Virtuelle Beteiligung, reale Teilhabe

Jana Tröge ist wissenschaftliche Mitarbeiterin in Drittmittelprojekten an der Hochschule für Wirtschaft und Recht Berlin.

Jan Stepczynski ist wissenschaftlicher Mitarbeiter in Drittmittelprojekten an der Hochschule für Wirtschaft und Recht Berlin.

Heike Wiesner ist Professorin für Betriebliche Informations- und Kommunikationssysteme an der Hochschule für Wirtschaft und Recht Berlin.

Christoph Runde ist Leiter des Virtual Dimension Center sowie Obmann des DIN-Normungsgremiums »Metaverse und Extended Reality«.

Jana Tröge, Jan Stepczynski, Heike Wiesner,
Christoph Runde (Hg.)

Virtuelle Beteiligung, reale Teilhabe

Transformative Technologien für eine inklusivere Gesellschaft

Campus Verlag
Frankfurt/New York

Die Erstellung des Sammelbands wurde im Rahmen der Förderrichtlinie »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft (VAR2)« zwischen 2020 und 2025 vom Bundesministerium für Bildung und Forschung (zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt) gefördert.

Gefördert durch:



Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Der Text dieser Publikation wird unter der Lizenz »Creative Commons Namensnennung-Nicht-kommerziell-Weitergabe unter gleichen Bedingungen 4.0 International« (CC BY-NC-SA 4.0) veröffentlicht. Den vollständigen Lizenztext finden Sie unter: https://creativecommons.org/licenses/by-nc-sa/4.0/deed.de

Verwertung, die den Rahmen der CC BY-NC-SA 4.0 Lizenz überschreitet, ist ohne Zustimmung des Verlags unzulässig. Die in diesem Werk enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Quellenangabe/Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



ISBN 978-3-593-52131-2 Print
ISBN 978-3-593-46260-8 E-Book (PDF)
DOI 10.12907/978-3-593-46260-8

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung des Verlags unzulässig. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.
Trotz sorgfältiger inhaltlicher Kontrolle übernehmen wir keine Haftung für die Inhalte externer Links. Für den Inhalt der verlinkten Seiten sind ausschließlich deren Betreiber verantwortlich.
© 2025. Einige Rechte im Campus Verlag in der Verlagsgruppe Beltz
Werderstr. 10, 69469 Weinheim, info@campus.de.
Umschlaggestaltung: Verlagsgruppe Beltz
Druck und Bindung: Beltz Grafische Betriebe GmbH, Bad Langensalza
Beltz Grafische Betriebe ist ein Unternehmen mit finanziellem Klimabeitrag (ID 15985-2104-1001).
Printed in Germany

www.campus.de

Inhalt

Vorwort 9

Einleitung 11
Jana Tröge, Heike Wiesner, Jan Stepczynski und Christoph Runde

I. Interaktive Technologien für die digitale Gesellschaft

Zur Verschmelzung der Natürlichkeit und Künstlichkeit von Intelligenz und Realität 23
Frank Steinicke

Virtuelle Welten: Chancen und Herausforderungen aus Verbrauchersicht .. 33
Tatjana Halm

II. Transformative Technologien im Bildungsbereich

Lernen mit intelligentem Stift in der erweiterten Realität – FederLeicht ... 55
André Uhl, Nick Wittig, Addythia Saphala, Noro Schlorke, Christian Prison, Boris Kantzow, Michaela Evers-Wölk, Rolf Jürgens, Miriam Schmitz, Ricarda Steinmayr, Oliver Amft und Stefan Schneegass

Novel Haptic Output for Spatial Input Devices and Controllers – HapticIO . 75
Alexander Kalus, Juan F. Olaya-Figueroa, Ferdinand Streicher, Thomas Weiss, Matthias Uhl, Sebastian Krügel, Niels Henze and Katrin Wolf

Development of a Mixed Reality Interactive Training System to Support Professional Socio-Emotional Interactions – MITHOS 95
Lara Chehayeb, Chirag Bhuvaneshwara, Manuel Anglet, Bernhard Hilpert, Ann-Kristin Meyer, Dimitra Tsovaltzi, Patrick Gebhard, Antje Biermann, Sinah Auchtor, Nils Lauinger, Julia Knopf, Andreas Kaiser, Fabian Kersting, Gregor Mehlmann, Florian Lingenfelser, Elisabeth André, Anna Welker, Kai Karren, Ruben Segundo Garcia Ucharima, Dorothea Irmisch, Timo Sowa and Massimo Romanelli

Stresssensitives, adaptives Mixed-Reality-Sicherheitstraining – NEXT-Reality: Entwicklung einer Lernplattform zum Training der Selbstrettungsmaßnahmen bei einer Notwasserung eines Helikopters unter physischer und psychischer Belastung 127
Nicola Drüeke, Anna Larissa Rauschelbach, Thomas Jahn, Andreas Probst, Marvin Künsken, Michael Stöber und Steve Dübel

Understanding Neurodiverse Cognitive-Affective Functioning in a Multimodal Technology-Mediated Framework: Ethical Guidelines and Human-Centered Design Implications – UFO 143
Laura Bareiß, Laura Wansitler, Alina Schwarten and Maria Wirzberger

Was ist Neurodiversität und was bedeutet das für die VR-Entwicklung? – UFO 167
Marc Ristau

Virtual Professors in Mixed Reality – VoluProf 185
Sergio Tejeda Pastor, Jangwoo Son, Johann-Christian Põder, Andy Räder, Juliane Wegner, Armin Keller, Tanja Bäumel, Bernd Kiefer, Michael Dietrich, Torsten Schimmer, Atilla Azgin, Benjamin Körner, Sven Wischnowsky, Thomas Buchholz, Mustafa Tevfik Lafci, Birgit Nierula, Peter Eisert, Anna Hilsmann, Sebastian Bosse and Cornelius Hellge

III. Transformative Technologien im Öffentlichen Raum

Beam me up, Goethe! Gemeinsam im hybriden Museum – Goethe-Live-3D 219
Annika Schlimm, Kristina Johannes, Dirk Wintergrün, Sophia Gröschke und Alexander Kulik

Skalierbare hybride Avatar-Agenten-Technologie für alltagstaugliche soziale Interaktion in XR – HiAvA 237
Christian Merz, Jonathan Tschanter, Niklas Krome, Nicolas Wagner, Mario Botsch, Stefan Kopp, Marc Erich Latoschik und Carolin Wienrich

Die Stadt von morgen mitgestalten! Partizipative Stadtplanung mit AR- und VR-Technologien – Chancen und Herausforderungen für die Bürgerbeteiligung – INSPIRER 263
Jan Stepczynski, Christoph Holtmann, Sebastian Keppler, Thomas Jung, Heike Wiesner und Johann Habakuk Israel

Den Klimawandel am eigenen Körper erfahren – Klima-ACT! 277
Maik Helfrich, Barbro Scholz, Sabine Gärtner und Jasmin Röseler

Kulturelle und soziale Teilhabe aus der Ferne – Gestaltung und Entwicklung eines immersiven und sozialen virtuellen Theatererlebnisses – SocialSTAGE-VR 315
Frank Dittrich, Franziska Legler, Selina Palige, Angelika C. Bullinger, Joachim Keinert, Tobias Jaschke, Andree Josef, Anne Zeiler, Martin Krabbe

XR-Technologie für die partizipative Stadtentwicklung – XR-Part: Potenziale und Herausforderungen am Beispiel der Städte Mannheim und Rostock 333
Svenja Rogoll, Heidi Sinning und Anja Wolter

Nutzerzentrierte Gestaltung von XR-Technologien für die Transformation urbaner Partizipationsprozesse – XR-Part: Herausforderungen und Lösungen für Geolokalisierung, Interaktion und Visualisierung zur Förderung der Bürgerbeteiligung 361
Yvonne Brandenburger, Sebastian Damek, Rolf Kruse, Mohammadreza Maleki Moghadam und Patrick Reipschläger

IV. Transformative Technologien im Gesundheitsbereich

Human-Centered Design of Mixed Reality Applications in Medical Education – GreifbAR 385
Robert Luzsa, Moritz Queisner, Christopher Remde, Igor Sauer, Nadia Robertini and Susanne Mayr

Mixed Reality zum Anfassen: Innovation für virtuelle Haptik durch Gewichts- und Kraftsimulation – interaktionsKRAFT 409
Apostolos Vrontos, Sophie Lambert, Martin Klasen, Saša Sopka, Lina Vogt, Alexander Mertens und Christopher Brandl

Zusammen Erleben, Immersiv Teilhaben: Ein Praxisbeispiel zur Entwicklung immersiver VR-Technologien zur Steigerung der sozialen Teilhabe älterer Menschen – ZEIT 431
Felix Krooß, Danchen Zhang, Simon Kimmel, Wilko Heuten, Eric Landwehr, Yuliya Sobol, Fabian Hemmert, May-Britt Rohden; Anja Härtel, Clemens Winter und Arya Dhar

V. Tranformative Technologien für eine inklusivere Gesellschaft

V/AR-Technologie für die digitale Gesellschaft – Living Lab »XR-Interakt« . 455
Jana Tröge, Heike Wiesner und Christoph Runde

Normung für XR und Metaverse – Living Lab »XR-Interakt« 481
Christoph Runde

Autorinnen und Autoren 495

Vorwort



Sibylle Quenett
© Jens Oellermann

»Innovative Technologien für die digitale Gesellschaft« – diesen zweiten Teil der Förderrichtlinie »Interaktive Systeme in virtuellen und realen Räumen« haben wir Anfang 2020 veröffentlicht. Damals ahnten wir nicht, dass die hier entstehenden Technologien kurze Zeit später derart relevant werden würden: Denn just als die Projekte ihre Forschungsarbeit aufnahmen, führte die Corona-Pandemie zu den ersten Lockdowns in Deutschland. Und nichts hätte uns die Notwendigkeit, an virtuellen Systemen zu forschen, die ein gemeinsames Erleben über Distanzen hinweg ermöglichen, eindrucksvoller vor Augen führen können!

Wir waren aber nicht nur in dieser Hinsicht unserer Zeit voraus, sondern hatten offenbar auch in der Forschungscommunity einen Nerv getroffen. Das spiegelte sich in der Resonanz auf die Bekanntmachung wider: Über 180 Projektkonsortien bewarben sich um eine Förderung. Ursprünglich hatten wir eine Fördergesamtsumme von 15 Millionen Euro eingeplant. Weil jedoch so viele gute Projektanträge eingingen, wurde diese Summe auf 25 Millionen Euro aufgestockt.

Fünfzehn thematische Verbundprojekte und ein Begleitprojekt, die allesamt in diesem Sammelband vorgestellt werden, konnten wir letztendlich fördern. Und wir sind sehr froh und dankbar, dass diese Konsortien trotz aller Widrig-

keiten, die die Corona-Lockdowns für die Forschungsarbeit mit sich brachten, so großartige Ergebnisse erzielen konnten! Großartig auch angesichts der sich rasant entwickelnden technischen Möglichkeiten. Denn die neuartigen virtuellen Technologien kommen langsam auch bei den Bürgerinnen und Bürgern an. Head-Mounted Displays (HMDs) und Augmented-Reality(AR)-Brillen sind inzwischen in vielen Haushalten zu finden. So erreichen die entwickelten Technologien immer mehr Menschen. Und das sollte ja immer das Ziel von Forschung sein.

Ich wünsche Ihnen interessante Einblicke in die virtuellen Technologien von morgen!

Sibylle Quenett

Leiterin des Referats »Interaktive Technologien für Gesundheit und Lebensqualität, Medizintechnik«, Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR)

Einleitung

Jana Tröge, Heike Wiesner, Jan Stepczynski und Christoph Runde

Transformative Technologien sind disruptive Querschnittstechnologien, die branchenübergreifend zum Einsatz kommen. Dazu zählen u. a. Cloudtechnologien, Drohnen, Roboter, Nanotechnologie, aber auch V/AR-Technologien, Wearables und Sensorik (Köhler-Schute, 2016). Als digitale Bausteine durchdringen und gestalten transformative Technologien den Arbeits- und Lebensalltag der Menschen. Die damit verbundene Politik steht vor der Herausforderung, die individuelle, aber auch digitale Souveränität im Umgang mit diesen neuen Technologien zu gewährleisten (BITKOM, 2015: 5). Technische Neuerungen bieten modernen Gesellschaften große Chancen, stellen sie aber auch vor ebenso viele Herausforderungen. Die Befürworter:innen und auch Kritiker:innen gehen davon aus, dass die Einführung von (transformativen) Technologien und Innovationen auch einen entscheidenden verändernden Einfluss auf gesellschaftliche Entwicklungen haben (Köhler-Schute, 2016; MacKenzie & Wajcman, 1999: 3–6).

Mit der Förderrichtlinie des Bundesministeriums für Bildung und Forschung (BMBF)[1] »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« (2020–2024) wurden Institutionen, Firmen, Forschende und Entwickler:innen aufgefordert, innovative Forschungs- und Entwicklungsprojekte der Mensch-Technik-Interaktion im Bereich der gemischten Realität (»Mixed Reality« [MR]) einschließlich der virtuellen Realität (VR) und der erweiterten Realität (»Augmented Reality« [AR]) zu kreieren und umzusetzen. Die Adressierung der *digitalen Gesellschaft* sowie der *sozialen Teilhabe* standen dabei im besonderen Förderfokus der V/AR-Schlüsseltechnologien. Neben einem sehr hohen Innovationsindex sollten neue Anwendungsgebiete und -szenarien für die digitale Gesellschaft erprobt, entwickelt und möglichst nachhaltig implementiert werden. Kernziel war es, die Immersion der MR-Technologie neben der visuellen und auditiven Wahrnehmung durch innovative Interaktionsmodalitäten, wie etwa haptische oder taktile Schnittstellen oder

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

auch olfaktorische und gustatorische Schnittstellen für eine realistische Darstellung und ein intensives Erleben, zu ermöglichen (Bundesministerium für Bildung und Forschung, 2020).

Es wurden neben dem übergreifenden Begleitprojekt *Living Lab »XR-Interakt«* insgesamt 15 V/AR-Projekte im Rahmen dieser Förderrichtlinie gefördert. Der vom Living Lab »XR-Interakt« verantwortete vorliegende Sammelband umfasst besondere Inneneinsichten aller 15 geförderten Projekte, die sich ausnahmslos dem Forschungs- und Entwicklungsfeld »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« gewidmet haben. Darüber hinaus wurden mehrere Protagonist:innen in dem V/AR-Entwicklungsfeld gewonnen, die im Projektförderzeitraum als Keynotes eingeworben werden konnten.

Alle Beiträge adressieren die digitale Gesellschaft auf unterschiedliche Art und Weise. Der Aufbau und die inhaltliche Struktur des Sammelbandes spiegeln diese Vielfalt in fünf Schwerpunkten wider. Eine alphabetische Sortierung der Autor:innen wurde umgesetzt, um das Auffinden der Einzelbeiträge zu erleichtern.

Der erste übergreifende Schwerpunkt *Interaktive Technologien für die digitale Gesellschaft* beginnt mit dem Beitrag »Zur Verschmelzung der Natürlichkeit und Künstlichkeit von Intelligenz und Realität« von *Frank Steinicke*. In einer Welt, in der physische und digitale Realitäten zunehmend miteinander verschmelzen, verschwinden die Grenzen zwischen Natürlichkeit und Künstlichkeit merklich. Technologien wie Künstliche Intelligenz (KI) und Extended Reality (XR) bieten einerseits neue Potenziale für menschliche Interaktion, werfen jedoch auch gleichzeitig (neue) ethische und gesellschaftliche Fragen auf.

Tatjana Halm nimmt diesen fragilen Faden direkt in ihrem Beitrag »Virtuelle Welten: Chancen und Herausforderungen aus Verbrauchersicht« auf. Diese Technologien fördern Inklusion und ermöglichen einen verbesserten Zugang zu wichtigen Dienstleistungen. Gleichzeitig werfen XR-Anwendungen erhebliche Verbraucherschutz-, Datenschutz- und Sicherheitsfragen auf, da sie eine Vielzahl personenbezogener Daten sammeln, darunter biometrische Informationen und Verhaltensmuster. Der Beitrag fokussiert dabei die Auswirkungen und Folgen von V/AR-Technologien aus der Perspektive der Verbraucherinnen und Verbraucher.

Im zweiten Schwerpunkt *Transformative Technologien im Bildungsbereich* sind die Beiträge von insgesamt sechs Projekten angesiedelt:

In dem Beitrag »Lernen mit intelligentem Stift in der erweiterten Realität« werden die ersten Ergebnisse aus dem Forschungsprojekt *FederLeicht* präsentiert. Ziel des Projekts war die Erforschung und Entwicklung eines innovativen Lernsystems für die erweiterte Realität, das direkte Rückmeldung an die Schülerinnen

und Schüler ermöglicht und dadurch die Lehrerschaft, aber auch die Eltern entlastet.

Der Beitrag »Novel Haptic Output for Spatial Input Devices and Controllers« spiegelt die Arbeit des Projekts *HapticIO* wider. Dabei handelt es sich um ein Forschungsprojekt, das darauf abzielt, haptische Feedback-Technologien für räumliche Eingabegeräte und Steuerungen in der erweiterten Realität voranzutreiben. Der Einsatz dieser spezifischen Technologien ermöglicht die Verwendung von virtuellen Werkzeugen, um z. B. eine formverändernde haptische Führung für präzise manuelle Aufgaben nachvollziehen zu können oder auch durch Gewichtssimulation ein besseres Verständnis von Werkzeugen in mobilen Szenarien zu ermöglichen.

In dem Beitrag »Development of a Mixed Reality Interactive Training System to Support Professional Socio-Emotional Interactions« geht es um das Vorhaben des *MITHOS*-Projekts. MITHOS ist ein System, das darauf abzielt, die Konfliktlösungsfähigkeiten von Lehrerinnen und Lehrern durch realistische, situative Lerngelegenheiten bei Konflikten im Klassenzimmer zu trainieren. Das Projekt leistet somit einen sehr wichtigen Beitrag für eine menschenzentrierte, sozial interaktive und verallgemeinerbare erweiterte Realität (XR), die zur Unterstützung von Lehrkräften entworfen und implementiert wurde.

Psychisch und physisch belastende Situationen können mit immersiven Technologien zu Lehrzwecken realitätsnah, aber in einem geschützten Umfeld trainiert werden. Dieser Herausforderung hat sich das Projekt *NEXT-Reality* mit seinem Beitrag »Stresssensitives, adaptives Mixed-Reality-Sicherheitstraining« gestellt. Das Ziel ist die optimale Vorbereitung der Trainees auf die Stressbelastung des Realtrainings »Helicopter Underwater Escape Training« (HUET) und die Erhöhung der Sicherheit und Handlungskompetenz in gefahrengeneigten Arbeitsumfeldern und Notfallsituationen unter Berücksichtigung ethischer und rechtlicher Fragestellungen.

Der Beitrag »Understanding Neurodiverse Cognitive-Affective Functioning in a Multimodal Technology-Mediated Framework: Ethical Guidelines and Human-Centered Design Implications« vom *UFO*-Projekt rankt sich um das Thema Neurodiversität. Das Projekt verfolgt das Ziel, ein gegenseitiges Bewusstsein und Verständnis für neurotypische und neurodivergente Muster kognitiv-affektiver Funktionsweisen zu fördern. Durch die Kombination multimodaler Wearable-Technologien für Sensorik, Interaktion und Feedback wird ein Trainingsansatz mit Teilaufgaben verfolgt, um kognitive Kernfähigkeiten aufzubauen, die anschließend auf realistische Anwendungsszenarien übertragen werden.

In dem zweiten Artikel aus dem *UFO*-Projekt »Was ist Neurodiversität und was bedeutet das für die VR-Entwicklung?« wird eine ungewohnte Perspektive eingenommen. Aus der Perspektive eines neurodiversen Mitarbeiters wird der

gesamte Entwicklungsprozess von der Idee bis hin zur Erprobung des UFO-Projekts begleitet. Der Artikel enthält einen echten Appell, die Inklusion neurodivergenter Personen in der VR-Gestaltung als fortlaufenden Prozess zu verstehen. Neben der expliziten (neurodiversen) Zielgruppen-Perspektive wurde auch Barrierefreiheit von Anfang an im Design- und Entwicklungsprozess priorisiert, um eine vielfältigere und benutzungsfreundlichere digitale Zukunft zu gestalten. Dieser außergewöhnliche partizipative Gestaltungsansatz im UFO-Projekt fördert nicht nur die Teilhabe von neurodiversen Nutzer:innen, sondern verbessert auch die Flexibilität und Benutzerfreundlichkeit für diverse Bürgerinnen und Bürger insgesamt.

Hervorzuheben ist sicherlich auch der Beitrag »Virtual Professors in Mixed Reality« des Projekts *VoluProf*. VoluProf möchte die Online-Bildung verändern, indem es die wichtigsten Herausforderungen herkömmlicher E-Learning-Methoden, wie begrenzte Interaktivität, Isolation und Zugangsbarrieren, angeht. VoluProf bietet eine immersive Mixed-Reality-Erfahrung, bei der die Studierenden mit volumetrischen Avataren von Professor:innen interagieren, die direkt in ihre physischen Räume integriert sind. Diese Avatare ahmen die menschliche Mimik und Gestik nach und schaffen so eine dynamische und ansprechende Lernumgebung.

Der dritte Schwerpunkt *Transformative Technologien im Öffentlichen Raum* vereinigt alle Projekte, die sich um echte Teilhabe von Bürgerinnen und Bürgern bemühen, z. B. Stadtplanungsprojekte, Museen und Konzertangebote.

Das Projekt *Goethe-Live-3D* ist mit dem Beitragstitel »Beam me up, Goethe! Gemeinsam im hybriden Museum« diesem Schwerpunktthema zugeordnet. Ausgelöst durch die Notwendigkeit der räumlichen Distanzierung während der COVID-19-Pandemie rückt die Bedeutung digitaler Museumsangebote seitdem verstärkt in den Fokus. Die digitale Rekonstruktion des Goethe Wohnhauses in Weimar bietet die Möglichkeit, das Museumserlebnis neu zu gestalten und die Erwartungen der Besucher:innen zu verändern. Das historische Gebäude wurde mittels der Verknüpfung von digitalen und analogen Strategien zu Experimentalräumen eines gesellschaftlichen Austauschs über Kulturgeschichte. Besucher:innen können Goethes (hybrides) Wohnhaus barrierefrei und trotz unterschiedlicher Standorte gemeinsam besuchen. Das Projekt stellt sich die Frage, wie es gelingen kann, ein tatsächlich hybrides Museum zu erschaffen, das nicht nur reale und digitale Ausstellung nebeneinanderstellt, sondern diese intelligent verbindet, um sie so für lokale und nichtlokale Besuchende gemeinsam und auch gleichzeitig erkundbar zu machen.

Der Beitrag »Skalierbare hybride Avatar-Agenten-Technologie für alltagstaugliche soziale Interaktion in XR« vom *HiAvA*-Projekt ermöglicht einen Qualitätssprung für diverse kooperative und kommunikative XR-Umgebungen eines

zukünftigen Cyberspace. Es erlaubt recht niedrigschwellig die Teilnahme an sozialen VR-Anwendungen auch mit eingeschränkter Hardwareausstattung. Im Rahmen des Projekts wurden Lösungen für die berührungslose Verfolgung von Körpern und Gesichtern entwickelt und evaluiert. HiAvA ermöglicht durch den Einsatz von immersiven Technologien die Teilnahme an natürlichen sozialen Interaktionen in XR-Umgebung, egal ob Nutzer:innen ein VR-Headset, ein Tablett, ein Smartphone oder einen Laptop zur Teilnahme nutzen.

»Die Stadt von morgen mitgestalten! Partizipative Stadtplanung mit AR- und VR-Technologien – Chancen und Herausforderungen für die Bürgerbeteiligung« lautet der Titel des Beitrags von *INSPIRER* (PartizipatIoN in StadtPlanungsprozessen In viRtuEllen und Realen Räumen). Die Forschenden des Projekts INSPIRER haben eine Anwendung entwickelt, mit der mehrere Benutzerinnen und Benutzer per Mixed-Reality-Technologie kooperieren können. Virtuelle Planungsstände lassen sich somit immersiv erleben. Die Interaktion funktioniert über intuitive, virtuelle Bedienelemente. Somit erlaubt das System breiten Schichten der Bevölkerung, sich an demokratischen Entscheidungsprozessen besser als bisher in der Stadtplanung zu beteiligen. Das INSPIRER-System für partizipative urbane Planungsprozesse beinhaltet grundlegend neue Verfahren, beispielsweise Outdoortracking. Durch den Einsatz von Mixed Reality und neuartiger Bedienkonzepte eröffnen sich neue Möglichkeiten für eine bessere bürgerzentrierte Stadtplanung.

Das Forschungsprojekt *Klima-ACT!* geht in dem Beitrag »Den Klimawandel am eigenen Körper erfahren« der Frage nach, wie eine immersive Erfahrung geschaffen werden kann, um emotionale Betroffenheit und folglich ein positiveres Klimahandeln bei Nutzenden hervorzurufen. Klima-ACT! basiert auf körper- und materialbasierten Ansätzen aus den Bereichen Human-Computer Interaction und Game Design, die auf spielerische Motivation und Verbundenheit abzielen. In einem designgetriebenen Prozess wurden explorative Studien und Workshops durchgeführt, um eine Kombination aus Textilien und VR zu erforschen. Diese auf Textilien basierten sogenannten Wearables sind das Schlüsselmaterial, um eine Brücke zwischen der physischen und der virtuellen Realität zu schlagen und um eine körpernahe Erfahrung zu vermitteln.

Für den Beitrag »Kulturelle und soziale Teilhabe aus der Ferne – Gestaltung und Entwicklung eines immersiven und sozialen virtuellen Theatererlebnisses« zeichnet sich das Projekt *SocialSTAGE-VR* aus. Im Rahmen des Projekts wurden Ansätze für die virtuelle Teilnahme an realen Bühnenveranstaltungen entwickelt und beforscht. Im Beitrag wird insbesondere das technische Konzept für das Erleben von Theateraufführungen mit Virtual-Reality-Headsets vorgestellt. Hierbei wurden reale Bühneninhalte von Theateraufführungen in ein virtuelles Theatermodell übertragen.

Der Beitrag »XR-Technologie für die partizipative Stadtentwicklung – Potenziale und Herausforderungen am Beispiel der Städte Mannheim und Rostock« vom Projekt *XR-Part* adressiert die digitale Gesellschaft mit Blick auf die kommunale Stadtentwicklung. Werden partizipative Verfahren in der Stadtentwicklung mit VR-Technologien kombiniert, erhofft man sich mehr Akzeptanz, mehr demokratische Teilhabe und eine Legitimation von stadtplanerischen Maßnahmen in der Bevölkerung. Die Verbesserung der Vorstellungskraft, eine kreative, informierte Mitwirkung und eine konkretere Darstellung des Beteiligungsgegenstands werden durch den Einsatz von XR-Technologien überhaupt erst ermöglicht. Diese wurden durch einen iterativen Prozess von technischer Entwicklung von XR-Part-Partizipationsformaten, deren Erprobung in Maßnahmen der Stadtentwicklung in Rostock und Mannheim und einer anschließenden Evaluation generiert.

Der zweite Beitrag aus dem Projekt *XR-Part* mit dem Titel »Nutzerzentrierte Gestaltung von XR-Technologien für die Transformation urbaner Partizipationsprozesse – Herausforderungen und Lösungen für Geolokalisierung, Interaktion und Visualisierung zur Förderung der Bürgerbeteiligung« schließt den dritten Themenschwerpunkt pointiert ab. Ein besonderer (Forschungs-)Fokus des Beitrags liegt auf der Benutzerfreundlichkeit und Barrierefreiheit der Anwendungen, einschließlich der Unterstützung durch intuitive räumliche Bedienkonzepte, Sprachausgabe und adaptive Visualisierungstechniken. Technisch-gestalterische Herausforderungen wie präzises Tracking, Lokalisierung und die Einflüsse von Witterungsbedingungen auf die Bedienbarkeit wurden ebenso bearbeitet wie organisatorische Herausforderungen bei der Erstellung immersiver 3D-Modelle sowie der Integration neuer Technologien in bestehende Beteiligungsformate. Das Projekt bestätigt das hohe Potenzial von XR-Technologien, Bürgerbeteiligung nachhaltiger und inklusiver zu gestalten.

Der vierte Schwerpunkt *Transformative Technologien im Gesundheitsbereich* vereinigt alle Projekte, die sich um das Thema Gesundheit ranken.

Der erste Beitrag mit dem Titel »Human-Centered Design of Mixed Reality Applications in Medical Education« beinhaltet die wichtigsten Ergebnisse aus dem Projekt *GreifbAR*. Das Projekt untersucht, wie die Hände von Nutzenden in Virtual-Reality-Anwendungen integriert werden können, um arbeitsbezogene Aufgaben mit physischen Objekten erlernen und trainieren zu können. Als Anwendungsfeld wurde ein medizinischer Anwendungsbereich gewählt. Es wurde festgestellt, dass räumliche und greifbare Interaktion in Mixed-Reality-Systemen den Auszubildenden hilft, sich praktisches Wissen in der medizinischen Ausbildung anzueignen, wenn sie Aufgaben wie das Binden von Knoten oder den Umgang mit Instrumenten mithilfe von MR-Technologien erlernen.

Der Beitrag »Mixed Reality zum Anfassen: Innovation für virtuelle Haptik durch Gewichts- und Kraftsimulation« wurde von dem Projekt *interaktionsKRAFT* zusammengestellt. Die Immersion bei Anwendungen der virtuellen (VR), erweiterten (AR) und gemischten Realität (MR) ist häufig dadurch eingeschränkt, dass Einflüsse und Effekte auf die Nutzenden nicht auch haptisch wahrnehmbar sind. Dieser Herausforderung hat sich das Projekt angenommen. Anhand von zwei konkreten Nutzungsszenarien wird die Umsetzung einer solchen innovativen Haptikkomponente dargestellt. Zum einen wird die Simulation virtueller Gewichte mittels Elektromyostimulation (EMS) für eine VR-Anwendung dargestellt. Hierzu sind die Antagonisten der bei der Haltearbeit beteiligten Muskeln zu stimulieren. Zum anderen wird die Korrektur der Bewegungsausführung bei Kniebeugen in einer virtuellen Physiotherapie beschrieben, bei der die EMS um Bewegungsinformationen durch Motion Capture erweitert wird.

Der Beitrag vom Projekt *ZEIT* »Zusammen Erleben, Immersiv Teilhaben: Ein Praxisbeispiel zur Entwicklung immersiver VR-Technologien zur Steigerung der sozialen Teilhabe älterer Menschen« beschäftigt sich damit, wie die soziale Verbundenheit von getrennt bzw. entfernt lebenden Menschen durch Virtual Reality (VR) und soziale Signale gestärkt werden kann. Im Mittelpunkt stehen dabei die emotionalen Aspekte der zwischenmenschlichen Kommunikation, wie z. B. die Übertragung von Freude und Glück, Zufriedenheit, Unwohlsein, Stress und Sorgen, aber auch die sonst in der Kommunikation per Video und Audio fehlenden typischen physischen Interaktionen, wie die Berührung, Umarmung, der Handschlag zur Begrüßung oder das Schulterklopfen zum Lob und zur Bestätigung. Das Projektziel ist die Erforschung und Entwicklung einer multimodalen VR-Anwendung zum Brückenschlag zwischen alleinlebenden älteren Menschen und ihren Angehörigen.

