrels*. Game [Oculus Quest, iOS / Android]. https://www.resolutiongames.com/acron Resolution Games, Stockholm, Sweden.

[75] Katja Rogers, Jana Funke, Julian Frommel, Sven Stamm, and Michael Weber. 2019. Exploring Interaction Fidelity in Virtual Reality: Object Manipulation and Whole-Body Movements. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300644

[76] Katja Rogers, Giovanni Ribeiro, Rina R. Wehbe, Michael Weber, and Lennart E. Nacke. 2018. Vanishing Importance: Studying Immersive Effects of Game Audio Perception on Player Experiences in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173902

[77] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68. https://doi.org/10.1037110003-066X.55.1.68

[78] Pejman Sajjadi, Edgar Omar Cebolledo Gutierrez, Sandra Trullemans, and Olga De Troyer. 2014. Maze Commander: A Collaborative Asynchronous Game Using the Oculus Rift & the Sifteo Cubes. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play* (Toronto, Ontario, Canada) *(CHI PLAY '14)*. Association for Computing Machinery, New York, NY, USA, 227–236. https://doi.org/10.1145/2658537.2658690

[79] Corina Sas and Rohit Chopra. 2015. MeditAid: a wearable adaptive neurofeedback-based system for training mindfulness state. *Personal and Ubiquitous Computing* 19, 7 (2015), 1169–1182. https://doi.org/10.1007/s00779-015-0870-z

[80] Jeff Sinclair, Philip Hingston, and Martin Masek. 2007. Considerations for the Design of Exergames. In *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia* (Perth, Australia) *(GRAPHITE '07)*. Association for Computing Machinery, New York, NY, USA, 289–295. https://doi.org/10.1145/1321261.1321313

[81] Jeff Sinclair, Philip Hingston, and Martin Masek. 2009. Exergame Development Using the Dual Flow Model. In *Proceedings of the Sixth Australasian Conference on Interactive Entertainment* (Sydney, Australia) *(IE '09)*. Association for Computing Machinery, New York, NY, USA, Article 11, 7 pages. https://doi.org/10.1145/1746050.1746061

[82] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 3, 2 (1994), 130–144. https://doi.org/10.1162/pres.1994.3.2.130

[83] Michael Smilovitch and Richard Lachman. 2019. BirdQuestVR: A Cross-Platform Asymmetric Communication Game. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Barcelona, Spain) *(CHI PLAY '19 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 307–313. https://doi.org/10.1145/3341215.3358246

[84] Misha Sra, Xuhai Xu, and Pattie Maes. 2018. *BreathVR: Leveraging Breathing as a Directly Controlled Interface for Virtual Reality Games*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173914

[85] Aaron Stafford, Wayne Piekarski, and Bruce Thomas. 2006. Implementation of God-like Interaction Techniques for Supporting Collaboration Between Outdoor AR and Indoor Tabletop Users. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '06)*. IEEE Computer Society, Washington, DC, USA, 165–172. https://doi.org/10.1109/ISMAR.2006.297809

[86] Aaron Stafford, Bruce H. Thomas, and Wayne Piekarski. 2008. Efficiency of Techniques for Mixed-space Collaborative Navigation. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '08)*. IEEE Computer Society, Washington, DC, USA, 181–182. https://doi.org/10.1109/ISMAR.2008.4637356

[87] Frank Steinicke. 2016. *Being Really Virtual: Immersive Natives and the Future of Virtual Reality* (1st ed.). Springer Publishing Company, Incorporated, New York, NY, USA.

[88] Team Future LLC. 2016. *Black Hat Cooperative*. Game [VR, PC]. https://store.steampowered.com/app/503100/Black_Hat_Cooperative/ Team Future LLC, Bangor, ME, USA.

[89] Unity Technologies. 2019. Unity Real-Time Development Platform. https://unity3d.com/get-unity/download/archive

[90] Valve. 2009. *Left 4 Dead 2*. Game [PC]. Valve Corporation, Bellevue, Washington, United States.

[91] Valve. 2010. *Alien Swarm*. Game [PC]. Valve Corporation, Bellevue, Washington, United States.

[92] Amy Voida and Saul Greenberg. 2009. Wii All Play: The Console Game As a Computational Meeting Place. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. ACM, New York, NY, USA, 1559–1568. https://doi.org/10.1145/1518701.1518940

[93] Wouter Walmink, Danielle Wilde, and Florian 'Floyd' Mueller. 2014. Displaying Heart Rate Data on a Bicycle Helmet to Support Social Exertion Experiences. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction* (Munich, Germany) *(TEI '14)*. Association for Computing Machinery, New York, NY, USA, 97–104. https://doi.org/10.1145/2540930.2540970

[94] Jacob O Wobbrock and Matthew Kay. 2016. Nonparametric statistics in human–computer interaction. In *Modern statistical methods for HCI*. Springer, New York, NY, USA, 135–170.

[95] Keng-Ta Yang, Chiu-Hsuan Wang, and Liwei Chan. 2018. ShareSpace: Facilitating Shared Use of the Physical Space by Both VR Head-Mounted Display and External Users. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. ACM, New York, NY, USA, 499–509. https://doi.org/10.1145/3242587.3242630

[96] Zhuoming Zhou, Elena Márquez Segura, Jared Duval, Michael John, and Katherine Isbister. 2019. Astaire: A Collaborative Mixed Reality Dance Game for Collocated Players. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 5–18. https://doi.org/10.1145/3311350.3347152

11.4 PLAYING WITH FRIENDS OR STRANGERS? THE EFFECTS OF FAMILIARITY BETWEEN PLAYERS IN AN ASYMMETRIC MULTIPLAYER VIRTUAL REALITY GAME

[CorePub9] **Sukran Karaosmanoglu**, Tom Schmolzi, and Frank Steinicke. Playing with friends or strangers? the effects of familiarity between players in an asymmetric multiplayer virtual reality game. In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY Companion '23), page 76–82, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9798400700293. DOI 10.1145/3573382.3616079. URL https://doi.org/10.1145/3573382.3616079

Supplementary materials for this article are available here: https://doi.org/10.1145/3573382.3616079.

Playing with Friends or Strangers? The Effects of Familiarity between Players in an Asymmetric Multiplayer Virtual Reality Game

Sukran Karaosmanoglu
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
sukran.karaosmanoglu@uni-hamburg.de

Tom Schmolzi
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
tom.schmolzi@uni-hamburg.de

Frank Steinicke
Human-Computer Interaction, Universität Hamburg
Hamburg, Germany
frank.steinicke@uni-hamburg.de

ABSTRACT

Many people play games to socialize. Previous studies have shown that asymmetric virtual reality (VR) games —games that utilize and consider differences of various kinds— can improve users' social and player experience. However, the impact of existing social factors on users' experiences is still unknown in asymmetric games. In this paper, we designed and implemented an asymmetric game in which both players are in VR. In the game, players had to exchange asymmetric information to complete tasks, resulting in a strategic interdependence. We studied how familiarity (friends vs. strangers) affects social and player experience, as well as game performance. In our preliminary between-participants study (N=14), we did not find significant differences between the friend and stranger teams in social and player experience, or game performance. We discuss how asymmetries can be used to create social VR games, and how and why familiarity does (not) affect the experience of players.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; **Virtual reality**; • **Software and its engineering → Interactive games**.

KEYWORDS

asymmetric games, virtual reality, familiarity, player experience, social experience, game performance

ACM Reference Format:
Sukran Karaosmanoglu, Tom Schmolzi, and Frank Steinicke. 2023. Playing with Friends or Strangers? The Effects of Familiarity between Players in an Asymmetric Multiplayer Virtual Reality Game. In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '23 Companion), October 10–13, 2023, Stratford, ON, Canada.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3573382.3616079

1 INTRODUCTION

Multiplayer games are social worlds for players; they can engage in communication and even form friendships. Research explores ways to improve multiplayer player experience (PX) of users, such as with asymmetric games. Asymmetric games focus on the differences between players in their design to provide engaging experiences for all parties involved [21]. The asymmetries in games can address differences in a multitude of aspects (e.g., hardware, player skill level). Alternatively, the asymmetric games could be simply aiming to elicit a richer social experience (SX) between players (e.g., social connectedness [20]). While current research on asymmetric games focuses on how these games affect PX and SX, it is still unknown how existing social bonds affect PX and SX in asymmetric games. To this end, our work investigates the effects of familiarity (friends vs. strangers) between players in an asymmetric game.

In recent years, virtual reality (VR) technology has become more common. In VR, users can synchronously interact and socialize with each other. Asymmetric games have also received attention in VR research [31]. These games were mostly explored with the aim of providing similar levels of PX between VR and non-VR users (who often have less opportunities to interact in VR worlds) [17, 18, 25]. Nevertheless, given the increasing usage of VR [2], using asymmetries in games where both players wear head-mounted displays (HMDs) requires attention and further research. However, asymmetric VR games are still an underexplored research topic.

We conducted a preliminary between-participants study (N=14) with two familiarity conditions. In one condition (friends), two befriended participants played together, whereas in the second condition (strangers), two randomly assigned participants played together. We investigated the effects of familiarity between players on PX, SX, and game performance in an asymmetric game in which two players are in VR. The game featured information asymmetry that resulted in a strategic interdependence between the players. In the game, the players were dependent on each other in a mirrored form (i.e., dependency of players were the same) [21] and exchanged their asymmetric information with their co-player to complete the game. Using this game, our research specifically answered the following research question: **(RQ)** *How does familiarity between players impact PX, SX, and game performance in an asymmetric multiplayer VR game?*

Our results did not reveal any significant effect of familiarity on PX, SX, and performance in our asymmetric multiplayer VR game. However, the asymmetric gameplay helped to improve social closeness in the strangers group. We discuss how asymmetries can be used to create social VR games regardless of the familiarity factor between players, and how familiarity does (not) affect these experiences. We summarize our contributions as follows:

CHI PLAY '23 Companion, October 10–13, 2023, Stratford, ON, Canada

- We present a preliminary investigation of the effects of an existing social dynamic—familiarity—in an asymmetric VR game.
- We demonstrate an asymmetric VR game without using interface asymmetry to support PX and SX of VR users.

2 RELATED WORK

Many people play multiplayer games [10]; these games have a social meaning for players. Although multiplayer games can affect PX and SX, existing social bonds, such as familiarity, can play a role in how these outcomes are shaped.

Several researchers explored the effects of familiarity between players on PX, SX, and game performance [12, 24, 26, 29, 36]. Playing against a co-located friend compared to playing against a stranger yielded to a greater spatial presence [29] and fun [26], but did not have a significant effect on immersion [4]. A survey study [27] found that teams of friends outperformed teams of strangers in *Halo: Reach* [3]. Hudson et al. [24]'s findings suggested that the interaction between game performance and familiarity might not be strong and could be dependent on a game, and familiarity and cooperative social presence are positively correlated. In contrast, two studies [36] showed that playing with a friend or stranger did not significantly affect PX (e.g., enjoyment) and SX (e.g., connection). Nonetheless, while the games used in Vella et al. [36]'s studies were cooperative, players were not allowed to communicate, which is not typically the case for asymmetric games [20, 25]. These games mostly require verbal communication (sometimes even gestural [33]) between partners [20, 25], which can impact PX and SX further. Overall, there is little work on the effects of familiarity in games, and past studies presented mixed results. Despite the exploration of familiarity in competitive and cooperative games, there is no exploration of this dynamic in the asymmetric VR games literature [31].

Asymmetric games try to serve for identified differences between many aspects (e.g., users, interface) [21]. A multiplayer game for a pair of a grandparent and grandchild can be designed using different game mechanics and elements for each player to create an enjoyable experience for both players. These games have also been shown to support PX and SX. A video game study by Harris and Hancock [20] showed that symmetric games led to less social connectedness and social presence between players than asymmetric games. Nevertheless, Harris and Hancock [20] only recruited players with pre-existing relationships, leaving a research gap to explore: the familiarity factor in asymmetric games.

Asymmetry research in VR mainly focused on interface asymmetry. They mainly looked for ways to enable interaction between VR and non-VR users, and used asymmetric game design as a potential solution [17, 18]. Gugenheimer et al. [18] presented a modified VR headset with three touch displays attached to allow non-VR users to interact with the VR world. A study by Karaosmanoglu et al. [25] showed that an asymmetric game with resulting strategic and biometric interdependences leads to a comparable level of high PX for both VR and non-VR players. However, VR technology has begun to enter our homes [2], and there are online remote options for playing VR games and experiencing virtual worlds (e.g., Rec Room [30], VRChat [37]). Therefore, it is worth exploring the potential of asymmetric VR games without using interface asymmetry (i.e., both users in VR) to provide socially rich gameplays for multiple VR users.

To address the research gaps outlined above, we designed an asymmetric game in which two players are in VR. Harris and Hancock [20] only tested their game with people with existing social bonds. An asymmetric VR game study [5] which mostly recruited teams of strangers suggested that playing with people with existing relationships compared to strangers could contribute to PX, but it did not evaluate this. We argue that asymmetric VR games should be explored with a focus on the familiarity factor because (1) these games by offering social gameplays might not be affected by this factor and can lead to similar levels of PX, SX, and game performance for friends and strangers dyads. Alternatively, (2) the verbal communication required to be successful in asymmetric VR games could be an advantage for friends who have interacted before and can potentially communicate more efficiently than strangers.

3 ASYMMETRIC VIRTUAL REALITY GAME

To answer our research question, we implemented a game in Unity (v2020.3.3f1) [35]. The SteamVR plugin (v2.7.3) [6], and Photon Unity Networking 2 (v2.30) [13] and Photon Voice 2 (v2.25.1) [14] libraries were used. Each player was equipped with an HMD (i.e., HTC Vive Pro [22]) and two controllers. Two computers were used to run the multiplayer game: one with an i7-4790K CPU and an NVIDIA Quadro RTX 8000 graphics card, the other with an i7-11700K CPU and an NVIDIA GeForce RTX 3090 graphics card.

Game Design Rationale. The asymmetric game is played with two VR players. The game was inspired by the commercial multiplayer game *Overcooked 2!* [15]. However, we decided to use a self-implemented game over a pre-existing game because we wanted to control two aspects. Firstly, we did not want to introduce roles. Since players usually need to share the tasks with their co-players, *Overcooked 2!* implicitly features roles. As different roles typically involve different abilities and information (e.g., [16]), we think that different roles can affect PX of players differently. To overcome this issue, we designed a low level of asymmetry and used one type of asymmetry— *asymmetry of information*, which creates a *strategic interdependence* between players. The strategic interdependence was used in a prior work [25] and led to positive results in an asymmetric VR game. Thus, we employed this interdependence in a mirrored form [21]; both players had different but the same type of information that they had to exchange with their partner in the same steps of the game. Secondly, using a self-implemented game, we aimed to obtain game performance measures easily. Further, we decided to place two players in two different physical rooms (separated by a glass window) due to the external validity of playing multiplayer VR games (typically remote e.g., in VRChat [37]). In the game, players communicated via voice chat and used teleportation for navigation. We also used genderless cartoon VR avatars to represent players (see Figure 2a).

Gameplay. The game begins with an interactive VR tutorial to familiarize players with interaction in VR (e.g., teleporting and grabbing objects). When players feel ready, they can join the actual gameplay scene and meet with their co-player in VR. Once both

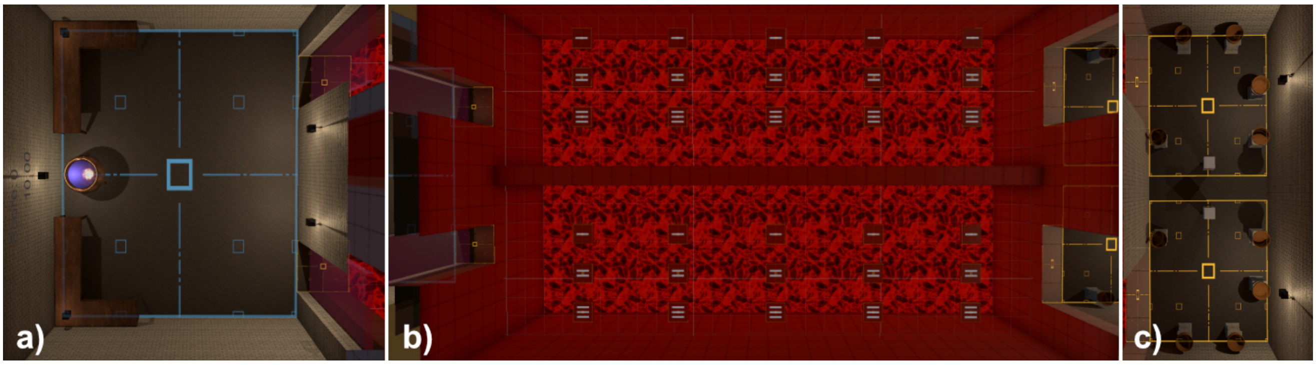

Figure 1: The game has three blocks. The start block (a) has a cauldron and opening gates for the maze. In the maze block (b), players have to exchange asymmetric information with their co-player, which helps them move through the tiles. In the ingredients block (c), players tell their asymmetric information (an ingredient) to their partner. After the dyad has their ingredients, they need to return from this area (c) to the start block (a) by passing through the maze (b) to put their ingredients in the cauldron in a synchronized manner.

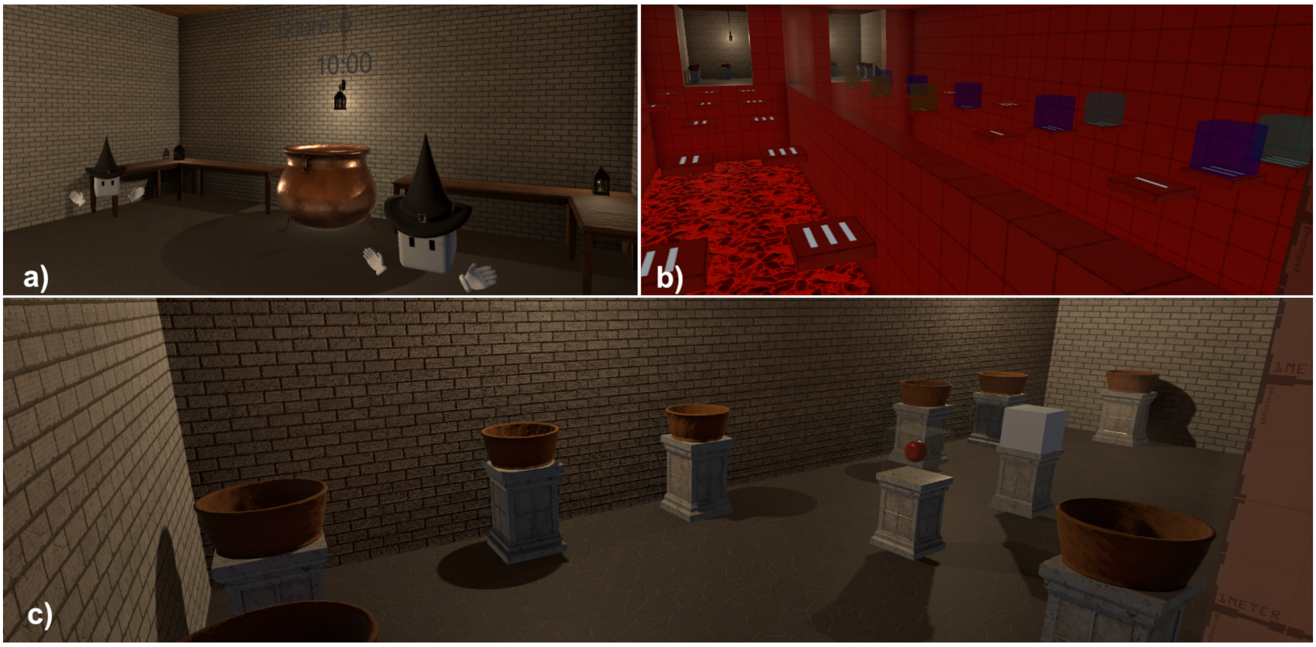

Figure 2: (a) The start block with VR avatars. (b) The maze block from a player's view. The player needs to communicate the information of the empty tile that the co-player needs to move forward. (c) In the ingredients block, the player needs to tell their co-player which ingredient (i.e., apple) they see on the middle block.

players join the game scene, the 10-minute playing phase begins. The main goal of the game is to prepare as many recipes as possible. To prepare a recipe, players need to manage three special blocks (see Figure 1): (a) start, (b) maze, and (c) ingredients blocks. Players are restricted to moving within their assigned area in each block.

In the *start block* (see Figure 1a), players could see how much time was left in the game. Players have a cauldron where they must bring their ingredients from the ingredients area to prepare a recipe. To maximize their points, they must put the ingredients in the cauldron in a synchronized manner. The score for each recipe was calculated using the following equation: *max(1000 - 200 * [Δseconds], 50)*. Finally, this block has gates that lead players to the maze block.

The *maze block* contains three tiles per row (15 tiles in total; see Figure 1b). Players are separated by a glass window in VR. Each player sees the other players' tile information (asymmetric) and guides them to move through the maze by telling the empty tile to teleport (see Figure 2b). Players reach the ingredients block after completing five rows.

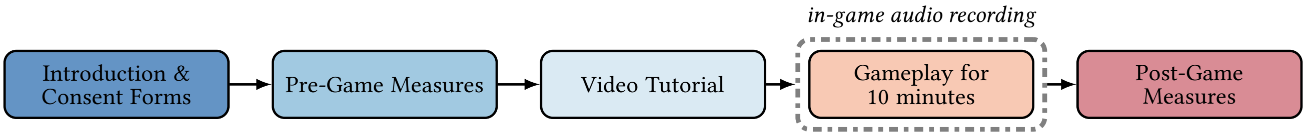

Figure 3: We followed a between-participants design (friends vs. strangers) and assessed questionnaire, performance, and audio data.

In the *ingredients block* (see Figure 1c), players were shown a required ingredient that their co-player (asymmetric) must collect. Once a player communicates this information to their co-player (apple, see Figure 2c), the co-player can find and grab the ingredient from the trays in their area of the block. Afterwards, players return through the maze area (communicating asymmetric information of tiles) to reach the cauldron in the start block to complete the task.

4 PRELIMINARY USER STUDY

We conducted a between-participants (one independent variable with two levels: friends and strangers) study and recruited 18 participants. However, due to technical problems, the data of two pairs (four players) were discarded, yielding 14 participants (8 female, 6 male, age: M=21.07, SD=3.79 years). Four dyads were strangers, while three dyads were friends. Eleven participants had prior experience with games, while three had no experience (frequency of playing: 2 rarely; 2 sometimes; 5 very often; 2 always). From these participants, almost all of them (10) had previous experience with multiplayer games (frequency of playing: 1 never; 3 rarely; 1 sometimes; 5 very often). Half of the participants did not have prior VR experience, while half (7) did (frequency of using VR: 1 never; 5 rarely; 1 sometimes).

Pre-Game Measures. We used a demographic questionnaire to collect information about sample characteristics (see supplementary materials). We employed a single-item yes/no measure to control for players' familiarity: "*Are you familiar with your co-player (friends vs. strangers)?*" Additionally, to have a quantitatively more nuanced understanding of player's existing level of familiarity (similar to [24]), players rated a modified version (see supplementary materials) of the Inclusion of Self (IoS) [1, 19] pictorial item on a 7-point scale; higher scores indicates higher closeness.

In-Game Measures. We recorded the game events, which led us to determine the total score, the number of errors, and the number and duration (in seconds) of completing the game loop (i.e., completing a recipe) for each dyad. We also recorded the game's in-game audio to analyze the interesting interaction between players.

Post-Game Measures. After the gameplay, we used IoS to assess the change in players' feeling of social closeness. We obtained social presence scores using Social Presence in Gaming Questionnaire (SPGQ) [7] on a 5-point Likert scale (*0/not at all — 4/extremely*). We only used empathy (7 items) and behavioural engagement (8 items) sub-categories of this scale because its negative feelings sub-category involves items targeting competitive gameplay [23]. We used the interest/enjoyment sub-scale of Intrinsic Motivation Inventory (IMI) [32] to measure enjoyment of players (on a 7-point Likert scale, *1/not at all true — 7/very true*). Player Experience of Need Satisfaction (PENS) [32] questionnaire was employed to measure autonomy, competence, relatedness, and presence/immersion ratings of players (7-point Likert scale, *1/do not agree — 7/strongly agree*). Lastly, we used an open-ended question to get additional comments from participants: "*Is there anything you would like to mention or comment on the study/game?*"

Procedure. We announced the study via a convenience sampling method (e.g., emailing lists). The participants registered for multiple time slots with a friend. Then, they were invited to the study with their friend or a stranger. If players were assigned to a stranger group, they were asked if they knew each other to control the familiarity factor.

At the beginning of the study, players were placed in different rooms. They filled out a consent form and the pre-game measures. They then watched a video tutorial explaining the gameplay (see supplementary). Upon completion, the game began with a tutorial scene that gave players the chance to get familiar with interactions (e.g., teleporting and grabbing objects) in VR. In their assigned conditions, participants played the game with their co-players for 10 minutes. During the gameplay, audio was recorded. After the gameplay, they filled out the post-game measures. The study took approximately 45–60 minutes for each dyad. If applicable, as a remuneration, some participants received a participation certificate which is necessary to complete their degree. The overall study procedure is illustrated in Figure 3.

Analysis Methods. For quantitative data, we performed parametric tests when the data were normally distributed (tested with Shapiro-Wilk tests) and homogeneity of variance was not violated (Levene type test), and otherwise we used non-parametric tests. For between factor cases, we employed unpaired t-tests (t) and Wilcoxon rank-sum tests (W), and for within factor cases, we used paired t-tests (t) and Wilcoxon signed-rank tests (V). For IoS measure, since it is ordinal, we used non-parametric tests, and reported both pre- and post-value comparison for each group (within factor, via IoS-friends and IoS-strangers), and the comparison of IoS score change between the groups (between factor, via IoS-cha.). We note that we experienced minor technical problems during the study, which affected the performance data logs of two dyads; these logs were manually corrected based on unreasonable times in the file and audio-recordings.

We followed an informal approach to analyze qualitative data. The lead author translated the German data (i.e., audio recordings and answers to the open-ended question; seven gameplay, N=14) into English (via DeepL Pro [8]) and used the Dovetail tool [9] for the qualitative analysis process. Then, they inductively and iteratively coded (e.g., laughing and clarifying the game) the interactions

Table 1: The table shows the results of employed statistical tests for the study measures. The bold text indicates significance.

Questionnaire	Friends (M)	Friends (SD)	Strangers (M)	Strangers (SD)	Test	P Value	Effect Size
IoS - Cha.	1.33	1.86	1.63	1.06	W=16.5	p=0.343	r=-0.25
SPGQ - Emp.	3.10	0.86	3.07	0.45	t(12)=0.067	p=0.947	d=0.04
SPGQ - Beh. Eng.	3.67	0.20	3.08	0.84	t(12)=1.670	p=0.121	d=0.90
IMI - Int./Enj.	6.52	0.58	6.36	0.48	t(12)= 0.586	p=0.569	d=0.32
PENS - Aut.	4.5	1.75	5.33	1.26	t(12)=-1.041	p=0.319	d=-0.56
PENS - Com.	5.94	0.71	5.42	0.92	t(12)=1.162	p=0.268	d=0.63
PENS - Rel.	5.72	1.08	4.83	1.39	t(12)=1.294	p=0.220	d=0.70
PENS- Pre./Imm.	5.72	0.73	5.55	0.85	W=25	p=0.948	r=-0.02
Questionnaire	Pre (M)	Pre (SD)	Post (M)	Post (SD)	Test	P Value	Effect Size
IoS - Friends	3.33	1.51	4.67	1.86	V=0	p=0.089	r=-0.49
IoS - Strangers	1	0	2.63	1.06	V=0	**p=0.021**	r=-0.58
Game Performance	Friends (M)	Friends (SD)	Strangers (M)	Strangers (SD)	Test	P Value	Effect Size
Score	5633.33	860.72	4475	1721.19	t(5)=1.053	p=0.341	d=0.80
# of Errors	2.67	3.06	4.25	3.30	t(5)=-0.646	p=0.547	d=-0.49
# of Loop Comp.	6.67	1.53	5	1.41	t(5)=1.494	p=0.195	d=1.14
Time of Loop Comp.	87.71	20.99	122.01	43.67	t(5)=-1.236	p=0.271	d=-0.94

between dyads. For the reporting purposes, we used *(FX_PY)* for the friends and *(SX_PY)* for the strangers groups. For example, *(F3_P2)* indicates the second player from the third friend dyad.

5 RESULTS

Quantitative Findings. The results of questionnaire and game performance measures can be found in Table 1. Since our data were generally normally distributed, we report mean (*M*) and standard deviation (*SD*) values. Overall, our quantitative results show that almost all metrics did not differ significantly between groups. However, the social closeness scores of the strangers group improved significantly after playing the game (*M*=2.63, *SD*=1.06) compared to their pre-measurement scores (*M*=1, *SD*=0, *V*=0, *p*=0.021, *r*=-0.58).

Qualitative Findings. Overall, regardless of their groups, all dyads enjoyed their experience; they laughed and most of them made jokes. In both groups, they were also a few occurrences of swearing and apologizing for their actions. A stranger team member explicitly reported that how the gameplay affected their SX: "*It felt awesome connecting with a stranger and building some kind of relationship form within bare minutes!*"-*(S2_P1)*. Though, a friend team member *(F2_P1)* specifically noted the positive impact of playing with their best friend on their experience. Occasionally, a few group members, in particular a stranger team, positively reacted to their achievements: "*Oha, we are getting better [laughing]*"-*(S1_P1)*.