Der letzte Schwerpunkt *Transformative Technologien für eine inklusivere Gesellschaft* vereinigt zwei Beiträge, die sich mit der Frage der nachhaltigen Verstetigung von V/AR-Technologien für eine inklusivere Gesellschaft beschäftigen.

Der Beitrag des Begleitprojekts *Living Lab »XR-Interakt«* lautet »V/AR-Technologie für die digitale Gesellschaft«. Die Zielsetzung bestand darin, die V/AR-Programmlinie des BMBF durchgängig zu begleiten. Das Living Lab »XR-Interakt« initiierte, moderierte und stärkte die Zusammenarbeit zwischen den Einzelprojekten im Rahmen der Förderlinie »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft«. Die Expertise aus den verschiedenen V/AR-Anwendungsbereichen sollte durch die Initiative des Begleitprojekts so gebündelt und verteilt werden, dass einerseits die digitale Gesellschaft adressiert wird und andererseits den projektübergreifenden Herausforderungen durch einen gemeinsamen – in Teilen auch informellen – Austausch begegnet werden sollte. Der Beitrag fasst die wichtigsten Ergebnisse der

Workshopreihe, der Expert:innenenbefragung und der empirischen Exploration mit Bürger:innen zusammen. Durch die direkte Mitwirkung von Bürgerinnen und Bürgern konnten Verbundprojekte mit einer größeren und diversen Gruppe erprobt werden, um wertvolle Rückmeldungen für künftige Forschungs- und Entwicklungsstrategien zu erhalten. Neue Technologien finden so schnell ihren Weg in den Alltag der Menschen.

In dem letzten Beitrag aus dem Begleitprojekt *Living Lab »XR-Interakt«* geht es um »Normung für XR und Metaverse«. Bei allen in den Einzelprojekten der BMBF-Fördermaßnahme zu entwickelnden Lösungen ist die Bestrebung, alltagstaugliche und erfolgreiche Angebote an die digitale Gesellschaft zu adressieren. Das Begleitprojekt versteht sich somit auch als Entwicklungsprojekt, das Unternehmen und Institutionen eine wichtige Informationsbasis liefert, um V/AR-Technologien erfolgreich in gesellschaftliche Digitalisierungsprozesse und -strategien zu implementieren und gegebenenfalls auch zu vermarkten. Der inhaltliche Schwerpunkt rankt sich somit um Standards und Normen im Bereich V/AR-Technologien.

In dem vorliegenden Sammelband *Virtuelle Beteiligung, reale Teilhabe – Transformative Technologien für eine inklusivere Gesellschaft* sind die Ergebnisse aller 15 beteiligten Projekte aus den Anwendungsfeldern Bildung, Gesundheit und dem Öffentlichen Raum zusammengeführt worden. Die in den Beiträgen dargestellten vielfältigen Ansätze und Lösungen können als transdisziplinäres Wissen neue grundlegende, aber auch ungewohnte Perspektiven und Inneneinsichten in neue Einsatzgebiete eröffnen und auch nachhaltig sichern.

Die Herausgeberinnen und Herausgeber danken allen beteiligten Personen, die zur Realisierung des Sammelbandes beigetragen haben. Die Verantwortung für den Inhalt sowie die Wahrung der Urheber- und Nutzungsrechte an Text- und Bildmaterial der einzelnen Beiträge liegt bei den jeweiligen Autor:innen und Beiträger:innen. Jede Autorin und jeder Autor ist für den Inhalt ihres/seines Beitrags selbst verantwortlich und versichert, dass durch die Veröffentlichung keine Rechte Dritter verletzt werden. Die Herausgeber:innen übernehmen hierfür keine Haftung.

Literatur

BITKOM (2015). *Digitale Souveränität, Positionsbestimmung und erste Handlungsempfehlungen für Deutschland und Europa,* Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e. V., Berlin, S. 1–24.

Bundesministerium für Bildung und Forschung (2020). Bekanntmachungen. Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft (VAR2), Bundesministerium für Bildung und Forschung, Berlin. https://www.interaktive-technologien.de/foerderung/bekanntmachungen/var2 [Zugriff 17.12.2024]

Bundesministerium für Bildung und Forschung (2020). Bekanntmachungen. Richtlinie zur Förderung und Forschung und Entwicklung auf dem Gebiet »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft«, Bundesanzeiger vom 02.04.2020, Bundesministerium für Bildung und Forschung, Berlin. https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2020/04/2912_bekanntmachung.html#searchFacets [Zugriff 17.12.2024]

Köhler-Schute, Christiana (2016). Digitalisierung und Transformation in Unternehmen: Strategien und Konzepte, Methoden und Technologien, Praxisbeispiele. KS-Energy Verlag, Berlin.

MacKenzie, Donald & Wajcman, Judy (1999). *The social shaping of technology* (2nd ed.), Open University Press, Buckingham, S. 3–6.

I.
Interaktive Technologien für die digitale Gesellschaft

Zur Verschmelzung der Natürlichkeit und Künstlichkeit von Intelligenz und Realität

Frank Steinicke

Zusammenfassung

Die Trennung zwischen Natürlichkeit und Künstlichkeit wird im Kontext moderner Technologien zunehmend fragil. Als *künstlich* werden alle vom Menschen geschaffenen Dinge bezeichnet, während *natürlich* jene sind, die ohne menschliches Zutun entstanden sind. In einer Welt, in der physische und digitale Realitäten zunehmend miteinander verschmelzen, verschwinden diese Grenzen merklich. Technologien wie Künstliche Intelligenz (KI) und Extended Reality (XR) bieten neue Potenziale für menschliche Interaktion, werfen jedoch ethische und gesellschaftliche Fragen auf. Um die Auswirkungen dieser Entwicklungen zu verstehen, sind eine interdisziplinäre Diskussion und aktives Handeln erforderlich, um sicherzustellen, dass diese Technologien und deren Auswirkungen auf unsere Gesellschaft unseren Wertvorstellungen und Bedürfnissen entsprechen. Der vorliegende Beitrag befasst sich mit diesen Möglichkeiten und Herausforderungen, die in der Abschluss-Keynote im Juni 2024 an der Hochschule für Wirtschaft und Recht Berlin (HWR Berlin) während des Abschlusstreffens der vom Bundesministerium für Bildung und Forschung (BMBF)[1] geförderten Maßnahme »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft (VAR2)« diskutiert wurden.

1. Einleitung

Der deutsche Philosoph Andreas Birnbacher hat die Dichotomie von Künstlichkeit und Natürlichkeit im Kontext moderner Technologien sowie deren Einfluss auf unsere Wahrnehmung der Welt hinterfragt (vgl. Birnbacher, 2014: 94). Als *künstlich* gelten alle vom Menschen geschaffenen Dinge, wohingegen alle Dinge

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

der Welt, die ohne Zutun des Menschen entstanden sind, als *natürlich* bezeichnet werden. Unsere traditionellen Vorstellungen von natürlich gelten dabei häufig als ideal, im Gegensatz zu künstlich. Birnbacher argumentiert jedoch, dass viele menschliche Errungenschaften, die als künstlich gelten, wesentliche Aspekte unseres Lebens deutlich verbessern und insofern nicht als inferior gelten sollten. Außerdem drängt sich aufgrund des menschlichen Einflusses auf unsere Erde grundsätzlich die Frage auf, welche Objekte unserer realen Umgebung überhaupt noch natürlich sind bzw. lassen sich auch diverse Objekte nicht eindeutig als künstlich oder natürlich einordnen. Ist beispielsweise ein durch den Menschen gepflanzter Wald künstlich oder natürlich? Diese Überlegungen regen an, die Grenzen zwischen dem Naturgegebenen und dem vom Menschen Geschaffenen neu zu definieren und die Rolle des Menschen in einem zunehmend technologisierten Umfeld zu reflektieren (vgl. Birnbacher, 2014: 94).

Wir leben in einer komplexen Welt, in der die physische Welt, digitale Medien und virtuelle Umgebungen miteinander verschmelzen (vgl. Steinicke, 2016: 89 ff.). Die Grenzen zwischen Natürlichkeit und Künstlichkeit, Realität und Virtualität sowie menschlicher und maschineller Intelligenz scheinen immer mehr zu verschwimmen. Wir beobachten, dass diese ursprünglich klar zu unterscheidenden Aspekte nicht mehr eindeutig unterscheidbar sind. Stattdessen bilden sie ein *Kontinuum*, in dem die Begriffe lediglich Endpunkte einer Linie definieren. Der viel größere Raum wird durch die Strecke zwischen den Endpunkten gebildet.

Im Folgenden ist eine Reihe von Beispielen aufgeführt, die solche Kontinuen für unterschiedliche Anwendungsfelder und Technologien verdeutlichen:

Vor etwa 300.000 Jahren bestand unsere Welt nur aus natürlichen Umgebungen. Dann begann der Homo sapiens, künstliche Objekte zu produzieren, die von Steinwerkzeugen über das Rad bis hin zu Taschenrechnern reichen (vgl. Harari, 2024). Heutzutage sind die meisten Objekte in unseren urbanen Lebensgebieten von Menschen – und daher künstlich – angelegt. Die vom Menschen gepflanzten und beschnittenen Bäume sind ein typisches Beispiel, welches sich entlang dieses oben bereits erwähnten Kontinuums bewegt.

In klassischen Computersystemen (basierend auf binärer Darstellung) werden Daten in Bits kodiert, wobei jedes Bit einen Zustand darstellt, der entweder 1 oder 0 ist. Mit dieser Technologie werden seit jeher digitale Systeme realisiert. Wir beobachten jedoch aktuell auch hier einen Wandel von diesem klassischen Modell der Berechnung hin zu anderen Ansätzen, wie dem Quantencomputing. Hier werden Daten in Quantenbits (Qubits) kodiert, die eine Eins, eine Null oder eine kohärente Überlagerung beider Zustände gleichzeitig darstellen können. Diese Superpositionszustände und Quantenverschränkung, die es Qubits erlauben, den Zustand anderer Qubits zu beeinflussen, versprechen enor-

me Leistungsverbesserungen für diverse komplexe Aufgaben (vgl. Benioff, 1982). Auch hier können wir den Bereich zwischen 0 und 1 als Kontinuum verstehen.

Ein weiteres Beispiel findet sich in der Technologie der virtuellen Realität (VR). VR bezeichnet eine immersive, computergenerierte, multisensorische Stimulation, durch die Benutzer:innen in eine vollständig künstliche Realität eintauchen können. Auch hier kann VR als semantisches Differenzial zur realen physischen Welt betrachtet werden. Andere Techniken wie Augmented Reality (AR) oder Augmented Virtuality (AV) kombinieren jedoch virtuelle und reale Objekte in einem kohärenten und konsistenten Raum – das sogenannte Realitäts-Virtualitäts-Kontinuum –, entlang dessen die verschiedenen Technologien der gemischten Realität (MR) beschrieben werden können (vgl. Milgram & Kishino, 1994). Heutzutage wird der Raum, der alle diese überlagerten Realitäten sowie die vollständige Virtualität beinhaltet, als Extended Reality (XR) bezeichnet.

Auch im Bereich der Künstlichen Intelligenz (KI) lässt sich eine ähnliche Situation beobachten. Hier stand jahrelang die *natürliche* Intelligenz des Menschen im Gegensatz zur *künstlichen* Intelligenz von Computer- und IT-Systemen. Aktuelle Entwicklungen im Bereich der generativen KI stellen jedoch auch diese Trennung infrage (vgl. Jebara, 2012). Mittels generativer KI können Computersysteme beispielsweise Texte, Musik, Bilder und Videos erstellen. Dies verdeutlicht ebenfalls, dass gerade die Verschmelzung von menschlicher Intelligenz mit KI neue Potenziale beispielsweise in der gemischten Kreativität erlauben.

In diesem Artikel wird zunächst in Abschnitt 2 die Verschmelzung der beiden Technologien KI und XR betrachtet, bevor einige Beispiele in Abschnitt 3 erläutert werden. Abschnitt 4 schließt mit einem kurzen Fazit.

2. Verschmelzung von KI und XR

Gerade die Verschmelzung von realen und künstlichen Umgebungen sowie von natürlicher und künstlicher Intelligenz erlaubt neuartiges Potenzial für das Zusammenwirken von Menschen in einer zunehmend technologisierten Welt. Allerdings erfordert diese neue Welt der gemischten Realitäten und gemischten Intelligenz neue Paradigmen für die Interaktion und Kommunikation zwischen Menschen und Technologie (vgl. Steinicke, 2016: 155 ff.). Mit der wachsenden Verschmelzung der realen und virtuellen Welt sowie der zunehmenden Verbreitung von intelligenten Brillen (engl. Smartglasses), erleben wir derzeit einen Paradigmenwechsel hin zum Spatial Computing und dem Metaverse. Aktuell sind Apple Vision Pro, Meta Quest und Microsoft HoloLens die prominentesten Modelle solcher Smartglasses mit insgesamt über 50 Millionen Nutzer:innen. Bei den konstant wachsenden Nutzerzahlen ist somit von einer signifikanten Verbreitung

in den kommenden Jahren auszugehen. Smartglasses können, im Gegensatz zu Smartphones, Informationen direkt im Sichtbereich der Benutzer:innen darstellen und die Ansicht auf die reale Welt erweitern. Dazu erfassen sie die reale Umgebung mit einer Reihe von Sensoren wie z. B. RGB- und IR-Kameras, Tiefensensoren sowie Mikrofonen. So kann nicht nur die Umgebung erfasst, sondern es können auch reale und virtuelle Umgebungen manipuliert werden, z. B. über persönliche Assistenten, die in die Brillen integriert sind. In der Regel erfolgt dies durch einfache, direkte Instruktionen der Benutzer:innen (z. B. »Hey Meta, turn off the music!«). Obwohl Smartglasses derzeit nur begrenzte »intelligente« oder »kognitive« Funktionalität bieten, sind sie aufgrund der körpernahen Sensorik und Technologie, der Vielzahl möglicher Anwendungsbereiche sowie ihrer zukünftig zu erwartenden Verbreitung eine ideale Basistechnologie, um die Zusammenarbeit zwischen Menschen und KI zu revolutionieren. Die beiden konvergierenden Technologiefelder – XR und KI – bilden somit die Grundlage für zukünftige, gesellschaftlich relevante kognitive Systeme. Dies ist auch durch das Strategiepapier der Europäischen Kommission zu virtuellen Welten dokumentiert.[2] Aufgrund der sensiblen und personenbezogenen Daten, die hierfür notwendig sind, ist eine verantwortungsvolle Verarbeitung unerlässlich. Die Möglichkeit, (KI-generierte) Informationen in solchen Brillen darzustellen, die kaum noch von der Realität zu unterscheiden sind, erfordert darüber hinaus Vertrauenswürdigkeit und Transparenz.

3. Beispiele von XR- und KI-Systemen

In der Keynote auf dem Abschlusstreffen der durch das BMBF geförderten XR-Forschungsverbünde in Berlin im Juni 2024 wurden einige Beispiele aufgeführt, die aufzeigen, wie sich KI und XR verbinden lassen. Die Beispiele in diesem Artikel, beziehen sich allesamt auf intelligente virtuelle Agenten (engl. Intelligent Virtual Agents [IVAs]). Solche IVAs sind in der Regel humanoide 3D-Modelle in der virtuellen Umgebung, die KI verkörpern und mit denen Benutzer:innen in der XR interagieren können (vgl. Yee et al., 2007; vgl. Keller et al., 2024; vgl. Joshi et al., 2022; Nowak & Biocca, 2003).

2 https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/13757-Virtual-worlds-metaverses-a-vision-for-openness-safety-and-respect_en

3.1 Anthropomorphe IVAs

Virtuelle Agenten unterstützen uns bereits in einer Vielzahl von alltäglichen Aufgaben, wie der Terminvereinbarung, der Überwachung unserer Fitness und der Organisation von Nachrichten. Das Hinzufügen einer humanoiden Körperdarstellung zu diesen größtenteils sprachbasierten virtuellen Agenten hat enormes Potenzial, den Kommunikationsprozess zwischen Mensch und Agent zu bereichern, wirft jedoch gleichzeitig Erwartungen an das soziale, räumliche und intelligente Verhalten des Agenten auf. Die dann resultierenden IVAs können als weniger menschenähnlich wahrgenommen werden, wenn sie beispielsweise keinen Blickkontakt herstellen oder kein plausibles Kollisionsverhalten mit der physischen Umgebung zeigen. Dazu wurde ein Modell entwickelt, welches die Interaktion zwischen Menschen auf die Interaktion mit IVAs erweitert und verschiedene multimodale und multisensorische Kanäle abdeckt, die erforderlich sind, um glaubwürdig verkörperte IVAs zu (vgl. Schmidt et al., 2020) schaffen.



Abbildung 1: Beispiele von IVAs (links) ohne und (rechts) mit Ausdruck von Unsicherheit
Quelle: Eigene Darstellung

Im Rahmen des EU-Projekts PRESENCE wird eine Pipeline zur effizienten Erstellung von intelligenten virtuellen Menschmodellen entwickelt, die entweder als Avatare, d. h. zu Repräsentation von realen Personen in der virtuellen Umgebung, oder als IVAs, also zur Präsentation von KI, verwendet werden können (siehe Abbildung 1 links).[3] Dazu werden verschiedene KI-Bibliotheken eingesetzt, die

3 PRESENCE wird durch die Europäische Union im Rahmen des »Horizon Europe Research and Innovation Programme« unter der Fördernummer 101135025 gefördert. Weitere Informationen zu den IVAs in PRESENCE finden sich hier: https://presence-xr.eu/introducing-presences-intelligent-virtual-humans/.

es erlauben, die Sprache der Benutzer:innen in Text zu transkribieren, diese Texte dann als Bestandteile von Prompts in große Sprachmodelle (z. B. GPT-4o) zu überführen und die resultierenden Texte dann über Sprachausgabe den Benutzer:innen wieder zu präsentieren. Dadurch lässt sich dann ein natürlich wirkendes Gespräch auf verbale Art und Weise mit einer KI führen.

3.2 Vermenschlichung der KI

Alan Turing stellt in seinem Papier »Computing Machinery and Intelligence« die Frage: »Können Maschinen denken?« (vgl. Turing, 1950). Nach Turing können die Begriffe »denken« und »Maschine« nicht klar definiert werden, sodass er stattdessen die Frage aufwirft, ob eine KI einen Menschen imitieren kann. Im von ihm dafür vorgeschlagenen »Turing-Test« muss eine KI eine Kommunikation wie ein Mensch führen. Bei genauerer Betrachtung wird allerdings deutlich, dass die KI folglich auch Unzulänglichkeiten, Limitierungen und Schwächen von Menschen imitieren müsste, um den Turing-Test zu bestehen. Um die technische Machbarkeit sowie die Auswirkungen auf die Mensch-Maschine-Interaktion zu untersuchen, haben wir exemplarisch das Thema Unsicherheit in der Kommunikation untersucht. Unsicherheit wird teilweise bewusst und teilweise unbewusst in der natürlichen menschlichen Kommunikation signalisiert, z. B. durch Ausdrücke wie »vielleicht« oder »ich denke« sowie auch nonverbal (z. B. durch das Abwenden des Blicks). Obwohl aktuelle KI-Algorithmen, die auf Modellen des maschinellen Lernens (ML) basieren, inhärent probabilistisch sind, offenbaren KI-basierte Dienste und maschinelle Lernmodelle wie ChatGPT normalerweise nicht die Zuverlässigkeit der Antworten gegenüber ihren Nutzenden. In einer Studie wurde das Potenzial untersucht, ML-Modelle als leistungsstarke Informationsquellen mit menschlichen Mitteln zur Ausdrucksweise von Unsicherheit zu kombinieren (vgl. Schmidt et al., 2024). Die Ergebnisse zeigen, wie IVAs durch verbale und nonverbale Signale Unsicherheit kommunizieren können (vgl. Abbildung 1 rechts).

3.3 Blended Intelligent Agents

XR-Technologien erlauben es realen Benutzer:innen und IVAs, mittels Smartbrillen und Trackingsystemen innerhalb desselben virtuell erweiterten physischen Raumes zu koexistieren. Während durch das Tracken verschiedener Körperteile, wie dem Kopf und den Händen der Benutzenden, die virtuellen Umgebungen und Objekte plausible Reaktionen auf die Handlungen der realen Benutzer:innen

zeigen können, haben virtuelle Agenten nur einen sehr begrenzten Einfluss auf ihre physikalische Umgebung. Um diese Trennung aufzuheben, haben wir das Konzept der *Blended Intelligent Agents* eingeführt, die in der Lage sind, ihre physikalische Umgebung zu manipulieren (vgl. Schmidt et al., 2019; vgl. Schmidt et al., 2020). Dazu wurden zwei prototypische Implementierungen von virtuellen-physikalischen Interaktionen unter Verwendung von robotischen Aktuatoren und thermochromatischer Tinte vorgestellt. Im Kontext eines Golfszenarios interagierten die Teilnehmenden mit einem gemischten Agenten, der in der Lage war, virtuelle-physikalische Manipulationen wie das Schlagen eines Golfballs und das Schreiben auf physikalischem Papier durchzuführen. Da beide Interaktionen deutlich unterschiedliche Merkmale aufweisen, z. B. hinsichtlich ihrer Persistenz, Erklärbarkeit und Beobachtbarkeit, haben wir eine Benutzerstudie durchgeführt, um ihre Auswirkungen auf subjektive Maße wie die wahrgenommene soziale und räumliche Präsenz des Agenten zu untersuchen. Eine statistische Analyse der quantitativen Daten ergab keine signifikanten Unterschiede zwischen gemischten Agenten und virtuellen Agenten ohne physikalische Fähigkeiten. Allerdings deutet das qualitative Feedback der Teilnehmenden darauf hin, dass persistente Manipulationen sowohl die wahrgenommene Realitätsnähe des Agenten als auch das gesamte Benutzererlebnis verbessern.

4. Fazit

In dieser Arbeit wurde aufgezeigt, wie sich aktuelle Technologien wie KI und XR miteinander verschmelzen und welche Potenziale daraus entstehen. Gerade die exponentiellen Entwicklungen in diesem Bereich zeigen allerdings auch auf, dass sich schon bald virtuelle Umgebungen darstellen lassen, die wir nicht mehr von der Realität unterscheiden können. Außerdem werden auch die Entwicklungen in KI zunehmend intelligentere Systeme hervorbringen, die dann dem Menschen in vielen Bereichen überlegen sein werden. Während die Kombination von XR und KI immenses Potenzial birgt, wirft die Verschmelzung insbesondere ethische, moralische und gesellschaftliche Fragen auf. Diese zu klären und auszuhandeln, ist nur in einem interdisziplinären Prozess zwischen Wissenschaft, Wirtschaft und Gesellschaft möglich. Es ist allerdings von enormer Bedeutung, dass wir nicht nur die Diskussion zur Verantwortung in diesem Kontext führen, sondern uns auch an den technischen Entwicklungen aktiv beteiligen. Andernfalls gibt es keine andere Option, als die Entwicklungen anderer zu verwenden, die gegebenenfalls nicht nach unseren Wertvorstellungen und Bedürfnissen realisiert wurden.

Förderhinweis

Die hier vorgestellten Arbeiten wurden von verschiedenen Institutionen gefördert, u. a. von der Deutschen Forschungsgemeinschaft, dem Bundesministerium für Bildung und Forschung (BMBF)[4]*, der Europäischen Union und der Universität Hamburg.*

Literatur

Ball, Matthew (2022). *The Metaverse: And How It Will Revolutionize Everything,* Liveright Publishing, W.W. Norton & Company, New York.

Birnbacher, Dieter (2014). *Naturalness: Is the »Natural« Preferable to the »Artificial«?* University Press of America, Lanham, Maryland. S. 94.

Benioff, Paul (1982). Quantum mechanical Hamiltonian models of discrete processes that erase their own histories: Application to Turing machines. In: *International Journal of Theoretical Physics*, Vol. 21, S. 177–201.

Harari, Yuval Noha (2024). SAPIENS – Eine kurze Geschichte der Menschheit: Der legendäre Weltbestseller aktualisiert und mit neuem Nachwort, Penguin Verlag, München.

Jebara, Tony (2012). *Machine Learning: Discriminative and Generative,* Springer Science+Business Media, Berlin.

Joshi, Nidhi et al. (2022). Advanced Emotion Analytics of Virtual Group Meetings involving Intelligent Virtual Agents. In*: IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, Christchurch, New Zealand. S. 344–350.

Keller, Judith et al. (2024). ACHOO – Bless you! Sense of Presence can provoke Proactive Mucosal Immune Responses in Immersive Human-Agent Interactions. In: *IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, Orlando, Florida, USA. S. 557–567.

Korkmaz, Leon et al. (2024). I Talk-You Write. Exploring Asymmetric Text-and Voice-Based Communication Between Humans and Virtual Agents. In: *Proceedings of the 12th International Conference on Human-Agent Interaction (CHAI)*, Swansea, United Kingdom, S. 25–33.

Milgram, Paul & Kishino, Fumio (1994). *A taxonomy of mixed reality visual displays.* In: *IEICE Transaction on Information and Systems*, Vol. E77-D 77, No. 12(12), The Institut of Electronics, Information and Systems, Tokyo, Japan. S. 1321–1329.

Nowak, Kristine & Biocca, Frank (2003). The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. In: *Teleoperators and Virtual Environments*, Vol. 12, S. 481–494.

Schmidt, Susanne, Ariza, Oscar & Steinicke, Frank (2019). Blended Agents: Manipulation of Physical Objects within Mixed Reality Environments and Beyond. In: *ACM Symposium on Spatial User Interaction*, Vol. 6, S. 1–10.

4 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Schmidt, Susanne, Ariza, Oscar & Steinicke, Frank (2020). Intelligent Blended Agents: Reality–Virtuality Interaction with Artificially Intelligent Embodied Virtual Humans. In: *Multimodal Technologies and Interaction*, Vol. 4(4), S. 85.

Schmidt, Susanne et al. (2024). Natural expression of a machine learning model's uncertainty through verbal and non-verbal behavior of intelligent virtual agents. In: *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, Vol. 14, S. 1–15.

Steinicke, Frank (2016). *Being Really Virtual*. Springer International Publishing AG, Basel.

Turing, Alan Mathison (1950). Computing Machinery and Intelligence. In: *Mind*, Edinburgh, Vol. 59, No. 236, S. 433–460.

Yee, Nick, Bailenson, Jeremy & Rickertsen, Kathryn (2007). A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In: *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, S. 1–10.

Virtuelle Welten: Chancen und Herausforderungen aus Verbrauchersicht

Tatjana Halm

Zusammenfassung

Extended Reality (XR) bietet bedeutende Chancen unter dem Aspekt sozialer Teilhabe insbesondere in Bereichen wie Bildung, Gesundheit und dem Öffentlichen Raum. Diese immersiven Technologien fördern Inklusion und ermöglichen einen verbesserten Zugang zu wichtigen Dienstleistungen, indem sie V/AR-erweiterte Lernumgebungen, medizinische Innovationen und/oder einen barrierefreien Zugang zu virtuellen XR-Veranstaltungsangeboten speziell für Bürgerinnen und Bürger schafft. Gleichzeitig werfen XR-Anwendungen erhebliche Verbraucherschutz-, Datenschutz- und Sicherheitsfragen auf, da sie eine Vielzahl personenbezogener Daten sammeln, darunter biometrische Informationen und Verhaltensmuster. Dies führt zu unklaren rechtlichen Situationen und zu Risiken für die Privatsphäre und Sicherheit. Zudem stellen sich ethische und soziale Herausforderungen wie eine digitale Exklusion und die Gefahr der Manipulation durch gezielte Werbung oder politische Einflussnahme. Eine verantwortungsvolle Nutzung von XR erfordert daher klare rechtliche Regelungen, Standards und eine kritische Auseinandersetzung mit den damit verbundenen Risiken und Chancen.

1. Einleitung

Die Entwicklung von Extended Reality (XR) wird momentan als eine technologische Neuerung betrachtet, die erhebliche Möglichkeiten zur Veränderung alltäglicher Prozesse bietet. XR umfasst Anwendungen in wesentlichen gesellschaftlichen Bereichen wie Bildung, Gesundheit, Nachhaltigkeit und Inklusion und geht dabei über die Möglichkeiten herkömmlicher digitaler Technologien hinaus. Diese Systeme ermöglichen eine Verknüpfung physischer und digitaler Welten und schaffen immersive Erfahrungen, die das menschliche Erleben erweitern und individuell anpassen können. Eine Studie von Deloitte aus dem Jahr 2022 zeigt, dass

sich altersübergreifend damals schon fast 40 % der Befragten in Deutschland vorstellen konnten, selbst Teil des »Metaverse« zu werden (Deloitte, 2022).

Der Verbraucherschutz gewinnt daher in virtuellen Welten zunehmend an Bedeutung, da digitale Technologien und virtuelle Plattformen immer stärker in das alltägliche Leben integriert werden. So finden sich bereits heute Beispiele insbesondere im Gamingbereich, u. a. Meta wirbt mit den neuen »Ray-Ban Meta«-Brillen. Mit dem Aufkommen des Metaversums, virtueller Realität und digitaler Marktplätze entstehen für Verbraucherinnen und Verbraucher neue Chancen, aber zeitgleich auch (neue) Risiken. Themen wie Datenschutz, Transparenz bei digitalen Dienstleistungen, Schutz vor Missbrauch oder Manipulation sowie die Gewährleistung von Sicherheit und Fairness stehen im Fokus. Verbraucherschutz zielt in diesem Kontext darauf ab, Nutzerinnen und Nutzer in einer zunehmend digitalisierten Welt vor rechtlichen, finanziellen und ethischen Herausforderungen zu bewahren und ihre Rechte zu stärken.

Im folgenden Artikel werden die Potenziale von XR für Verbraucherinnen und Verbraucher näher betrachtet. Der Schwerpunkt liegt auf den verschiedenen Anwendungsbereichen, in denen XR-Technologien das tägliche Leben bereichern können. Besonders in den Bereichen Bildung, Gesundheit und dem Öffentlichen, wie z. B. soziale Teilhabe in Museen, eröffnen XR-Anwendungen neue Möglichkeiten, bestehende Herausforderungen zu bewältigen und gleichzeitig Inklusion sowie gesellschaftliche Partizipation zu fördern. So ermöglichen virtuelle Lernumgebungen interaktives und individuelles Lernen, medizinische Anwendungen verbessern die Versorgung und Diagnosemöglichkeiten, während barrierefreie digitale Räume die Teilhabe von Menschen mit Behinderungen an der digitalen Welt fördern.

Im vorliegenden Beitrag sollen insbesondere die Verbraucherperspektive und der Verbraucherschutz im Fokus stehen. Nach einer kurzen Darstellung der besonderen Potenziale von XR-Angeboten für Verbraucherinnen und Verbraucher werden die damit verbundenen Risiken eingehend erörtert. Trotz der vielen Chancen, die XR-Technologien bieten, treten Bedenken hinsichtlich des Datenschutzes, der Datensicherheit und des Schutzes der Verbraucherrechte in virtuellen Welten auf. Die Erhebung großer Mengen sensibler Daten wirft Fragen zu Privatsphäre und Sicherheitsrisiken auf, die zum einen ein Bewusstsein, aber auch eine Regulierung und klare rechtliche Rahmenbedingungen erforderlich machen. Darüber hinaus werden ethische und soziale Herausforderungen thematisiert, wie etwa die Gefahr der digitalen Kluft, die durch fehlenden Zugang zu XR-Technologien verstärkte Ungleichheiten verursachen könnte. Ebenso wird auf die Risiken manipulativer Werbung und der Verbreitung von Fehlinformationen, etwa durch Deepfakes, eingegangen. Diese Fragen verlangen eine kritische

Auseinandersetzung, um die Vorteile von XR verantwortungsvoll und nachhaltig zu nutzen.