Through verbal communication, both groups of dyads strategized to complete their tasks: "*I would say we always say the first 3 [referring to maze block tiles]*"-*(F3_P1)* and "*Ok well let's try something new ok? We say the announcement at the same time each per step*"-*(S2_P1)*. In some instances, figuring out what is the ingredient sparked conversation between dyads: "*Green cane? I would have thought it's a dragon's tail, dragon's tail somehow*"-*(F2_P2)*. While we designed the game using mirrored interdependence between players (i.e., players were dependent on each other to the same degree), a few players took the role of a leader in their team as their co-player were asking explicit questions about the game tasks: "*Yes? But there are different things, where do I see that now?*"-*(F1_P2)* and "*Where should it go now?*"-*(S4_P2)*. However, there were also some natural occurrences in which players clarified some tasks to each other without being asked: "*I would have remembered the numbers*"-*(F2_P2)*, and replying: "*No that changes*"-*(F2_P1)* and "*You can also hold it with the button on your index finger, if that helps*"-*(S4_P1)*. Finally, a few players noted technical problems: for instance, they "*felt the audio had [...] delay which made the coordinated dropping of items much more challenging than expected in the first place*"-*(F2_P2)*.

6 DISCUSSION

This section discusses how an existing social bond, i.e., familiarity, can affect SX, PX, and performance of players in an asymmetric VR game. Finally, we address the limitations of our study.

Regardless of group type, the game resulted in high scores on social presence constructs; our findings reveal that neither empathy nor behavioural scores of the groups were significantly different from each other. Similarly, for both groups, there was an increase in IoS scores, however, this increase was not significantly different between the two groups. These findings suggest that asymmetric games, due to the nature of their social gameplay, can provide high SX not only when played with people with existing social relationships [20], but also when played with strangers.

Our findings also point out that after playing the game, strangers groups' IoS scores were significantly higher than their baseline scores. Although the IoS scores for the friends teams also improved, their post-IoS ratings were not significantly higher than their baseline. We think that since the friends dyads already had a social connection through their real-world experience, the gameplay might have been a short event for the friends teams to influence their established social closeness. However, for the strangers groups, the asymmetric game was an—first—opportunity for social connection, as also noted by one player: "*It felt awesome connecting with a stranger and building some kind of relationship form within bare minutes!*"-*(S2_P1)*. Thus, even a limited time of social interaction (e.g., communication) could have an impact on their social closeness.

Both groups had a similarly high level of enjoyment. For the PENS' sub-categories, both groups had moderately high autonomy, competence, relatedness, and presence/immersion scores. However, there were no significant differences on these constructs between the familiarity groups either. This complements Vella et al. [36]'s work on cooperative games, which showed that familiarity had no significant effect on enjoyment and the PENS sub-category scores. We transfer their results to asymmetric VR games and show that regardless of the familiarity group, players enjoy these games and have positive PX. Similarly, our qualitative results also match the players' positive experience shown by our quantitative data (e.g., laughs and jokes). However, these results do not support Cheng et al. [5]'s suggestion; playing the asymmetric VR game with a friend did not contribute significantly to the PX compared to playing with a stranger. We attribute this to design differences between our and Cheng et al. [5]'s game; their game features physical proximity between players, where a pre-existing social bond can play a role.

Contrary to Mason and Clauset [27]'s work on a commercial game, our findings show that there is no significant difference in game performance between teams of friends and strangers. Relying on a previous study [24], we think that the familiarity may be a factor affecting game performance depending on a game. Our preliminary findings suggest that game performance in asymmetric games is not significantly affected by the familiarity factor. However, these results need to be generalized with larger sample size and various kinds of asymmetric games.

Overall, our findings reveal that PX, SX, and game performance are not significantly affected by the familiarity factor in an asymmetric VR game. These suggest that asymmetries can be used to provide high SX and PX regardless of familiarity between players, even in virtual worlds such as VR, where cues of sociability between players are limited (e.g., people physically cannot high five). Moreover, our results show that by employing asymmetry (and the resulting interdependence between players) people's closeness can increase even if they are not physically in the same place and do not know each other; our results contribute to ongoing efforts to understand and support social interaction in social VR mediums [11, 34]. Given the increasing number of VR users [2], this opens up the possibilities to explore different aspects of asymmetric game design in VR beyond the typical explorations that focus solely on the asymmetry of the interface [31].

Limitations. Due to COVID-19 regulations, the sample size of the study was low. Further, participants were only allowed to register to participate in the study with a friend (who they specified), and they registered in multiple available time slots. We then assigned people to teams of friends and strangers. We did this to avoid having a biased sample because potentially, people could have preferred their favorite way of playing games (e.g., with friends) to enroll in the study. However, consequently, these assignments were affected by the participants' volunteered time slots and the matching of those time slots with the other participants' volunteered time slots. In terms of the sample characteristics, the pre-IoS scores of our friends teams were between the number of two and five out of seven. But, a different IoS score distribution could have affected the results.

Our game design features information asymmetry, but different types of asymmetry (e.g., ability) might have a different impact on the results. Further, our game offers only digital interaction, not physical. We speculate that PX, SX, and game performance metrics of physical interaction-based asymmetric games may be affected by the familiarity factor. Therefore, to inform game designers, developers, and researchers about creating engaging multiplayer VR games, we suggest that future studies consider these aspects, and explore various types of asymmetry and the effects of different level of social factors on PX, SX, and game performance.

For the qualitative analysis, we follow the best practices [28] and provide a positionality statement. The lead author works on the intersection of asymmetric game design and VR research, and plays asymmetric games. This may have both biased and enriched the data analysis, as they could be more attentive to asymmetric game design elements.

7 CONCLUSION

This work investigated how familiarity (friends vs. strangers) impacts PX, SX, and game performance in an asymmetric VR game. Our preliminary results (N=14) suggest that familiarity has no significant effect on PX, SX, or game performance. However, the asymmetric game improved social closeness between players on the strangers teams. We recommend that multiplayer game designers and developers use asymmetric design to provide VR users with a socially rich and positive PX without being significantly affected by the familiarity factor. We believe asymmetric VR games have the potential to be explored beyond the use of asymmetric interfaces to enrich the experience of millions of VR users.

ACKNOWLEDGMENTS

We would like to thank Niklas Wilczek for his help with asset creation and play testing. This work was supported by the European Union, the German Federal Ministry of Education and Research (BMBF), and the German Research Foundation (DFG).

REFERENCES

[1] Arthur Aron, Elaine N Aron, and Danny Smollan. 1992. Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology* 63, 4 (1992), 596. https://doi.org/10.1037/0022-3514.63.4.596 pages #4

[2] ARtillery Intelligence. 2021. Virtual reality (VR) headset unit sales worldwide from 2019 to 2024 (in million units). https://arinsider.co/ The Statistic Portal, Statista from https://www.statista.com/statistics/677096/vr-headsets-worldwide/. Accessed: 10.01.2022. pages #1, #2, #6

[3] Bungie. 2010. *Halo: Reach.* Game. https://store.steampowered.com/app/1064220/Halo_Reach/. Accessed: 17.01.2023. pages #2

[4] Paul Cairns, Anna L. Cox, Matthew Day, Hayley Martin, and Thomas Perryman. 2013. Who but not where: The effect of social play on immersion in digital games. *International Journal of Human-Computer Studies* 71, 11 (2013), 1069–1077. https://doi.org/10.1016/j.ijhcs.2013.08.015 pages #2

[5] Lung-Pan Cheng, Patrick Lühne, Pedro Lopes, Christoph Sterz, and Patrick Baudisch. 2014. Haptic Turk: A Motion Platform Based on People. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3463–3472. https://doi.org/10.1145/2556288.2557101 pages #2, #6

[6] Valve Corporation. 2021. *SteamVR Plugin*. https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647 SteamVR Plugin v2.7.3. Accessed: 14.04.2022. pages #2

[7] Yvonne AW De Kort, Wijnand A IJsselsteijn, and Karolien Poels. 2007. Digital games as social presence technology: Development of the Social Presence in Gaming Questionnaire (SPGQ). *Proceedings of PRESENCE* 195203 (2007), 1–9. pages #4

[8] DeepL SE. 2017. DeepL Translator. DeepL Pro version. Cologne, Germany. https://www.deepl.com/pro?cta=header-pro/. pages #4

[9] Dovetail Research PTY. LTD. 2016. Dovetail. Surry Hills, Australia. https://dovetailapp.com/. pages #4

[10] Entertainment Software Association (ESA). 2021. 2021 Essential Facts about the Video Game Industry. https://www.theesa.com/resource/2021-essential-facts-about-the-video-game-industry/ Washington, D.C., USA. pages #2

[11] Guo Freeman and Dane Acena. 2021. Hugging from A Distance: Building Interpersonal Relationships in Social Virtual Reality. In *ACM International Conference on Interactive Media Experiences* (Virtual Event, USA) *(IMX '21)*. Association for Computing Machinery, New York, NY, USA, 84–95. https://doi.org/10.1145/3452918.3458805 pages #6

[12] Brian J. Gajadhar, Yvonne A. W. de Kort, and Wijnand A. IJsselsteijn. 2008. Shared Fun Is Doubled Fun: Player Enjoyment as a Function of Social Setting. In *Fun and Games*, Panos Markopoulos, Boris de Ruyter, Wijnand IJsselsteijn, and Duncan Rowland (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 106–117. pages #2

[13] Exit Games. 2021. *Photon Unity Networking 2*. https://www.photonengine.com/ v2.30. 13 April 2021. https://assetstore.unity.com/packages/tools/network/pun-2-free-119922. Accessed: 24.05.2022. pages #2

[14] Exit Games. 2021. *Photon Voice 2*. https://www.photonengine.com/ v2.25.1. 9 June 2021. https://assetstore.unity.com/packages/tools/audio/photon-voice-2-130518. Accessed: 24.05.2022. pages #2

[15] Ghost Town Games Ltd. and Team17. 2018. *Overcooked! 2*. Game [PC]. https://store.steampowered.com/app/728880/Overcooked_2/. Accessed: 16.10.2022. pages #2

[16] David Gonçalves, André Rodrigues, Mike L. Richardson, Alexandra A. de Sousa, Michael J. Proulx, and Tiago Guerreiro. 2021. Exploring Asymmetric Roles in Mixed-Ability Gaming. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 114, 14 pages. https://doi.org/10.1145/3411764.3445494 pages #2

[17] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683 pages #1, #2

[18] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. 2018. FaceDisplay: Towards Asymmetric Multi-User Interaction for Nomadic Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173628 pages #1, #2

[19] Simon Gächter, Chris Starmer, and Fabio Tufano. 2015. Measuring the Closeness of Relationships: A Comprehensive Evaluation of the 'Inclusion of the Other in the Self' Scale. *PLOS ONE* 10, 6 (06 2015), 1–19. https://doi.org/10.1371/journal.pone.0129478 pages #4

[20] John Harris and Mark Hancock. 2019. To Asymmetry and Beyond! Improving Social Connectedness by Increasing Designed Interdependence in Cooperative Play. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300239 pages #1, #2, #5

[21] John Harris, Mark Hancock, and Stacey D. Scott. 2016. Leveraging Asymmetries in Multiplayer Games: Investigating Design Elements of Interdependent Play. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) *(CHI PLAY '16)*. Association for Computing Machinery, New York, NY, USA, 350–361. https://doi.org/10.1145/2967934.2968113 pages #1, #2

[22] HTC and Valve Corporation. 2018. HTC Vive Pro. https://www.vive.com/us/product/vive-pro/ pages #2

[23] Matthew Hudson and Paul Cairns. 2014. Measuring social presence in team-based digital games. *Interacting with Presence: HCI and the Sense of Presence in Computer-mediated Environments* 83 (2014). https://doi.org/10.2478/9783110409697.6 pages #4

[24] Matthew Hudson, Paul Cairns, and A Imran Nordin. 2015. Familiarity in team-based online games: the interplay between player familiarity and the concepts of social presence, team trust, and performance. In *International Visual Informatics Conference*. Springer, 140–151. https://doi.org/10.1007/978-3-319-25939-0_13 pages #2, #4, #6

[25] Sukran Karaosmanoglu, Katja Rogers, Dennis Wolf, Enrico Rukzio, Frank Steinicke, and Lennart E. Nacke. 2021. Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 443, 15 pages. https://doi.org/10.1145/3411764.3445492 pages #1, #2

[26] Regan L. Mandryk, M. Stella Atkins, and Kori M. Inkpen. 2006. A Continuous and Objective Evaluation of Emotional Experience with Interactive Play Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 1027–1036. https://doi.org/10.1145/1124772.1124926 pages #2

[27] Winter Mason and Aaron Clauset. 2013. Friends FTW! Friendship and Competition in Halo: Reach. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13)*. Association for Computing Machinery, New York, NY, USA, 375–386. https://doi.org/10.1145/2441776.2441820 pages #2, #6

[28] Benjamin John Newton, Zuzana Rothlingova, Robin Gutteridge, Karen LeMarchand, and Jon Howard Raphael. 2012. No room for reflexivity? Critical reflections following a systematic review of qualitative research. *Journal of Health Psychology* 17, 6 (2012), 866–885. https://doi.org/10.1177/1359105311427615 pages #6

[29] Niklas Ravaja, Timo Saari, Marko Turpeinen, Jari Laarni, Mikko Salminen, and Matias Kivikangas. 2006. Spatial Presence and Emotions during Video Game Playing: Does It Matter with Whom You Play? *Presence: Teleoperators and Virtual Environments* 15, 4 (08 2006), 381–392. https://doi.org/10.1162/pres.15.4.381 arXiv:https://direct.mit.edu/pvar/article-pdf/15/4/381/1624472/pres.15.4.381.pdf pages #2

[30] Rec Room. 2016. *Rec Room*. Game [Mobile, PC, VR]. https://recroom.com/. Accessed: 04.01.2023. pages #2

[31] Katja Rogers, Sukran Karaosmanoglu, Dennis Wolf, Frank Steinicke, and Lennart E. Nacke. 2021. A Best-Fit Framework and Systematic Review of Asymmetric Gameplay in Multiplayer Virtual Reality Games. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.694660 pages #1, #2, #6

[32] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion* 30, 4 (2006), 344–360. https://doi.org/10.1007/s11031-006-9051-8 The IMI items are available online under this link: http://selfdeterminationtheory.org/intrinsic-motivation-inventory/. Accessed: 02.09.2022. pages #4

[33] Michael Smilovitch and Richard Lachman. 2019. BirdQuestVR: A Cross-Platform Asymmetric Communication Game. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Barcelona, Spain) *(CHI PLAY '19 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 307–313. https://doi.org/10.1145/3341215.3358246 pages #2

[34] Philipp Sykownik, Sukran Karaosmanoglu, Katharina Emmerich, Frank Steinicke, and Maic Masuch. 2023. VR Almost There: Simulating Co-Located Multiplayer Experiences in Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 789, 19 pages. https://doi.org/10.1145/3544548.3581230 pages #6

[35] Unity Technologies. 2021. Unity Real-Time Development Platform. https://unity.com/ LTS Release 2020.3.3f1. 7 April 2021. https://unity3d.com/get-unity/download/archive. Accessed: 14.04.2022. pages #2

[36] Kellie Vella, Christopher James Koren, and Daniel Johnson. 2017. The Impact of Agency and Familiarity in Cooperative Multiplayer Games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) *(CHI PLAY '17)*. Association for Computing Machinery, New York, NY, USA, 423–434. https://doi.org/10.1145/3116595.3116622 pages #2, #6

[37] VRChat Inc. 2017. *VRChat*. Game [VR]. https://hello.vrchat.com/. Accessed: 04.01.2023. pages #2

Received 2023-06-22; accepted 2023-08-03

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. Development and validation of the player experience inventory: A scale to measure player experiences at the level of functional and psychosocial consequences. *International Journal of Human-Computer Studies*, 135:102370, 2020. ISSN 1071-5819. DOI 10.1016/j.ijhcs.2019.102370. URL https://doi.org/10.1016/j.ijhcs.2019.102370.

[2] Vero Vanden Abeele, Brenda Schraepen, Hanne Huygelier, Celine Gillebert, Kathrin Gerling, and Raymond Van Ee. Immersive virtual reality for older adults: Empirically grounded design guidelines. *ACM Transactions on Accessible Computing (TACCESS)*, 14(3), 2021. ISSN 1936-7228. DOI 10.1145/3470743. URL https://doi.org/10.1145/3470743.

[3] Caroline Adam and Veit Senner. Which motives are predictors for long-term use of exergames? *Procedia Engineering*, 147:806–811, 2016. ISSN 1877-7058. DOI 10.1016/j.proeng.2016.06.310. URL https://doi.org/10.1016/j.proeng.2016.06.310. The Engineering of SPORT 11.

[4] Lena Fanya Aeschbach, Sebastian A.C. Perrig, Lorena Weder, Klaus Opwis, and Florian Brühlmann. Transparency in measurement reporting: A systematic literature review of chi play. *Proceedings of the ACM on Human-Computer Interaction*, 5 (CHI PLAY), 2021. DOI 10.1145/3474660. URL https://doi.org/10.1145/3474660.

[5] Maximilian Altmeyer, Vladislav Hnatovskiy, Katja Rogers, Pascal Lessel, and Lennart E. Nacke. Here comes no boom! the lack of sound feedback effects on performance and user experience in a gamified image classification task. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, (CHI '22), New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391573. DOI 10.1145/3491102.3517581. URL https://doi.org/10.1145/3491102.3517581.

[6] Alzheimer's Society. Risk factors for dementia, August 2024. URL https://www.alzheimers.org.uk/about-dementia/managing-the-risk-of-dementia/risk-factors-for-dementia. Accessed: 07.03.2025.

[7] Toni C Antonucci, Kristine J Ajrouch, and Jasmine A Manalel. Social relations and technology: Continuity, context, and change.

Innovation in Aging, 1(3):igx029, 2017. ISSN 2399-5300. DOI 10.1093/geroni/igx029. URL https://doi.org/10.1093/geroni/igx029.

[8] AR Insider. Virtual reality (vr) headset unit sales worldwide from 2019 to 2024 (in million units) [graph]. Statista. Online. https://www.statista.com/statistics/677096/vr-headsets-worldwide/, Accessed: 13.02.2025, 2021.

[9] Oscar Ariza, Gerd Bruder, Nicholas Katzakis, and Frank Steinicke. Analysis of proximity-based multimodal feedback for 3d selection in immersive virtual environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces*, (IEEE VR '18), pages 327–334, Los Alamitos, CA, USA, 2018. IEEE. DOI 10.1109/VR.2018.8446317. URL https://doi.org/10.1109/VR.2018.8446317.

[10] Oliver Assad, Robert Hermann, Damian Lilla, Björn Mellies, Ronald Meyer, Liron Shevach, Sandra Siegel, Melanie Springer, Saranat Tiemkeo, Jens Voges, Jan Wieferich, Marc Herrlich, Markus Krause, and Rainer Malaka. Motion-based games for parkinson's disease patients. In Junia Coutinho Anacleto, Sidney Fels, Nicholas Graham, Bill Kapralos, Magy Saif El-Nasr, and Kevin Stanley, editors, *Entertainment Computing – ICEC 2011*, pages 47–58, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg. ISBN 978-3-642-24500-8. DOI 10.1007/978-3-642-24500-8_6. URL https://doi.org/10.1007/978-3-642-24500-8_6.

[11] Anna Barenbrock, Marc Herrlich, and Rainer Malaka. Design lessons from mainstream motion-based games for exergames for older adults. In *2014 IEEE Games Media Entertainment*, (GEM '14), pages 1–8, 2014. DOI 10.1109/GEM.2014.7048096. URL https://doi.org/10.1109/GEM.2014.7048096.

[12] Anna Barenbrock, Marc Herrlich, Kathrin Maria Gerling, Jan David Smeddinck, and Rainer Malaka. Varying avatar weight to increase player motivation: Challenges of a gaming setup. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, (CHI EA '18), page 1–6, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356213. DOI 10.1145/3170427.3188634. URL https://doi.org/10.1145/3170427.3188634.

[13] Beat Games. *Beat Saber*. Game [VR], May 2019. Beat Games. Los Angeles, CA, USA. https://beatsaber.com/.

[14] Bemani. *Dance Dance Revolution*. Game [Arcade], September 1998. https://www.konami.com/games/asia/en/products/ddr_a/.

[15] David Benyon, Phil Turner, and Susan Turner. *Designing interactive systems: People, activities, contexts, technologies*. Pearson Education, 2005.

[16] Hugh Beyer and Karen Holtzblatt. *Contextual Design: Defining Customer-Centered Systems*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1997. ISBN 9780080503042. DOI 10.5555/2821566. URL https://doi.org/10.5555/2821566.

[17] Ananya Bhattacharjee, Joseph Jay Williams, Jonah Meyerhoff, Harsh Kumar, Alex Mariakakis, and Rachel Kornfield. Investigating the role of context in the delivery of text messages for supporting psychological wellbeing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. DOI 10.1145/3544548.3580774. URL https://doi.org/10.1145/3544548.3580774.

[18] Pauline Bimberg, Tim Weissker, and Alexander Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, (IEEE VR VRW '20), pages 464–467, 2020. DOI 10.1109/VRW50115.2020.00098. URL https://doi.org/10.1109/VRW50115.2020.00098.

[19] Frank Biocca, Chad Harms, and Jenn Gregg. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. *4th annual international workshop on presence, Philadelphia, PA*, pages 1–9, 2001.

[20] Frank Biocca, Chad Harms, and Judee K. Burgoon. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators and Virtual Environments*, 12(5):456–480, 2003. DOI 10.1162/105474603322761270. URL https://doi.org/10.1162/105474603322761270.

[21] Daniel Boland and Mark McGill. Lost in the rift: engaging with mixed reality. *XRDS*, 22(1):40–45, 2015. ISSN 1528-4972. DOI 10.1145/2810046. URL https://doi.org/10.1145/2810046.

[22] Felix Born, Sophie Abramowski, and Maic Masuch. Exergaming in vr: The impact of immersive embodiment on motivation, performance, and perceived exertion. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications*, (VS-Games '19), pages 1–8. IEEE, 2019. DOI 10.1109/VS-Games.2019.8864579. URL https://doi.org/10.1109/VS-Games.2019.8864579.

[23] Felix Born, Adrian Rygula, and Maic Masuch. Motivating players to perform an optional strenuous activity in a virtual reality

exergame using virtual performance augmentation. *Proceedings of the ACM on Human-Computer Interaction*, 5 (CHI PLAY), 2021. DOI 10.1145/3474652. URL https://doi.org/10.1145/3474652.

[24] Robert Bowman, Camille Nadal, Kellie Morrissey, Anja Thieme, and Gavin Doherty. Using thematic analysis in healthcare hci at chi: A scoping review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. DOI 10.1145/3544548.3581203. URL https://doi.org/10.1145/3544548.3581203.

[25] James V. Bradley. Complete counterbalancing of immediate sequential effects in a latin square design. *Journal of the American Statistical Association*, 53(282):525–528, 1958. DOI 10.1080/01621459.1958.10501456. URL https://doi.org/10.1080/01621459.1958.10501456.

[26] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006. DOI 10.1191/1478088706qp063oa. URL https://doi.org/10.1191/1478088706qp063oa.

[27] Virginia Braun and Victoria Clarke. Can i use ta? should i use ta? should i not use ta? comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, 21(1):37–47, 2021. DOI 10.1002/capr.12360. URL https://doi.org/10.1002/capr.12360.

[28] Tim Böttger, Thomas Rudolph, Heiner Evanschitzky, and Thilo Pfrang. Customer inspiration: Conceptualization, scale development, and validation. *Journal of Marketing*, 81(6):116–131, 2017. DOI 10.1509/jm.15.0007. URL https://doi.org/10.1509/jm.15.0007.

[29] Pablo Campo-Prieto, Gustavo Rodríguez-Fuentes, and José Mª Cancela-Carral. Immersive virtual reality exergame promotes the practice of physical activity in older people: An opportunity during covid-19. *Multimodal Technologies and Interaction*, 5(9), 2021. ISSN 2414-4088. DOI 10.3390/mti5090052. URL https://doi.org/10.3390/mti5090052.

[30] Christopher Carroll, Andrew Booth, and Katy Cooper. A worked example of" best fit" framework synthesis: a systematic review of views concerning the taking of some potential chemopreventive agents. *BMC medical research methodology*, 11:1–9, 2011. DOI 10.1186/1471-2288-11-29. URL https://doi.org/10.1186/1471-2288-11-29.

[31] Christopher Carroll, Andrew Booth, Joanna Leaviss, and Jo Rick. "best fit" framework synthesis: refining the method. *BMC medical research methodology*, 13(1):37, 2013. DOI 10.1186/1471-2288-13-37. URL https://doi.org/10.1186/1471-2288-13-37.

[32] Marcus Carter, John Downs, Bjorn Nansen, Mitchell Harrop, and Martin Gibbs. Paradigms of games research in hci: a review of 10 years of research at chi. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY '14), page 27–36, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450330145. DOI 10.1145/2658537.2658708. URL https://doi.org/10.1145/2658537.2658708.

[33] Sammi R Chekroud, Ralitza Gueorguieva, Amanda B Zheutlin, Martin Paulus, Harlan M Krumholz, John H Krystal, and Adam M Chekroud. Association between physical exercise and mental health in 1·2 million individuals in the usa between 2011 and 2015: a cross-sectional study. *The Lancet Psychiatry*, 5(9):739–746, 2018. ISSN 2215-0366. DOI 10.1016/S2215-0366(18)30227-X. URL https://doi.org/10.1016/S2215-0366(18)30227-X.

[34] Lung-Pan Cheng, Patrick Lühne, Pedro Lopes, Christoph Sterz, and Patrick Baudisch. Haptic turk: A motion platform based on people. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (CHI '14), page 3463–3472, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450324731. DOI 10.1145/2556288.2557101. URL https://doi.org/10.1145/2556288.2557101.

[35] Jesper Vang Christensen, Mads Mathiesen, Joakim Have Poulsen, Ea Ehrnberg Ustrup, and Martin Kraus. Player experience in a vr and non-vr multiplayer game. In *Proceedings of the Virtual Reality International Conference - Laval Virtual*, (VRIC '18), New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450353816. DOI 10.1145/3234253.3234297. URL https://doi.org/10.1145/3234253.3234297.

[36] Sharon Lynn Chu Yew Yee, Henry Been-Lirn Duh, and Francis Quek. Investigating narrative in mobile games for seniors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (CHI '10), page 669–672, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605589299. DOI 10.1145/1753326.1753424. URL https://doi.org/10.1145/1753326.1753424.

[37] Victoria Clarke and Virginia Braun. Thematic analysis: A practical guide. *Thematic Analysis*, pages 1–100, 2022.

[38] Sebastian Cmentowski and Jens Krüger. Playing with friends - the importance of social play during the covid-19 pandemic. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY '20), page 209–212, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450375870. DOI 10.1145/3383668.3419911. URL https://doi.org/10.1145/3383668.3419911.

[39] Sebastian Cmentowski and Jens Krüger. Effects of task type and wall appearance on collision behavior in virtual environments. In *2021 IEEE Conference on Games*, (CoG '21), pages 1–8, 2021. DOI 10.1109/CoG52621.2021.9619039. URL https://doi.org/10.1109/CoG52621.2021.9619039.

[40] Stanley J. Colcombe, Kirk I. Erickson, Naftali Raz, Andrew G. Webb, Neal J. Cohen, Edward McAuley, and Arthur F. Kramer. Aerobic Fitness Reduces Brain Tissue Loss in Aging Humans. *The Journals of Gerontology: Series A*, 58(2):M176–M180, 2003. ISSN 1079-5006. DOI 10.1093/gerona/58.2.M176. URL https://doi.org/10.1093/gerona/58.2.M176.

[41] Helena Cole and Mark D Griffiths. Social interactions in massively multiplayer online role-playing gamers. *Cyberpsychology & behavior*, 10(4):575–583, 2007. DOI 10.1089/cpb.2007.9988. URL https://doi.org/10.1089/cpb.2007.9988.

[42] Harris Cooper. *Research synthesis and meta-analysis: A step-by-step approach*, volume 2. Sage publications, 2015.

[43] Harris M Cooper. A taxonomy of literature reviews. 1985.

[44] Simon Cooper, Robyn Cant, Michelle Kelly, Tracy Levett-Jones, Lisa McKenna, Philippa Seaton, and Fiona Bogossian. An evidence-based checklist for improving scoping review quality. *Clinical Nursing Research*, 30(3):230–240, 2021. DOI 10.1177/1054773819846024. URL https://doi.org/10.1177/1054773819846024.

[45] CreateVR. *Spider-Man: Far From Home Virtual Reality*. Game [VR], June 2019. Sony Pictures Virtual Reality. https://www.oculus.com/experiences/rift/2190323657748572/.

[46] John W Creswell. *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications, 2017.

[47] John W Creswell and Vicki L Plano Clark. *Designing and conducting mixed methods research*. Sage publications, Thousand Oaks, California, United States, 2017.