2. Der Nutzen für Verbraucherinnen und Verbraucher von XR

XR-Anwendungen können in vielen Lebens- und Arbeitsbereichen sehr nützlich sein. Welche Technologien und welche Angebote sich im Alltag durchsetzen werden, wird zukünftig (neu) verhandelt werden müssen. Im Folgenden wird zunächst auf die Möglichkeiten und Potenziale einiger ausgewählter Anwendungsfelder eingegangen.

2.1 Bildung und XR: Neue immersive Lernerfahrungen

Im Bildungsbereich weist XR-Technologie ein erhebliches Potenzial zur Neugestaltung von Lernprozessen auf. Sie könnte es ermöglichen, Wissen auf interaktive und immersive Weise zu vermitteln, und würde somit die Optionen des traditionellen Lernens erweitern. »Immersives Lernen bezeichnet einen Lernansatz, der digitale Technologien wie Virtual Reality (VR) oder Augmented Reality (AR) nutzt, um Lernende beim Kompetenzaufbau sowie beim Erlernen von Prozessen zu unterstützen.« (Bitkom, 2020: 7) Hierzu müsste aber nicht nur die Integration von Technologie in den Unterricht erfolgen, sondern es müsste auch insgesamt eine Lernumgebung geschaffen werden, die das Lernen spannender und effektiver gestaltet.

Eine mögliche Anwendung von XR in der Bildung sind beispielsweise virtuelle Exkursionen. Diese erlauben es Schülerinnen und Schülern sowie Studierenden bereits jetzt schon, historische Stätten, berühmte Museen oder entlegene geografische Orte zu erkunden, ohne dass physische Reisen erforderlich sind. Solche Erfahrungen sind vor allem dann interessant, wenn Reisen durch finanzielle, geografische oder gesundheitliche Einschränkungen limitiert sind. »Extended Reality (XR) umfasst ein Spektrum von Technologien, die das Potenzial haben, die Prinzipien des zukunftsorientierten Lernens auf einzigartige Weise zu unterstützen, indem sie Gestaltungs-, Begegnungs- und Kommunikationsräume bieten, die die physische Welt mit der virtuellen Welt verbinden.« (Wössner 2024) So können Lernende beispielsweise durch die Ruinen von Pompeji spazieren oder den Grand Canyon aus der Nähe betrachten, ohne das Klassenzimmer zu verlassen. Außerdem hat es sich auch für virtuelle Umgebungen gezeigt, dass es den »Lernenden durch Exkursionen einfacher fällt, gelerntes Fachwissen mithilfe von Erinnerungen an Erlebnisse abzurufen« (Gryl & Kuckuck, 2023: 14).

Eine weitere Einsatzmöglichkeit könnte darin bestehen, XR-Technologien so zu gestalten, dass sie den spezifischen Bedürfnissen von Lernenden mit Lernschwierigkeiten gerecht werden. Durch gezielte Anpassungen in der Benutzeroberfläche, der Geschwindigkeit des Lernmaterials oder der Interaktivität könnten Lehrkräfte sicherstellen, dass jede Schülerin und jeder Schüler auf seinem individuellen Lernniveau unterstützt wird.

Alles in allem könnte ein Nutzen darin liegen, dass Schülerinnen und Schüler, statt passiv Informationen aufzunehmen, in interaktive Szenarien eintauchen können, die das Lernen emotional bereichern und ein tieferes Verständnis der Materie ermöglichen. Insbesondere in den Naturwissenschaften können sich interaktive Simulationen als besonders wertvoll erweisen. So eröffnen sie Schülerinnen und Schülern die Möglichkeit, komplexe naturwissenschaftliche Phänomene zu erforschen, ohne die Risiken oder den finanziellen Aufwand realer Experimente. »Virtual-Reality-Experimente (VRE) simulieren ausgewählte physikalische Versuche, die aufgrund von Gefahren oder wegen ihres Aufwands im Schulunterricht nicht (mehr) real durchgeführt werden, in einer realitätsgetreuen und inhaltlich korrekten Darstellung.« (Johannes-Gutenberg-Universität Mainz, Virtual-Reality-Experimente, Konzept Virtual-Reality-Experimente, 2024)

Ein Beispiel für eine solche Simulation ist das Programm »Labster«[1], das Schülerinnen und Schülern sowie Studierenden Zugang zu virtuellen Laboren bietet, in denen sie Experimente durchführen können, ohne physisch an einem Ort sein zu müssen (Labster, 2025). Solche Programme bieten nicht nur eine kostengünstige Alternative zu traditionellen Laborexperimenten, sondern auch die Möglichkeit, verschiedene Szenarien und Ergebnisse zu erforschen, ohne dass Materialien verloren gehen. Ein weiteres Beispiel ist das Projekt »FederLeicht« (Näheres in Teil II des Sammelbandes), das den Einsatz von AR-Technologie im Unterricht integriert. Hier lernen Schülerinnen und Schüler mithilfe eines intelligenten Stifts, der interaktive Inhalte auf Arbeitsblätter projiziert und so das Lernen unterstützt. Diese Art des Lernens verbindet digitale Inhalte mit physischen Materialien, wodurch die Lernenden auf neue Weise mit den Informationen interagieren können. Und schließlich soll durch das Projekt »Goethe-Live-3D« der Zugang zu Museen in hybriden Formaten ermöglicht werden. Dieses Projekt erlaubt es beispielsweise Lernenden, gemeinsam virtuelle Räume zu betreten und Kultur auf neuartige Weise zu erleben (Näheres in Teil III des Sammelbandes). Hier können sie nicht nur historische Artefakte betrachten, sondern auch an interaktiven Führungen teilnehmen, die von Expertinnen und

1 Vgl. uber-uns | Labster: Labster verfolgt das Ziel, Schülerinnen und Schülern durch immersive, wissenschaftlich fundierte Simulationen den Zugang zur Wissenschaft zu erleichtern und ihren Lernerfolg im und außerhalb des Klassenzimmers zu fördern (abgerufen am 11.02.2025).

Experten geleitet werden. Solche Formate fördern die Zusammenarbeit und den Austausch unter den Lernenden, auch wenn sie physisch nicht am selben Ort sind.

2.2 Gesundheit und XR: Ein Paradigmenwechsel in der Medizin

XR-Technologien könnten auch den Gesundheitssektor grundlegend verändern, insbesondere durch ihre Anwendung in der Telemedizin und in therapeutischen Settings. Diese Technologien verbessern langfristig gesehen nicht nur die Effizienz von medizinischen Dienstleistungen, sondern können auch einen leichteren Zugang zu Gesundheitsversorgung für Menschen ermöglichen, die in ländlichen oder unterversorgten Regionen leben, wo der Zugang zu medizinischen Fachkräften oft eingeschränkt ist. Die Einführung von XR-Technologien in der Telemedizin ermöglicht »virtuelle Konsultationen« (Grgic, 2024: 8) und die Fernüberwachung von Patientinnen und Patienten. Diese Ansätze könnten dabei helfen, die räumliche Distanz zwischen Ärztinnen und Ärzten und ihren Patientinnen und Patienten zu überbrücken und eine flexible, ortsunabhängige Lösung für die medizinische Versorgung anzubieten. Patientinnen und Patienten wäre es so möglich, bequem von zu Hause aus mit ihren Ärztinnen und Ärzten zu kommunizieren. Auch »im Bereich der Chirurgie werden bereits heute digitale Hilfsmittel wie Kameraaufnahmen auf Bildschirmen verwendet« (Büchel & Klös, 2022: 8).

In der Psychotherapie lassen sich ebenfalls vorteilhafte Nutzungsmöglichkeiten erkennen (Riva et al., 2021: 3 ff.). »VR-Behandlungen haben gezeigt, dass sie chronische Schmerzen, PTSD und Phobien lindern können.« (Büchel & Klös, 2022: 12) Und schließlich lassen sich XR-Technologien auch sehr sinnvoll in der medizinischen Ausbildung einsetzen, wo sie dazu beitragen, die Fähigkeiten von angehenden Fachkräften zu verbessern. »Ein wichtiger Benefit ist der Zuwachs an Praxiserfahrung durch realitätsnahes Erleben.« (Grgic, 2024: 8)

Ein gutes Beispiel im medizinischen Bereich ist die Anwendung »GreifbAR«, die medizinischem Fachpersonal die Möglichkeit bietet, ihre Feinmotorik in einer Mixed-Reality-Umgebung zu trainieren. Hier können sie chirurgische Verfahren und technische Fähigkeiten in einer sicheren und kontrollierten Umgebung üben, ohne dabei Patientinnen und Patienten einem Risiko auszusetzen. Durch die Verwendung von XR können Medizinstudierende komplexe Verfahren simulieren, die sie in der realen Welt möglicherweise noch nicht ausprobiert haben. Diese Form des Trainings ermöglicht es den Lernenden, praktische Fähigkeiten zu entwickeln und Selbstvertrauen zu gewinnen, bevor sie mit realen Menschen arbeiten. Zudem können sie in der virtuellen Umgebung jederzeit auf

zusätzliche Informationen zugreifen, was das Lernen vertieft und personalisiert (Näheres in Teil IV des Sammelbandes).

2.3 Nachhaltigkeit und Klimaschutz: Umweltfreundliche Potenziale von XR

XR-Technologien bieten außerdem Chancen, um die Herausforderungen im Bereich Nachhaltigkeit und Klimaschutz anzugehen. Diese Technologien können das Bewusstsein für Umweltfragen schärfen und somit mittelbar den CO_2-Fußabdruck reduzieren. Eine Möglichkeit, wie XR-Technologien zur Nachhaltigkeit beitragen könnten, ist die Reduzierung der physischen Reisen. Besonders im geschäftlichen Kontext könnten XR-Lösungen wie virtuelle Meetings und Konferenzen dazu führen, dass Unternehmen effektiv kommunizieren, auch ohne die Umwelt durch Reisen zu belasten. Insbesondere die Corona-Pandemie hat neue Möglichkeiten der digitalen Kommunikation und Durchführung von Businessevents aufgezeigt (Statista, 2024), auch wenn klassische Präsenztreffen wahrscheinlich aufgrund des Wunsches nach persönlichem Kontakt nicht vollständig ersetzt werden. Aber auch im privaten Umfeld könnte es für den einen oder anderen ausreichend sein, Reiseerlebnisse, insbesondere beispielsweise kulturelle Städteurlaube, in virtuellen Welten zu erleben, anstatt lange Flugreisen auf sich zu nehmen.

Darüber hinaus bietet XR die Möglichkeit, realistische Simulationen von Umweltauswirkungen zu erstellen, die Bildungs- und Entscheidungsprozesse in der Umweltpolitik unterstützen können. XR-Technologien ermöglichen es, komplexe ökologische Systeme und deren Veränderungen anschaulich darzustellen, wodurch das Verständnis für die Ursachen und Folgen von Umweltveränderungen vertieft wird. So können auch mit »Hilfe von immersiven 3D-Visualisierungen bereits bei der Planung Umweltaspekte besser berücksichtigen werden« (Bitkom, 2022: 45).

Ein konkretes Beispiel für eine solche Anwendung ist die Plattform »Klima-ACT!«, die es Nutzerinnen und Nutzern ermöglicht, die Auswirkungen des Klimawandels direkt zu erleben. Diese immersive Erfahrung bietet den Nutzenden die Möglichkeit, die komplexen Wechselwirkungen zwischen ihrem Handeln und den ökologischen Auswirkungen nachzuvollziehen. Durch diese interaktive Herangehensweise kann das Bewusstsein für die Notwendigkeit von Umweltschutzmaßnahmen geschärft werden (Näheres in Teil IV des Sammelbandes).

2.4 Barrierefreiheit: XR als Chance für mehr Teilhabe

Virtuelle Welten bieten zudem innovative Lösungen zur Förderung der Barrierefreiheit. Sie schaffen damit für Menschen mit Behinderungen neue Möglichkeiten, am gesellschaftlichen Leben teilzuhaben. Diese Technologien können Barrieren abbauen und den Zugang zu sozialen, kulturellen und bildungsbezogenen Erlebnissen erleichtern, die oft schwer zugänglich sind. So kann XR besonders für Menschen mit sensorischen oder motorischen Einschränkungen Vorteile bieten. Beispielsweise wäre es möglich, dass durch den Einsatz von virtuellen und erweiterten Realitäten assistive Technologien entwickelt werden, die die Interaktion und Teilhabe in einer Weise ermöglichen, die in der physischen Welt oft nicht umsetzbar ist. Ein Beispiel ist der Einsatz von XR in der beruflichen Ausbildung und am Arbeitsplatz. Virtuelle Trainingsumgebungen können speziell auf die Bedürfnisse von Menschen mit unterschiedlichen Behinderungen zugeschnitten werden, sodass sie Fähigkeiten entwickeln und erlernen können, die in herkömmlichen Ausbildungsformaten möglicherweise nicht ausreichend berücksichtigt werden.

Ein weiteres Beispiel ist das Projekt *Canetroller* von Microsoft. Hierbei wird ein virtueller Blindenstock eingesetzt, der mithilfe von haptischem Feedback und präzisen Klangwelten die räumliche Anordnung und Beschaffenheit von Objekten sowie die Struktur von Räumen in virtuellen Umgebungen erlebbar macht (Rosin, 2023).

Eine ebenfalls interessante Anwendung von XR-Technologien im Kontext der Teilhabe ist die Plattform »SocialSTAGE-VR«, die es Nutzenden ermöglicht, gemeinsam an kulturellen Veranstaltungen in einer virtuellen Umgebung teilzunehmen (Näheres in Teil III des Sammelbandes). Nutzerinnen und Nutzer können beispielsweise Theateraufführungen, Konzerte oder Ausstellungen erleben, ohne physisch anwesend sein zu müssen. Dies ist besonders vorteilhaft für Personen mit Behinderungen, die möglicherweise Schwierigkeiten haben, öffentliche Veranstaltungen zu besuchen, sei es aufgrund von Mobilitätseinschränkungen oder anderen gesundheitlichen Herausforderungen. Durch die Möglichkeit, soziale und kulturelle Erlebnisse in einer virtuellen Realität zu teilen, wird nicht nur der Zugang zu Kunst und Kultur erleichtert, sondern auch das Gemeinschaftsgefühl gestärkt. »Menschen, die geografisch oder aus anderen Gründen isoliert sind, können mit anderen interagieren, Freundschaften schließen und an einem gemeinsamen kulturellen Erlebnis teilhaben.« (Bitkom, 2022: 38) Unity und Insomniac Entertainment, die Electronic-Dance-Music-Sparte von Live Nation, verfolgen gemeinsam das Ziel, eine persistente Metaverse-Welt für Fans elektronischer Musik zu schaffen. Damit sollen Millionen Menschen weltweit die Möglichkeit erhalten, virtuell an Events wie dem Electric Daisy

Carnival (EDC) in Las Vegas teilzunehmen, unabhängig von ihrem physischen Standort. Diese 3D-Echtzeittechnologie kann auch auf andere Bereiche der Live-Unterhaltung ausgeweitet werden, sodass etwa Sportübertragungen interaktiver gestaltet oder Broadway-Shows wie »Hamilton« für ein globales Publikum virtuell zugänglich gemacht werden könnten, sogar mit der Option, als Zuschauerin und Zuschauer virtuell auf der Bühne zu stehen (Bitkom, 2022: 38).

3. Herausforderungen und Risiken von Extended Reality

Virtuelle Welten haben also das Potenzial, zahlreiche Sektoren nachhaltig zu verändern. Doch damit einher gehen auch signifikante Herausforderungen und Risiken. Diese betreffen insbesondere den Datenschutz, die Datensicherheit sowie den Verbraucherschutz. Die zunehmende Integration von XR-Technologien in den Alltag erfordert daher eine kritische Auseinandersetzung mit diesen Themen, um sicherzustellen, dass die Vorteile der Technologie nicht durch negative Konsequenzen überschattet werden.

3.1 Die Herausforderungen für Datenschutz und Privatsphäre

Die Datenschutzgrundverordnung (DSGVO) der Europäischen Union stellt die Rahmenbedingung für den Umgang mit personenbezogenen Daten. Sie setzt dabei hohe Anforderungen an die Erhebung, Speicherung und Verarbeitung personenbezogener Daten und bietet einen besonders hohen Schutz (vgl. zur DSGVO der Europäischen Union: https://datenschutz-grundverordnung.eu/).

XR-Anwendungen wie VR-Headsets und AR-Apps benötigen in der Regel aber einen umfassenden Zugriff auf personenbezogene Daten wie Standortinformationen, biometrische Daten und Verhaltensmuster. »Dazu zählt die kontinuierliche Erfassung von Gesichtserkennung, Mimik, Körperhaltung, Stimme und Pupillenreaktionen, um Emotionen zu analysieren und die Interaktivität zu steigern.« (Mensch und Medien GmbH, 2023: 1) Diese Technologien bieten somit viele faszinierende Möglichkeiten, da sie ein maßgeschneidertes Nutzererlebnis ermöglichen und die Interaktivität erhöhen können. Allerdings wirft ihre Nutzung auch erhebliche Datenschutzbedenken auf, da sie etwa »unter anderem den Datensammlern Aufschluss über den Gesundheitsstatus [und] die Lebensführung des Akteurs« geben (Mensch und Medien GmbH, 2023: 1). Ein Beispiel ist der Einsatz von Gesichtserkennungssoftware in VR-Anwendungen, die dazu dient, emotionale Reaktionen zu erfassen und Bewegungsabläufe zu analysieren. Dabei werden besonders sensible biometrische Daten erfasst, die eine Identifizierung der

Nutzenden ermöglichen und tief in die persönliche Privatsphäre eingreifen. Trotz der spannenden Möglichkeiten dieser Technologien zeigt sich darin ein zentraler datenschutzrechtlicher Knackpunkt: »Die enorme Anzahl an zusätzlichen Datenverarbeitungsprozessen« (Mensch und Medien GmbH, 2023: 1) stellt eine große Herausforderung für den Datenschutz dar.

Ein weiteres zentrales Problem im Zusammenhang mit den Datenschutzfragen in XR-Anwendungen ist auch, dass viele Nutzende sich nicht bewusst sind, welche und wie viele Daten überhaupt erfasst werden und in welchem Umfang diese verarbeitet werden. »Bei einer 20-minütigen Nutzung eines VR-Headsets werden bis zu zwei Millionen Datenpunkte erfasst.« (Heynike & Wipper, 2023: 1) Diese Daten umfassen nicht nur einfache Nutzungsstatistiken, sondern auch sehr detaillierte und sensible Informationen, die tiefe Einblicke in das Verhalten und die physischen Reaktionen der Nutzenden ermöglichen. Dieses mangelnde Wissen über das Ausmaß der Datensammlung führt zu einer immer größeren Bewusstseinslücke. »Es ist zu erwarten, dass die digitale Asymmetrie zwischen Anbieter- und Verbraucherseite durch den Einsatz immersiver Technologien weiter zunimmt.« (SVRV, 2024: 20)

Zudem betrifft der Datenschutz nicht nur die Nutzenden selbst, sondern auch unbeteiligte Dritte, deren Daten möglicherweise ebenfalls erfasst werden, ohne dass diese es bemerken oder dem zustimmen. Wenn eine VR-Brille beispielsweise genutzt wird, können auch die Daten von Personen in der Umgebung erfasst und gespeichert werden, ohne dass diese einwilligen oder sich dessen bewusst sind. Diese unbemerkte Erfassung von Daten unbeteiligter Dritter stellt eine erhebliche Datenschutzlücke dar, da diese Personen keine Kontrolle darüber haben, wie ihre unbewusst erfassten Informationen verwendet werden (SVRV, 2024: 21).

Diese gesammelten Daten ermöglichen es den Unternehmen allerdings, sehr tiefe Einblicke in das Leben der Nutzenden zu gewinnen und dadurch potenziell in Bereiche vorzudringen, die nicht nur das private Leben betreffen, sondern auch die Arbeitswelt oder den Versicherungsbereich. Und es ist immer damit zu rechnen, dass negative Folgen eintreten. So könnten beispielsweise Versicherungen eingeschränkte Körperbewegungen als Anzeichen für Erkrankungen werten und höhere Tarife verlangen oder Arbeitgeber könnten den Zugang zu bestimmten Stellen aufgrund von Gesundheitsdaten verweigern. In solchen Fällen zeigt sich, wie wichtig der Schutz personenbezogener Daten und die Wahrung der Privatsphäre der Nutzenden sind. Diese zu wahren, wird ein zu lösender Aspekt sein, um das Vertrauen der Verbraucherinnen und Verbraucher in die Angebote zu schaffen.

3.2 Datensicherheitsrisiken und notwendige Schutzmaßnahmen

Neben den datenschutzrechtlichen Herausforderungen stellen auch die Datensicherheitsrisiken eine erhebliche Bedrohung dar. XR-Anwendungen sammeln nicht nur eine enorme Menge an Daten, sondern diese Daten enthalten auch besonders sensible Informationen, die bei unbefugtem Zugriff missbraucht werden können. Erschwerend kommt hinzu, dass diese Daten besonders schwer zu anonymisieren sind, und sie könnten, falls sie in die falschen Hände geraten, für Identitätsdiebstahl oder betrügerische Aktivitäten missbraucht werden (u. a. IEEE Digital Reality, 2022). Cyberangriffe, Datenlecks und der unbefugte Zugriff auf persönliche Informationen gehören zu den gravierendsten Risiken, die mit der Nutzung dieser Technologien verbunden sind. XR-Plattformen könnten etwa Ziel von Cyberangriffen wie Malware oder Ransomware werden, was zu schwerwiegenden Sicherheitslücken und Datenlecks führen kann (Bradley Intelligence Report, 2023: 1). Angesichts der umfangreichen Datensammlung durch XR-Anwendungen ist das Risiko, dass Nutzende Opfer von Identitätsdiebstahl werden, daher besonders hoch. Wenn sensible Daten wie biometrische Informationen oder emotionale Reaktionen in die falschen Hände geraten, können diese für betrügerische Aktivitäten oder zur gezielten Manipulation von Verbraucherinnen und Verbraucher genutzt werden. Dies könnte nicht nur das Vertrauen in die Technologie untergraben, sondern auch das Sicherheitsgefühl der Nutzenden massiv beeinträchtigen.

Darüber hinaus zeigt sich durch medial bekanntgemachte Datenlecks regelmäßig, dass Unternehmen nicht ausreichend auf die Risiken im Bereich der Datensicherheit vorbereitet sind (so waren Unternehmen wie Facebook, Infoscore, Bonify, Verivox, Check24 und viele weitere bereits betroffen). Es werden nicht ausreichend präventive Maßnahmen getroffen (Bradley Intelligence Report, 2023: 1 ff.). Häufig investieren diese Unternehmen nicht genug in robuste Sicherheitsmaßnahmen, um ihre Systeme gegen Angriffe oder Datenlecks zu schützen. Das führt dazu, dass nicht nur finanzielle Schäden entstehen können, sondern auch die Privatsphäre und die persönliche Integrität der Nutzenden gefährdet werden. Wenn darüber hinaus auch Daten von Dritten ohne deren Wissen erfasst und gespeichert werden und diese Daten dann aufgrund von Sicherheitslücken in falsche Hände geraten, verstärkt dies das Mistrauen der Nutzerinnen und Nutzer in die Technologiebranche.

Um diesen Gefahren zu begegnen, sind eine vorausschauende Regulierung und ein klarer rechtlicher Rahmen erforderlich, der sicherstellt, dass Unternehmen angemessene Sicherheitsstandards einhalten. Es sollte gewährleistet werden, dass sowohl die Daten der nutzenden Personen als auch die von unbeteiligten Dritten ausreichend geschützt sind. Unternehmen müssen also nicht

nur die datenschutzrechtlichen Anforderungen der DSGVO erfüllen. Sie müssen auch in der Lage sein, effektive Sicherheitsvorkehrungen zu treffen, um ihre Systeme gegen die wachsende Bedrohung durch Cyberangriffe zu verteidigen. Nur durch einen soliden rechtlichen Rahmen und durch die Implementierung robuster Sicherheitsmaßnahmen können diese Risiken minimiert und das Vertrauen der Nutzerinnen und Nutzer in XR-Technologien gewahrt werden.

3.3 Verbraucherschutz in virtuellen Welten

Regulierung im Allgemeinen und Verbraucherschutz im Besonderen sind in virtuellen Umgebungen ein noch vernachlässigtes, jedoch entscheidendes Thema, insbesondere im Kontext der zunehmenden Verbreitung von XR-Technologien in Alltagsbereichen. Daher werden sich künftig »weitere verschiedene und zahlreiche Regulierungsfragen ergeben« (Wenzel et al., 2024: 33). Angesichts der steigenden Popularität dieser Technologien wird es für Verbraucherinnen und Verbraucher immer wichtiger, die Produkte und Dienstleistungen, die sie nutzen, überprüfen zu können. Gleichzeitig gestaltet sich die Identifizierung vertraglich relevanter Aspekte zunehmend schwieriger, insbesondere hinsichtlich verbraucherrelevanter Fragen wie die nach bestehenden Widerrufsrechten. Denn neben den Datenschutz- und Datensicherheitsfragen »stellen sich für das Zivil- respektive Vertragsrecht im Metaversum Fragen, die Formalitäten oder auch Haftung und Gewährleistung betreffen« (Wenzel et al., 2024: 22). Außerdem wird zu klären sein, inwieweit beispielsweise Avatare rechtswirksame Erklärungen abgeben können. Es zeigt sich also, dass »Anwendungen für VR und AR aus Sicht der Verbraucherpolitik zahlreiche Fragen aufwerfen« werden (SVRV, 2024: 8).

3.3.1 Verbraucherschützende Rechte und die Herausforderungen ihrer Durchsetzung

In virtuellen Welten sind Verbraucherschutzrechte, insbesondere das Widerrufsrecht und die Informationspflichten, von zentraler Bedeutung. Diese Rechte bieten Verbraucherinnen und Verbrauchern wichtige Schutzmechanismen, die Transparenz und Sicherheit in digitalen Umgebungen gewährleisten sollen (verbraucheranwalt-online.de, 2024). Das Widerrufsrecht ermöglicht es Verbraucherinnen und Verbrauchern, sich innerhalb einer bestimmten Frist und ohne Angabe von Gründen von Verträgen zu lösen. In virtuellen Welten ist dies besonders relevant, da viele Käufe und Dienstleistungen sofort digital geliefert werden, wie z. B. virtuelle Güter für Avatare oder der Zugang zu Events. Solche digitalen Produkte sind jedoch oft schwer rückgängig zu machen. Wenn das Widerrufsrecht bei digitalen Inhalten eingeschränkt wird, sobald diese genutzt

werden, sind klare und leicht verständliche Informationen erforderlich, damit Verbraucherinnen und Verbraucher die Bedingungen vor dem Kauf verstehen und bewusst zustimmen können. Die Informationspflichten der Anbieter müssen also dafür sorgen, dass Verbraucherinnen und Verbraucher umfassend über die Vertragsinhalte und über ihre Rechte aufgeklärt werden. Da in virtuellen Welten Entscheidungen oft in Echtzeit getroffen und Transaktionen sofort abgeschlossen werden, ist es entscheidend, dass Nutzende vorab transparent und verständlich über den Kaufvertrag und die Widerrufsregelungen informiert werden, um fundierte Entscheidungen treffen zu können.

In anderen Fällen haben Verbraucherinnen und Verbraucher auch Gewährleistungsrechte für den Fall, dass die Leistungen mangelhaft sind. So kann beispielsweise eine virtuelle Stadtführung oder Tour technisch Probleme aufweisen oder es fehlen versprochene Inhalte (verbraucheranwalt-online.de, 2024). Wird das Angebot direkt nach der Buchung genutzt, wird in der Regel das 14-tägige Widerrufsrecht ausgeschlossen. Allerdings stehen den Nutzenden dann noch Gewährleistungsrechte zu, also gegebenenfalls muss das Unternehmen die Leistung nochmal nachbessern oder die Kosten müssen zurückerstattet werden. Es ist daher wichtig, dass Verbraucherinnen und Verbraucher ihre Rechte kennen.

Allerdings stellt die Durchsetzung dieser Rechte in global operierenden XR-Umgebungen häufig eine Herausforderung dar. »Aufgrund der Internationalität des Metaversums und seiner verschiedenen (interoperablen) Plattformen, die von allen Nationalstaaten dieser Welt erreicht werden können, ist die sich zunächst aufdrängende Frage die des anwendbaren Rechts.« (Wenzel et al., 2024: 17) Viele dieser Plattformen haben ihren Sitz also in anderen Rechtsräumen, was die Geltendmachung von Verbraucherrechten verkomplizieren kann. Dies kann dazu führen, dass Verbraucherinnen und Verbraucher sich von der Durchsetzung ihrer Rechte abhalten lassen. Nutzerinnen und Nutzer brauchen daher einfache und klare Prozesse, um Beschwerden einzureichen oder ihre Rechte geltend zu machen.

3.3.2 Marketing, Werbung und manipulative Praktiken

Immersive Umgebungen bieten darüber hinaus neue Möglichkeiten für Marketing und Werbung, erhöhen damit aber auch das Risiko unlauterer Methoden. In einer zukünftigen virtuellen Umgebung könnten Unternehmen beispielsweise »vielfältigere Kommunikationsmöglichkeiten mit Kunden nutzen oder ihre Produkte und Dienstleistungen anders [als bisher] vermarkten« (Büchel & Klös, 2022: 7). Werbeinhalte könnten als Teil der Nutzererfahrung getarnt werden, wodurch es aufgrund des immersiven Charakters für Verbraucherinnen und Verbraucher schwieriger wird, zwischen Werbung und echtem Inhalt zu unterscheiden. Diese

Form des »Native Advertising« (Ritzel et al., 2013) kann problematisch sein, wenn Verbraucherinnen und Verbraucher unwissentlich für Produkte oder Dienstleistungen gewonnen werden, ohne über die Werbeinhalte informiert zu sein. Die Zunahme von Werbung in immersiven virtuellen Welten, ohne dass klare und transparente Informationen gegeben werden, stellt somit eine erhebliche Herausforderung für den Verbraucherschutz dar. Es wird entscheidend sein, dass auch in diesen neuen digitalen Räumen Verbraucherinnen und Verbraucher ihre Rechte wahren können, um nicht durch versteckte oder irreführende Werbung manipuliert zu werden.

Zusätzlich wird es Unternehmen aufgrund noch umfassender Datenanalysen und Künstlicher Intelligenz (KI) ermöglicht, gezielte, personalisierte Werbung zu schalten, die noch intensiver als bisher auf individuelle Vorlieben und Verhaltensmuster abgestimmt ist. Dadurch entstehen »neue Möglichkeiten der Personalisierung und zielgerichteten Werbung« (SVRV, 2024: 21). Obwohl dies in gewissem Maße auch zu positiven Nutzererfahrungen führen kann, besteht doch eine erhebliche Gefahr, dass diese personalisierte Werbung manipulativ wirken kann. Wenn Werbung auf so präzise Weise auf die persönlichen Präferenzen und das Verhalten der Konsumentinnen und Konsumenten zugeschnitten wird, könnte sie deren Entscheidungen auf subtile Weise beeinflussen, ohne dass diese sich dessen umfassend bewusst sind. Dies wirft Bedenken auf, ob Verbraucherinnen und Verbraucher in einer zunehmend digitalen und datengetriebenen Welt noch immer in der Lage sind, fundierte und freie Entscheidungen (selbst und bewusst) zu treffen.