[48] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. Surround-screen projection-based virtual reality: the design and

implementation of the cave. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, (SIGGRAPH '93), page 135–142, New York, NY, USA, 1993. Association for Computing Machinery. ISBN 0897916018. DOI 10.1145/166117.166134. URL https://doi.org/10.1145/166117.166134.

[49] Conor Cunningham, Roger O' Sullivan, Paolo Caserotti, and Mark A. Tully. Consequences of physical inactivity in older adults: A systematic review of reviews and meta-analyses. *Scandinavian Journal of Medicine & Science in Sports*, 30(5):816–827, 2020. DOI 10.1111/sms.13616. URL https://doi.org/10.1111/sms.13616.

[50] Yvonne AW De Kort, Wijnand A IJsselsteijn, and Karolien Poels. Digital games as social presence technology: Development of the social presence in gaming questionnaire (spgq). *Proceedings of PRESENCE*, 195203:1–9, 2007. URL https://www.researchgate.net/publication/228091680_Digital_Games_as_Social_Presence_Technology_Development_of_the_Social_Presence_in_Gaming_Questionnaire_SPGQ.

[51] Ansgar E. Depping and Regan L. Mandryk. Cooperation and interdependence: How multiplayer games increase social closeness. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY '17), page 449–461, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450348980. DOI 10.1145/3116595.3116639. URL https://doi.org/10.1145/3116595.3116639.

[52] Ansgar E. Depping, Colby Johanson, and Regan L. Mandryk. Designing for friendship: Modeling properties of play, in-game social capital, and psychological well-being. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY '18), page 87–100, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356244. DOI 10.1145/3242671.3242702. URL https://doi.org/10.1145/3242671.3242702.

[53] Kristen K. Dergousoff, Regan L. Mandryk, and Kathrin Gerling. Is movement better? Comparing sedentary and motion-based game controls for older adults. 2013. URL https://repository.lincoln.ac.uk/articles/conference_contribution/Is_movement_better_Comparing_sedentary_and_motion-based_game_controls_for_older_adults/25166216.

[54] Heather Desurvire, Martin Caplan, and Jozsef A. Toth. Using heuristics to evaluate the playability of games. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, (CHI EA '04),

page 1509–1512, New York, NY, USA, 2004. Association for Computing Machinery. ISBN 1581137036. DOI 10.1145/985921.986102. URL https://doi.org/10.1145/985921.986102.

[55] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. From game design elements to gamefulness: defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, (MindTrek '11), page 9–15, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450308168. DOI 10.1145/2181037.2181040. URL https://doi.org/10.1145/2181037.2181040.

[56] Andrew T. Dilanchian, Ronald Andringa, and Walter R. Boot. A pilot study exploring age differences in presence, workload, and cybersickness in the experience of immersive virtual reality environments. *Frontiers in Virtual Reality*, 2, 2021. ISSN 2673-4192. DOI 10.3389/frvir.2021.736793. URL https://doi.org/10.3389/frvir.2021.736793.

[57] Mary Dixon-Woods. Using framework-based synthesis for conducting reviews of qualitative studies. *BMC medicine*, 9:1–2, 2011. DOI 10.1186/1741-7015-9-39. URL https://doi.org/10.1186/1741-7015-9-39.

[58] Benjamin D. Douglas, Patrick J. Ewell, and Markus Brauer. Data quality in online human-subjects research: Comparisons between mturk, prolific, cloudresearch, qualtrics, and sona. *PLOS ONE*, 18(3):1–17, 2023. DOI 10.1371/journal.pone.0279720. URL https://doi.org/10.1371/journal.pone.0279720.

[59] Jeff Dyck, David Pinelle, Barry AT Brown, and Carl Gutwin. Learning from games: Hci design innovations in entertainment software. In *Graphics interface*, volume 2003, pages 237–246. Citeseer, 2003. DOI 10.20380/GI2003.28.

[60] Mahzar Eisapour, Shi Cao, Laura Domenicucci, and Jennifer Boger. Participatory design of a virtual reality exercise for people with mild cognitive impairment. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, (CHI EA '18), page 1–9, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356213. DOI 10.1145/3170427.3174362. URL https://doi.org/10.1145/3170427.3174362.

[61] Enhanced Well-being (WEL) and Health Promotion (HPR). Geneva charter for well-being, December 2021. URL https://www.who.int/publications/m/item/the-geneva-charter-for-well-being. Accessed: 15.03.2025.

[62] Entertainment Software Association. 2024 essential facts about the u.s. video game industry, 2024. https://www.theesa.com/wp-content/uploads/2024/05/Essential-Facts-2024-FINAL.pdf. Accessed: 24.12.2024.

[63] European Commission. Horizon 2020 - The Framework Programme for Research and Innovation, 2011. URL https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0808:FIN:en:PDF. https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-2020_en.

[64] Peer Eyal, Rothschild David, Gordon Andrew, Evernden Zak, and Damer Ekaterina. Data quality of platforms and panels for online behavioral research. *Behavior research methods*, pages 1–20, 2021. DOI 10.3758/s13428-021-01694-3. URL https://doi.org/10.3758/s13428-021-01694-3.

[65] Nuša Faric, Henry W W Potts, Adrian Hon, Lee Smith, Katie Newby, Andrew Steptoe, and Abi Fisher. What players of virtual reality exercise games want: Thematic analysis of web-based reviews. *J Med Internet Res*, 21(9):e13833, 2019. ISSN 1438-8871. DOI 10.2196/13833. URL https://doi.org/10.2196/13833.

[66] Andy P Field, Jeremy Miles, and Zoë Field. *Discovering statistics using R*. Sage, 2012.

[67] Flying Mollusk. *Nevermind*. Game [PC], September 2015. URL https://store.steampowered.com/app/342260/Nevermind/. Flying Mollusk, Los Angeles, CA, United States.

[68] Lawrence Hoc Nang Fong, Long Wai. Lam, and Rob Law. How locus of control shapes intention to reuse mobile apps for making hotel reservations: Evidence from chinese consumers. *Tourism Management*, 61:331–342, 2017. DOI 10.1016/j.tourman.2017.03.002. URL https://doi.org/10.1016/j.tourman.2017.03.002.

[69] Guo Freeman and Dane Acena. Hugging from a distance: Building interpersonal relationships in social virtual reality. In *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences*, (IMX '21), page 84–95, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450383899. DOI 10.1145/3452918.3458805. URL https://doi.org/10.1145/3452918.3458805.

[70] Tracy Fullerton. *Game design workshop: a playcentric approach to creating innovative games*. CRC press, 2014.

[71] Borko Furht, editor. *Immersive Virtual Reality*, pages 345–346. Springer US, Boston, MA, 2008. ISBN 978-0-387-78414-4. DOI 10.1007/978-0-387-78414-4_85. URL https://doi.org/10.1007/978-0-387-78414-4_85.

[72] Alessandro Gabbiadini, Cristina Baldissarri, Federica Durante, Roberta Rosa Valtorta, Maria De Rosa, and Marcello Gallucci. Together apart: The mitigating role of digital communication technologies on negative affect during the covid-19 outbreak in italy. *Frontiers in Psychology*, 11, 2020. ISSN 1664-1078. DOI 10.3389/fpsyg.2020.554678. URL https://doi.org/10.3389/fpsyg.2020.554678.

[73] Hong Gao, Efe Bozkir, Lisa Hasenbein, Jens-Uwe Hahn, Richard Göllner, and Enkelejda Kasneci. Digital transformations of classrooms in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, (CHI '21), New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380966. DOI 10.1145/3411764.3445596. URL https://doi.org/10.1145/3411764.3445596.

[74] Azucena Garcia-Palacios, Hunter Hoffman, Albert Carlin, Thomas A Furness III, and Cristina Botella. Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour Research and Therapy*, 40(9):983–993, 2002. ISSN 0005-7967. DOI 10.1016/S0005-7967(01)00068-7. URL https://doi.org/10.1016/S0005-7967(01)00068-7.

[75] Kathrin Gerling, Patrick Dickinson, Kieran Hicks, Liam Mason, Adalberto L. Simeone, and Katta Spiel. Virtual reality games for people using wheelchairs. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, (CHI '20), page 1–11, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. DOI 10.1145/3313831.3376265. URL https://doi.org/10.1145/3313831.3376265.

[76] Kathrin Maria Gerling, Jonas Schild, and Maic Masuch. Exergame design for elderly users: The case study of silverbalance. In *Proceedings of the 7th International Conference on Advances in Computer Entertainment Technology*, (ACE '10), page 66–69, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605588636. DOI 10.1145/1971630.1971650. URL https://doi.org/10.1145/1971630.1971650.

[77] Maria J. Grant and Andrew Booth. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2):91–108, 2009. DOI 10.1111/j.1471-1842.2009.00848.x. URL https://doi.org/10.1111/j.1471-1842.2009.00848.x.

[78] Rachel Grieve, Michaelle Indian, Kate Witteveen, G. Anne Tolan, and Jessica Marrington. Face-to-face or facebook: Can social connectedness be derived online? *Computers in Human Behavior*, 29 (3):604–609, 2013. ISSN 0747-5632. DOI 10.1016/j.chb.2012.11.017. URL https://doi.org/10.1016/j.chb.2012.11.017.

[79] Kristen Grinyer, Sara Czerwonka, Adrian Alvarez, Victoria McArthur, Audrey Girouard, and Robert J. Teather. Massively multiplayer online role-playing games on promoting social well-being in the covid-19 pandemic. In *Proceedings of the 17th International Conference on the Foundations of Digital Games*, (FDG '22), New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450397957. DOI 10.1145/3555858.3555871. URL https://doi.org/10.1145/3555858.3555871.

[80] Timothy C Guetterman. Basics of statistics for primary care research. *Family medicine and community health*, 7(2), 2019. DOI 10.1136/fmch-2018-000067. URL https://doi.org/10.1136/fmch-2018-000067.

[81] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. Sharevr: Enabling co-located experiences for virtual reality between hmd and non-hmd users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, (CHI '17), page 4021–4033, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450346559. DOI 10.1145/3025453.3025683. URL https://doi.org/10.1145/3025453.3025683.

[82] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. Facedisplay: Towards asymmetric multi-user interaction for nomadic virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, (CHI '18), page 1–13, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356206. DOI 10.1145/3173574.3173628. URL https://doi.org/10.1145/3173574.3173628.

[83] Stuart Hallifax, Maximilian Altmeyer, Kristina Kölln, Maria Rauschenberger, and Lennart E. Nacke. From points to progression: A scoping review of game elements in gamification research with a content analysis of 280 research papers. *Proceedings of the ACM on Human-Computer Interaction*, 7 (CHI PLAY), 2023. DOI 10.1145/3611048. URL https://doi.org/10.1145/3611048.

[84] Juho Hamari, Jonna Koivisto, and Harri Sarsa. Does gamification work? – a literature review of empirical studies on gamification. In *2014 47th Hawaii International Conference on System Sciences*, pages 3025–3034, 2014. DOI 10.1109/HICSS.2014.377. URL https://doi.org/10.1109/HICSS.2014.377.

[85] Yu Han, Yu Miao, Jie Lu, Mei Guo, and Yi Xiao. Exploring intervention strategies for distracted students in vr classrooms. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, (CHI EA '22), New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391566. DOI 10.1145/3491101.3519627. URL https://doi.org/10.1145/3491101.3519627.

[86] John Harris and Mark Hancock. To asymmetry and beyond! improving social connectedness by increasing designed interdependence in cooperative play. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, (CHI '19), page 1–12, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. DOI 10.1145/3290605.3300239. URL https://doi.org/10.1145/3290605.3300239.

[87] John Harris, Mark Hancock, and Stacey D. Scott. Leveraging asymmetries in multiplayer games: Investigating design elements of interdependent play. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY '16), page 350–361, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450344562. DOI 10.1145/2967934.2968113. URL https://doi.org/10.1145/2967934.2968113.

[88] Andrew F Hayes. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications, 2017.

[89] David Healy, Aisling Flynn, Owen Conlan, Jenny McSharry, and Jane Walsh. Older adults' experiences and perceptions of immersive virtual reality: Systematic review and thematic synthesis. *JMIR Serious Games*, 10(4):e35802, 2022. ISSN 2291-9279. DOI 10.2196/35802. URL https://doi.org/10.2196/35802.

[90] Patricia Heyn, Beatriz C Abreu, and Kenneth J Ottenbacher. The effects of exercise training on elderly persons with cognitive impairment and dementia: a meta-analysis. *Archives of physical medicine and rehabilitation*, 85(10):1694–1704, 2004. DOI 10.1016/j.apmr.2004.03.019. URL https://doi.org/10.1016/j.apmr.2004.03.019.

[91] Teresa Hirzle, Maurice Cordts, Enrico Rukzio, Jan Gugenheimer, and Andreas Bulling. A critical assessment of the use of ssq as a measure of general discomfort in vr head-mounted displays. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, (CHI '21), New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380966. DOI 10.1145/3411764.3445361. URL https://doi.org/10.1145/3411764.3445361.

[92] Teresa Hirzle, Florian Müller, Fiona Draxler, Martin Schmitz, Pascal Knierim, and Kasper Hornbæk. When xr and ai meet - a scoping review on extended reality and artificial intelligence. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. DOI 10.1145/3544548.3581072. URL https://doi.org/10.1145/3544548.3581072.

[93] Anne Hösch. *Simulator Sickness in Fahrsimulationsumgebungen-drei Studien zu Human Factors*. PhD thesis, Technische Universität Ilmenau, 2018. URL https://www.db-thueringen.de/receive/dbt_mods_00037865.

[94] Samory Houzangbe, Olivier Christmann, Geoffrey Gorisse, and Simon Richir. Fear as a biofeedback game mechanic in virtual reality: Effects on engagement and perceived usability. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*, (FDG '18), New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450365710. DOI 10.1145/3235765.3235787. URL https://doi.org/10.1145/3235765.3235787.

[95] Samory Houzangbe, Olivier Christmann, Geoffrey Gorisse, and Simon Richir. Effects of voluntary heart rate control on user engagement and agency in a virtual reality game. *Virtual Reality*, 24:665–681, 2020. DOI 10.1007/s10055-020-00429-7. URL https://doi.org/10.1145/3235765.3235787.

[96] Dennis Howitt and Duncan Cramer. *Introduction to research methods in psychology*. Pearson Education, 2007. ISBN 978-0-273-73499-4.

[97] Robin Hunicke, Marc LeBlanc, and Robert Zubek. MDA: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI*, volume 4, page 1722, Palo Alto, CA, USA, 2004. AAAI Press. URL https://cdn.aaai.org/Workshops/2004/WS-04-04/WS04-04-001.pdf.

[98] Wijnand Ijsselsteijn, Henk Herman Nap, Yvonne de Kort, and Karolien Poels. Digital game design for elderly users. In *Proceedings of the 2007 Conference on Future Play*, (Future Play '07), page 17–22, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595939432. DOI 10.1145/1328202.1328206. URL https://doi.org/10.1145/1328202.1328206.

[99] International Organization for Standardization. Iso 9241-210:2019 ergonomics of human-system interaction — part 210: Human-centred design for interactive systems, July 2019. URL https://

www.iso.org/standard/77520.html. International Organization for Standardization. Geneva, Switzerland.

[100] Christos Ioannou, Patrick Archard, Eamonn O'Neill, and Christof Lutteroth. Virtual performance augmentation in an immersive jump & run exergame. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, (CHI '19), page 1–15, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. DOI 10.1145/3290605.3300388. URL https://doi.org/10.1145/3290605.3300388.

[101] Aki Sakari Järvinen. *Games without Frontiers.: Theories and Methods for Game Studies and Design*. PhD thesis, 2009.

[102] R. Burke Johnson, Anthony J. Onwuegbuzie, and Lisa A. Turner. Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2):112–133, 2007. DOI 10.1177/1558689806298224. URL https://doi.org/10.1177/1558689806298224.

[103] Raine Kajastila, Leo Holsti, and Perttu Hämäläinen. Empowering the exercise: a body-controlled trampoline training game. *International Journal of Computer Science in Sport*, 1(13):6–23, 2014. URL https://users.aalto.fi/~hamalap5/publications/IJCSS_2014_personal_copy_Kajastila.pdf.

[104] Sukran Karaosmanoglu. Don't sweat it!: Exploring asymmetrical biometric feedback in collaborative gaming. Master's thesis, Universität Ulm, 2019.

[105] Tuomas Kari and Markus Makkonen. Explaining the usage intentions of exergames. In *International Conference on Information Systems*. Association for Information Systems (AIS), 2014. URL https://aisel.aisnet.org/icis2014/proceedings/HumanBehavior/39/.

[106] Linda K. Kaye. Exploring flow experiences in cooperative digital gaming contexts. *Computers in Human Behavior*, 55:286–291, 2016. DOI 10.1016/j.chb.2015.09.023. URL https://doi.org/10.1016/j.chb.2015.09.023.

[107] Kevin Keeker, Randy Pagulayan, Jonathan Sykes, and Nicole Lazzaro. The untapped world of video games. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, (CHI EA '04), page 1610–1611, New York, NY, USA, 2004. Association for Computing Machinery. ISBN 1581137036. DOI 10.1145/985921.986170. URL https://doi.org/10.1145/985921.986170.

[108] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. Simulator sickness questionnaire: An en-

hanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, 1993. DOI 10.1207/s15327108ijap0303_3. URL https://doi.org/10.1207/s15327108ijap0303_3.

[109] Hanan Khalil, Micah Peters, Christina M. Godfrey, Patricia McInerney, Cassia Baldini Soares, and Deborah Parker. An evidence-based approach to scoping reviews. *Worldviews on Evidence-Based Nursing*, 13(2):118–123, 2016. DOI 10.1111/wvn.12144. URL https://doi.org/10.1111/wvn.12144.

[110] Sebastian Klöß. Die zukunft der consumer technology 2021: Marktentwicklung & mediennutzung, trends & technologien, 2021. URL https://www.bitkom.org/sites/main/files/2021-09/210817_ct_studie_2021.pdf.

[111] Jonna Koivisto and Juho Hamari. The rise of motivational information systems: A review of gamification research. *International Journal of Information Management*, 45:191–210, 2019. ISSN 0268-4012. DOI 10.1016/j.ijinfomgt.2018.10.013. URL https://doi.org/10.1016/j.ijinfomgt.2018.10.013.

[112] Tanja Kojić, Lan Thao Nugyen, and Jan-Niklas Voigt-Antons. Impact of constant visual biofeedback on user experience in virtual reality exergames. In *2019 IEEE International Symposium on Multimedia*, (ISM '19), pages 307–3073, Los Alamitos, CA, USA, 2019. IEEE. DOI 10.1109/ISM46123.2019.00068. URL https://doi.org/10.1109/ISM46123.2019.00068.

[113] Bastian Kordyaka, Samuli Laato, Sebastian Weber, Juho Hamari, and Bjoern Niehaves. Exploring the association between engagement with location-based game features and getting inspired about environmental issues and nature. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, (CHI '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703300. DOI 10.1145/3613904.3642786. URL https://doi.org/10.1145/3613904.3642786.

[114] Robert Kraut, Michael Patterson, Vicki Lundmark, Sara Kiesler, Tridas Mukophadhyay, and William Scherlis. Internet paradox: A social technology that reduces social involvement and psychological well-being? *American psychologist*, 53(9):1017, 1998. DOI 10.1037/0003-066X.53.9.1017. URL https://doi.org/10.1037/0003-066X.53.9.1017.

[115] Dennis Krupke, Paul Lubos, Lena Demski, Jonas Brinkhoff, Gregor Weber, Fabian Willke, and Frank Steinicke. Control methods in a supernatural flight simulator. In *2016 IEEE Virtual Reality*, (IEEE VR '16), pages 329–329, 2016. DOI

10.1109/VR.2016.7504788. URL https://doi.org/10.1109/VR.2016.7504788.

[116] Simone Kühn, Tobias Gleich, Robert C Lorenz, Ulman Lindenberger, and Jürgen Gallinat. Playing super mario induces structural brain plasticity: gray matter changes resulting from training with a commercial video game. *Molecular psychiatry*, 19(2):265–271, 2014. DOI 10.1038/mp.2013.120. URL https://doi.org/10.1038/mp.2013.120.

[117] Eugene Kukshinov. It's (not) me: Dynamic nature of immersive experiences in video game play. *Human Studies*, pages 1–22, 2024. DOI 10.1007/s10746-024-09768-9. URL https://doi.org/10.1007/s10746-024-09768-9.

[118] Joseph J. LaViola. A discussion of cybersickness in virtual environments. *SIGCHI Bull.*, 32(1):47–56, 2000. ISSN 0736-6906. DOI 10.1145/333329.333344. URL https://doi.org/10.1145/333329.333344.

[119] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017.

[120] Lauri Lehtonen, Maximus D. Kaos, Raine Kajastila, Leo Holsti, Janne Karsisto, Sami Pekkola, Joni Vähämäki, Lassi Vapaakallio, and Perttu Hämäläinen. Movement empowerment in a multiplayer mixed-reality trampoline game. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY '19), page 19–29, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450366885. DOI 10.1145/3311350.3347181. URL https://doi.org/10.1145/3311350.3347181.

[121] Ting Li, Ye Yao, Yan Cheng, Bing Xu, Xinyi Cao, David Waxman, Wei Feng, Yuan Shen, Qingwei Li, Jijun Wang, et al. Cognitive training can reduce the rate of cognitive aging: a neuroimaging cohort study. *BMC Geriatrics*, 16(1):1–12, 2016. DOI 10.1186/s12877-016-0194-5. URL https://doi.org/10.1186/s12877-016-0194-5.

[122] Suk Ling Ma, Katsie Tung Tang, Niki Cheuk Ting Lau, Clement Lek Hin Chiu, Cuichan Lin, Linda Chiu Wa Lam, and Allen Ting Chun Lee. Effect of computerized cognitive training on mood, cognition, and serum brain-derived neurotrophic factor level in late-life depression — a pilot randomized controlled trial. *Frontiers in Psychiatry*, 14, 2024. ISSN 1664-0640. DOI 10.3389/fpsyt.2023.1287822. URL https://doi.org/10.3389/fpsyt.2023.1287822.

[123] Jake Mace. 7 tai chi moves for beginners | 15 minute daily taiji routine, 2018. URL https://www.youtube.com/watch?v=DmlsB_KOGtM.

[124] Rainer Malaka. How computer games can improve your health and fitness. In Stefan Göbel and Josef Wiemeyer, editors, *Games for Training, Education, Health and Sports*, pages 1–7, Cham, 2014. Springer International Publishing. ISBN 978-3-319-05972-3. DOI 10.1007/978-3-319-05972-3_1. URL https://doi.org/10.1007/978-3-319-05972-3_1.

[125] Regan L. Mandryk, M. Stella Atkins, and Kori M. Inkpen. A continuous and objective evaluation of emotional experience with interactive play environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (CHI '06), page 1027–1036, New York, NY, USA, 2006. Association for Computing Machinery. ISBN 1595933727. DOI 10.1145/1124772.1124926. URL https://doi.org/10.1145/1124772.1124926.

[126] Anna Lisa Martin-Niedecken and Elisa D. Mekler. The exercube: Participatory design of an immersive fitness game environment. In Stefan Göbel, Augusto Garcia-Agundez, Thomas Tregel, Minhua Ma, Jannicke Baalsrud Hauge, Manuel Oliveira, Tim Marsh, and Polona Caserman, editors, *Serious Games*, pages 263–275, Cham, 2018. Springer International Publishing. ISBN 978-3-030-02762-9. DOI 10.1007/978-3-030-02762-9_28. URL https://doi.org/10.1007/978-3-030-02762-9_28.

[127] Anna Lisa Martin-Niedecken, Katja Rogers, Laia Turmo Vidal, Elisa D. Mekler, and Elena Márquez Segura. Exercube vs. personal trainer: Evaluating a holistic, immersive, and adaptive fitness game setup. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, (CHI '19), page 1–15, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. DOI 10.1145/3290605.3300318. URL https://doi.org/10.1145/3290605.3300318.

[128] Liam Mason, Kathrin Gerling, Patrick Dickinson, Jussi Holopainen, Lisa Jacobs, and Kieran Hicks. Including the experiences of physically disabled players in mainstream guidelines for movement-based games. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, (CHI '22), New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391573. DOI 10.1145/3491102.3501867. URL https://doi.org/10.1145/3491102.3501867.

[129] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Sys-*

tems, 77(12):1321–1329, 1994. URL https://search.ieice.org/bin/summary.php?id=e77-d_12_1321.

[130] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. Spie, 1995. DOI 10.1117/12.197321. URL https://doi.org/10.1117/12.197321.

[131] Sahar Mirhadi, Ioanna Iacovides, and Alena Denisova. Playing through tough times: Exploring the relationship between game aspects and coping strategies during difficult life challenges. *Proceedings of the ACM on Human-Computer Interaction*, 8, 2024. DOI 10.1145/3677097. URL https://doi.org/10.1145/3677097.

[132] Briar Lee Mitchell. *Game design essentials*. John Wiley & Sons, 2012.

[133] Karyn Moffatt. Older-adult hci: why should we care? *Interactions*, 20(4):72–75, 2013. ISSN 1072-5520. DOI 10.1145/2486227.2486242. URL https://doi.org/10.1145/2486227.2486242.

[134] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, and Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *PLoS medicine*, 6(7):e1000097, 2009. DOI 10.1371/journal.pmed.1000097. URL https://doi.org/10.1371/journal.pmed.1000097.

[135] Antonio Moreno-Llamas, Jesús García-Mayor, and Ernesto De la Cruz-Sánchez. The impact of digital technology development on sitting time across europe. *Technology in Society*, 63:101406, 2020. ISSN 0160-791X. DOI 10.1016/j.techsoc.2020.101406. URL https://doi.org/10.1016/j.techsoc.2020.101406.

[136] David L Morgan, Richard A Krueger, and Jean A King. *The focus group guidebook*. Sage, 1998. DOI 10.4135/9781483328164. URL https://doi.org/10.4135/9781483328164.

[137] Anna Moschny, Petra Platen, Renate Klaaßen-Mielke, Ulrike Trampisch, and Timo Hinrichs. Barriers to physical activity in older adults in germany: a cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity*, 8(1):1–10, 2011. DOI 10.1186/1479-5868-8-121. URL https://doi.org/10.1186/1479-5868-8-121.

[138] Fariba Mostajeran, Frank Steinicke, Oscar Javier Ariza Nunez, Dimitrios Gatsios, and Dimitrios Fotiadis. Augmented reality for older adults: Exploring acceptability of virtual coaches for home-based balance training in an aging population. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing*

Systems, (CHI '20), page 1–12, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. DOI 10.1145/3313831.3376565. URL https://doi.org/10.1145/3313831.3376565.

[139] Florian Mueller and Katherine Isbister. Movement-based game guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (CHI '14), page 2191–2200, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450324731. DOI 10.1145/2556288.2557163. URL https://doi.org/10.1145/2556288.2557163.

[140] Florian Mueller, Rohit Ashok Khot, Kathrin Gerling, and Regan Mandryk. Exertion games. *Foundations and Trends® in Human–Computer Interaction*, 10(1):1–86, 2016. ISSN 1551-3955. DOI 10.1561/1100000041. URL http://dx.doi.org/10.1561/1100000041.

[141] Florian 'Floyd' Mueller, Darren Edge, Frank Vetere, Martin R. Gibbs, Stefan Agamanolis, Bert Bongers, and Jennifer G. Sheridan. Designing sports: A framework for exertion games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (CHI '11), page 2651–2660, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450302289. DOI 10.1145/1978942.1979330. URL https://doi.org/10.1145/1978942.1979330.

[142] Zachary Munn, Micah DJ Peters, Cindy Stern, Catalin Tufanaru, Alexa McArthur, and Edoardo Aromataris. Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach. *BMC medical research methodology*, 18:1–7, 2018. DOI 10.1186/s12874-018-0611-x. URL https://doi.org/10.1186/s12874-018-0611-x.

[143] John Muñoz, Samira Mehrabi, Yirou Li, Aysha Basharat, Laura E Middleton, Shi Cao, Michael Barnett-Cowan, and Jennifer Boger. Immersive virtual reality exergames for persons living with dementia: User-centered design study as a multistakeholder team during the covid-19 pandemic. *JMIR Serious Games*, 10 (1):e29987, 2022. ISSN 2291-9279. DOI 10.2196/29987. URL https://doi.org/10.2196/29987.

[144] Joschka Mütterlein and Thomas Hess. Immersion, presence, interactivity: Towards a joint understanding of factors influencing virtual reality acceptance and use. 2017.

[145] Lennart Nacke, Anders Drachen, Kai Kuikkaniemi, Joerg Niesenhaus, Hannu J Korhonen, Wouter M Hoogen, Karolien Poels, Wijnand A IJsselsteijn, and Yvonne AW De Kort. Playability and

player experience research. In *Proceedings of digra 2009: Breaking new ground: Innovation in games, play, practice and theory*. DiGRA, 2009. URL https://urn.kb.se/resolve?urn=urn:nbn:se:bth-7956.

[146] Namco and Bandai Namco Studios. *Pac-Man*. Game, May 1980. https://en.bandainamcoent.eu/games. Accessed: 10.01.2025.

[147] Meng Ni, Kiersten Mooney, Anoop Balachandran, Luca Richards, Kysha Harriell, and Joseph F. Signorile. Muscle utilization patterns vary by skill levels of the practitioners across specific yoga poses (asanas). *Complementary Therapies in Medicine*, 22(4):662–669, 2014. ISSN 0965-2299. DOI https://doi.org/10.1016/j.ctim.2014.06.006.