Damit wirft die Entwicklung und Implementierung von XR-Technologien in der Konsequenz auch ethische Fragen auf, insbesondere im Hinblick auf den Einsatz von KI. Diese wird häufig eingesetzt, um personalisierte Erlebnisse zu schaffen, kann jedoch auch algorithmische Vorurteile erzeugen, die bestimmte Gruppen diskriminieren (Beck et al., 2019, 1–24). Dies ist besonders besorgniserregend in sensiblen Bereichen wie Bildung oder Gesundheit, in denen solche Ungleichheiten gravierende Auswirkungen haben können. Das macht deutlich, dass Bedenken hinsichtlich der Manipulation und Täuschung in virtuellen Umgebungen berechtigt sind. Der Einsatz von Technologien wie »Deepfakes« oder manipulierten Inhalten kann zudem das Vertrauen in Informationen untergraben und die Verbreitung von Fehlinformationen fördern. In einer Zeit, in der Fake News bereits eine Herausforderung darstellen, könnten XR-Technologien das Problem weiter verschärfen, indem sie es einfacher machen, täuschende Inhalte zu erstellen und zu verbreiten. »Der Einsatz eines KI-generierten fotorealistischen Avatar-Bots, der über die Bedürfnisse, Wünsche und Schwächen seines menschlichen Gegenübers bestens informiert ist, könnte dabei eine Manipulation von Ver-

braucherinnen und Verbrauchern in bisher unbekanntem Ausmaß ermöglichen.« (SVRV, 2024: 21)

3.3.3 Digitale Kluft und Ungleichheit

Unabhängig von den Inhalten der virtuellen Welten besteht ein weiteres Risiko darin, dass es zu einer ungleichen Verteilung des Zugangs zu XR-Technologien kommen kann. Während einige Nutzerinnen und Nutzer von den Vorteilen dieser neuen Technologien profitieren werden, werden andere möglicherweise keinen Zugang zu den erforderlichen Geräten oder der entsprechenden Infrastruktur haben. »Digitalität stellt armutsbetroffene Personen vor Herausforderungen: Zugangshürden, ungenügendes Wissen und Unterstützung oder kein stabiles Mobilenetzwerk. Dadurch spitzt sich der digitale Ausschluss zu und die gesellschaftliche Teilhabe erschwert sich.« (Chiapparini, 2024) Diese digitale Kluft kann zu sozialen Ungleichheiten führen, die sich negativ auf Bildungs- und Beschäftigungsmöglichkeiten auswirken könnten.

Ein Beispiel hierfür ist der mangelnde Zugang zu hochwertigen Bildungsressourcen, die durch XR-Technologien bereitgestellt werden. Während Schulen in wohlhabenden Gegenden möglicherweise Zugang zu VR-Labors und AR-Tools haben, könnte eine Schule in einem benachteiligten Viertel Schwierigkeiten haben, die notwendige Technologie bereitzustellen. Wie bereits ausgeführt, kann der Einsatz solcher Technologien potenziell bildungsförderlich sein. Allerdings können diese Unterschiede bestehende soziale Ungleichheiten vertiefen und den Zugang zu Bildungschancen einschränken. Besonders benachteiligte Gruppen sind in hohem Maße von digitaler Ungerechtigkeit und den negativen Auswirkungen der Digitalisierung betroffen. Sie profitieren dadurch weniger von der digitalen Transformation und sind nicht in der Lage, deren Potenziale zu ihrem Vorteil zu nutzen. Digitale Vulnerabilität bedeutet somit nicht nur den Ausschluss vom Erfolg der Digitalisierung, sondern auch ein höheres Risiko, negativ von ihren Folgen beeinflusst zu werden (Trkulja & Stiller 2024: 247–252).

Diese Problematik verdeutlicht, wie wichtig es ist, digitale Inklusion und den breiten Zugang zu digitalen Technologien in allen gesellschaftlichen Schichten zu fördern, um die Teilhabechancen zu sichern und digitale Ungerechtigkeit zu überwinden.

4. Fazit

Zusammenfassend lässt sich festhalten, dass XR-Anwendungen einerseits nützlich sein können, aber auch große Herausforderungen und Risiken bergen und

dass eine sorgfältige Auseinandersetzung mit diesen Technologien unerlässlich ist. XR hat das Potenzial, die Art und Weise, wie wir leben, nachhaltig zu verändern. Damit dieses Potenzial jedoch voll ausgeschöpft werden kann, müssen die Rechte und das Wohlergehen der Verbraucherinnen und Verbraucher an erster Stelle stehen. Ein proaktiver Ansatz, der Datenschutz, Datensicherheit und Verbraucherschutz in den Fokus rückt, ist entscheidend für eine verantwortungsvolle Nutzung von XR-Technologien. Bei der Frage, wie alltagstauglich die neuen Technologien werden können, sollte der D21-Digital-Index[2] berücksichtigt werden. Faktoren wie Alter, Geschlecht, kultureller und sozialer Hintergrund sowie andere demografische Merkmale sollten also bei der Entwicklung, Testung und Verbreitung von VR/AR-Technologien berücksichtigt werden. Nur so lässt sich sicherstellen, dass die Technologien nicht nur eine spezifische, technikaffine Nutzergruppe ansprechen, sondern für eine breite und diverse Bevölkerung zugänglich sind.

Ein solcher Ansatz fördert die Inklusion und Chancengleichheit, da er potenzielle Barrieren abbaut, die durch unterschiedliche digitale Kompetenzen, kulturelle Unterschiede oder soziale Ungleichheiten entstehen können. Zudem wird durch die Berücksichtigung der Vielfalt der Nutzergruppen sichergestellt, dass die entwickelten Technologien den tatsächlichen Bedürfnissen und Herausforderungen aller Gesellschaftsschichten gerecht werden. Insgesamt führt dies zu einer größeren Teilhabe und Integration diverser Bevölkerungsgruppen in die digitale Transformation.

Entwicklerinnen und Entwickler sowie Unternehmen sollten daher nicht nur auf bestehende gesetzliche Rahmenbedingungen reagieren, sondern auch vorausschauend handeln. Dazu gehört die kontinuierliche Überprüfung von Praktiken, um zu gewährleisten, dass die gesammelten Daten sicher sind und der Datenschutz gewährleistet bleibt. Die Implementierung fortschrittlicher Verschlüsselungstechnologien und anonymisierter Datenspeicherungen ist essenziell, um personenbezogene Informationen zu schützen. Insbesondere Informatikerinnen und Informatiker sollten sich zudem von Anfang an an ethische Standards und Normen halten, die sicherstellen, dass die Technologie nicht missbraucht wird. Ein Beispiel könnte die Auseinandersetzung mit ELSI-Fragen (ELSI-SAT, 2025) sein, die sich im Zusammenhang mit der Gestaltung und Nutzung solcher Wel-

2 Der D21-Digital-Index umfasst, um nur einige Aspekte zu nennen, die Kategorien Alter, Geschlecht, kultureller und sozialer Hintergrund bei der Entwicklung, Testung und Verbreitung von VR/AR-Technologien. Diese – auch intersektional-verstärkenden Aspekte – müssen berücksichtigt werden, um mehr Teilhabe von diversen Bevölkerungsschichten zu ermöglichen und auszubauen. Vgl. dazu u. a. D21-Digital-Index 2023/24: https://initiative21.de/uploads/03_Studien-Publikationen/D21-Digital-Index/2023-24/d21digitalindex_2023-2024.pdf (abgerufen am 10.02.2025).

ten ergeben. Entwicklerinnen und Entwickler müssen etwa die ethischen, rechtlichen und sozialen Auswirkungen, der von ihnen geschaffenen virtuellen Umgebungen berücksichtigen, wie z. B. den Schutz der Privatsphäre der Nutzenden und die Vermeidung von Manipulation. Durch die frühzeitige Auseinandersetzung mit diesen Aspekten kann sichergestellt werden, dass virtuelle Welten verantwortungsbewusst und im Einklang mit gesellschaftlichen Werten gestaltet werden. Darüber hinaus sollten Produkte benutzerfreundlich gestaltet und rechtliche Standards eingehalten werden, um Diskriminierung und die Missachtung von Verbraucherrechten zu vermeiden.

Nutzerinnen und Nutzer müssen zudem transparent über die Art der (berechtigt) gesammelten Daten und deren Verwendung informiert werden. Nur ein tatsächlich informierter Nutzender kann eigenverantwortliche Entscheidungen treffen. Das Bewusstsein über eigene Rechte und eine aktive Auseinandersetzung mit den Technologien sind dabei entscheidend. Dabei sollte die Gelegenheit genutzt werden, innovative Ansätze zu entwickeln, auf welche Art Informationen so dargestellt werden können, damit Verbraucherinnen und Verbraucher sie auch tatsächlich wahrnehmen. »Der Status quo, der sich dadurch auszeichnet, dass Informationen überwiegend langen Texten zu entnehmen sind, verfehlt diesen Zweck.« (SVRV, 2023: 8) Hier besteht eine Chance, mit technologischen Mitteln neue und zielführende Informationsvermittlungswege zu schaffen.

Aber auch der Gesetzgeber sollte geeignete rechtliche Rahmenbedingungen schaffen, die den Schutz der Verbraucherinnen und Verbraucher gewährleisten und gleichzeitig Innovationen im XR-Bereich nicht behindern. Anreize für Unternehmen, die verantwortungsvolle Praktiken verfolgen, sind eine mögliche Option, damit die richtigen Weichen gestellt werden. Schulen und Universitäten sollten XR in ihre Lehrpläne integrieren und den Studierenden die notwendigen Fähigkeiten vermitteln, um in einer von XR geprägten Zukunft teilhaben zu können. Sie könnten auch als Vermittler zwischen Entwicklerinnen und Entwicklern, Verbraucherinnen und Verbrauchern sowie Entscheidungsträgern fungieren, um einen offenen Dialog zu fördern.

Und schließlich wird es unerlässlich sein, dass der Zugang zu XR-Technologien gerecht gestaltet wird, sodass niemand ausgeschlossen wird. Es ist wichtig, Barrieren zu identifizieren und abzubauen, um den Zugang zu fördern, insbesondere in benachteiligten Gemeinschaften. Denn »hohe Zugangshürden, wie komplizierte Anwendungen und Funktionen für Einsteigerinnen und Einsteiger […] oder die mit dem Kauf der, für ein immersives Erlebnis erforderlichen, Endgeräte« (Wenzel et al., 2024: 32), die hohe Kosten verursachen, werden ansonsten einen großen Teil der Bevölkerung von der Nutzung abhalten. Förderprogramme könnten hierbei unterstützen.

Insgesamt verdeutlicht die Betrachtung der XR-Technologien, dass diese eine Transformation in vielen gesellschaftlichen Bereichen entfalten können. Um jedoch die damit verbundenen Herausforderungen und Risiken adäquat zu bewältigen, sind umfassende Maßnahmen aller Stakeholder erforderlich. Die Zusammenarbeit zur Förderung transparenter Praktiken, ethischer Standards und ein gerechter Zugang werden entscheidend sein, um XR zu einem attraktiven Angebot zu machen, welches die Lebensqualität der Menschen verbessert und gleichzeitig aktuelle Herausforderungen angeht. Nur durch eine gemeinsame Anstrengung kann eine Zukunft gestaltet werden, in der XR-Technologien nicht nur innovativ und aufregend, sondern auch gerecht, sicher und inklusiv sind.

Literatur

Beck, Susanne et al. (2019). *Künstliche Intelligenz und Diskriminierung, Herausforderungen und Lösungsansätze,* AG IT-Sicherheit, Privacy, Recht und Ethik Herausforderungen und Lösungsansätze, Whitepaper aus der Plattform Lernende Systeme, München, S. 1–24. https://www.plattform-lernende-systeme.de/publikationen-details/kuenstliche-intelligenz-und-diskriminierung-herausforderungen-und-loesungsansaetze.html [Zugriff 10.02.2025]

Bitkom (2020). *Lernen in immersiven Welten,* Autoren AK Learning Solutions, Bitkom e. V., Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e. V., Berlin, S. 7. https://www.bitkom.org/Bitkom/Publikationen/Lernen-in-immersiven-Welten [Zugriff 10.02.2025]

Bitkom (2022). Wegweiser in das Metaverse. In: *Technologische und rechtliche Grundlagen, geschäftliche Potenziale, gesellschaftliche Bedeutung*, Bitkom-Gremium AG Metaverse Forum unter der Projektleitung von Sebastian Klöß, Herausgeber Bitkom e. V., Berlin, S. 1–84. https://www.bitkom.org/sites/main/files/2022-07/220714_LF_Metaverse.pdf [Zugriff 10.02.2025]

Bradley (2023). Extended Reality: Risks and Opportunities on the Cutting Edge. In: *Bradley Intelligence Report,* Vol. 9/14/2023, Bradley Arant Boult Cummings LLP, Nashville. https://www.bradley.com/insights/publications/2023/09/extended-reality-risks-and-opportunities-on-the-cutting-edge [Zugriff 10.02.2025]

Büchel, Jan & Klös, Hans-Peter (2022). Metaverse. Hype oder »next big thing«? Potenziale und Erfolgsbedingungen. In: *IW-Report*, Nr. 42, Institut der Deutschen Wirtschaft, Köln, S. 1–34. https://www.iwkoeln.de/fileadmin/user_upload/Studien/Report/PDF/2022/IW-Report_2022-Metaverse.pdf [Zugriff 11.02.2025]

Chiapparini, Emanuela (2024). Digitalität und Armut: Wege zur Überwindung der digitalen Kluft. In: Caritas Schweiz Thementag »Digitale Kompetenzen für alle: Wege zur Überwindung der digitalen Kluft?«, Konferenzbeitrag, University of Applied Sciences, Bern. https://arbor.bfh.ch/entities/publication/ba36ca7c-fe75-471b-9aff-bba71e6ef445 [Zugriff 07.02.2025]

Deloitte (2022). Media Consumer Survey 2022. Das Metaverse aus Konsumentensicht. In: Deloitte Media Consumer Sumery 2022, Vol. 04/2022, Deloitte Touche Tohmatsu Limited,

London. https://www.deloitte.com/de/de/Industries/tmt/media-consumer-survey.html [Zugriff 14.02.2025]

ELSI-SAT (2025). ELSI-SAT Mehr Ethik und Recht in Ihrem Forschungsprojekt. Automatisiertes ELSI-Screening- und Assessment-Tool für MTI-Forschungsvorhaben, Hochschule der Medien, Stuttgart, Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg, Stuttgart, Institut für digitale Ethik, Stuttgart. https://www.elsi-sat.de [Zugriff 10.2.2025]

Grgic, Ivica (2024). Besser gewappnet für den Ernstfall – XR-Simulationstraining für Mediziner im Management von Notfallpatienten. In: Heinemann et al. (Hg.), *Digitale Realitäten in Medizin und Gesundheit,* MWV Medizinisch Wissenschaftliche Verlagsgesellschaft mbH & Co. KG., Berlin.

Gryl, Inga & Kuckuck, Miriam (Hg.) (2023). *Exkursionsdidaktik. Geographische Bildung in der Grundschule,* Verlag Julius Klinkhardt, Bad Heilbrunn. S. 14 ff.

Heynike, Francois & Wipper, Anna (2023). Virtuelle Welt, reale Probleme: Rechtliche Fragen im Metaverse. In: *KPMG Klardenker – das Portal für Wirtschaftsthemen,* KPMG AG Wirtschaftsprüfungsgesellschaft, Berlin. https://klardenker.kpmg.de/virtuelle-welt-reale-probleme-rechtliche-fragen-im-metaverse/ [Zugriff 11.02.2025]

IEEE Digital Reality (2022). *Virtual Reality Security*. IEEE Corporate Headquarter, New York. https://digitalreality.ieee.org/publications/virtual-reality-security (Zugriff 15.12.2024)

Johannes-Gutenberg-Universität Mainz (2024). *Virtual-Reality-Experimente, Konzept | Virtual-Reality-Experimente*, Johannes-Gutenberg-Universität, Mainz. https://www.vre.uni-mainz.de/ [Zugriff 11.02.2025]

Labster (2025). *Inspire Students with Immersive Learing,* Labster, Cambridge (MA) und Labster ApS, Coppenhagen. https://www.labster.com [Zugriff 11.02.2025]

Datenschutz & Datensicherheit (2023). Metaverse vs. Datenschutz: Die virtuellen Welten dürfen keine rechtsfreien Räume sein. In: *DSB Ratgeber,* Mensch und Medien GmbH, Augsburg. https://www.dsb-ratgeber.de/artikel/Metaverse-versus-Datenschutz.html [Zugriff 11.02.2025]

Ritzel, Lukas, Goodman, Steven & Van der Schaar, Cem (2013). Native Advertising: Das Trojanische Pferd der Marketing Strategen um das ultimative Gewinnmodell, Diplomica Verlag, Hamburg.

Riva, Giuseppe et al. (2021). Virtual reality therapy in the Metaverse: Merging VR for the outside with VR for the inside. In: Wiederhold, Brenda & Riva, Guiseppe (Hg.) *Annual Review of Cybertherapy and Telemedicine,* Vol. 19, 2021, Interactive Media Institute, Sent-Lambrechts-Woluwe, Belgien. S. 3–9. https://www.researchgate.net/publication/361582897_Virtual_Reality_Therapy_in_the_Metaverse_Merging_VR_for_the_Outside_with_VR_for_the_Inside/link/62bab97b93242c74cad37e82/download [Zugriff 10.2.2025]

Rosin, Jakob (2023). How extended reality tech could help visually impaired people. In: *The Parliament,* Political Holdings Limited, Hull, United Kingdom. https://www.theparliamentmagazine.eu/news/article/virtual-reality-technology-visually-impaired-people-help [Zugriff 15.12.2024]

Statista (2024): *Statistiken zu Geschäftsreisen*, Statista GmbH. Hamburg. https://de.statista.com/themen/821/geschaeftsreisen/ [Zugriff 30.10.2024]

SVRV (2024): *Effektive Verbraucherkommunikation. Veröffentlichungen des Sachverständigenrats für Verbraucherfragen*. Sachverständigenrat für Verbraucherfragen (Hg.), Berlin. https://www.

svr-verbraucherfragen.de/wp-content/uploads/2024/11/SVRV_Policy-Brief-Effektive-Verbraucherkommunikation.pdf [Zugriff 11.02.2025]

SVRV (2023). *Entwicklung der Informationspflichten. Verbraucherinformationen als Instrument des Verbraucherschutzes,* Sachverständigenrat für Verbraucherfragen beim Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz, Berlin, S. 1–11.

Trkulja, Violeta & Stiller, Juliane (2024). Digitale Teilhabe und Chancengleichheit: Wege zu gerechten Bildungsangeboten für Erwachsene. In: *Information. Wissenschaft & Praxis 2024.* Vol. 2024–2030, S. 243–252. https://doi.org/10.1515/iwp-2024-2030

verbraucheranwalt-online.de (2024). *Rechte und Schutz bei virtuellen Reisen und digitalen Reiseerlebnissen*, mak Anwaltskanzlei Menzel – Rechtsanwaltsgesellschaft mbH, Erfurt. https://www.verbraucheranwalt-online.de/blog/artikel/rechte-bei-virtuellen-reisen [Zugriff 11.02.2025]

Wenzel, Günter, Manca, Tiziana & Braun, Steffen (2024). *Regulierung des Metaversums. Analyse rechtlicher Aspekte virtueller und erweiterter Realitäten aus Innovationsperspektive,* Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Fraunhofer Publica, Stuttgart, S. 1–42. https://publica-rest.fraunhofer.de/server/api/core/bitstreams/97ad1e7e-a974-4cff-8598-77b523c6930e/content [Zugriff 10.02.2025]

Wössner, Stephanie (2024). *Next-level Learning: Die Potenziale von XR als Wegbereiter für eine Zukunft des Lernens*, Mixed.de, Deep Content GmbH. Hannover. https://mixed.de/next-level-learning-die-potenziale-von-xr-als-wegbereiter-fuer-eine-zukunft-des-lernens/ [Zugriff 07.02.2025]

II.
Transformative Technologien im Bildungsbereich

Lernen mit intelligentem Stift in der erweiterten Realität – FederLeicht

André Uhl, Nick Wittig, Addythia Saphala, Noro Schlorke, Christian Prison, Boris Kantzow, Michaela Evers-Wölk, Rolf Jürgens, Miriam Schmitz, Ricarda Steinmayr, Oliver Amft und Stefan Schneegass

Zusammenfassung

Der folgende Artikel präsentiert erste Ergebnisse aus dem Forschungsprojekt »FederLeicht – Lernen mit intelligentem Stift in der erweiterten Realität«, welches im August 2021 startete und im Dezember 2024 abgeschlossen wurde. Ziel des Verbundprojektes war die Erforschung und Entwicklung eines innovativen Lernsystems für die erweiterte Realität, das direkte Rückmeldung an die Schülerinnen und Schüler ermöglicht und dadurch die Lehrerschaft und Eltern entlastet. Mit diesem neuartigen Verfahren lassen sich Lernmaterialien, die mit Stift und Papier bearbeitet werden, um virtuelle Informationen erweitern. Der Beitrag beleuchtet die Funktionsweise von ProLamp, Smartpen und Triple-A-Anwendungen als technologische Kernelemente des FederLeicht-Lernsystems. Zudem werden erste Evaluationsergebnisse aus dem Einsatz des Lernsystems im Klassenzimmer vorgestellt sowie die im Projekt erarbeiteten Leitlinien für eine verantwortungsvolle Technologieentwicklung und -anwendung digitaler Lernassistenzsysteme zusammenfassend dargestellt.

1. Einleitung

Die Herausforderungen, mit denen sich Schülerinnen und Schüler täglich konfrontiert sehen, sind in den letzten Jahrzehnten stark gestiegen. Je nachdem welches Vorwissen die Kinder und Jugendlichen mitbringen, können sie neue, komplexe Lerninhalte unterschiedlich gut verarbeiten. Das führt dazu, dass sich individuelle Lernerfolge immer stärker voneinander unterscheiden. Hinzu kommt, dass Lernen längst nicht mehr ausschließlich im Klassenzimmer stattfindet: Daher wird das häusliche Umfeld zunehmend relevant. Als Antwort auf diese Herausforderungen hat das FederLeicht-Forschungsteam ein Lehr- und Lernsystem entwickelt, das erweiterte Realität nutzt, um Schülerinnen und Schülern interak-

tives Feedback beim Schreiben auf Papier zu geben. Diese individuelle Rückmeldungsfunktion entlastet Lehrende und Eltern.

Das FederLeicht-Lernsystem besteht aus drei integrativ kombinierten technologischen Kernelementen: einer intelligenten »Lampe« bestehend aus Computer, Lampenständer und Projektor, einem intelligenten Stift, der über Sensoren unterschiedliche Parameter messen kann, und einer Software-Anwendung, welche Autorensysteme, Avatare und Animationen in den Lernprozess einbezieht. Alle drei Bereiche werden in den nachfolgenden Kapiteln im Einzelnen beschrieben. Der interaktive Stift, hier Smartpen genannt, dient nicht nur zum Schreiben, sondern integrierte Sensoren messen auch Herzschlag, Druck und Beschleunigung und erfassen so die kognitive Belastung sowie den physischen und emotionalen Zustand der Lernenden. Das Lernsystem zeichnet auf, was die Schülerinnen und Schüler schreiben. Wenn es feststellt, dass etwa eine Aufgabe nicht richtig gelöst wurde, können virtuelle Korrekturhinweise ergänzt werden. Die intelligente Lampe, hier ProLamp genannt, projiziert diese Informationen räumlich präzise über den geschriebenen Text. Ein interaktiver Avatar kommuniziert mit den Schülerinnen und Schülern und dient als Ansprechpartner bei Fragen.

In enger Zusammenarbeit mit Expertinnen und Experten hat das Forschungsteam zudem ethische Leitlinien entwickelt, um den verantwortungsvollen Umgang mit der Technologie zu gewährleisten. Die Entwicklung bildet die Grundlage für eine neue Generation von Lehr- und Lernsystemen, die traditionelle und innovative Lehrmethoden vereinen und sich an die individuellen Fähigkeiten und Bedürfnisse der Schülerinnen und Schüler anpassen.

2. ProLamp

Die ProLamp ist ein projektionsbasiertes Augmented-Reality(AR)-Projektor-Kamera-System, das als eigenständiges, quelloffenes und 3D-druckbares Gerät entwickelt wurde. Dieses innovative System wurde mit dem Schwerpunkt auf Zugänglichkeit und Reproduzierbarkeit entwickelt, wodurch es sich für verschiedene Anwendungen, insbesondere im Bildungsbereich, eignet. Die ProLamp bietet eine praktische Lösung, um AR in alltägliche Umgebungen zu bringen, wie z. B. in Schulen oder Lernräume zu Hause, indem sie Bildungsinhalte, Hinweise und Feedback direkt in die physische Welt projiziert.

2.1 Konzept und Anwendungsfall

Die ProLamp wurde entwickelt, um Bildungserfahrungen zu verbessern, sowohl individuell als auch gemeinschaftlich. Sie ermöglicht es den Schülerinnen und Schülern, selbstgesteuert zu lernen, indem sie räumliches AR-Feedback in Form von Anleitungen oder Hinweisen projiziert. Das Gerät kann automatisch Fehler erkennen und personalisiertes Feedback geben, um eine rechtzeitige Unterstützung bei Lernaktivitäten zu gewährleisten. Die Einsatzmöglichkeiten reichen vom Bearbeiten von Einzelaufgaben bis hin zur Zusammenarbeit in Gruppen.

Die ProLamp ermöglicht eine interaktive Lernerfahrung in Umgebungen, in denen Lehrkräfte nicht immer anwesend sind, um sofortige Hilfe zu leisten.



Abbildung 1: ProLamp projiziert auf den Arbeitsplatz
Quelle: Projekteigene Darstellung

2.2 Entwurf und Herstellung der ProLamp

Die ProLamp, zu sehen in Abbildung 1, wurde als individuelle, verstellbare Schreibtischlampe entworfen, die sich nahtlos in verschiedene Umgebungen wie Schulen oder Arbeitsbereiche integrieren lässt. Die Lampe besteht aus 3D-gedruckten Komponenten, die einen Lampenfuß, einen Arm und ein Kopfstück kombinieren, die alle so gestaltet sind, dass sie ein Gleichgewicht zwischen Funktionalität und Ästhetik herstellen. Die Projektions- und Kamerakomponenten sind diskret im Kopfstück untergebracht, das von einem Lampenschirm verdeckt wird, um ein sauberes, unauffälliges Erscheinungsbild zu gewährleisten. Der Sockel der Leuchte sorgt für Stabilität und beherbergt die notwendigen Verarbeitungskomponenten sowie die Stromversorgung des Systems. Der verstellbare Arm ermöglicht eine flexible Positionierung, da er sowohl in der Höhe als auch in der Tiefe verstellt werden kann, um verschiedenen Aufstellungen gerecht zu werden. Die Armsegmente sind mit Gelenken verbunden, die auch der effizienten Kabelführung dienen und für ein aufgeräumtes und organisiertes Erscheinungsbild sorgen. Mit ihrem Design fügt sich die ProLamp nahtlos in jede Schreibtischumgebung ein. Das vielseitige Design macht die ProLamp zu einem benutzerfreundlichen räumlichen AR-Tool, das sich sowohl für den Bildungsbereich als auch für professionelle Umgebungen eignet, in denen interaktives, projektionsbasiertes Feedback von Vorteil sein kann.

2.3 AR-Feedback-Anwendung

Die für die ProLamp entwickelte AR-Feedback-Anwendung ermöglicht AR-Interaktionen in Echtzeit, sodass die Benutzer:innen sofortiges, personalisiertes Feedback in Bildungs- oder Berufsumgebungen erhalten können. Die Software wurde entwickelt, um die interaktive Lernerfahrung zu verbessern, indem Feedback direkt auf physische Oberflächen projiziert wird, wie z. B. handschriftliche Antworten auf Übungsblättern, wodurch der Lernprozess ansprechender und intuitiver wird. Ein Hauptmerkmal der Anwendung ist die Fähigkeit, die genaue Ausrichtung der AR-Projektionen zu gewährleisten, was eine nahtlose Interaktion zwischen den digitalen Inhalten und der physischen Welt ermöglicht. Die Anwendung bietet visuelles Feedback, wie z. B. farbkodierte Signale für richtige oder falsche Antworten sowie zusätzliche Symbole, um die Zugänglichkeit für alle Nutzenden zu gewährleisten.

Die Software arbeitet in zwei Modi: zum einen für die Verwaltung und Bereitstellung von Feedback über eine intuitive Schnittstelle und zum anderen für die Projektion dieses Feedbacks in Echtzeit. Diese beiden Modi arbeiten über ein

Netzwerk zusammen, sodass der Feedbackprozess zwischen den Nutzenden und dem System synchronisiert werden kann.

In Kombination mit der ProLamp-Hardware bietet diese AR-Feedback-Anwendung ein hochgradig interaktives, flexibles Tool, das ein breites Spektrum von Lernumgebungen unterstützt, indem es zeitnahes und zugängliches Feedback liefert.

2.4 Evaluation der ProLamp

Die ProLamp wurde im Rahmen des FederLeicht-Projekts einer umfassenden Evaluation unterzogen, die im Dezember 2024 abgeschlossen wurde. Bei dieser wurde in einer Stichprobe von 150 Kindern der sechsten Jahrgangsstufe in Nordrhein-Westfalen die Benutzerfreundlichkeit sowie die Auswirkungen auf die Motivation und die Leistung der Benutzer:innen erfasst und mit dem Lernen ohne die ProLamp verglichen. Ziel dieser Evaluierung war es, herauszufinden, wie sich die Lampe im Vergleich zu traditionellem Lernen auf die aktuelle Motivation und Leistung der Kinder auswirkt. Es wurden bereits mehrere Veröffentlichungen auf der Grundlage von Studien mit der ProLamp vorgelegt, die sich auf die Leistung der ProLamp bei der Bereitstellung von AR-Feedback konzentrieren. Die Forscher:innen waren besonders daran interessiert, herauszufinden, welche Feedback-Modalitäten am nützlichsten sind und wie das Gerät die Lernerfahrung verbessert. Auch im Kontext von Fernlehre wurde eine Studie durchgeführt (Wittig et al., 2024: 255–263). Diese Auswertungen werden wertvolle Erkenntnisse darüber liefern, wie räumliche AR-Technologie, wie die ProLamp, für den Einsatz im Bildungsbereich optimiert werden kann und welches Potenzial sie hat, die Motivation und die Lernergebnisse zu verbessern.

3. Smartpen

Der Smartpen wurde entwickelt, um relevante Informationen über das Verhalten und die Anspannung von Schülerinnen und Schülern in Lernsituationen zu erfassen. Insbesondere erfassen Sensoren im Smartpen physiologische und physische Informationen, während der Smartpen benutzt wird. Somit kann der Smartpen zusammen mit der ProLamp im AR-Lernsystem eingesetzt werden. Die Integration von Sensoren zur Erfassung der Herzaktivität im Smartpen bietet sich aufgrund des direkten Hautkontakts an. Weitere Sensoren im Smartpen ermöglichen es, die Bewegung während der Nutzung in der Lernsituation sowie die aufgewendete Kraft beim Schreiben zu erfassen. Trotz der Sensoren kann der Smart-

pen wie ein normaler Kugelschreiber genutzt werden, ohne dass die Lernaktivitäten der Schülerinnen und Schüler beeinträchtigt werden.