[148] Robert C Nickerson, Upkar Varshney, and Jan Muntermann. A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, 22(3):336–359, 2013. DOI 10.1057/ejis.2012.26. URL https://doi.org/10.1057/ejis.2012.26.

[149] Nintendo EAD. *Wii Fit*. Game [Wii], December 2007.

[150] Pedro A Nogueira, Vasco Torres, Rui Rodrigues, Eugénio Oliveira, and Lennart E Nacke. Vanishing scares: biofeedback modulation of affective player experiences in a procedural horror game. *Journal on Multimodal User Interfaces*, 10(1):31–62, 2016. DOI 10.1007/s12193-015-0208-1. URL https://doi.org/10.1007/s12193-015-0208-1.

[151] Don Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[152] Odders Lab. *OhShape*. Game [VR], December 2019. Odders Lab. Seville, Spain. https://ohshapevr.com/.

[153] Daniel M. Oppenheimer, Tom Meyvis, and Nicolas Davidenko. Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45(4):867–872, 2009. ISSN 0022-1031. DOI 10.1016/j.jesp.2009.03.009. URL https://doi.org/10.1016/j.jesp.2009.03.009.

[154] World Health Organization. Dementia, March 2023. URL https://www.who.int/news-room/fact-sheets/detail/dementia. Accessed: 18.01.2025.

[155] World Health Organization. Promoting well-being, March 2025. URL https://www.who.int/activities/promoting-well-being. Accessed: 15.03.2025.

[156] World Health Organization. Who commission on social connection, March 2025. URL https://www.who.int/groups/commission-on-social-connection. Accessed: 15.03.2025.

[157] Yang-Kun Ou, Yu-Lin Wang, Hua-Cheng Chang, Shih-Yin Yen, Yu-Hua Zheng, and Bih-O. Lee. Development of virtual reality rehabilitation games for children with attention-deficit hyperactivity disorder. *Journal of Ambient Intelligence and Humanized Computing*, 11(11):5713 – 5720, 2020. DOI 10.1007/s12652-020-01945-9. URL https://doi.org/10.1007/s12652-020-01945-9.

[158] A Ant Ozok. Survey design and implementation in hci. In *The human-computer interaction handbook*, pages 1177–1196. CRC Press, 2007. DOI 10.1201/b11963. URL https://doi.org/10.1201/b11963.

[159] Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, Roger Chou, Julie Glanville, Jeremy M Grimshaw, Asbjørn Hróbjartsson, Manoj M Lalu, Tianjing Li, Elizabeth W Loder, Evan Mayo-Wilson, Steve McDonald, Luke A McGuinness, Lesley A Stewart, James Thomas, Andrea C Tricco, Vivian A Welch, Penny Whiting, and David Moher. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, 2021. DOI 10.1136/bmj.n71. URL https://doi.org/10.1136/bmj.n71.

[160] Federica Pallavicini and Alessandro Pepe. Comparing player experience in video games played in virtual reality or on desktop displays: Immersion, flow, and positive emotions. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, (CHI PLAY '19 Extended Abstracts), page 195–210, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368711. DOI 10.1145/3341215.3355736. URL https://doi.org/10.1145/3341215.3355736.

[161] Federica Pallavicini, Alessandro Pepe, and Maria Eleonora Minissi. Gaming in virtual reality: What changes in terms of usability, emotional response and sense of presence compared to non-immersive video games? *Simulation & Gaming*, 50(2): 136–159, 2019. DOI 10.1177/1046878119831420. URL https://doi.org/10.1177/1046878119831420.

[162] Seong-Hi Park, Kuem Sun Han, and Chang-Bum Kang. Effects of exercise programs on depressive symptoms, quality of life, and self-esteem in older people: A systematic review of randomized controlled trials. *Applied Nursing Research*, 27(4):219–226, 2014.

ISSN 0897-1897. DOI 10.1016/j.apnr.2014.01.004. URL https://doi.org/10.1016/j.apnr.2014.01.004.

[163] Russell R Pate, Ruth P Saunders, Jennifer R O'Neill, and Marsha Dowda. Overcoming barriers to physical activity: helping youth be more active. *ACSM's Health & Fitness Journal*, 15(1):7–12, 2011. DOI 10.1249/FIT.0b013e318201c8ee. URL https://doi.org/10.1249/FIT.0b013e318201c8ee.

[164] Jorge Perez-Gomez and JA Calbet. Training methods to improve vertical jump performance. *The Journal of sports medicine and physical fitness*, 53(4):339–357, 2013.

[165] Sebastian A. C. Perrig, Nicolas Scharowski, Florian Brühlmann, Nick von Felten, Klaus Opwis, and Lena Fanya Aeschbach. Independent validation of the player experience inventory: Findings from a large set of video game players. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, (CHI '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703300. DOI 10.1145/3613904.3642270. URL https://doi.org/10.1145/3613904.3642270.

[166] Micah DJ Peters, Christina Godfrey, Patricia McInerney, Hanan Khalil, Palle Larsen, Casey Marnie, Danielle Pollock, Andrea C Tricco, and Zachary Munn. Best practice guidance and reporting items for the development of scoping review protocols. *JBI evidence synthesis*, 20(4):953–968, 2022. DOI 10.11124/JBIES-21-00242. URL https://doi.org/10.11124/JBIES-21-00242.

[167] Catherine Pope, Sue Ziebland, and Nicholas Mays. Analysing qualitative data. *BMJ*, 320(7227):114–116, 2000. ISSN 0959-8138. DOI 10.1136/bmj.320.7227.114. URL https://doi.org/10.1136/bmj.320.7227.114.

[168] Niklas Ravaja, Timo Saari, Marko Turpeinen, Jari Laarni, Mikko Salminen, and Matias Kivikangas. Spatial Presence and Emotions during Video Game Playing: Does It Matter with Whom You Play? *Presence: Teleoperators and Virtual Environments*, 15(4):381–392, 2006. DOI 10.1162/pres.15.4.381. URL https://doi.org/10.1162/pres.15.4.381.

[169] Anke V. Reinschluessel, Thomas Muender, Daniela Salzmann, Tanja Döring, Rainer Malaka, and Dirk Weyhe. Virtual reality for surgical planning – evaluation based on two liver tumor resections. *Frontiers in Surgery*, 9, 2022. ISSN 2296-875X. DOI 10.3389/fsurg.2022.821060. URL https://doi.org/10.3389/fsurg.2022.821060.

[170] Anke Verena Reinschluessel, Joern Teuber, Marc Herrlich, Jeffrey Bissel, Melanie van Eikeren, Johannes Ganser, Felicia Koeller,

Fenja Kollasch, Thomas Mildner, Luca Raimondo, Lars Reisig, Marc Ruedel, Danny Thieme, Tobias Vahl, Gabriel Zachmann, and Rainer Malaka. Virtual reality for user-centered design and evaluation of touch-free interaction techniques for navigating medical images in the operating room. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, (CHI EA '17), page 2001–2009, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450346566. DOI 10.1145/3027063.3053173. URL https://doi.org/10.1145/3027063.3053173.

[171] Daniel Reis, Pedro Pais, David Gonçalves, Kathrin Gerling, and André Rodrigues. Exploring asymmetry of information in co-operative games. In Anabela Marto, Rui Prada, Patrícia Gouveia, Ruth Contreras Espinosa, Alexandrino Gonçalves, Eduarda Abrantes, and Roberto Ribeiro, editors, *Videogame Sciences and Arts*, pages 155–170, Cham, 2025. Springer Nature Switzerland. ISBN 978-3-031-81713-7. DOI 10.1007/978-3-031-81713-7_11. URL https://doi.org/10.1007/978-3-031-81713-7_11.

[172] H. Rex Hartson. Human–computer interaction: Interdisciplinary roots and trends. *Journal of Systems and Software*, 43(2):103–118, 1998. ISSN 0164-1212. DOI 10.1016/S0164-1212(98)10026-2. URL https://doi.org/10.1016/S0164-1212(98)10026-2.

[173] Renato Alexandre Ribeiro, Inês Gonçalves, Manuel Piçarra, Letícia Seixas Pereira, Carlos Duarte, André Rodrigues, and João Guerreiro. Investigating virtual reality locomotion techniques with blind people. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, (CHI '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703300. DOI 10.1145/3613904.3642088. URL https://doi.org/10.1145/3613904.3642088.

[174] Nathalie Ringgenberg, Sarah Mildner, Marcia Hapig, Sarah Hermann, Katharina Kruszewski, Anna Lisa Martin-Niedecken, Katja Rogers, Alexandra Schättin, Frank Behrendt, Sonja Böckler, et al. Exerg: adapting an exergame training solution to the needs of older adults using focus group and expert interviews. *Journal of NeuroEngineering and Rehabilitation*, 19(1):89, 2022. DOI 10.1186/s12984-022-01063-x. URL https://doi.org/10.1186/s12984-022-01063-x.

[175] Sebastian Rings, Frank Steinicke, Tobias Picker, and Caspar Prasuhn. Enabling patients with neurological diseases to perform motor-cognitive exergames under clinical supervision for everyday usage. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, (IEEE VR VRW '20), pages 360–364, New York, USA, 2020. IEEE. DOI

10.1109/VRW50115.2020.00078. URL https://doi.org/10.1109/VRW50115.2020.00078.

[176] Sebastian Rings, Frank Steinicke, Tobias Picker, and Caspar Prasuhn. Memory journalist: Creating virtual reality exergames for the treatment of older adults with dementia. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, (IEEE VR VRW '20), pages 686–687, New York, USA, 2020. IEEE. DOI 10.1109/VRW50115.2020.00194. URL https://doi.org/10.1109/VRW50115.2020.00194.

[177] Raquel Breejon Robinson, Elizabeth Reid, James Collin Fey, Ansgar E. Depping, Katherine Isbister, and Regan L. Mandryk. Designing and evaluating 'in the same boat', a game of embodied synchronization for enhancing social play. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, (CHI '20), page 1–14, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450367080. DOI 10.1145/3313831.3376433. URL https://doi.org/10.1145/3313831.3376433.

[178] Katja Rogers and Michael Weber. Audio habits and motivations in video game players. In *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound*, (AM '19), page 45–52, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450372978. DOI 10.1145/3356590.3356599. URL https://doi.org/10.1145/3356590.3356599.

[179] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, 30(4):344–360, 2006. DOI 10.1007/s11031-006-9051-8. URL https://doi.org/10.1007/s11031-006-9051-8. The IMI items are available online under this link: http://selfdeterminationtheory.org/intrinsic-motivation-inventory/. Accessed: 05.03.2025.

[180] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoperators and Virtual Environments*, 10(3):266–281, 2001. DOI 10.1162/105474601300343603. URL https://doi.org/10.1162/105474601300343603.

[181] Marc Schubhan, Maximilian Altmeyer, Katja Rogers, Donald Degraen, Pascal Lessel, and Antonio Krüger. Auditory, visual, or both? comparing visual and auditory representations of game elements in a gamified image-tagging task. *Proceedings of the ACM on Human-Computer Interaction*, 8 (CHI PLAY), 2024. DOI 10.1145/3677059. URL https://doi.org/10.1145/3677059.

[182] Valentin Schwind, Pascal Knierim, Nico Haas, and Niels Henze. Using presence questionnaires in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, (CHI '19), page 1–12, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. DOI 10.1145/3290605.3300590. URL https://doi.org/10.1145/3290605.3300590.

[183] Yasaman S. Sefidgar, Carla L. Castillo, Shaan Chopra, Liwei Jiang, Tae Jones, Anant Mittal, Hyeyoung Ryu, Jessica Schroeder, Allison Cole, Natalia Murinova, Sean A. Munson, and James Fogarty. Migrainetracker: Examining patient experiences with goal-directed self-tracking for a chronic health condition. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, (CHI '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703300. DOI 10.1145/3613904.3642075. URL https://doi.org/10.1145/3613904.3642075.

[184] L. Shamseer, D. Moher, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart and. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-p) 2015: elaboration and explanation. *BMJ*, 349(jan02 1):g7647–g7647, 2015. DOI 10.1136/bmj.g7647. URL https://doi.org/10.1136/bmj.g7647.

[185] Lindsay Alexander Shaw, Burkhard Claus Wuensche, Christof Lutteroth, Jude Buckley, and Paul Corballis. Evaluating sensory feedback for immersion in exergames. In *Proceedings of the Australasian Computer Science Week Multiconference*, (ACSW '17), New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450347686. DOI 10.1145/3014812.3014823. URL https://doi.org/10.1145/3014812.3014823.

[186] Richard Skarbez, Missie Smith, and Mary C. Whitton. Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2, 2021. ISSN 2673-4192. DOI 10.3389/frvir.2021.647997. URL https://doi.org/10.3389/frvir.2021.647997.

[187] Mel Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535): 3549–3557, 2009. DOI 10.1098/rstb.2009.0138. URL https://doi.org/10.1098/rstb.2009.0138.

[188] Mel Slater and Maria V. Sanchez-Vives. Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, 3, 2016. ISSN 2296-9144. DOI 10.3389/frobt.2016.00074. URL https://doi.org/10.3389/frobt.2016.00074.

[189] Mel Slater, Anthony Steed, John McCarthy, and Francesco Maringelli. The influence of body movement on subjective presence in virtual environments. *Human Factors*, 40(3):469–477, 1998. DOI 10.1518/001872098779591368. URL https://doi.org/10.1518/001872098779591368.

[190] Mel Slater, Domna Banakou, Alejandro Beacco, Jaime Gallego, Francisco Macia-Varela, and Ramon Oliva. A separate reality: An update on place illusion and plausibility in virtual reality. *Frontiers in Virtual Reality*, 3, 2022. ISSN 2673-4192. DOI 10.3389/frvir.2022.914392. URL https://doi.org/10.3389/frvir.2022.914392.

[191] Jan David Smeddinck, Marc Herrlich, and Rainer Malaka. Exergames for physiotherapy and rehabilitation: A medium-term situated study of motivational aspects and impact on functional reach. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, (CHI '15), page 4143–4146, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450331456. DOI 10.1145/2702123.2702598. URL https://doi.org/10.1145/2702123.2702598.

[192] Michael Smilovitch and Richard Lachman. Birdquestvr: A cross-platform asymmetric communication game. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, (CHI PLAY '19 Extended Abstracts), page 307–313, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368711. DOI 10.1145/3341215.3358246. URL https://doi.org/10.1145/3341215.3358246.

[193] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. What is mixed reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, (CHI '19), page 1–15, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. DOI 10.1145/3290605.3300767. URL https://doi.org/10.1145/3290605.3300767.

[194] Statista. Games - worldwide. Statista. Online. https://www.statista.com/outlook/amo/app/games/worldwide, Accessed: 24.12.2024, 2024.

[195] Statista. Virtual reality (vr) hardware b2c market revenue worldwide from 2019 to 2029 (in billion u.s. dollars) [graph]. Statista. Online. https://www.statista.com/forecasts/1337171/vr-hardware-b2c-market-revenue-worldwide, Accessed: 13.02.2025, 2024.

[196] SteamDB. Steamdb, 2024. URL https://steamdb.info/instantsearch/?refinementList%5Btags%5D%5B0%5D=VR. https:

//steamdb.info/instantsearch/?refinementList%5Btags%5D%5B0%5D=VR, Accessed: 21.08.2024.

[197] Steel Crate Games. *Keep Talking And Nobody Explodes*. Game [VR, PC], October 2015. Steel Crate Games, Ottawa, Canada. https://keeptalkinggame.com/.

[198] Frank Steinicke. *Being really virtual*. Springer, Berlin, Germany, 2016. ISBN 978-3-319-43076-8. DOI 10.1007/978-3-319-43078-2. URL https://doi.org/10.1007/978-3-319-43078-2.

[199] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. Estimation of detection thresholds for redirected walking techniques. *IEEE Transactions on Visualization and Computer Graphics*, 16(1):17–27, 2010. DOI 10.1109/TVCG.2009.62. URL https://doi.org/10.1109/TVCG.2009.62.

[200] Tessa Strain, Seth Flaxman, Regina Guthold, Elizaveta Semenova, Melanie Cowan, Leanne M Riley, Fiona C Bull, and Gretchen A Stevens. National, regional, and global trends in insufficient physical activity among adults from 2000 to 2022: a pooled analysis of 507 population-based surveys with 5· 7 million participants. *The Lancet Global Health*, 12(8):e1232–e1243, 2024. DOI 10.1016/S2214-109X(24)00150-5. URL https://doi.org/10.1016/S2214-109X(24)00150-5.

[201] Ivan E Sutherland. The ultimate display. In *Proceedings of the IFIP Congress*, volume 2, pages 506–508. New York, 1965.

[202] Ivan E. Sutherland. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, (AFIPS '68 (Fall, part I)), page 757–764, New York, NY, USA, 1968. Association for Computing Machinery. ISBN 9781450378994. DOI 10.1145/1476589.1476686. URL https://doi.org/10.1145/1476589.1476686.

[203] Anthea Sutton, Mark Clowes, Louise Preston, and Andrew Booth. Meeting the review family: exploring review types and associated information retrieval requirements. *Health Information & Libraries Journal*, 36(3):202–222, 2019. DOI 10.1111/hir.12276. URL https://doi.org/10.1111/hir.12276.

[204] Philipp Sykownik, Katharina Emmerich, and Maic Masuch. Exploring patterns of shared control in digital multiplayer games. In *International Conference on Advances in Computer Entertainment (ACE 2017), London, UK, December 14-16, 2017*, pages 847–867, Cham, Switzerland, 2017. Springer. DOI 10.1007/978-3-319-76270-8_57. URL https://doi.org/10.1007/978-3-319-76270-8_57.

[205] Luma Tabbaa, Chee Siang Ang, Vienna Rose, Panote Siriaraya, Inga Stewart, Keith G. Jenkins, and Maria Matsangidou. Bring the outside in: Providing accessible experiences through vr for people with dementia in locked psychiatric hospitals. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, (CHI '19), page 1–15, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. DOI 10.1145/3290605.3300466. URL https://doi.org/10.1145/3290605.3300466.

[206] Jamie L. Tait, Rachel L. Duckham, Catherine M. Milte, Luana C. Main, and Robin M. Daly. Influence of sequential vs. simultaneous dual-task exercise training on cognitive function in older adults. *Frontiers in Aging Neuroscience*, 9:368, 2017. ISSN 1663-4365. DOI 10.3389/fnagi.2017.00368. URL https://doi.org/10.3389/fnagi.2017.00368.

[207] Katie Salen Tekinbas and Eric Zimmerman. *Rules of play: Game design fundamentals*. MIT press, 2003.

[208] Todd M Thrash and Andrew J Elliot. Inspiration as a psychological construct. *Journal of personality and social psychology*, 84(4):871, 2003. DOI 10.1037/0022-3514.84.4.871. URL https://doi.org/10.1037/0022-3514.84.4.871.

[209] Todd M Thrash and Andrew J Elliot. Inspiration: core characteristics, component processes, antecedents, and function. *Journal of personality and social psychology*, 87(6):957, 2004. DOI 10.1037/0022-3514.87.6.957. URL https://doi.org/10.1037/0022-3514.87.6.957.

[210] Allison Tong, Kate Flemming, Elizabeth McInnes, Sandy Oliver, and Jonathan Craig. Enhancing transparency in reporting the synthesis of qualitative research: Entreq. *BMC medical research methodology*, 12:1–8, 2012. DOI 10.1186/1471-2288-12-181. URL https://doi.org/10.1186/1471-2288-12-181.

[211] Andrea C Tricco, Erin Lillie, Wasifa Zarin, Kelly K O'Brien, Heather Colquhoun, Danielle Levac, David Moher, Micah DJ Peters, Tanya Horsley, Laura Weeks, et al. Prisma extension for scoping reviews (prisma-scr): checklist and explanation. *Annals of internal medicine*, 169(7):467–473, 2018. DOI 10.7326/M18-0850. URL https://doi.org/10.7326/M18-0850.

[212] Manfred Tscheligi. The importance of hci for solving societal challenges. *Interactions*, 14(5):24–25, 2007. ISSN 1072-5520. DOI 10.1145/1288515.1288533. URL https://doi.org/10.1145/1288515.1288533.

[213] Alan M Turing. Digital computers applied to games. *Faster than thought*, 1953.

[214] United Nations. Transforming our world: the 2030 Agenda for Sustainable Development, 2015. URL https://docs.un.org/en/A/RES/70/1.

[215] United Nations, Department of Economic and Social Affairs, Population Division. World population ageing 2019: Highlights (st/esa/ser.a/430), December 2019. URL https://www.un-ilibrary.org/content/books/9789210045537.

[216] United Nations, Department of Economic and Social Affairs, Population Division. World population ageing 2019 (st/esa/ser.a/444), October 2020. URL https://www.un-ilibrary.org/content/books/9789210045544.

[217] Athanasios Vanezis and Adrian Lees. A biomechanical analysis of good and poor performers of the vertical jump. *Ergonomics*, 48 (11-14):1594–1603, 2005. DOI 10.1080/00140130500101262. URL https://doi.org/10.1080/00140130500101262.

[218] Luke Vecchio, Hays Daewoud, and Shannon Green. The health and performance benefits of the squat, deadlift. *and bench press. MOJ Yoga Phys. Ther*, 3:40–47, 2018. DOI 10.15406/mojypt.2018.03.00042. URL https://doi.org/10.15406/mojypt.2018.03.00042.

[219] Kellie Vella, Christopher James Koren, and Daniel Johnson. The impact of agency and familiarity in cooperative multiplayer games. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, (CHI PLAY '17), page 423–434, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450348980. DOI 10.1145/3116595.3116622. URL https://doi.org/10.1145/3116595.3116622.

[220] Virtual Reality Institute of Health and Exercise. Vr ratings, 2024. URL https://vrhealth.institute/vr-ratings/. Accessed: 10.08.2024.

[221] Jinliang Wang, Haizhen Wang, et al. The predictive effects of online communication on well-being among chinese adolescents. *Psychology*, 2(04):359, 2011. DOI 10.4236/psych.2011.24056. URL https://doi.org/10.4236/psych.2011.24056.

[222] Marilyn Domas White and Emily E Marsh. Content analysis: A flexible methodology. *Library trends*, 55(1):22–45, 2006. DOI 10.1353/lib.2006.0053. URL https://doi.org/10.1353/lib.2006.0053.

[223] Josef Wiemeyer, Lennart Nacke, Christiane Moser, and Florian 'Floyd' Mueller. *Player Experience*, pages 243–271. Springer International Publishing, Cham, 2016. ISBN 978-3-319-40612-1. DOI 10.1007/978-3-319-40612-1_9. URL https://doi.org/10.1007/978-3-319-40612-1_9.

[224] Dmitri Williams, Nicolas Ducheneaut, Li Xiong, Yuanyuan Zhang, Nick Yee, and Eric Nickell. From tree house to barracks: The social life of guilds in world of warcraft. *Games and Culture*, 1(4):338–361, 2006. DOI 10.1177/1555412006292616. URL https://doi.org/10.1177/1555412006292616.

[225] Maximiliane Windl, Albrecht Schmidt, and Sebastian S. Feger. Investigating tangible privacy-preserving mechanisms for future smart homes. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, (CHI '23), New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394215. DOI 10.1145/3544548.3581167. URL https://doi.org/10.1145/3544548.3581167.

[226] Jacob O. Wobbrock. Statistical analysis and reporting in r, 2025. URL https://depts.washington.edu/acelab/proj/Rstats/index.html. Accessed: 05.03.2025.

[227] Jacob O. Wobbrock and Julie A. Kientz. Research contributions in human-computer interaction. *Interactions*, 23(3):38–44, 2016. ISSN 1072-5520. DOI 10.1145/2907069. URL https://doi.org/10.1145/2907069.

[228] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (CHI '11), page 143–146, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450302289. DOI 10.1145/1978942.1978963. URL https://doi.org/10.1145/1978942.1978963.

[229] Mary N. Woessner, Alexander Tacey, Ariella Levinger-Limor, Alexandra G. Parker, Pazit Levinger, and Itamar Levinger. The evolution of technology and physical inactivity: The good, the bad, and the way forward. *Frontiers in Public Health*, 9, 2021. ISSN 2296-2565. DOI 10.3389/fpubh.2021.655491. URL https://doi.org/10.3389/fpubh.2021.655491.

[230] Dennis Wolf, Katja Rogers, Christoph Kunder, and Enrico Rukzio. *JumpVR: Jump-Based Locomotion Augmentation for Virtual Reality*, page 1–12. (CHI '20). Association for Computing Machinery, New York, NY, USA, 2020. ISBN 9781450367080. DOI

10.1145/3313831.3376243. URL https://doi.org/10.1145/3313831.3376243.

[231] World Health Organization. Active ageing: A policy framework, April 2002. URL https://extranet.who.int/agefriendlyworld/wp-content/uploads/2014/06/WHO-Active-Ageing-Framework.pdf. Accessed: 17.01.2025.

[232] World Health Organization. *Global status report on physical activity 2022*. World Health Organization, 2022. https://www.who.int/publications/i/item/9789240059153 Accessed: 25.12.2024.

[233] World Health Organization. Physical activity, June 2024. URL https://www.who.int/news-room/fact-sheets/detail/physical-activity. Accessed: 22.12.2024.

[234] World Health Organization. Coronavirus disease (covid-19) pandemic, January 2025. URL https://www.who.int/europe/emergencies/situations/covid-19. Accessed: 09.01.2025.

[235] Wenge Xu, Hai-Ning Liang, Kangyou Yu, and Nilufar Baghaei. Effect of gameplay uncertainty, display type, and age on virtual reality exergames. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, (CHI '21), New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380966. DOI 10.1145/3411764.3445801. URL https://doi.org/10.1145/3411764.3445801.

[236] Wenge Xu, Hai-Ning Liang, Kangyou Yu, Shaoyue Wen, Nilufar Baghaei, and Huawei Tu. Acceptance of virtual reality exergames among chinese older adults. *International Journal of Human–Computer Interaction*, 0(0):1–15, 2022. DOI 10.1080/10447318.2022.2098559. URL https://doi.org/10.1080/10447318.2022.2098559.

[237] Wenge Xu, Kangyou Yu, Xuanru Meng, Diego Monteiro, Dominic Kao, and Hai-Ning Liang. Exploring the effect of the group size and feedback of non-player character spectators in virtual reality exergames. *Frontiers in Psychology*, 14, 2023. ISSN 1664-1078. DOI 10.3389/fpsyg.2023.1079132. URL https://doi.org/10.3389/fpsyg.2023.1079132.

[238] Anna Xygkou, Chee Siang Ang, Panote Siriaraya, Jonasz Piotr Kopecki, Alexandra Covaci, Eiman Kanjo, and Wan-Jou She. Mindtalker: Navigating the complexities of ai-enhanced social engagement for people with early-stage dementia. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, (CHI '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703300. DOI 10.1145/3613904.3642538. URL https://doi.org/10.1145/3613904.3642538.

[239] Minoru Yamada, Y Kimura, D Ishiyama, Y Otobe, M Suzuki, S Koyama, T Kikuchi, H Kusumi, and H Arai. Effect of the covid-19 epidemic on physical activity in community-dwelling older adults in japan: a cross-sectional online survey. *The Journal of Nutrition, Health & Aging*, 24:948–950, 2020. DOI 10.1007/s12603-020-1501-6. URL https://doi.org/10.1007/s12603-020-1501-6.

[240] Georgios N Yannakakis and Julian Togelius. *Artificial intelligence and games*, volume 2. Springer, 2018.

[241] Jungpil Yoon, Seungwoo Lee, and Taiwoo Park. Jediflight: Design and evaluation of wing-based flying experience in virtual reality. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, (CHI PLAY '18 Extended Abstracts), page 309–320, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450359689. DOI 10.1145/3270316.3273043. URL https://doi.org/10.1145/3270316.3273043.

[242] Mark Young. How to do breaststroke arms, 2022. URL https://www.swim-teach.com/how-to-do-breaststroke-arms.html.

[243] Kangyou Yu, Shaoyue Wen, Wenge Xu, Maurizio Caon, Nilufar Baghaei, and Hai-Ning Liang. Cheer for me: effect of non-player character audience feedback on older adult users of virtual reality exergames. *Virtual Reality*, 27(3):1887–1903, 2023. DOI 10.1007/s10055-023-00780-5. URL https://doi.org/10.1007/s10055-023-00780-5.

[244] Yixuan Zhang, Yimeng Wang, Nutchanon Yongsatianchot, Joseph D Gaggiano, Nurul M Suhaimi, Anne Okrah, Miso Kim, Jacqueline Griffin, and Andrea G Parker. Profiling the dynamics of trust & distrust in social media: A survey study. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, (CHI '24), New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400703300. DOI 10.1145/3613904.3642927. URL https://doi.org/10.1145/3613904.3642927.

[245] Wei Zhao, Steven Baker, and Jenny Waycott. Challenges of deploying vr in aged care: A two-phase exploration study. In *32nd Australian Conference on Human-Computer Interaction*, (OzCHI '20), page 87–98, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450389754. DOI 10.1145/3441000.3441018. URL https://doi.org/10.1145/3441000.3441018.

DECLARATION

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

Hamburg, June 16, 2025

Sukran Karaosmanoglu

colophon

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*".

Final Version as of June 16, 2025 (classicthesis v4.8).