3.1 Hintergrund

In den letzten Jahren wurden zahlreiche Untersuchungen und Entwicklungen im Zusammenhang mit der Messung der kognitiven Belastung durchgeführt. Zhou et al. erläuterten die theoretischen Grundlagen für die verhaltensbasierte und physiologische Messung kognitiver Belastung, einschließlich des multimodalen Modells zur Messung kognitiver Belastung für robuste Messungen kognitiver Belastung, Operationen, Datenfluss und Datenfusionstechniken (Zhou et al., 2018: 287–329). Ahmad et al. schlugen einen Rahmen zur Erkennung kognitiver Belastung durch nichtinvasive Messung physiologischer Daten von Augen und Herz vor. Das System der Autor:innen verwendete überwachtes maschinelles Lernen mit Techniken zur Merkmalseliminierung und zeigte die Klassifizierung von drei verschiedenen Stufen kognitiver Belastung. Die Ergebnisse zeigten, dass die Herzfrequenzvariabilität (HRV) mit zunehmender kognitiver Belastung leicht abnahm und die Herzfrequenz (HR) mit zunehmender kognitiver Belastung leicht anstieg (Ahmad et al., 2020: 1–15). Muthukrishnan untersuchten an 24 gesunden Proband:innen den Zusammenhang zwischen HRV und kognitiver Leistungsfähigkeit anhand eines visuell-räumlichen Arbeitsgedächtnisparadigmas. Die Korrelationsanalyse nach Pearson wurde verwendet, um die Korrelation zwischen der HRV und der Leistung bei jeder Gedächtnisaufgabe zu untersuchen. Die Ergebnisse zeigten, dass die HRV-Messungen mit zunehmender Gedächtnisbelastung abnahmen (Muthukrishnan et al., 2017: 14–22). Im Vorhaben FederLeicht wird erstmals ein Smartpen für die Schätzung von kognitiver Belastung entwickelt und erprobt.

3.2 Smartpen-Entwicklung

Als Plattform für die Integration der notwendigen elektronischen Komponenten wurden 3D-druckbare Smartpen-Rahmen entworfen (siehe Abbildung 2). Der Griffbereich des Stiftes wurde anhand von Beispielen und Anregungen des Projektpartners LAMY entwickelt. Als Mikrocontroller wurde der Seeed Xiao BLE Sense mit integriertem Batteriemanagement, Bluetooth und einer Inertialmesseinheit (IMU) ausgewählt. Die Sensorplatine unter der transparenten Abdeckung des Griffstücks enthält einen Photoplethysmographen(PPG)-Sensor zur Erfassung des Blutvolumenpulses (BVP) und der HRV. Der Smartpen ist für

den kabellosen Betrieb mit Bluetooth-Datenübertragung und wiederaufladbarer Batterie ausgelegt. Die Sensordaten werden zusammen mit dem Zeitstempel über Bluetooth mit einer Abtastrate von 20 Hz an das Computersystem übertragen, wo sie dann weiterverarbeitet werden.

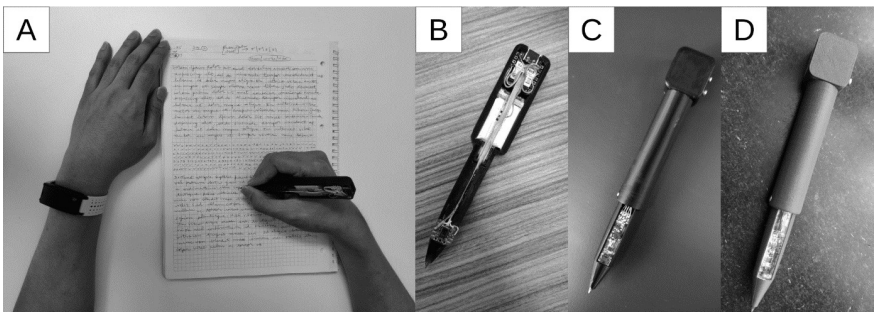


Abbildung 2: FederLeicht-Smartpen.
(A) Studienteilnehmende beim Schreiben mit dem FederLeicht-Smartpen-Prototyp. (B) Erster funktionierender Prototyp. (C) Alter Prototyp. (D) Aktuelle Version
Quelle: Projekteigene Darstellung

3.3 Smartpen-Studie

Das Ziel der Smartpen-Studie war es, die Fähigkeit des Smartpen zu bewerten und die kognitive Belastung der Nutzer:innen einzuschätzen. Für das Vorversuchsexperiment wurden die Daten von sechs freiwilligen Teilnehmer:innen aus dem Intelligent Embedded Systems Lab der Universität Freiburg verwendet. Die Studie dauerte durchschnittlich 55 Minuten pro Teilnehmer:in. Die Daten der PPG- und IMU-Sensoren des Smart Pen wurden während der Studie aufgezeichnet.

Das Experiment war in drei Teile mit unterschiedlichen kognitiven Belastungsstufen organisiert. Jeder Abschnitt enthielt eine Schreibaufgabe und Fragebögen, die ausgefüllt werden mussten. Der erste Abschnitt war eine Baseline-Aufnahme, bei der die Teilnehmer:innen einen Absatz auf ein leeres Blatt Papier schreiben sollten. Im zweiten Abschnitt wurden die Teilnehmer:innen durch Ereignisse mit hoher kognitiver Belastung, wie z. B. Interviews und Rechentests, unter Stress gesetzt. Im dritten Abschnitt wurde ein Ereignis mit geringer kognitiver Belastung simuliert, indem Langeweile durch das Drehen von Bleistiften erzeugt wurde.

Der Perceived Stress Questionnaire (PSQ) (Fliege et al., 2009: 142–152) und die Multidimensional State Boredom Scale (MSBS) (Hunter et al., 2015: 1–10; Donati et al., 2021: 928–941) wurden verwendet, um die induzierte kognitive Belastung durch das Experiment zu überprüfen. Der PSQ (Fragebogen zur Stresswahrnehmung) ist ein Selbstbeurteilungsinstrument zur Messung der Stresswahrnehmung und des wahrgenommenen Stressniveaus einer Person. Der PSQ besteht aus Fragen, die Personen auf der Grundlage ihrer Erfahrungen und Gefühle in einem bestimmten Zeitraum beantworten sollen. Die MSBS ist ebenfalls ein Selbstbeurteilungsinstrument, mit dem das Erleben von Langeweile bei Einzelpersonen gemessen und bewertet werden kann. Die MSBS enthält Fragen und Aussagen zu verschiedenen Dimensionen von Langeweile.

3.4 Schätzung von mentaler Belastung

Das Ziel der KI-Analyse ist die binäre Bewertung der kognitiven Belastung der Benutzer:innen, z. B. hoch oder niedrig. Dazu werden die Sensordaten in Segmente zerlegt. Die Datensegmente werden anhand der Selbsteinschätzungen von Studienteilnehmenden als hohe oder niedrige kognitive Belastung gekennzeichnet. Die Sensordaten werden anschließend mit Softwarefiltern vorverarbeitet, um Bewegungsartefakte zu reduzieren. Außerdem werden Fenstertechniken eingesetzt, um die Zeitreihendaten zu segmentieren. Die PPG-Daten wurden im Zeitbereich, im Frequenzbereich und im nichtlinearen Bereich analysiert und Merkmale der HRV extrahiert, um die kognitive Belastung anzuzeigen. IMU-Merkmale, wie Mittelwert, Median, Standardabweichung, Varianz, Kurtosis, Schiefe und mittlere absolute Abweichung, wurden nach dem Vorbild der Auswahl von Handgesten-Merkmalen ausgewählt (Suri & Gupta, 2018: 45–50; Siddiqui & Chan, 2020: 1–12).

Es wurden verschiedene Algorithmen, einschließlich Support Vector Machine (SVM), Gradient Boosting (GB) und Gaussian Naive Bayes (GNB), verwendet, um KI-Lernmodelle zu trainieren. Die KI-Modelle wurden mit der Teilnehmer:in-Auslassen-Strategie validiert, wobei die KI-Modelle mit den Daten eines/einer Testteilnehmenden überprüft wurden, während die Daten der restlichen Teilnehmenden als Trainingsdatensatz verwendet wurden. Mit dem beschriebenen Validierungsverfahren konnte eine Genauigkeit von ca. 80 % über alle Fenstersegmente für die richtige Erkennung von hoher bzw. geringer Belastung erreicht werden.

3.5 Erfahrungen aus der Smartpen-Studie

In der Studie zur Abschätzung der kognitiven Belastung kann die Korrelation zwischen physiologischen Daten und dem kognitiven Zustand aus den Sensordaten des intelligenten Stifts während der Schreibaufgaben abgeleitet werden. Weitere Experimente und Entwicklungen sind notwendig, um die Zuverlässigkeit und Leistung bei der Ableitung der kognitiven Belastung zu verbessern. Während das KI-Lernmodell für strukturierte Labordaten gut funktioniert, ergeben sich bei realen Situationen unterschiedliche Herausforderungen. Technische Herausforderungen, wie unsachgemäße Handhabung, Energiebeschränkungen und Bewegungsartefakte, werden durch die Weiterentwicklung gelöst. Darüber hinaus müssen individuelle Unterschiede in der Reaktion auf kognitive Belastung, die sich sowohl in der Stiftbewegung als auch in der Herzaktivität zeigen, weiter untersucht werden.

4. Triple-A-Anwendungen: Autoren – Avatare – Animationen

Um den resultierenden Prototyp für die Zielgruppe Schülerinnen und Schüler nicht nur didaktisch wertvoll, sondern auch möglichst ansprechend und unterhaltsam zu gestalten, was aus didaktischer Sicht ebenfalls von großer Bedeutung ist, lagen zusätzliche Entwicklungsschwerpunkte auf den Bereichen Autorensysteme, Gamification und Design von Avataren.

4.1 Autorensysteme

Eine der größten Herausforderungen beim Einsatz augmentierender Systeme ist die Erstellung der Inhalte. Oft müssen diese mithilfe von Game Developers und 3D Artists gestaltet werden, was den Prozess komplex und zeitintensiv macht. Die Zielsetzung war es daher, Werkzeuge zu identifizieren, die es Konzeptionist:innen ermöglichen, Inhalte eigenständig und mit minimaler Unterstützung durch Entwickler:innen zu erstellen.

Ein Beispiel für ein solches Tool ist »Articy Draft«. Der Vorteil dieses Produkts liegt in seiner grafischen, intuitiven Bedienbarkeit und seinen Schnittstellen (API) zu den bekannten Game Engines Unity und Unreal. Während der Evaluation zeigte sich jedoch, dass viel manuelle Arbeit notwendig ist, was die wertvolle Zeit von Lehrenden bindet. Zudem sind die Schnittstellen zu Unity und Unreal eher rudimentär und wenig geeignet für die komplexe Anwendungsentwicklung.

Ein weiteres Tool, »miro«, ist browserbasiert und unter Konzeptionierenden weit verbreitet. Die Freiheit in der Gestaltung ist ein großer Vorteil, ebenso wie die umfangreiche API, die sich gut in eine Production Pipeline integrieren lässt. Allerdings birgt genau diese Freiheit auch das Risiko unstrukturierter Konzepte, die schwer maschinell lesbar sind. Außerdem erfordert die API eine lange Einarbeitungszeit, besonders für Game Developer, die ihre Fähigkeiten an anderer Stelle haben. Wichtige, für die Integration notwendige Mechanismen (wie die Vergabe von IDs für Konzeptbausteine) müssen zudem manuell gepflegt werden.

Das dritte Tool, »storyboarder.ai«, setzt auf browserbasierte KI-Unterstützung und ermöglicht sogar automatische Bebilderung und Produktionsplanung. Allerdings liegt der Fokus hier stark auf der Filmproduktion, was es für die Anwendungsentwicklung wenig nutzbar macht. Es könnte für Anwendungen mit stark linearem Spielverlauf geeignet sein, vor allem der Schritt von der Idee zum cinematischen Spielverlauf lässt sich durch dieses Tool gut vollziehen und visuell umsetzen. Aus Sicht der Autor:innen existiert derzeit im Bereich der Anwendungsentwicklung noch kein Markt für Produkte, die Konzeptionist:innen wirklich unterstützen.

4.2 Avatar-Entwicklung

Ein zentrales Ziel der Projektarbeit war die emotionale Bindung der Lernenden an eine digitale Begleitfigur. Diese sollte nicht nur unterstützende Funktionen übernehmen, sondern auch als personalisierte Figur agieren, um das Lernsetting so realitätsnah wie möglich zu gestalten. Gamification spielte in diesem Zusammenhang eine wichtige Rolle. Viele der dafür notwendigen Assets – wie Belohnungssysteme, Level-Fortschritte oder versteckte »Easter Eggs« – wurden bereits außerhalb des Forschungsprojekts entwickelt. Zusätzlich wurden Minispiele konzipiert, die bei Bedarf als Belohnung oder zur Steigerung der Aufmerksamkeitsspanne der Lernenden eingebaut werden können. Eine besonders spannende Idee war die eines »Sidekicks«, wie man sie aus Superhelden- oder Talkshow-Kontexten kennt. Entwickelt wurden zwei Avatare, die sowohl miteinander als auch mit den Anwender:innen interagieren.

Das Ergebnis waren »Doc Jenkins«, ein zerstreuter Professor, der den Fokus auf Lernziele legt, und »Joy-CE«, ein sympathischer, fliegender Roboter, der für Spaß, Engagement und Interesse sorgt (siehe Abbildung 3). Für beide Charaktere wurden mehrere grafische Iterationen durchlaufen, um ihren Look und ihre Persönlichkeit zu verfeinern. Parallel dazu wurden verschiedene Lernumgebungen entwickelt, die flexibel eingesetzt werden können.

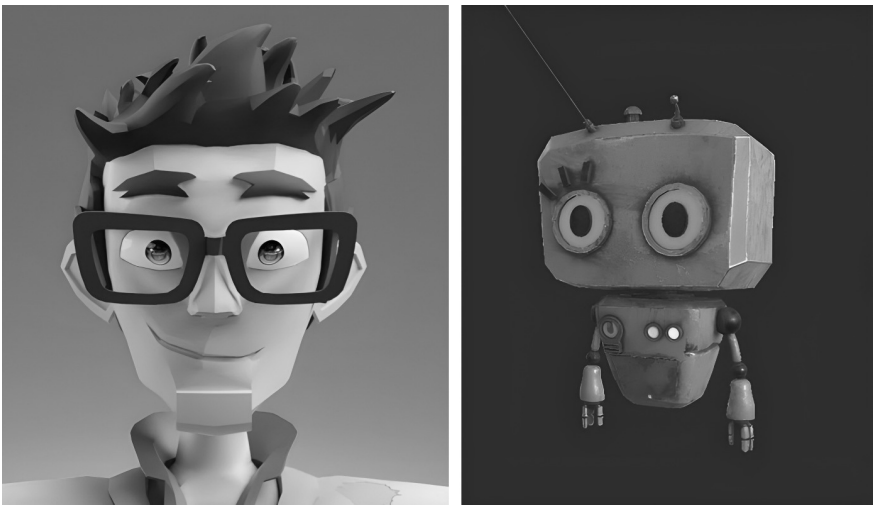


Abbildung 3: Avatare »Doc Jenkins« und »Joy-CE«
Quelle: Projekteigene Darstellung

4.3 Automatische Animation

Beim Abschlusstreffen des Fördermittelträgers wurde dem Projekt FederLeicht auch ein weiteres Projekt vorgestellt, das die automatische Erstellung von Animationen aus Audiodateien ermöglichte. Dadurch inspiriert entstand die Idee, dass durch die direkte Konvertierung von Audio zu Animation der zeitaufwendige manuelle Prozess der Animationserstellung deutlich verkürzt werden könnte. Es folgte eine intensive Auseinandersetzung mit verschiedenen Code-Bibliotheken. Als vielversprechend stellte sich NVIDIAs »Omniverse«-Plattform heraus. Das Programm »Machinima« besitzt ein Plugin namens »Audio2Gesture«, welches genau die Funktionalität der Animationserstellung aus Audiodateien bereitstellt, nach der gesucht wurde. Das Programm selbst ist kostenfrei, erfordert jedoch eine NVIDIA-Grafikkarte. »Audio2Gesture« funktioniert, indem es Audioaufnahmen analysiert und auf Basis des Sprachflusses Bewegungen für einen Avatar generiert. Die erstellten Animationen passen sich gut an den Redefluss an, was bedeutet, dass der Charakter eine natürliche Körpersprache entwickelt, während er spricht. Die Bedienung ist allerdings nicht einfach genug für die Zielgruppe. Ein weiterer Nachteil ist, dass das System nur humanoide Charaktere unterstützt. Dies bedeutet, dass der Roboter-Avatar »Joy-CE«, der weder Beine noch einen Mund hat und dessen Arme nicht mit dem Körper verbunden sind, nicht von »Audio2Gesture« animiert werden kann. Trotz dieser Einschränkun-

gen konnte die Technik erfolgreich auf den humanoiden Avatar »Doc Jenkins« angewendet und eine Vielzahl natürlicher Bewegungen generiert werden.

4.4 Ergebnis

Ein zentrales Ergebnis war die Erkenntnis, dass es bei der Erstellung von Lerninhalten nach wie vor an benutzerfreundlichen Tools mangelt, die ohne umfassende technische Vorkenntnisse bedient werden können. Dennoch wurden mit der Nutzung von Tools wie »miro« oder »Machinima« vielversprechende Ansätze identifiziert, die weiterentwickelt und ausgebaut werden sollen. Gerade zum Ende des Projektes war deutlich der Impact von KI auf die Arbeitswelt zu spüren: Übersetzungen, Text to Speech, automatisierte Animationen und vieles mehr sind bereits möglich. Täglich erscheinen neue Tools, bereits heute sind sie aus unserem Arbeitsalltag nicht mehr wegzudenken. Im Rahmen dieser Erkenntnis hat das Team sich auch intensiv mit der automatisierten Erstellung von Videos mit »realen Avataren«, also z. B. gesprochene Erklärvideos, die niemals von ihnen aufgenommen wurden, beschäftigt. Die Ergebnisse sind erstaunlich gut und wirken real. Es besteht begründete Hoffnung, dass auch die Erstellung von Animationen in wenigen Monaten durch eine No-Code-KI möglich sein werden.

5. Akzeptanz im Klassenzimmer

Innovative Technologie wie das FederLeicht-Lernsystem und neue Lernkonzepte müssen im Schulkontext eine entscheidende Barriere überwinden: Sie müssen von Schüler:innen und Lehrkräften angenommen werden. Erfahrungsgemäß wird am besten angenommen, was sich möglichst ohne Widerstände mit dem Bestehendem verbinden lässt (Hallfors & Godette, 2002: 463). Sind die geplanten Veränderungen des Bestehenden klein, ist es wahrscheinlicher, dass die Lehrkräfte die Änderungen akzeptieren und umsetzen. Interventionen im Schulkontext sollten daher immer an bestehende Strukturen anschließen, auch dann, wenn anders möglicherweise größere positive Effekte erzielt werden könnten. Doch nicht nur die Distanz zum Bestehenden ist entscheidend dafür, ob Interventionen im Schulkontext angenommen werden. Auch persönliche Einstellungen von Lehrkräften und Schulmanagement spielen eine entscheidende Rolle und nicht zuletzt auch die Einstellung der Schüler:innen (Domitrovich et al., 2008: 16; Fixsen et al., 2009: 535).

Als erster Schritt in der Evaluation der im Projekt geplanten Interventionen hat sich das Team aus diesen Gründen mit der Akzeptanz der Technologien bei

Schüler:innen befasst. Parallel zur Erprobung schon der ersten Technologien im Klassenzimmer – dem von der Universität Duisburg-Essen entwickelten Übungsprogramm mittels Projektor und dem von der Universität Freiburg entwickelten interaktiven Stift – wurde jeweils die Akzeptanz der Technologien unter Schüler:innen erhoben. Quantitativ ergaben sich bei der Befragung von 115 Schüler:innen hierbei stets ein hohe Akzeptanz und eine breite Begeisterung über Geschlechter und sozioökonomischen Status hinweg. Allerdings ist bei einer solchen Erfragung zu beachten, dass es sich in aller Regel um den ersten Kontakt eines Kindes mit der Technologie handelt. Darüber hinaus ist eine Untersuchung im Klassenraum auch immer eine Unterbrechung des oft als langweilig empfundenen Unterrichts und ermöglicht es, Schulaufgaben durch Teilnahme an der Studie zu umgehen. Es darf also angenommen werden, dass Kinder dem Untersuchungsgegenstand bei dieser Befragung mit einer deutlich positiveren Grundhaltung gegenübergetreten sind, als sie es nach Monaten regelmäßiger Benutzung im Schulkontext täten.

Neben der quantitativen Erhebung der Technologieakzeptanz wurden Schüler:innen daher auch offen nach ihrer Meinung über die im Projekt entwickelte Technologie gefragt. Auffällig hierbei ist, dass sich die befragten Kinder bei der Verwendung des Stifts vor allem zu Eigenschaften des Stifts äußerten. Das Gewicht und die Form des Stiftes hatten für die Kinder große Bedeutung, während die im Stift verbauten Sensoren für viele Kinder keine weiteren Kommentare wert waren. Aus der Perspektive der befragten Kinder ist es die zentrale Funktion eines Stiftes, zu schreiben. Anders stellt sich das bei dem Projektor dar. Die Funktionalität des Projektors im Rahmen unseres Forschungsprojekts war für die Kinder sehr viel leichter zu erkennen. Entsprechend vielfältig fielen Ideen über Veränderungen und mögliche Verbesserungen aus. Zentrale Rolle in den Antworten spielte die Raum-Frage. Ein eigener Beamer auf dem Tisch jedes Kindes? Das war vielen Kindern zu viel und gleichzeitig wurde die Funktionalität – schnelle Rückmeldung zu Lösungen, ohne jedes Mal die Lehrkraft zu fragen – hoch gelobt. Die Vorschläge der Kinder zur Lösung dieser Spannung reichten von »ein Beamer pro Tisch« über »Vielleicht könnte man das irgendwie an die Decke machen« bis hin zu »In jedem Klassenraum sollte es einen geben, zu dem dann gegangen werden kann, um Ergebnisse zu checken«. Die Antworten der Kinder zeigen zum einem, dass sie gegenüber neuen Technologien im Zimmer sehr aufgeschlossen sind, und zum anderen, dass sie sich mehr Feedback zu ihren Arbeiten wünschen.

Diese Antworten verdeutlichen darüber hinaus eine weitere Dimension, in der Technologien im Klassenraum miteinander konkurrieren: der Raum. Lehrbuch, Arbeitsblatt, Projektor, Federmappe, Trinkflasche und Arme des Kindes müssen sich alle den gleichen Tisch teilen. Entsprechend begründet waren die Anstrengungen im Projekt über Aufgabenauswahl mithilfe eines Recommender-

Systems und automatisierter Hilfestellungen und Hinweise über die AR-Umgebung und Begleitung der Schüler:innen den Platz auf dem Arbeitstisch zu rechtfertigen. So wichtig wie die Frage danach, wie ein Projektor zur Benutzung der Kinder am Arbeitsplatz räumlich ins Klassenzimmer passt, so deutlich ist der erkennbare Wunsch von Schüler:innen nach schneller und unkomplizierter Rückmeldung zu den Ergebnissen ihrer Arbeit. Andere Technologien wie Tablets ermöglichen dies bereits, limitieren aber dafür in anderen Bereichen – so erlaubt eine Projektion auf den Arbeitsplatz beispielsweise das gleichzeitige Kritzeln auf verwendetem Papier, ein oft missverstandener, aber deutlich lernunterstützender Prozess (Sundararaman, 2020: 31).

Unterstützend für den Lernprozess wäre es auch, wenn Kinder sich etwas weniger im Unterricht langweilen würden. Ziel des im Projekt entwickelten Stifts ist es, mit verschiedenen physiologischen Indikatoren den emotionalen Zustand der Kinder zu diagnostizieren. Dazu gehört u. a. die Emotion der Langeweile. Langeweile ist nicht nur weit verbreitet im Klassenraum, Langeweile hat auch deutliche negative Effekte auf Lernleistungen (Sundararaman, 2020: 31). Der entwickelte Stift soll Langeweile erkennen und dann eine Mikro-Intervention auslösen, mit der Kinder an die Vorteile möglicher Lernerfolge erinnert werden und so die Motivation zur aktiven Unterrichtsteilnahme wieder gestärkt werden soll. Darüber hinaus sind weitere Einsatzmöglichkeiten denkbar. Die Fähigkeit des Stifts, verschiedene physiologische Maße wie Herzfrequenz und Hautleitfähigkeit der Kinder beim Lernen zu erfassen, macht es möglich, zu erkennen, ob ein Kind gestresst ist. Ein hoher Stresspegel kann das Lernen beeinträchtigen, indem er z. B. die Konzentrationsfähigkeit und das Gedächtnis beeinflusst (Li et al., 2021: 8). Durch die Überwachung dieser physiologischen Maße können Lehrer:innen oder Betreuer:innen z. B. erkennen, wann ein Kind gestresst ist, und entsprechende Maßnahmen ergreifen, wie z. B. Pausen einzulegen oder Entspannungsübungen durchzuführen. Darüber hinaus kann bei emotionaler Erregung durch die Anpassung der Lernumgebung oder der Aufgaben der emotionale Zustand des Kindes verbessert und damit das Lernen erleichtert werden.

6. Leitlinien für eine verantwortungsvolle Technologieentwicklung und -anwendung

Begleitend zur Konzeption und technischen Entwicklung des FederLeicht-Lernsystems wurde untersucht, welche Faktoren berücksichtigt werden sollten, um eine verantwortungsvolle Entwicklung und Anwendung digitaler Lernsysteme nach ethischen Grundsätzen zu gewährleisten. Die gewonnenen Erkenntnisse

speisen sich aus den Ergebnissen zweier Arbeitsschritte: Im ersten Schritt wurde eine umfassende Literaturanalyse durchgeführt, um die unterschiedlichen Aspekte des Lernens mit digitalen Systemen in die Betrachtung einzubeziehen sowie den aktuellen Kenntnisstand der Debatte um Leitlinien zur Entwicklung digitaler Technologien zu erfassen. In einem zweiten Schritt wurden leitfadengestützte Interviews mit neun Expert:innen aus dem deutschen Bildungssystem sowie drei Fokusgruppen mit insgesamt 38 Schüler:innen im Alter zwischen 14 und 17 Jahren durchgeführt. Dieser zweite Schritt diente dazu, die Ergebnisse aus der Literaturanalyse zu reflektieren, um aktuelle empirische Befunde zu ergänzen sowie diese auf das speziell im Kontext des FederLeicht-Projektes zu entwickelnde Lernsystem hin zu überprüfen und anzupassen. Ziel war die Ableitung von Leitlinien für die Entwicklung des FederLeicht-Lernsystems, die als Orientierung für ein Design nach ethischen Kriterien dienen können. Diese werden hier nachfolgend kurz skizziert.

Prüfung möglicher Diskriminierungspotenziale (Gleichheit)

Die Prüfung von möglichen Risiken der Diskriminierung hat das Ziel, eine Benachteiligung von Nutzenden des Lernsystems aufgrund von Einkommen, Bildungshintergrund oder anderer individueller Gegebenheiten möglichst auszuschließen. Ein Ansatz zur Minimierung von Diskriminierungspotenzialen steckt bereits in der Möglichkeit der Anpassung der Aufgabenstellungen an das individuelle Lernniveau der Nutzenden. Ein weiterer Ansatz wird in einem niederschwelligen Zugang zum FederLeicht-Lernsystems nach dessen Einführung, also etwa unabhängig von der finanziellen Situation, bestehen. Weitere Diskriminierungspotenziale sind zu überprüfen und gegebenenfalls zu adressieren.

Einhaltung des Datenschutzes (Privatheit)

Die zentrale Rolle des Datenschutzes wurde in allen betrachteten Leitlinien zur verantwortungsvollen Entwicklung digitaler Technologien hervorgehoben. Auch im Kontext des FederLeicht-Lernsystems spielen Datenschutz und Aspekte der Privatheit eine besondere Rolle, nicht zuletzt wegen der Erhebung sensibler physischer Daten zur Verfasstheit der Lernenden. Entsprechende Bedenken und Forderungen zum verantwortungsvollen Umgang mit erhobenen Daten wurden sowohl im Rahmen der Interviews mit den Expert:innen als auch von Schüler:innen während der Fokusgruppen geäußert. So sollte etwa eine Weitergabe an Dritte

ausgeschlossen werden (insbesondere zu kommerziellen Zwecken), ebenso geklärt werden muss die Frage des Speicherortes.

Berücksichtigung menschlicher Handlungs- und Entscheidungsfreiheit (Autonomie)

Die Notwendigkeit zur Berücksichtigung menschlicher Handlungs- und Entscheidungsfreiheit betrifft unterschiedliche Anspruchsgruppen des Lernsystems. Zum einen die Nutzenden selbst, etwa wenn ihnen die Wahl ermöglicht wird, je nach Wunsch die Funktionen zur Überprüfung der Lernbefindlichkeit zu aktivieren oder zu deaktivieren. Zum anderen ist aber auch die Autonomie der Lehrenden betroffen, und zwar einerseits in ihrer Analyse und Entscheidung zur Bewertung der Leistungen der Schüler:innen, andererseits in ihrer Entscheidung zum speziellen Einsatzfeld des Stiftes. Je größer das mögliche Anwendungsfeld des Lernsystems, desto größer ist die Möglichkeit der Adaption des Systems für ein individuelles und offenes Unterrichtskonzept.

Stärkung von Vertrauen in die Technologie (Legitimität)

Um die Legitimität des Einsatzes der Technologie zu erhöhen, muss Vertrauen bei den unterschiedlichen Anspruchsgruppen aufgebaut werden. Dies gilt besonders für den Kreis der Elternschaft, aber auch für die Lehrenden sowie für die das Lernsystem nutzenden Schüler:innen selbst. Eine Stärkung des Vertrauens in das FederLeicht-Lernsystem ist vor allem auch durch ein möglichst großes Maß an Transparenz zu erreichen, sowohl über die genaue Funktionsweise der Technologie als auch über die Verwendung der erhobenen Daten. Zudem können kommunikative Maßnahmen, etwa durch begleitende Informationsmaterialien oder informatorische Veranstaltungen vor Einführung des Lernsystems dabei helfen, eventuell bestehende Sorgen bezüglich der Technologie abzubauen.

Wahrung der Offenheit und Erklärbarkeit (Transparenz)

Der Aspekt der Transparenz hängt wie oben beschrieben unmittelbar mit dem Aspekt der Legitimität und dem Schaffen von Vertrauen zusammen. Im Zusammenhang mit der Entwicklung und dem Einsatz des Lernsystems ergeben sich Fragen der Transparenz und Offenheit in mehreren Bereichen: Welche Daten werden während des Lernprozesses erhoben? Wo werden sie gespeichert und wer

hat Zugriff darauf? Auf welcher Forschungsgrundlage wurden die Technologie und der verwendete Algorithmus entwickelt? Welche Materialien werden bei der Herstellung verwendet? Welche Anspruchsgruppen haben auf welche Weise am Entwicklungsprozess partizipiert? Dies sind nur einige Beispiele für Fragen, die sich im Rahmen der Technologieentwicklung stellen und die im Zweifelsfall kritisch adressiert werden sollten. Das Ziel kann hier nicht eine kaum zu realisierende absolute Transparenz in allen Bereichen der Technologieentwicklung sein, sondern vielmehr eine Annäherung an eine größtmögliche Offenheit im Sinne des Vertrauensaufbaus in besonders kritischen Bereichen.

Prüfung ökologischer Aspekte (Nachhaltigkeit)

Die Notwendigkeit der Berücksichtigung von Aspekten der ökologischen Nachhaltigkeit wurde als ethische Kategorie im Rahmen der durchgeführten Fokusgruppen mit den Schüler:innen besonders betont. Viele der Teilnehmenden haben Bedenken geäußert, was die Verwendung von Rohstoffen und Materialien zur Herstellung der Technologie betrifft. Die ökologische Dimension betrifft dabei sowohl den Beschaffungs- und Herstellungsprozess als auch die Frage zukünftiger Entsorgung und eventueller Wiederverwertung von Materialien oder einzelner Komponenten.

Unterstützung der Technikgestaltung durch empirische Ergebnisse (kontinuierliche Fundierung)

Schließlich spielt die wissenschaftliche Fundierung durch empirische Ergebnisse eine zentrale Rolle bei der Entwicklung des FederLeicht-Lernsystems – und dies nicht nur, um eine möglichst fehlerfreie Funktionsweise des Systems und einen größtmöglichen Mehrwert für die Nutzenden zu erzielen, sondern gerade auch, um eben jene oben beschriebenen ethischen Aspekte mit den unterschiedlichen Anspruchsgruppen zu diskutieren und die entsprechenden Ergebnisse einfließen zu lassen. Nur so kann eine Überprüfung dieser Aspekte bezogen auf den konkreten Anwendungsfall des FederLeicht-Lernsystems erfolgen und gegebenenfalls eine Priorisierung im Falle auftretender Spannungsfelder vorgenommen werden.

7. Fazit

Mit dem FederLeicht-Lernsystem wurde ein neuartiges digitales Lernassistenzsystem entwickelt, welches das Lernen mit Stift und Papier und die damit verbundenen kognitiven Vorteile mit den Möglichkeiten der Ergänzung um Elemente der erweiterten Realität verbindet. Auf diese Weise kann die Lernunterstützung den individuellen Lernbedarfen und Lerngeschwindigkeiten der Schülerinnen und Schüler besser angepasst werden und nach ersten Erkenntnissen auch zu einer höheren Lernmotivation beitragen. Aktuell wird das FederLeicht-Lernsystem weiteren Evaluationsprozessen unterzogen, um die Benutzerfreundlichkeit und die Effekte auf die Motivation und die Leistung der Lernenden im Vergleich zu konventionellen Lernmethoden besser einschätzen zu können. Die Autor:innen des Beitrags gehen davon aus, dass unter Berücksichtigung der im Projekt formulierten Leitlinien für eine verantwortungsvolle Technologieentwicklung und -anwendung digitale Lernassistenzsysteme wie das FederLeicht-Lernsystem einen entscheidenden Beitrag zu einer Schule der Zukunft im Kontext einer digitalen Gesellschaft leisten kann.

Förderhinweis

Das Projekt FederLeicht wurde vom Bundesministerium für Bildung und Forschung (BMBF)[1] gefördert.

Literatur

Ahmad, Muneeb Imtiaz et al. (2020). A framework to estimate cognitive load using physiological data. In: *Personal and Ubiquitous Computing*, Vol. 09, Springer Nature, London, S. 1–15.

Domitrovich, Celene E. et al. (2008). Maximizing the Implementation Quality of Evidence-Based Preventive Interventions in Schools: A Conceptual Framework, Advances in school mental health promotion, Vol. 1(3), Pennsylvania, S. 6–28. https://pmc.ncbi.nlm.nih.gov/articles/PMC4865398/pdf/nihms327282.pdf [Zugriff 01.03.2025]

Donati, Maria Anna et al. (2021). Using the short form of the msbs to assess state boredom among adolescents: Psychometric evidence by applying item response theory. In: *Assessment*, Vol. 28, No. 3, S. 928–941. https://doi.org/10.1177/1073191119864655

Fixsen, Dean L. et al. (2009). Core Implementation Components. In: *Research on Social Work Practice*, Vol. 19(5), S. 531–540. https://doi.org/10.1177/1049731509335549

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Fliege, Herbert et al. (2009). PSQ. Perceived Stress Questionnaire [Verfahrensdokumentation, PSQ20-Skalenberechnung, PSQ20-Fragebogen Englisch, Deutsch, Deutsch (letzte 2 Jahre), PSQ30-Skalenberechnung, PSQ30-Fragebogen Englisch Französisch, Deutsch, Italienisch und Spanisch], Leibnitz-Institut für Psychologie (ZPID), Trier, Open Test Archive (ZPID), Trier. http://dx.doi.org/10.23668/psycharchives.5138

Hallfors, Denise D &, Godette, D. (2002). Will the ›Principles of Effectiveness‹ improve prevention practice? Early findings from a diffusion study. In: *Health Education Research: Theory & Practice*, Vol. 17, S. 461–470. https://academic.oup.com/her/article/17/4/461/590927?login=true [Zugriff 01.03.2025]

Heo, Seongsil, Kwon, Sunyoung & Lee, Jaekoo (2021). Stress detection with single ppg sensor by orchestrating multiple denoising and peak-detecting methods. In: *IEEE Access*, Vol. 9, S. 47 777–47 785. https://ieeexplore.ieee.org/document/9358140 [Zugriff 01.03.2025]

Hunter, Jennifer A. et al. (2015): Exploring the utility of the multidimensional state boredom scale. In: *European Journal of Psychological Assessment*, Vol. 32, S. 1–10.

Lehmann, Stephanie, Ruf, Esther & Misoch, Sabine (2021). Emotions and attitudes of older adults toward robots of different appearances and in different situations. In: *International Conference on Information and Communication Technologies for Ageing Well and E-Health*, Springer Verlag, S. 21–43. https://doi.org/10.1007/978-3-030-70807-8_2

Li, Yuntao et al. (2021): The relationship among trait mindfulness, attention, and working memory in junior school students under different stressful situations. In: *Frontiers in psychology*, Vol. 12, Frontiers Media S.A.. Lausanne, S. 558690. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2021.558690/full [Zugriff 01.03.2025]

Muthukrishnan, Suriya et al. (2017). Does heart rate variability predict human cognitive performance at higher memory loads? In*: Indian journal of physiology and pharmacology*, Vol. 61, New Delhi, S. 14–22.

Siddiqui, Nabeel & Chan, Rosa H. M. (2020). Multimodal hand gesture recognition using single IMU and acoustic measurements at wrist. In: *PLoS ONE*, Vol. 15, PLOS, San Francisco, S. 1–12. https://doi.org/10.1371/journal.pone.0227039

Sundararaman, Deekshita (2020). Doodle away: Exploring the effects of doodling on recall ability of high school students. In: *International Journal of Psychological Studies*, Vol. 12(2), John Wiley and Sons Ltd., Montreal, S. 31–44.

Suri, Karush, Gupta, Rinki (2018). Classification of hand gestures from wearable IMUs using deep neural network. In: *Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, India. S. 45–50.

Wittig, Nick et al. (2024). LeARn at Home: Comparing Augmented Reality and Video Conferencing Remote Tutoring. In: *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia*, Association for Computing Machinery, New York, S. 255–263 https://doi.org/10.1145/3701571.3701577

Zhou, Jianlong (2018). Multimodal Behavioral and Physiological Signals as Indicators of Cognitive Load. In: *The Handbook of Multimodal-Multisensor Interfaces: Signal Processing, Architectures, and Detection of Emotion and Cognition*, Vol. 2, Association for Computing Machinery and Morgan & Claypool, New York, S. 287–329. https://doi.org/10.1145/3107990.3108002

Novel Haptic Output for Spatial Input Devices and Controllers – HapticIO

Alexander Kalus, Juan F. Olaya-Figueroa, Ferdinand Streicher, Thomas Weiss, Matthias Uhl, Sebastian Krügel, Niels Henze and Katrin Wolf

Abstract

This article presents HapticIO, a research project aimed at advancing haptic feedback technologies for spatial input devices and controllers in Extended Reality. We present four prototypes: HapticCollider, providing ungrounded force feedback for rigid collisions in virtual tool use; MorphGrip, offering shape-changing haptic guidance for precise manual tasks; PumpVR, rendering object and avatar weight using liquid mass transfer; and MobileGravity, enabling weight simulation in mobile scenarios. Furthermore, we derive a predictive model for optimizing weight rendering through a study assessing the effects of absolute mass and weight distribution. Additionally, each system is evaluated through empirical user studies, revealing significant improvements in realism. We discuss integrating these technologies into unified devices while addressing associated challenges.

1. Introduction

The current input for spatial interaction, such as in Augmented, Mixed, Extended and Virtual Reality (AR, MR, XR, VR), is mainly limited to camera-based systems and controllers for gesture-, button- and touch-based input. More complex manual actions, such as those required to inspect elements (e. g., touching a pump and detecting vibration and heat) or modeling 3D objects, are not yet possible due to the limited haptic input options. Haptic feedback in a more natural form is missing to make physical interactions with objects intuitive even over a distance.

Previous solutions can be improved in terms of the user experience as they only enable very simplified and therefore imprecise interaction.

As vibration cannot realistically simulate the entire range of haptic sensations we can perceive, developing controllers that provide haptic feedback beyond vibrotactile stimuli is therefore the vision of the project HapticIO (10/2021–08/2024). HapticIO is a joint project with the following partners: Berlin University of Applied Sciences and Technology (BHT), University of Regensburg, Konstruktiv GmbH, Intuity Media Lab GmbH, and Technical University Ingolstadt of Applied Sciences (THI).

The objective is to create haptic controllers that can be perceived as an analog user interface, thereby enabling users to utilize digital tools in a manner consistent with their traditional predecessors. Consequently, users who know how to use a drill should also be able to drill with a haptic controller through intuition and knowledge transfer. They should be able to do this just as well as with the traditional tool because the virtual tool feels and behaves exactly like its reference tool. The acquisition of skills is therefore transferred from the knowledge of how we handle traditional tools to their virtual and intelligent equivalents. The integration of sensors, actuators, and networking technologies also makes it possible to use virtual and intelligent tools in joint projects and to, for example, enable the transfer of knowledge between experts and learners.

Through such technological innovation, we can better develop technology to control digital reality, intelligent objects, and intelligent spaces around us. Haptic feedback devices were therefore developed in this project, for example, to enable

- the use of tools for collaborative work or training in virtual workshops,
- the use of multimodal AR, XR, and VR control elements for intelligent environments and
- interaction with intelligent tools in a mixed reality environment.

To explore how such systems should be developed, we designed and implemented artifacts and demonstrators that were evaluated in user studies, and which are presented in the following paragraphs. These include HapticCollider, a VR controller providing ungrounded force feedback for simulating collision forces; MorphGrip, a shape-changing device delivering haptic guidance for manual tasks; PumpVR, which adjusts the weight of VR controllers to match the properties of virtual objects through liquid mass transfer; and MobileGravity, a mobile proxy that simulates weight changes while ensuring user mobility. Finally, a study on weight perception explores how absolute mass and weight distribution impact realism, leading to predictive models for optimizing weight rendering. We conclude with a discussion on the potential for combining these technologies and the challenges involved.

2. HapticCollider

Researchers have been exploring VR controllers that offer more advanced haptic feedback beyond basic vibration, aiming to enhance realism in virtual tool use. However, providing realistic force and weight feedback, especially for heavy tools, remains a challenge. HapticCollider addresses this gap by simulating collision force feedback when using weighted tools, such as a hammer, without requiring external apparatus. HapticCollider aims to improve immersion in VR applications like hammering, baseball, or fishing, while offering a more realistic experience than standard controllers. This section is based on our corresponding publication at the Mensch und Computer conference (Olaya-Figueroa et al., 2023).

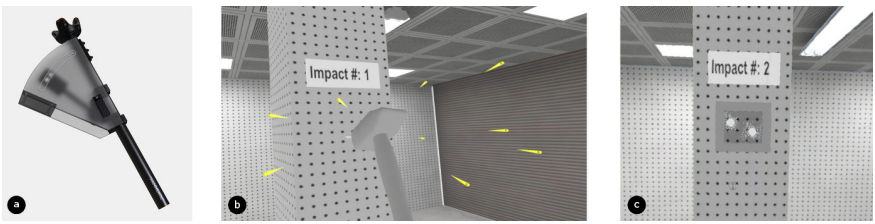


Figure 1: a) 3D model of HapticCollider created in the design phase and before assembly, b) HapticCollider in the virtual environment while hitting a nail and c) after breaking a glass
Source: Own representation based on Olaya-Figueroa et al., 2023: 116

2.1 Enhancing Force Feedback for Heavy Tools in VR

In VR, immersion refers to the degree to which the user feels present within a virtual environment (Slater et al., 1999: 1–13). Consequently, haptic feedback is essential for allowing users to perceive physical responses from the virtual world they experience. To achieve realistic haptic feedback, the concept of haptic fidelity outlines three key components that are relevant to our objective of designing a novel VR controller with force feedback: Sensing, Hardware, and Software (Muender et al., 2022: 1–17). Sensing refers to the ability to generate natural haptic sensations, while Hardware encompasses the elements responsible for generating these stimuli and the necessity of minimizing any potential delays caused by mechanical parts. Software refers to algorithms that regulate haptic sensations. For example, prediction algorithms can anticipate events and trigger haptic feedback ahead of time, reducing latency and ensuring that the haptic feedback remains synchro-

nized with the visual cues. HapticCollider integrates these three components to create a realistic experience.

Despite the advancements in VR haptics, there remains a gap in delivering force feedback for heavy tools, such as a hammer or baseball bat, when striking rigid objects in VR. Addressing this gap, *HapticCollider* integrates these three components—Sensing, Hardware, and Software—to provide a realistic and immersive experience.

2.2 Prototype and Application

HapticCollider, a VR controller designed for realistic haptic experiences, utilizes a kinetic mechanism to simulate the impact of hammering a nail, integrating an Arduino interface and a powerful servomotor. (a) To protect participants' fingers a coverwas printed in 3D. (b) Key components of HapticCollider include a weighted mass that strikes a wooden handle when a collision is detected in the virtual environment, see Figure 2.

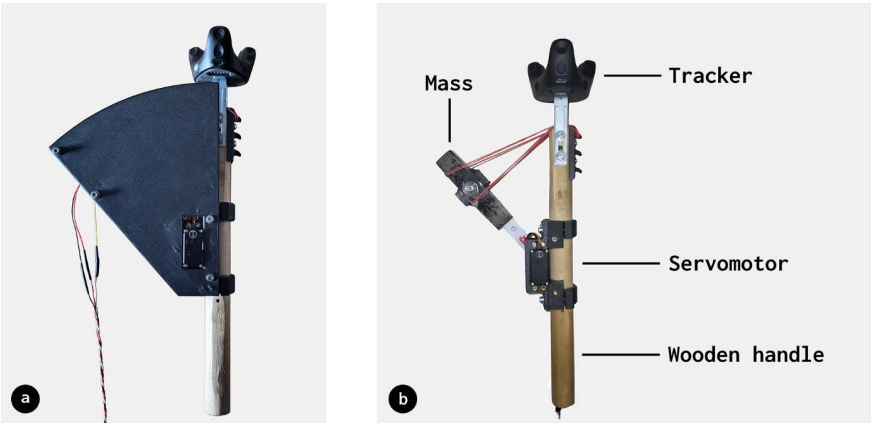


Figure 2: HapticCollider featuring (a) the cover to protect users and (b) its components
Source: Own representation based on Olaya-Figueroa et al., 2023: 119

Due to the mechanical elements of HapticCollider a slight latency is introduced to the force feedback, which may affect the realism of the experience. To address this, a collision prediction algorithm triggers force feedback just before a virtual collision, aligning haptic and visual cues in three steps (see Figure 3): (a) using an invisible ray to anticipate a collision by detecting if the direction of the hammer moves towards the target nail and that its speed is greater than a certain

threshold; (b) once detected, the system initiates the collision effect, (c) stopping the hammer at the collision point in the virtual environment while the physical controller continues briefly. This control-display offset helps simulate the illusion of impact. HapticCollider is highly adaptable for various VR applications. It mimics the weight, grip, and impact of tools like baseball bats, drumsticks, and fishing rods, addressing challenges in vibrotactile feedback and enabling realistic interactions. It also provides collision feedback for materials like wood and concrete in hammering tasks, thus enhancing the realism of activities such as drumming, fishing, and construction in VR.

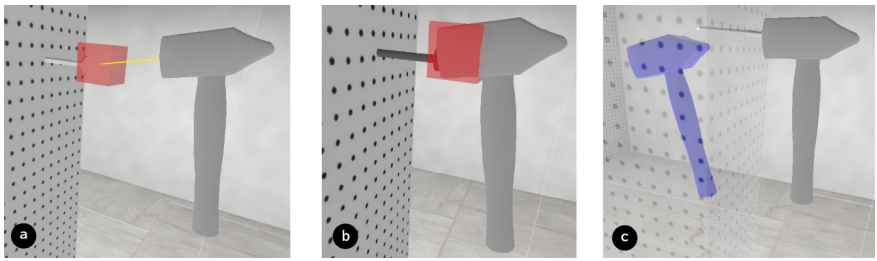


Figure 3: Steps of the collision prediction algorithm and the control-display ratio working while HapticCollider is hitting a nail.
Source: Own representation based on Olaya-Figueroa et al., 2023: 119

2.3 Results

In our study, 24 participants completed two tasks, once using HapticCollider and once using a standard VR controller. While HapticCollider provided force feedback, the standard VR controller delivered vibrotactile feedback. Participants stood in front of a VR column and used a virtual hammer to hit either a nail into a wall (Task 1) or smash glass (Task 2), with targets positioned at eye level. Upon each condition, participants completed questionnaires to obtain qualitative feedback and evaluate their perceived realism.

The user study demonstrated that HapticCollider significantly enhances the realism of virtual hammering tasks compared to a standard VR controller. Key factors included force feedback, controller weight, synchronized multimodal cues (visual, haptic) and the form factor since HapticCollider uses a hammer handle (Olaya-Figueroa et al., 2023: 120–125).

3. MorphGrip

Handheld devices, VR controllers, and tools often function as passive instruments, lacking responsiveness to user input and environmental events, which can limit their effectiveness in complex situations. While previous research has explored visual, auditory, and haptic feedback to guide users during manual motor tasks, shape-changing haptic feedback remains underexplored for tasks requiring precise translational and rotational movements. MorphGrip addresses this gap by providing tactile cues to the palm and fingers through dynamic grip adjustments. This design enables MorphGrip to guide users in performing manual motor tasks, such as translation and rotation, especially in industrial and therapeutic settings where auditory and visual cues may be overloaded or unavailable. This section is based on our corresponding publication at the International Conference on Mobile and Ubiquitous Multimedia (Olaya-Figueroa et al., 2024). MorphGrip is depicted in Figure 4. Figure 5 shows its inner mechanics.

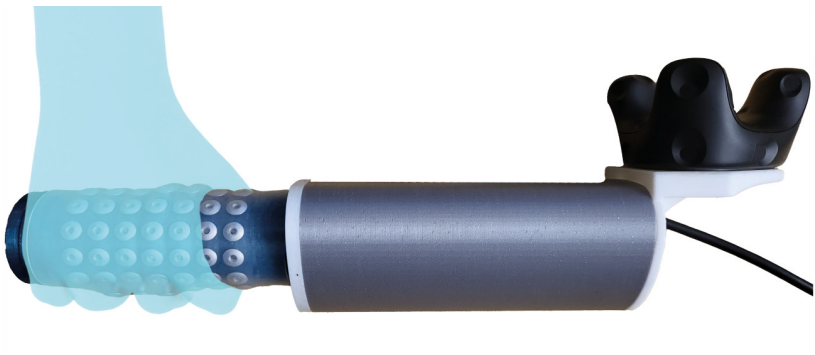


Figure 4: A user is grasping MorphGrip.
Source: Own representation based on Olaya-Figueroa et al., 2024: 196

3.1 Haptic Guidance

Research in guidance systems has evolved across visual, auditory, and haptic modalities, each with distinct advantages and limitations. Haptic guidance, offering directional cues, provides a viable alternative, especially with devices like wearable bands (Tsai et al., 2021: 1–13) and vibrotactile (Weber, 2011: 349–354). However, existing research on haptic guidance for manual motor tasks has predominantly relied on external devices to provide directional cues. To date, no

study has explored embedding guidance mechanisms directly into the grip of a device to assist users in such tasks. MorphGrip addresses this research gap by introducing a smart device with an integrated shape-changing grip that leverages the high sensitivity of the palm to deliver precise directional haptic cues. MorphGrip is specifically designed to provide self-guidance for manual motor tasks in eyes-free interactions. It is intended for scenarios where the visual channel is unavailable, impractical, or occupied by the task at hand.

3.2 Prototype and Application

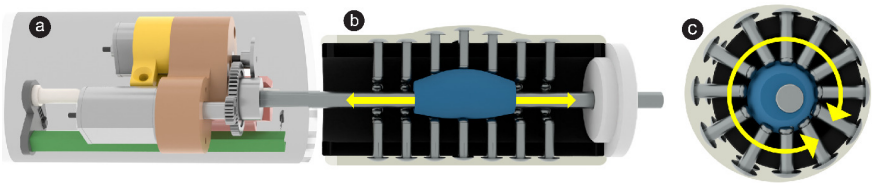


Figure 5: (a) The MorphGrip mechanism includes two motors: (b) one motor for generating the linear haptic cue and (c) the other for generating the circular haptic cue MorphGrip.
Source: Own representation based on Olaya-Figueroa et al., 2024: 200

MorphGrip is a shape-changing grip device designed to provide haptic guidance for manual motor tasks through two distinct haptic cues. It generates two types of feedback: a linear haptic cue (moving up/down) and a circular haptic cue (rotating clockwise/counterclockwise). The mechanism employs dual DC motors and an eccentric cam-follower system creating a movable wave crest pattern on the grip's surface by means of extendible pins. The design addresses three main challenges: implementing distinct haptic cues, embedding the mechanism within a cylindrical grip facilitating the integration into handheld devices, and maintaining mechanical simplicity for practical use in handheld devices and controllers. MorphGrip is engineered for diverse applications across multiple domains, including rehabilitation therapy, where it assists users in repetitive motor tasks to support physical recovery and training; smart tools in industrial settings, where it provides guidance for tools used in demanding environments; and specialized scenarios, such as underwater applications, where traditional visual and auditory cues are ineffective or unavailable. These applications highlight the versatility of MorphGrip as an innovative solution for scenarios requiring alternative guidance methods.

3.3 Results

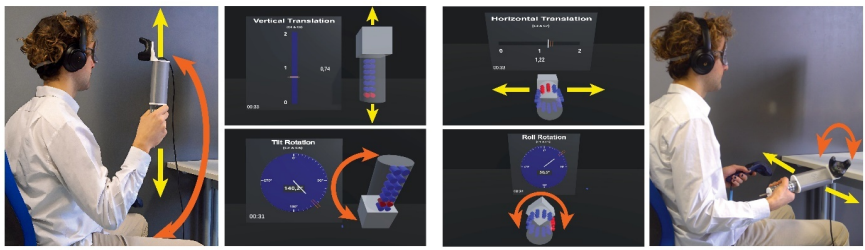


Figure 6: In the user study, participants performed 4 tasks being guided by MorphGrip: (Vertical, Horizontal) translations, and (Tilt, Roll) rotations MorphGrip.
Source: Own representation based on Olaya-Figueroa et al., 2024: 200

In our study, 24 participants performed four tasks, namely Vertical, Tilt, Roll, and Horizontal, guided by MorphGrip's linear and circular haptic cues. Each task involved reaching randomly assigned target points within specific ranges of height, tilt angle, rotation angle, or lateral position. The study compared the effectiveness of MorphGrip's haptic cues against verbal cues for guidance, evaluating performance based on success rate, task completion time, travel distance and overshooting. After completing each condition, participants filled out questionnaires to gather qualitative feedback and evaluate the task load based on their experiences. The procedure is illustrated in Figure 6.

The findings revealed that, although verbal cues outperformed haptic cues, participants found haptic cues helpful in completing tasks, with notably improved performance in the Horizontal and Vertical tasks. In contrast, the Roll and Tilt tasks demonstrated a higher mental demand, suggesting that these tasks may benefit from further refinement in haptic feedback design. Overall, MorphGrip shows strong potential for integration into handheld device grips, particularly in applications like rehabilitation therapy, industrial smart tools, and specialized environments such as underwater guidance, where visual or auditory cues may be limited or unavailable (Olaya-Figueroa et al., 2024: 196–209).

4. PumpVR

Perceiving the weight of virtual objects and avatars in virtual reality is important in order to understand their properties and to interact with them naturally. However, commercial VR controllers have a static, low weight, giving each virtual ob-

ject and avatar the same sensation of weight. PumpVR uses a bidirectional pump and solenoid valves to change the weight of the VR controllers so that the perceived weight corresponds to the virtual properties (see Figure 7). This allows users to feel the weight of virtual objects or avatars, enhancing the realism and enjoyment of VR experiences. This section is based on our corresponding publication at the CHI Conference on Human Factors in Computing Systems (Kalus et al., 2023b).



Figure 7: PumpVR at three different weight levels and exemplary virtual objects that these can represent
Source: Own representation based on Kalus et al., 2023b: 4

4.1 Weight Simulation in VR

When interacting with objects in the real world, humans naturally perceive their weight, which influences their understanding of the object's properties and affects how they handle or manipulate it. This is largely absent in VR, where users sense the weight of their input device, which is typically different from the weight of the object seen in the virtual environment. Providing realistic weight sensations is therefore a key challenge for the progress of VR. In addition to rendering the weight of virtual objects, simulating the body weight of avatars in VR has been

largely unexplored, yet it is crucial for enhancing the feeling of embodiment and thus the sense of being present in the virtual world.

Although previous research has attempted to address this gap by developing controllers capable of simulating weight, these systems suffer from several limitations, including single-handedness, a slow response time, or only being able to simulate small weight changes, as highlighted by Lim et al. (2021: 163270–163273) in their review on weight simulation in VR. These constraints hinder the effectiveness and practicality of such systems in a wide range of VR applications

4.2 Prototype and Application

The system renders up to 500 g per hand, reaching this weight by transferring water at a rate of 150 g/s into a reservoir located within each controller.

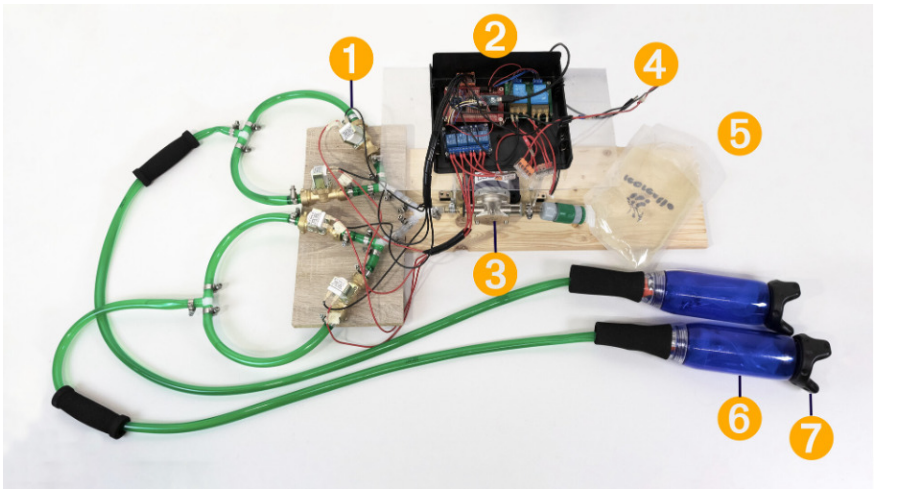


Figure 8: Components of PumpVR
Source: Own representation based on Kalus et al., 2023b: 5

Figure 8 shows PumpVR's components. The device employs a bi-directional pump (1) to allow reversing the water flow. Four solenoid valves (2) direct the flow of water, which is transferred between three water bags. Sealed foldable drinking bags were chosen, as they contract under negative pressure, avoiding the perception of liquid inertia. Two of them are placed inside the controllers' plastic housings (3). HTC Vive Trackers are screwed on top of these housings to enable positional and rotational tracking in the VR environment (4). A control unit (5),

consisting of an ARDUINO MICRO, a bluetooth module and six relay modules, communicates with the game engine via Bluetooth and controls the flow of water depending on the in-app event.

We implemented two use cases to demonstrate PumpVR's capability of rendering the weight of virtual objects and avatars. The first application involves the user interacting with game items of different weights, such as a hammer, a torch, a shield, or bare hands. The game requires the users to switch between these items and use them to combat approaching enemies and block arrows. The purpose of the second application is to explore the use of PumpVR to render body weight. In the application, the volume of the user's virtual body and the weight of both controllers are increased simultaneously until the controllers reach their maximum weight.

4.3 Results

The system can produce distinct weight levels within the controllers. A technical evaluation showed that PumpVR could target five different weight levels efficiently by controlling the water-pumping duration, achieving consistent weight changes with an average deviation of less than 5 %.

We also conducted a user study in within-subject design, in which 24 participants completed both test applications using either PumpVR or a regular VR controller. We assessed realism, enjoyment and virtual embodiment using established instruments such as the Reality Judgment and Presence Questionnaire, the Intrinsic Motivation Inventory and the Virtual Embodiment Questionnaire. Utilizing PumpVR to render object weight yielded significantly higher realism and enjoyment scores compared to the use of regular VR controllers. Additionally, participants perceived their virtual body significantly more strongly as their own when PumpVR was used to simulate body weight. The detailed analysis can be found in the accompanying publication (Kalus et al., 2023b: 1–13).

5. MobileGravity

In this section, we present MobileGravity, a VR proxy that adapts its weight to the virtual object being held in VR. The system separates bulky and heavy components from the weight-adjusted object and thus addresses the conflict between mobility and performance. Furthermore, by enabling changes of up to 1 kg in 235 g/s, it outperforms previous approaches in terms of weight range and responsiveness. This section is based on our corresponding publication at the CHI Conference on

Human Factors in Computing Systems (Kalus et al., 2024b). Figure 9 illustrates the interaction with MobileGravity.



Figure 9: Interacting with MobileGravity
Source: Own representation based on Kalus et al., 2024b: 1; Kalus et al., 2024c: 1

5.1 Balancing Performance and Mobility of Weight Interfaces

To facilitate weight change, a liquid-based weight interface, such as PumpVR, requires a connection to an external reservoir, which restricts the user's mobility (Kalus et al., 2023a: 1–3). However, a broad range of VR experiences requires unhindered movement, including VR fitness training, sports simulations, or motion-based learning applications. Furthermore, the ability to walk around freely in VR increases the effectiveness of locomotion, enhances cognitive map building and reduces motion sickness, whereas grounded cable connections were found to disturb the experience (Langbehn et al., 2018: 1–9). Lim et al. (2021: 163253–163283) reviewed 65 research prototypes and highlighted the trade-off between mobility and performance as a key challenge for advancing weight interfaces. Hence, there is a clear need for mobile weight simulation that does not compromise on other desired objectives.

5.2 Prototype and Application

To address the aforementioned trade-off, we developed MobileGravity, which uses water to change the weight of a handheld VR proxy object, but decouples it from the water supply and heavy components once the weight is adjusted. This allows the user to walk and turn in any direction. MobileGravity consists of a mobile proxy object equipped with a water reservoir, an HTC Vive Tracker and two handles, and a base station with a pump and water tank. The two units connect through dripless quick disconnect couplings that automatically cut off

the water flow when disconnected. A limit switch is used to continuously check if a connection is established. A recess in the top of the base station allows the user to insert the proxy object. To preserve immersion in VR, this insertion is facilitated via the user placing their handheld virtual object in a designated location in the virtual environment. This starts the weight adjustment. To separate the two units again after the weight adjustment, they are automatically unlocked by two solenoid actuators. The inner workings of MobileGravity are shown in Figure 10.

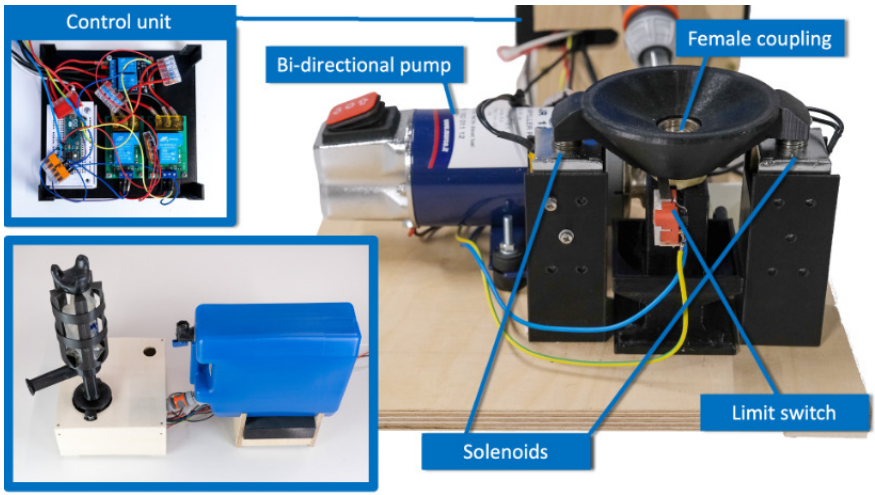


Figure 10: Inner workings of MobileGravity's base station and its connection to the handheld proxy (bottom left)
Source: Own representation based on Kalus et al., 2024b: 5

We created a virtual environment in Unity to test and demonstrate the concept. We chose a virtual workshop setting (see Figure 11), to feature objects with a broad range of weight in a coherent scenario. In the scene users can build a bird box using six virtual tools of different weight, i. e. a paint roller, a stapler, a hammer, a hacksaw, a blowtorch and a cordless drill.

Apart from the demonstrated application, we provide further use cases for MobileGravity (see Figure 12). Its adaptive weight feedback can be used to simulate different intensities of VR workouts. In our exemplary application, the user handles two MobileGravity proxy objects to hit objects of different colours that appear from the background, similar to popular fitness applications such as Beat Saber and FitXR. A further application consists of a co-operative, co-located game, in which users run through the virtual world and overcome enemies and obstacles. To do so, they are required to use specific virtual objects that require



Figure 11: Overview (left) and user's view (right) of the virtual workshop
Source: Own representation based on Kalus et al., 2024b: 6

different handling, which is perceptible through the weight feedback. Finally, MobileGravity can be used to perceive different avatar weights in room scale setups: In our third use case, the users wear two MobileGravity controllers on their bodies. The application adjusts the weight of the worn controllers to the expected body weight of the embodied avatar and decouples the user from the base station.

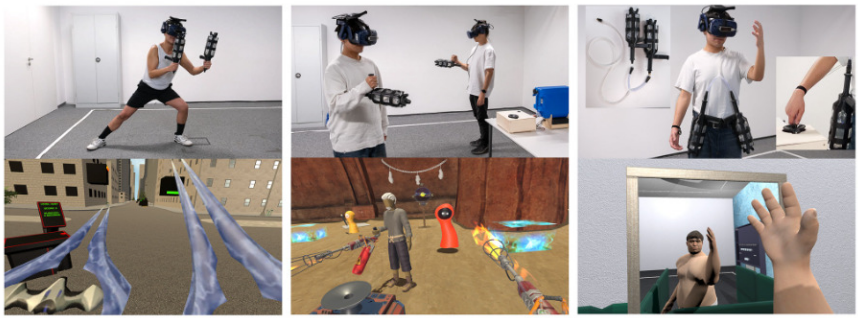


Figure 12: Further use cases for MobileGravity. VR workout (left), cooperative game (center), and avatar weight simulation (right)
Source: Own representation based on Kalus et al., 2024b: 10

5.3 Results

In a within-subject study, 30 participants were tasked with constructing a virtual bird box in the workshop environment using tools simulated with either MobileGravity or a standard VR controller. Measures included quantitative ratings of general and haptic realism, enjoyment, and qualitative feedback for deeper insights. Statistical analysis found that MobileGravity significantly enhanced gen-

eral and haptic realism, as well as enjoyment. Furthermore, participants reported that using MobileGravity made the virtual tools feel more like their real-life counterparts and that they adapted their movements to the virtual tools due to the haptic feedback. The complete analysis and implementation can be found in the accompanying publication (Kalus et al., 2024b: 1–13).

6. Simulating Object Weight in VR: The Role of Absolute Mass and Weight Distributions

To understand the optimal weight to be provided when using weight interfaces such as PumpVR or MobileGravity, we conducted a study in which 30 participants assigned the most realistic weight to differently configured virtual objects. There were 54 object configurations, varying in level of visual mass, visual weight distribution and center of physical mass. We found that participants underestimated the weight of heavy objects and overweighted light objects. While the visual weight distribution did not affect the weight estimation, the physical center of mass did. We derive guidance for the design and utilization of weight interfaces and provide predictive models to determine the weight to render based on visual mass and visual weight distribution. This section is based on our corresponding publication at the ACM Symposium on Virtual Reality Software and Technology (Kalus et al., 2024a).

6.1 Weight Perception

An object's perceived heaviness depends not only on its absolute mass but also on its weight distribution (Amazeen & Turvey, 1996: 213–232). Additionally, our brain relies on visual cues and expectations to estimate an object's weight. For instance, in the size-weight illusion when humans lift two equally weighted but different-sized objects, they perceive the bigger object as lighter (Saccone et al., 2019: 1195–1212). In the golf ball illusion, golfers who are familiar with the weight of golf balls perceive their weight differently compared to non-golfers, who lack this knowledge (Ellis & Lederman, 1998: 193–201). Hence, prompting a weight interface, such as PumpVR or MobileGravity, to replicate a virtual object's absolute weight does not necessarily provide the most realistic haptic representation. Consequently, there is a clear gap in understanding how to determine the weight to render for a given virtual object.

6.2 User Study

We conducted a study to understand how to best use such weight interfaces to simulate realistic object weight. In the study, 30 participants (14 female, 16 male, aged 18–34) were asked to choose the best-fitting weight for virtual objects in 54 different experimental trials. These trials varied in three factors (see Figure 13): the visual mass of the object (three levels), the distribution of the object's weight (six levels), and the position of the physical mass on the object's grip (three levels). For each of the objects, participants saw 13 instances in VR that looked identical but were co-located with different physical weights in the real world, ranging from 100 g to 1300 g. They were instructed to pick up, wield and heft the different objects and were permitted to try any object until they felt that the haptic sensation most closely aligned with the virtual object.

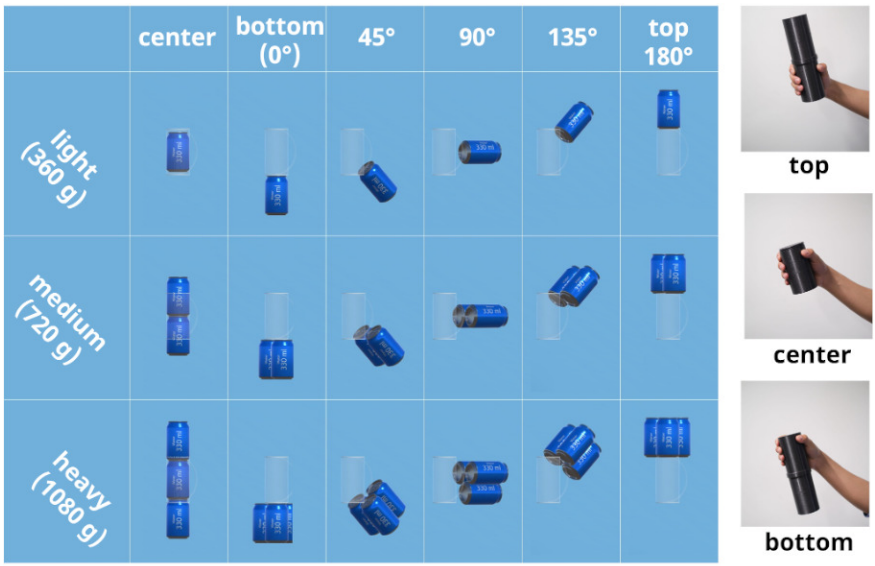


Figure 13: Experimental conditions
Source: Own representation based on Kalus et al., 2024a: 5

6.3 Results

Our Bayesian analysis revealed that participants underweighted heavier objects while overweighting lighter objects. This suggests that, for heavy objects, the ab-

solute weight rendered by VR systems can be reduced without sacrificing perceived realism. This finding is beneficial in reducing several issues that come with simulating heavy weights, such as motor noise, latency, high power consumption and bulkiness. Additionally, the study found that the position of the physical mass significantly impacted weight perception. Objects with the physical mass placed at the top of the grip were perceived as lighter than those with the mass placed at the bottom. Interestingly, the visual center of mass did not influence weight perception. Hence, we fitted a linear model to predict the optimal weight to render for virtual objects based on the virtual object's visual mass and the position on which the weight interface can provide the physical mass position relative to the handle. The regression equation was: $w = 163.7\,g + 0.68 \times v + 1.75 \times p$, where w is the weight to render, v is the visual mass in grams and p is the physical mass's distance to the handle in cm. Developers can use this model to estimate the weight to be rendered by the weight interface, balancing realism and performance. Details on the methodology and results can also be found in the accompanying publication (Kalus et al., 2024a: 1–11).

7. Discussion and Conclusion

In sum, this research explored novel approaches to enhancing haptic feedback in VR, XR and AR environments. The development and evaluation of systems like HapticCollider, MorphGrip, PumpVR, and MobileGravity offer promising advancements in simulating realistic physical sensations by addressing challenges related to weight perception, force feedback, and shape-shifting.

During a workshop, we engaged in an open discussion about combining multiple prototypes into a single device. The consensus from this discussion was that, while it is technically possible to merge some of the presented systems, it often does not make much sense from the user's perspective. Table 1 presents potential use cases for each combination of the prototype technologies.

Several key factors contributed to our conclusion. First, the spatial constraints and the increase in weight associated with adding more technology to a single device would likely diminish the usability of the prototype. Users could struggle with holding or using devices that become heavier and bulkier as more functionalities are integrated. Second, there was concern about haptic sensations overlaying and distracting from one another, which could lead to overstimulation. When multiple haptic sensations and mechanisms are combined in a single prototype, they could potentially interfere with one another, resulting in a confusing or overwhelming user experience. Finally, while each system has valid use cases on its own, finding a single use case where multiple sensations are required often ap-

	HapticCollider	**PumpVR**	**MobileGravity**
MorphGrip	A VR scenario; controlling a crane with feedback on how hard something hit; golfing/baseball with rotation guidance	Tool training with weight and rotational information	Tool training with weight and rotational information
HapticCollider		VR fishing, impact when pulling up hard	Hammering, sword fighting

Table 1: Potential use cases for combined prototypes
Source: Own representation

peared to be a »solution seeks problem« scenario. In other words, while the individual technologies serve their individual purposes well, finding a direct user need where combining them is critical is rare. However, despite these challenges, creating a unified prototype that combines multiple haptic properties holds enormous potential for immersion in VR. Achieving this would require overcoming the current limitations, such as spatial constraints and overweight. As such, more research is necessary to explore how various haptic properties can be effectively integrated into a single device, without sacrificing usability, comfort, or user experience.

Funding information

The Project HapticIO is funded by the German Federal Ministry of Education and Research (BMBF)[1] within the funding line »Interactive systems in virtual and real spaces – Innovative technologies for the digital society«.

Literature

Amazeen, E. L., & Turvey, M. T. (1996). Weight perception and the haptic size-weight illusion are functions of the inertia tensor. In: *Journal of Experimental Psychology: Human Perception and Performance*, vol. 22(1), American Psychological Association, Washington D.C., p. 213–232. https://doi.org/10.1037//0096-1523.22.1.213

1 At the time of printing: Federal Ministry of Research, Technology and Space (BMFTR)

Ellis, R. R., & Lederman, S. J. (1998). The golf-ball illusion: Evidence for top-down processing in weight perception. In: *Perception*, vol. 27(2), SAGE Publications, Thousand Oaks, p. 193–201. https://doi.org/10.1068/p270193

Kalus, Alexander, Klein, Johannes & Henze, Niels (2023a). Utilizing Liquid Transfer for Weight Simulation: Challenges and Future Directions. In: *Mensch Und Computer 2023 – Workshopband*, Gesellschaft für Informatik e.V., Rapperswil (SG), p. 1–3. https://doi.org/10.18420/muc2023-mci-ws09-410

Kalus, Alexander et al. (2024a). Simulating Object Weight in Virtual Reality: The Role of Absolute Mass and Weight Distributions. In: *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*, Association for Computing Machinery, New York, p. 1–11. https://doi.org/10.1145/3641825.3687732

Kalus, Alexander et al. (2024b). MobileGravity: Mobile Simulation of a High Range of Weight in Virtual Reality. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, p. 1–13. https://doi.org/10.1145/3613904.3642658

Kalus, Alexander et al. (2024c). Showcasing MobileGravity. In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, p. 1–2. https://doi.org/10.1145/3613905.3649128

Kalus, Alexander (2023b). PumpVR: Rendering the Weight of Objects and Avatars through Liquid Mass Transfer in Virtual Reality. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, p. 1–13. https://doi.org/10.1145/3544548.3581172

Langbehn, Eike, Lubos, Paul & Steinicke, Frank (2018). Evaluation of Locomotion Techniques for Room-Scale VR: Joystick, Teleportation, and Redirected Walking. In: *Proceedings of the Virtual Reality International Conference – Laval Virtual*, Association for Computing Machinery, New York, p. 1–9. https://doi.org/10.1145/3234253.3234291

Lim, Woan Ning et al. (2021). A Systematic Review of Weight Perception in Virtual Reality: Techniques, Challenges, and Road Ahead. In: *IEEE Access*, Institute of Electrical and Electronics Engineers, New York, p. 163253–163283. https://doi.org/10.1109/ACCESS.2021.3131525

Muender, Thomas et al. (2022). Haptic fidelity framework: Defining the factors of realistic haptic feedback for virtual reality. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, p. 1–17. https://doi.org/10.1145/3491102.3501953

Olaya-Figueroa, Juan F. et al. (2024). MorphGrip: Haptic Guidance Using a Shape-Changing Grip. In: *Proceedings of the International Conference on Mobile and Ubiquitous Multimedia (MUM '24)*, Association for Computing Machinery, New York, p. 196–209. https://doi.org/10.1145/3701571.3701573

Olaya-Figueroa, Juan F. et al. (2023). HapticCollider: Ungrounded Force Feedback for Rigid Collisions during Virtual Tool Use. In: *Proceedings of Mensch und Computer 2023*, Association for Computing Machinery, New York, p. 116–126. https://doi.org/10.1145/3603555.3603568

Saccone, Elisabeth J., Landry, Oriane & Chouinard, Philippe A. (2019). A meta-analysis of the size-weight and material-weight illusions. In: *Psychonomic Bulletin & Review*, vol. 26(4), Springer, New York, p. 1195–1212. https://doi.org/10.3758/s13423-019-01604-x

Slater, Mel (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. In: *Presence: teleoperators and virtual environments,* vol. 8(5). MIT Press, Cambridge, p. 560–565. https://doi.org/10.1162/10547469.9566477

Tsai, Hsin-Ruey et al. (2021). GuideBand: Intuitive 3D Multilevel Force Guidance on a Wristband in Virtual Reality. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems,* Association for Computing Machinery, New York, p. 1–13. https://doi.org/10.1145/3411764.3445262

Weber, Bernhard et al. (2011). Evaluation of a vibrotactile feedback device for spatial guidance. In: *2011 IEEE World Haptics Conference*, Institute of Electrical and Electronics Engineers, Piscataway, p. 349–354. https://doi.org/10.1109/WHC.2011.5945511

Development of a Mixed Reality Interactive Training System to Support Professional Socio-Emotional Interactions – MITHOS

Lara Chehayeb, Chirag Bhuvaneshwara, Manuel Anglet, Bernhard Hilpert, Ann-Kristin Meyer, Dimitra Tsovaltzi, Patrick Gebhard, Antje Biermann, Sinah Auchtor, Nils Lauinger, Julia Knopf, Andreas Kaiser, Fabian Kersting, Gregor Mehlmann, Florian Lingenfelser, Elisabeth André, Anna Welker, Kai Karren, Ruben Segundo Garcia Ucharima, Dorothea Irmisch, Timo Sowa and Massimo Romanelli

Abstract

Professionals facing challenging workplace conflicts often experience shame and self-blame, which relate to the feeling of incompetence but may externalise as anger. Sending mixed signals fails the contingency rule for affect co-regulation processes, result in confusion for the others' emotions, hinder emotion regulation, and lead to conflict escalation. Therefore, being able to constructively regulate one's emotions not only benefits individual experience of emotions but also influences conflict regulation. In classroom conflicts, teachers are mainly responsible for the co-regulation processes with students. MITHOS is a system aimed at training teachers' conflict resolution skills through realistic situative learning opportunities during classroom conflicts. In four stages, MITHOS supports teachers' socio-emotional self-awareness, perspective-taking and positive regard. It provides a) a safe virtual environment to train free social interaction and receive natural social feedback from reciprocal student-agent reactions, b) spatial situational perspective-taking through an avatar, c) individual virtual reflection guidance on emotional experiences through co-regulation processes, and d) expert feedback on professional behavioural strategies. This chapter presents the general concept of MITHOS with the four stages, as well as their implementation in a semi-automatic Wizard-of-Oz (WoZ) system. The WoZ system affords collecting data that are used for developing the fully automated hybrid (machine learning and model-based) system, and to validate the underlying psychological and

conflict resolution models. We present results validating the approach in terms of scenario realism, as well as of a systematic testing of the effects of external avatar similarity on antecedents of self-awareness with behaviour similarity. The chapter contributes to a common methodology of conducting interdisciplinary research for human-centered, socially interactive, and generalizable extended reality (XR) and presents a system designed and implemented to support it.

1. Introduction

In social interactions, people learn from each other through situative experiences. How others react to our behaviour can provide rich situative social feedback about what needs to be adjusted. Emotions comprise important indicators on how to assess a situation and which options for action are available to us (Gross et al., 2014). Teachers, similar to parents, hold a central role as multipliers of socio-emotional competence and behaviour thereof, which takes place largely also through unconscious socialisation processes (Dale & James, 2015; Eckstein & Rüth, 2015). However, people are not prepared for dealing with conflicts and fail to regulate their emotions at the workplace, where there is a lot at stake for them. Our goal is to strengthen professional communication and empathic competence. To this end, effective methods for emotion regulation and conflict resolution are required (Gross et al., 2014; G. Keller, 2001), which cannot be trained systematically and realistically in real human-human conflict situations. We focused on simulating classroom conflicts. Teachers often feel inadequately prepared for dealing with complex socio-emotional challenges, such as those triggered by heterogeneous learning groups or behavioural conflicts. Therefore, MITHOS[1], *an Interactive Mixed Extended Training to Support Professional Socio-Emotional Interactions at Schools*, examines the potential of an interactive extended reality training to support the ability of teachers to emotionally assess conflict situations in the classroom and re-evaluate possibilities for behaving with the students. Students are good at perceiving subtle emotions and underlying aggression in teacher behaviour, and recursive teacher-student behavioural influences disrupt classroom interaction (Frenzel et al., 2021) and result in conflict escalation. Emotion regulation, like self-compassion, and perspective-taking in learning contexts can support understanding and coping with one's own and others´ emotions (Park & Tsovaltzi, 2022). Training teachers' emotion regulation and perspective taking can thus improve subtle (non-)verbal behaviour and conflict resolution. However, teachers don't have opportunities to be trained using realistic conflict situations and mostly have only theoretical knowledge before they teach in real classrooms (Krause et al., 2013), leaving teachers with no available personal socio-emo-

tional resources. The MITHOS training system aims to enhance teachers' conflict resolution strategies. It employs realistic situative learning opportunities in a challenging classroom situation. Socially and emotionally challenging situations are realised in a virtual-immersive XR environment. In this environment, teachers interact in a safe virtual classroom environment with virtual student agents, who can simulate social interaction (Lugrin et al., 2021) and whose behaviour induces realistic conflict situations. The conflict interaction is implemented in an automatic modelling process of the agent providing immediate implicit and situative social feedback (Fonagy & Gergerly, 2005), much like in human-to-human interaction. The relative realism but reduced complexity of the training situation helps to bridge the gap between theory and classroom. The training supports the acquisition of professional social skills for teaching, focusing on perspective-taking and compassion.

The MITHOS system provides a complete training of four different stages (Figure 1):

- Training Stage 1 <Situative training stage>: a safe virtual classroom provided for the teacher to freely interact with a virtual student-agent.
- Training Stage 2 <Avatar-replay stage>: provides a spatial situative perspective of the teacher's behaviour during the interaction (from the perspective of the student).
- Training Stage 3 <Agent-feedback stage>: individual virtual reflection guidance on emotional experiences through co-regulation processes.
- Training Stage 4 <Expert-feedback stage>: expert feedback on professional conflict resolution in a classroom.

In an immersive XR interaction simulation (Training Stage 1), teachers may interact freely and experience different student reactions to their behaviour, as the conflict evolves in several turns (Hartmann et al., 2023). Through the interaction, the teacher becomes more aware of the conflict through an escalation of the behaviour. In the avatar-replay stage (Training Stage 2), the system plays back the teacher's behaviour using a virtual agent to simulate spatial perspective-taking (Kessler & Thomson, 2010) as a means to support emotional self-awareness (Gyurak et al., 2011a) and emotional perspective-taking (Batson et al., 1997). Thisthen promotes further self- awareness in the form of reflection on the teacher's own and the student's emotional experiences and behaviour (Hooi & Cho, 2014). Subsequent training stages focus on emotion co-regulation (Gergely & Watson, 1996) and positive regard (Lenske & Mayr, 2015). They are the mechanisms to enable emotional transformation and behavioural change (Moser & Von Zeppelin, 1996). They help the teacher to acquire a compassionate stance and implement differ-

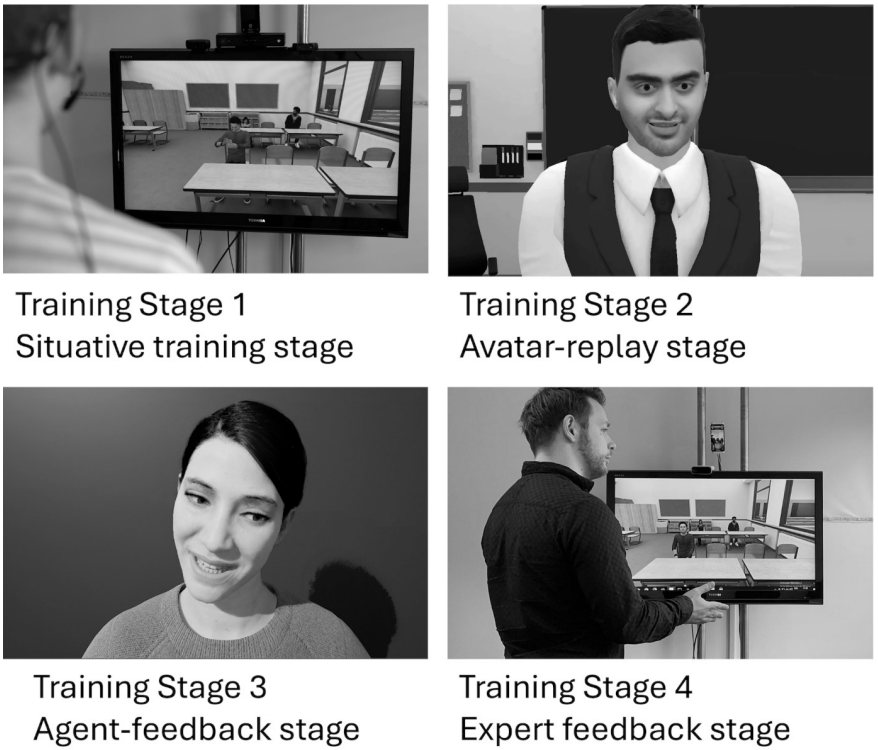


Figure 1: Four stages of the MITHOS training system
Source: Project's own representation

ent interaction patterns in the classroom. The new experiences gained in VR can become internalised and be transferred to real classroom situations.

2. Human-centered system development

The human-centered and generalizable concept includes a novel combination of virtual-immersive and collaborative 3D environments, social signal analysis and synthesis of verbal, non-verbal and paraverbal behaviours to simulate interactive virtual agents. Together, this allows for realistic training. MITHOS, as a first project, enables the real-time exchange of social signals between humans and agents in a teaching setting, including natural, implicit social feedback, as well as the coupling of virtual and real-world experiences. It thus simulates socio-emotional co-regulation.

The implementation is done in two versions: 1) a Wizard of Oz (WoZ) system (Figure 5), and 2) an Automated System. The WoZ system, an interface controlled mostly by a human, is designed to empirically test and validate separate aspects of the behaviour modeling and to collect data for developing the Automated System. In the current MITHOS WoZ system, the analysis is performed in real time by a wizard trained to evaluate the teacher's behaviour and the evaluation is automatically converted into suitable multimodal animation commands by the system. The collected multimodal audio-visual data is used to develop a hybrid affect model, which utilises both machine learning and theory-based reasoning models to decipher the situational emotional experience. This model is designed to process speech, gaze, pose, facial expression and natural language dialogue, represented as an interaction act, that is a representation structure defined and used to understand the teacher's behaviour and emotional experience using computational models of affect, namely ALMA (Gebhard, 2005) and structured psychodynamic diagnostic models (OPD Working Group, 2001). The planned affect model then automatically generates appropriate student responses over multiple dialogue turns. The main advantage of using psychodynamic models and deep modelling for co-regulation training lies in the long-term effectiveness for well-being and happiness, which goes beyond the treatment of symptoms. In this regard, psychodynamic methods are more suitable for interaction modelling, which does not target specific symptoms but generally observable, but relatively stable motives on a personal level, as described in the OPD (Operational Psychodynamic Diagnostic) (Pilling et al., 2020).

The MITHOS training, afforded by both the WoZ and the final automated system, is anchored in the teacher training university programme of studies for systematic empirical investigation. The training situations can be run through repeatedly and do not risk disadvantages for individuals. Both trainee teachers and experienced teachers can use these trainings to prepare purposefully for the challenges of teaching groups of learners who may have different cultural, emotional, and cognitive backgrounds.

3. XR in learning, teacher education, conflict and its resolution

Current research in the field of virtual 3D learning and simulation environments has shown that the use of this technology can have great potential for learning and training effects (Greenwald et al., 2017; Lerner et al., 2020). In teacher training, for example, XR is used to offer prospective teachers opportunities to practice strategic situational action in a protected environment (Dalinger et al., 2020). It has been shown that 3D learning and simulation environments with virtual in-

teractive agents have high potential for social and emotional learning. Fields of application include preparing for lessons with the help of a virtual audience (Batrinca et al., 2013).

Previous XR approaches in conflict training (Dalinger et al., 2020), in which student reactions are scripted and performed by actors, show advantages for authentic experiences but have problems with credibility and personalised training. The perception of one's own needs and emotional arousal play a very important role here (Bautista & Boone, 2015). In addition to AI methods, group awareness tools (GATs) combined with behavioural support can highlight communication characteristics, support group processes, and enhance perspective-taking in learning contexts (Puhl et al., 2015; Tsovaltzi et al., 2014, 2017).

In the training of teachers, knowledge about social interaction (Baumert & Kunter, 2011) and the perception and interpretation of cognitive and emotional processes (Praetorius et al., 2018) are particularly important. In order to successfully train these skills, an adequate theory-to-practice connection is essential. (Video-based) case vignettes are often used for corresponding training (Thiel et al., 2017), as is virtual training (Kervin et al., 2006; Lugrin et al., 2016). Another important aspect of training is feedback (Hattie, 2009), which consists of an immediate and a delayed component (Smith & Ragan, 1993). This allows teachers to learn to decide quickly and flexibly on a reaction (immediate feedback) as well as to rethink complex conditions within a challenging situation (Lotz & Lipowsky, 2015).

While there are previous existing approaches focusing on providing a safe learning environment, their focus is mostly on training knowledge acquisition and teaching strategies (Dieker et al., 2016; Dieker et al., 2015). TLE TeachLivE (Dieker et al., 2015) uses mixed-reality simulations to train teachers' instructional and classroom management skills through immersive, iterative practice and feedback, particularly in challenging or hard-to-reach educational settings. Their system is modelled based on the behavioural level of the student and teacher. Other studies show the importance of training teachers in a safe environment using computer games or serious games (Connolly et al., 2012). MITHOS stands out as it focuses on training socio-emotional skills and classroom management. MITHOS emphasises the importance of respectful and professional interactions in a classroom as a basis regardless of the actual teaching strategy or content used. It also emphasises teacher well-being. MITHOS provides individual reflective training stages to enhance the transfer of new experiences to day- to-day challenges.

4. Situational interaction for training professional respectful conflict resolution

4.1 Respectful social interaction in a classroom

Acquiring socio-emotional skills influences communication skills and professional behaviour (Anužienė et al., 2015) and contributes to health and well-being (Chang, 2009; M. M. Keller et al., 2014). During social interactions, reactions are determined based on the appraisal of perceived emotional reactions of others (Manstead et al., 2001). This may increase teachers' awareness of their behaviour during high arousal situations, i.e. the initiated conflict. Self-awareness draws one's attention not only to one's own emotions and needs but also those of the others. By increasing self-awareness, a person is more likely to adopt adaptive regulating strategies (Gross, 2015). In MITHOS, real-time interaction with a virtual agent provides teachers with immediate feedback on their performance, similar to human-to-human interaction. This feedback enhances teachers' awareness of whether their behaviour was respectful toward students. However, high arousal levels may reduce the ability of perspective-taking and trigger the switch of emotion regulation to automatic response through fight or flight. This switch may be influenced by different factors such as previous internal conflicts (Fonagy & Bateman, 2006). In MITHOS, we induce self-focused attention through the replay of the interaction using an avatar to foster self-awareness and perspective-taking (Training Stage 2). The teachers watch a replay of the interaction from the perspective of the student; watching their own expressions on an avatar.

Perspective-taking and self-awareness are directly influenced by one's internal conflict. Internal conflicts stem from experiences during development and are triggered by present interactions. Internal conflicts are experienced when one's desires or socio-emotional needs are not met. The dominant emotional state, the lead affect, influences how one behaves to reach the desired state in the interaction (Gross, 2002). Furthermore, the behaviour is influenced by the teacher's intention, which is shaped by generally accepted social norms (Ajzen, 1991). In MITHOS, we aim to track the lead affect of the teacher during the automated interaction, based on the OPD method (OPD Working Group, 2001). In the Wizard of Oz (WoZ) experiment, a trained interviewer conducted an interview with the participant to reflect on their experience in the VR classroom. The aim is to understand the participant's intentions behind specific behaviours and to identify the lead affect driving their actions. We analyse the connection between the self-reported internal conflict of the teacher, the lead affect and the teacher's intentions. In the final automated system, we utilise this data to model the automatic

responses of the virtual student agent, as well as the co-regulating feedback agent designed to train self-compassion.

Self-compassion improves emotion regulation (Neff et al., 2007), increasing the capacity to co-regulate. Training on self-compassion specifically may deactivate the threat system (associated with feelings of insecurity and defensiveness) and activate the self-soothing system (Gilbert, 2016). By training teachers to be self-compassionate, we expect to influence their ability to self-regulate and co-regulate students' emotional needs in a classroom, and in turn influence the students' ability to self-regulate their emotions (Mainhard et al., 2018). In MITHOS automated training system, the teacher will be supported by a feedback agent to reflect on the experience and provide support on emotion regulation. The feedback is specifically adapted to the personalised experience and responses as detected by the automated system. In an attempt to target the affect regulation system, the agent trains compassion to the self as a socio-emotional skill. Enhancing the self-soothing system allows the teacher to step back from the situation and perceive it objectively with consideration to both oneself and the other (Goetz et al., 2021). Reflecting on experiences with self-compassion may further enhance self-awareness and perspective-taking, promoting more respectful interactions and supporting overall well-being.

4.2 Professional conflict regulation in a classroom

Conflict regulation in the classroom – especially handling disturbances while teaching – is a crucial professional competence of teachers (Kunter et al., 2013). For novice teachers, dealing with conflicts in the classroom is a stress-related factor and they often don't feel well prepared (Dicke et al., 2015). To deal with disturbances and conflicts, the three dimensions control (rules and routines, behaviour control etc.), student-teacher relationship (positive regard, student participation etc.), and task focus (e. g. relevance of learning aims, cognitive activation) should be implemented in a balanced way (Lenske & Mayr, 2015; Marzano, 2005). A positive student-teacher relationship includes authenticity, positive regard, empathic understanding, and positive emotionality. There is a positive relationship between positive regard and empathic understanding of a teacher with less aggressive or disturbing behaviour from students, but also higher thinking levels and learning success in students (Aspy & Roebuck, 1974; Aspy & Roebuck, 1972; Hattie, 2008; Tausch, 2008). The effects can be explained with the behaviours of more empathic and compassionate teachers: For example, they tend to use more eye contact and smile more, are more interested in their students' feelings, and express appreciation toward them (Aspy & Roebuck, 1974;

Aspy & Roebuck, 1972). Despite the fact that positive regard and empathetic understanding of students are an important basis for conflict regulation, the task focus and control dimensions should not be overlooked. Students should be motivated and activated to participate in the lesson (cognitive activation, Kunter et al., 2013); the teacher should also manage the behaviour of the students (behaviour control), without neglecting the basic psychological needs of autonomy (cf. self-determination theory, Ryan & Deci, 2017). In MITHOS, we want to evolve teachers' ability to react in a balanced way to conflict behaviour of students, that means to focus on the task and simultaneously on a positive relationship to the student. Accordingly, the virtual students' reactions are composed of two parts,addressing both the relationship and the task dimension, to enable differentiated training.

5. Interaction simulation stage

The MITHOS training, whether WoZ system or Final Automated System, takes place in four main training stages: 1. Situative training stage, 2. Avatar-replay stage, 3. Agent-feedback stage and 4. Expert-feedback stage. All stages together aim at increasing the teacher's awareness and ability of perspective taking, and training social interaction with a conflict scenario in an interactive classroom.

In a situative training stage, teachers train social interaction with an interactive classroom conflict scenario. They hold a class in front of a virtual classroom where a conflict is initiated through the behaviour of one of the virtual students. The teacher is required to involve the student in the lesson and guide them to paying attention, in a professional and respectful way (Aspy & Roebuck, 1972; Baumert & Kunter, 2011; Kunter et al., 2013; Tausch, 2008). The virtual student reactions are implemented reciprocally and are multimodal, to reflect different communication channels that include verbal and non-verbal, as well as explicit and implicit signals in the interaction. Through the signals, the teacher's implicit emotional awareness increases. Implicit emotional awareness is non-effortful attention to one's emotional experiences which directly influences one's emotion regulation and in turn the conflict potential in the interaction. The next stages provide feedback (implicit and explicit) and train the teacher to be more professional. In the avatar-replay stage of the interaction, teachers use an explicit emotion awareness tool, implemented as a virtual agent who plays back their verbal and non-verbal behaviour during the interaction. This allows them to take the physical perspective of the students and observe their own, earlier reactions, to increase their awareness of the student's emotional perspective. In the agent-feedback stage, the teacher reflects on their experience with the help of a virtual

training agent in a process of co-regulation. The agent supports the teacher through compassionate reactions to their emotional experience. Compassion complements explicit emotional awareness (Gyurak et al., 2011b) as it reduces personal distress (Duarte et al., 2016). Finally, in the expert-feedback stage, the teacher gets verbal feedback from experts on positive regard as an important aspect of professional conflict regulation in the classroom (Prilop et al., 2020; Prilop et al., 2021; Weber et al., 2018).

5.1 Conflict situations in a heterogeneous classroom setting

The debate on heterogeneity in schools is primarily focused on academic performance and achievement, which is due to the fact that academic performance is often used as a measure of success in the education system (Glock et al., 2020). The internal conflict of the virtual student is defined by the OPD-model to represent a common internal conflict that surfaces in heterogeneous classrooms for students and teachers. Accordingly, the reactions of the virtual student are defined. Two dimensions of heterogeneity are additionally implemented: the linguistic heterogeneity and the visual representation of the virtual student. These address stereotypes that teachers may hold against combinations of representations of those dimensions. Teachers may be affected by their stereotypes in the conflict and more likely decide to use more drastic methods (Nichols, 2004; Rocque & Paternoster, 2011; Skiba et al., 2002; Townsend, 2000).

The conflict behaviour shown by the virtual student is thus designed by expert teachers and a therapist who were responsible for the consistency of the student behaviour and the represented internal conflict. The scenario is composed of four conflict stages (Darnon et al., 2006; G. Keller, 2001) that are each structured in seven levels of conflict for the task and relationship parts, mentioned in the previous chapter (Rattay & Wensing, 2011; Schwarz, 2013). Therefore, a mapping between the conflict stages and the conflict levels for task and relationship was created and filled with spoken text samples, gestures and body movements. As the conflict levels for task and relationship vary from each other, the behaviour was created separately and the combinations were validated in a separate pilot study. In total this makes a set of 86 different behaviours composed of a verbal and a non-verbal part for each scenario. Currently two scenarios are scripted, based on the experiences made in a higher academic track (Gymnasium) and a non-academic track (Realschule), both in Germany. The implementation of the linguistic heterogeneity is oriented on the three dimensions of academic language: the lexical, grammatical and discursive features (Heppt, 2016). The way that the text files are spoken will also be variable: Two students (male and female) are recorded for the

audio files, which are then rendered by an AI-voice model (https://elevenlabs.io/). The original sample files of the students are created in Saarland dialect and in standard German. This again provides the possibility to address another dimension of linguistic heterogeneity.

The linguistic heterogeneity, the gender and the skin color of the virtual student are randomised to not reinforce prejudices that may exist in form of combinations of the dimensions of the language and the visual representation of the virtual student.

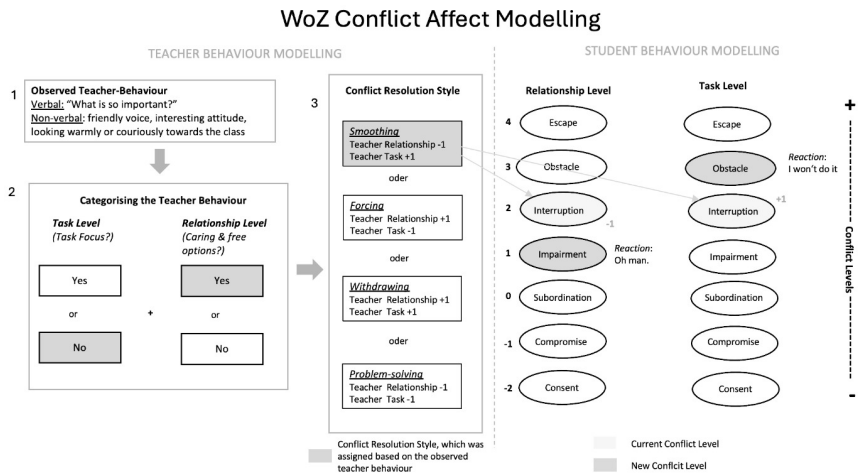


Figure 2: The mapping of the conflict resolution styles on to the different possible phases in the MITHOS Training Stage 1: Situative training stage
Source: Project's own representation

5.1.1 Training Stage 1 <Situative training stage>

This stage provides an interactive virtual platform for the teacher to experience dealing with complex socio-emotional challenges, such as those that can be triggered by heterogeneous learning groups or behavioural conflicts. During this stage, the teacher holds an introductory sequence in front of a virtual classroom where a conflict is initiated through the behaviour of one virtual student. A trained human-wizard evaluates the teacher's behaviour based on the semi-automatic model, which maps onto a student response. The interaction is structured in four conflict phases to account for different phases of conflict resolution (Darnon et al., 2006; G. Keller, 2001) (Figure 2). The virtual student's behaviour

during the whole interaction with the teacher is modelled, based on a turn-by-turn evaluation of the teacher behaviour on two dimensions, namely task and relationship conflict level (Darnon et al., 2006), and a matrix of conflict regulation potential ranging from -2 (low conflict potential) to +2 (high conflict potential) (Holt & DeVore, 2005) (Figure 4). The student reaction is based on the assessed conflict resolution style of the teacher. The conflict resolution style comprises two levels: task and relationship (G. Keller, 2001). The model penalises (-1) or rewards (+1) the teacher's behaviour at both levels locally for each turn, and a global conflict level status is updated accordingly (see Figure 2). The conflict level status is based on an ordered list which represents the conflict escalation potential of the conflict resolution style and connects the teacher behaviour to the student reaction. The concrete definitions of the levels of conflict resolution style are based on the OPD model and represent a type of internal conflict commonly observed in situations of diversity discrimination (OPD Working Group, 2001), and hence very likely to surface in relations of power in heterogenous classes: the control-submissiveness conflict. Therefore, at the task level, the teacher is required to involve the student in the lesson and guide them to paying attention, in a professional and respectful way. That is, the teacher should be able to maintain a balance between involving the student to a certain extent that does not violate their sense of control. At the same time, and crucial for the internal conflict, at the relationship level the teacher emotionally supports the student while respecting the student's personal space and limits. The virtual student reactions are implemented reciprocally and are multimodal, to reflect different communication channels that include verbal and non-verbal, as well as explicit and implicit signals in the interaction. The conflict behaviour of the teacher influences the conflict potential in the interaction. The teacher may successfully resolve the conflict in the final solution phase or fail to resolve it.

5.1.2 Training Stage 2 <Avatar-replay stage>

In this stage an avatar plays back the verbal and non-verbal behaviour of the teacher allowing them to observe their reactions and take the physical perspective of the student. The avatar-replay is considered an awareness tool, as the teacher learns about their own behaviour, such as facial expressions, from the others' perception. In this stage, the teacher is asked to watch the avatar-replay on a monitor (see Figure 3c). The visual similarity between the teacher and the avatar has been shown to play an important role in increasing self-awareness (Hooi & Cho, 2012). In a previous study, we investigated the effect of different degrees of similarity and the necessity to include expensive similar avatars for perspective-taking in MITHOS. We conducted a study defining three conditions

by manipulating specific facial features of the virtual avatar (Hilpert et al., 2024). The results show that perceived similarity is influenced by two degrees of manipulations of avatar-person similarity. This allows some room for error in studies when designing personalised avatars. However, the right degree of manipulation is required for explicit identification with the avatar. Therefore, here we further investigate the added value of personalised avatars for increasing self-awareness. We compare two different conditions, watching the replay of one's behaviour on a generic default avatar compared to a personalised avatar of maximum possible similarity using MetaHuman Creator (MH Creator), which can be used freely in the Unreal Engine 2 environment. Our empirical research indicates that generic avatars may not adequately represent the self when perceptions are involved (Alves da Silva et al., 2023). Hence, to further validate the importance of having a similar avatar and also assess the influence of behavioural contingency, we consider two groups. In the first group, the teacher watches the replay of their behaviour using a default avatar. In the second group, the teacher watches the replay of their behaviour on a personalised avatar (Figure 3c), bearing similar external facial features to the participant, based on three images received prior to the experiment day.



Figure 3: (a, b, c): The room setup for the MITHOS Wizard of Oz (WoZ) system used for Training Stage 1 (3a and 3b) and the personalised MetaHuman virtual avatar used for Training Stage 2 (3c)
Source: Project's own representation

5.1.3 Training Stage 3 <Agent-feedback stage>

This stage involves a virtual training agent supporting the teacher to emotionally assess their experience during the interaction with the virtual student. The stage starts with an introduction to establish trust between the teacher and the training agent (Lugrin et al., 2021). The training agent structures the question based on the detected lead affect and relevant internal conflict highlighted throughout the interaction. The intention of the teacher is questioned to ensure self-reflection to further increase their awareness of the experience. As the teacher reflects on

their experience, the training agent supports the teacher in a process of co-regulation. Co-regulation in this context is the training agent's ability to validate and understand the emotions that the teacher experienced and provide implicit or explicit empathic responses with the aim of supporting the teacher's emotion-regulation (Murray & Rosanbalm, 2017). We consider the empathic response to encompass both affective empathy, which is reflecting the ability to validate and share the feelings of the teacher, and cognitive empathy or perspective-taking, which is simply taking the view of the teacher into consideration. We further integrate compassion within the agent's empathic response to train the teacher in self-compassion.

Self-compassion has been positively correlated with self-regulation, prosocial behaviour and well-being (Sirois, 2015). By training teachers to be self-compassionate and supporting their ability to self-regulate during conflict situations, we increase their ability to focus on students' emotions and learning performance.

5.1.4 Training Stage 4 <Expert-feedback stage>

Finally, in the expert-feedback stage, the teacher gets verbal feedback from human experts on positive regard as an important aspect of professional conflict resolution in the classroom. Expert feedback plays an important role for the development of professional behaviour (Prilop et al., 2020, 2021; Scheeler et al., 2004), which also corresponds to Ericsson's theory of deliberate practice (Ericsson et al., 2003). In MITHOS, we support experts in giving feedback with a rating inventory regarding positive regard (adapted from Protzel, 2015). The computer-based rating inventory generated the feedback after the rating in a semi-automated way. This should support the expert, who observes the behaviour of the student teacher in the virtual classroom, to give informed and valuable feedback. We decided to add feedback from a human expert, due to the fact that the assessment of empathy or positive regard is actually too complex for a fully automated system.

6. Developing computational models for automating the system (used in both WoZ and automated systems)

6.1 Conceptual explanation of conflict model

The virtual student sequence behaviour has two end alternatives: escalation (the student gets up and leaves the classroom) and resolution (the student puts the mobile phone down and turns his attention to the lesson). This is dependent on

the internal conflict initially assigned to the virtual student and the behaviour of the teacher. Based on the conflict resolution strategies, the teacher may attempt to resolve the situation in 5 different ways: withdraw, force, problem solve, smoothe and compromise (Rahim, 2001).

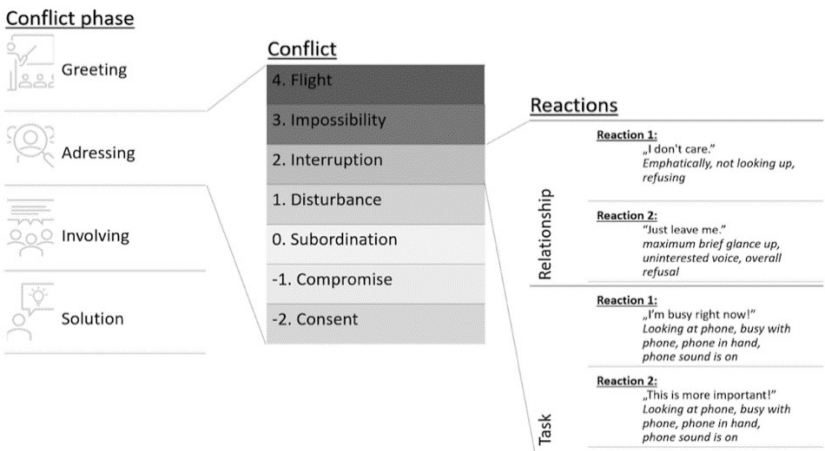


Figure 4: Modelling of teacher behaviour: task and relationship conflict
Source: Project's own representation

The student's reaction is always operationalised as a shift between the conflict levels by one unit as indicated in Figure 2. The top level represents the most suboptimal level and the lowest level represents the optimal level, where the student conforms to the request and attends to the classroom. For example, if the teacher displays a behaviour that has been assigned to the conflict resolution style »Smoothing« (Task Focus = No; Attachment & Freedom [Relationship] = Yes), the current conflict level moves up one level on the task level (towards »Flight«) and down one level on the relationship level (towards »Consent«). This shift in the current conflict level is then visible in the student's reaction. A stronger refusal to cooperate (task level) is shown in the student's next reaction while expressing a more emotionally agreeable approach to the teacher compared to the previous reaction (relationship level) (Figure 4).

6.2 Emotional state model, social obligations, intentions and internal conflict

In the WoZ study, a trained human expert, the wizard, watches the interaction on a live stream and rates the participant's performance on relationship level,

task level and conflict phase. In the final VR system, the process is to be automated based on computational models developed from the gathered data during the WoZ study. The emotion recognition module in the Automated System deploys what the participant decides to reveal and uses internal models such as ALMA (Gebhard, 2005) and MARSSI (Gebhard et al., 2018) to understand internal emotional experiences. This module is designed based on psychological models and theories, which are validated by self-reports and qualitative analysis of the recorded data in the WoZ system for fine-tuning in the final system.

The self-reported data from the WoZ study, although provided directly after the experience, provides room for self-reflection and can be strongly influenced by various biases (Choi et al., 2023). Therefore, to better understand the emotional experiences of the participants and to improve the computational models for an automated system, we conducted an interview with each participant after the experiment. The questions guided the participant to report rather personal and sensitive feelings they experienced during the interaction specifying what influenced their behaviour such as previous experiences and social norms. To support the participant to recall their interaction, they watched their interaction on a video. The video stopped at certain points that reflected expressions worthy of reflection. Understanding the intentions of the teacher provides a deeper understanding of the role of the perceived emotions in shaping their behaviour. The interviewer was trained to provide a safe and trusting environment and to encourage the participant to reflect on their feelings. The interviewer validated their understanding by repeating what the participant expressed. Furthermore, self-reported data on emotions was gathered at each point during the interview as this data is rather more specific and personal and could be sensitive to talk about. The self-reported data included questions on the emotional state model (Mehrabian & Russell, 1974) and lead affect (adapted from OPD Working Group, 2001). Qualitative analysis of the interview allowed us to assess the participant's reaction to gather the social obligations, intentions and the internal conflict of the teacher.

7. Technical implementation of MITHOS training system (DFKI)

7.1 MITHOS Wizard of Oz (WoZ) system

The MITHOS WoZ system consists of hardware and software components (Figure 5) that allow the teacher to interact with the virtual agent, while unknowingly being monitored by the wizard. The wizard is interpreting their actions »behind the scenes«. Based on this interpretation, VisualSceneMaker (VSM) automatically updates the conflict level for the teacher behaviour, as well as the conflict level and

phase of the student reaction, and sends corresponding, predefined animation commands to the social agents in the virtual classroom for rendering.

As seen in the interaction room (Figure 3a), one computer is used to run the Xbox Kinect, web camera, microphone and heart rate sensor. The data from these sensors are stored, as well as streamed in real time to the Wizard in the other room (Figure 3b), as indicated in Figure 5. Simultaneously, as shown in Figure 5, another computer is connected to the iPhone, which in turn detects the facial expressions of the participant during the interaction, which will be played back in the replay stage. At the same time, a third computer is used by the Wizard in another room (Figure 3b) to evaluate the behaviour, in real time, of the participant who is interacting with the agent student character. Detailed information about the individual software components, from Figure 5, is provided in Bhuvaneshwara et al. (2023).

In addition to enabling real-time and natural student-teacher interaction based on expert knowledge, this setup also allows for the recording of audio, visual, postural and heart rate features that are used to build a machine learning dataset. This data is being used to develop multimodal non-verbal and verbal affect analysis machine learning systems that are used to drive computational affect models for automatically determining the current emotional experience of the teacher and to generate appropriate student behaviour. Additionally, facial data captured during the interaction is used to replay the teacher's behaviour for the avatar-replay stage in the current WoZ system.

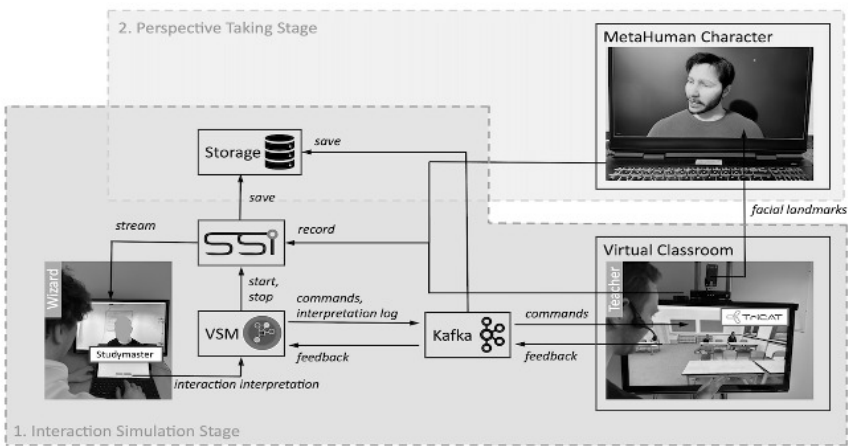


Figure 5: System architecture overview of MITHOS Wizard of Oz (WoZ) setup that depicts the setup spread over two rooms and the different system components that the participant and wizard, respectively, interact with

Source: Project's own representation



Figure 6: Participant using the VR classroom
Source: Project's own representation

7.2 MITHOS Automated Training System

The MITHOS Automated Training System is developed using the data collected from the study conducted using the MITHOS WoZ system. Besides the components in the WoZ data collection system (Bhuvaneshwara et al., 2023), the final automated MITHOS training system is set up in a single room with the HTC Vive and associated trackers, all connected to one computer. There are several software components that are required to enable the full functioning of this automated system and these components are as follows:

1. TriCAT spaces® VR enabled classroom environment for teachers: This is the central, virtual-immersive, 3D-environment which can be visited by the participating teachers, as in Figure 6, in Virtual Reality mode with the HTC Vive headset, through a desktop application.
2. AffectToolbox (Mertes et al., 2024): This is used in the context of MITHOS to generate an understanding of the teacher's observable emotional behvaiour within the Pleasure, Arousal and Dominance (PAD) dimensional space (Mehrabian & Russell, 1974). This information will be used by other components to further comprehend the affective state of the participant based on indicators in the verbal behaviour, classroom social norms and the decay of emotions with time.
3. Geni:OS is a software framework for multimodal intelligent assistants which is used in MITHOS for natural language understanding, integration of input modalities, and the evaluation of social norms (SNs) relevant in the classroom. This input processing replaces the phase and task focus interpretations by the wizard in the previously described WoZ system.
4. Interactive Co-regulation using VSM, ALMA and OPD: As in the WoZ study, all the final decisions to produce the agent student behaviour are made in the Visual SceneMaker (VSM). VSM receives multiple PAD updates from the AffectToolbox, and it also receives the interaction acts from Geni:OS. The updates from both of these components can be fed into the emotion model ALMA (Gebhard, 2005), which is running within the scope of VSM. ALMA uses all the AffectToolbox updates from within one interaction act and the information from the interaction act itself to simulate the teacher's regulated (i. e. perceivable to the outside) emotion in real time. The detected regulated emotion is then mapped to the concept of Lead Affect that is driving the person's behaviour. The mapping is based on the model developed using OPD and ALMA (Gebhard, 2005) and is designed to produce the information about the relationship level. The information about the task level and phase of the interaction is supplied by Geni:OS. VSM puts all of this information together to produce the conflict resolution style and therefore drive the student behaviour.

8. Empirical validations of the interaction scenario and evaluation of the avatar-user similarity

We conducted two evaluation studies for piloting (a) the authenticity of the interaction behaviour of the virtual student and (b) the effects of similarity of the virtual agent with the participant for the avatar-replay stage.

8.1 Study 1: Interaction scenario

For the effectiveness of a virtual training, the authenticity of the simulated scenario is of high importance for evoking comparable cognitive processes (like norms and expectations) and affective reactions (like stress) as an interaction with a real student in a real classroom setting would do (Poeschl & Doering, 2013). If a virtual scenario is seen as realistic and authentic by the users, it can result in a feeling of presence, meaning that people feel, think and behave the way they do in a similar real-world situation (Pertaub et al., 2002). To assess the authenticity of the scenario, we created three videos of a first demonstrator with reactions of the virtual student. The videos showed different conflict solutions and consisted of 10 to 11 reactions of the virtual student. The duration of the reactions varies between 7 and 15 seconds and is about 2.5 minutes for each of the three complete scenes. VSM automatically generates a combination of possible behaviours of the virtual student depending on the stage of the conflict and the rating of the Wizard. Therefore, over 500 behaviour combinations could be possible. This includes combinations with similar conflict levels on the task- and relationship-dimension and combinations with different conflict levels of the two dimensions. To evaluate further, whether the difference/similarity of conflict levels has an influence on authenticity, we selected reactions for (1) a similar conflict level and (2) a high difference of conflict level between task and relationship.

1) Participants and Design: We asked $N = 57$ participants to assess, first, each of the reactions of the student individually and, after that, the whole video scene. Participants were $n = 51$ pre-service and $n = 6$ in-service teachers. Pre-service teachers were on average in their 4th semester at university ($M = 4.10$, $SD = 2.94$) and had on average 14 hours of teaching experience during their internships ($M = 14.37$, $SD = 11.36$). The in-service teachers had on average 5 years of teaching experience ($M = 4.92$, $SD = 4.05$). We did not assess age and gender due to data protection reasons.

The assessment of the reactions of the interacting student includes three questions of authenticity (posture/gesture, language, and behaviour; example: »The behaviour of the virtual student corresponds to the behaviour of a real student«). After watching the whole scene, the participants rated the authenticity of the background students with four items regarding posture and gesture, behaviour, and comparability with real students (example: »The behaviour of the students in the class corresponds to the behaviour of real students«). Further, the whole scene should be rated with three items (Item stem: »I found the shown videos …«, »… meaningful«, »… untypical« and »… authentic«). All items were rated on a 5-point Likert scale from »not correct at all« to »fully correct«. The

items were adapted from (Fitrianie et al., 2020; Langer et al., 2022; Poeschl & Doering, 2013; Seidel et al., 2010; Seufert et al., 2022; Slater et al., 1999), Cronbach's alpha ranged from .89 to .96.

2) Results: The rating of the authenticity was for all scales above the mean scale value (see table 1 for details). The pre-service teachers assessed the scenario as more authentic than the in-service teachers. The difference was statistically significant for the authenticity of the whole scenario, but due to the very different sized groups, the results should be interpreted very carefully.

	All participants ($N = 56$)		**Pre-service teachers ($n = 51$)**		**In-service teachers ($n = 6$)**		**Mann-Whitney-U-Test**		
Authenticity of...	*M*	*SD*	*M*	*SD*	*M*	*SD*	*U*	*Z*	*p*
... interacting student	3.24	0.57	3.29	0.54	2.81	0.54	85.50	-1.76	.08
... background students	3.47	0.64	3.49	0.66	2.35	0.64	131.50	-0.50	.58
... whole scene	3.09	0.81	3.17	0.79	3.31	0.62	59.50	-2.44	.02

Table 1: Descriptive statistics and differences between pre-service and in-service teachers for the authenticity scales

Source: Own Representation

To evaluate, if we could find a difference of authenticity of the behaviour of the interacting student depending on the similarity of the conflict level of task and relationship dimension, we performed an ANOVA with repeated measurement on the factor similarity (high vs. low). We found a statistically significant difference, meaning that the reactions with a low similarity are less authentic than reactions with high similarity of the two dimensions (high similarity: $M = 3.34$, $SD = 0.62$; low similarity: $M = 3.19$, $SD = 0.56$; $F(1,56) = 8.01$, $p < .01$).

3) Discussion: A VR-Classroom should be authentic so that people feel, think and behave like they do in a similar situation in the real world, and therefore training can be effective. Our results showed a medium authenticity of the first scenario and slight differences between expertise groups; further we found a difference of authenticity depending on the similarity of the conflict level between task and relationship dimension. To optimise the authenticity of the MITHOS scenario, the movements of the virtual students will be smoothed; we will implement a natural voice and will adapt the content of the spoken sentences of the interacting student.

8.2 Study 2: Effects of agent-user similarity

As we use an avatar (a virtual agent representing the teacher-user) as an awareness tool during the avatar-replay stage to increase the teacher's self-awareness, we investigated the effect of varying degrees of avatar-person similarity and its influence on user-identification and affinity. Using a theory-based methodology, we looked beyond the distinction between a personalised vs. default avatar and included three conditions (Hilpert et al., 2024). We assessed the perceived facial similarity between avatar and user, the explicit-cognitive and implicit-emotional identification with the avatar and the perceived realism of the avatar.

1) Participants and Design: The study was conducted online and included 33 participants (13 males and 20 females) with an age range between 18 and 40 years ($M = 27.6$ years, $SD = 4.74$ years). The conditions ranged from the most similar having 0 % manipulation of facial features to two other conditions where manipulation increased by 50 % and 100 % and an additional control group (CG).

Participants filled out a questionnaire to collect their email address, three facial images (as required by FaceGen) and demographic data. The avatars for each condition were created in 3D and added in the form of animated images (GIF) representing idle behaviour. Then, the participants received the second questionnaire where the conditions and instruments were presented to the participants in random order. Among other measure, the instrument included 5 items on ›physical similarity‹ from the polythetic identification model (Downs et al., 2019).

2) Results: Our results partially confirmed the effect of manipulation of avatar-person similarity on user identification. The effect of avatar-person similarity manipulation on perceived similarity revealed a significant main effect, $F(1.87,59.72) = 10.13$, $p < .001$, $\eta 2 = .11$ Subsequent comparisons did not show a significant effect between the 0 % ($M = 3.34$, $SD = 1.26$) and the 50 %-condition ($M = 3.18$, $SD = 1.21$), $t(32) = 1.36$, $p = .18$. All other comparisons showed a significant difference, showing a steady decline in similarity with increasing similarity manipulation (0 % = highest – CG = lowest).

3) Discussion: Similar to previous research, the results indicate that avatar-person identification corresponds with perceived visual similarity. Including the different degrees of similarity shows that generic avatars are insufficient when studying concepts related to self-perception such as self-awareness. Moreover, the insignificance between 0 % and 50 % degree of manipulation gives researchers a margin of error to work with when personalising avatars for each participant. In MITHOS, we are creating personalised avatars as one condition in the avatar-replay stage to compare to default avatars. In this case, the avatar-replay simulates the behaviour of the participant during an interaction. The results would show

the influence of similarity on self-awareness within the context of behavioural contingency.

9. Conclusion and future work

Automated immersive XR environments developed based on social interaction and psychological models represent a novel combination to simulate interactive virtual agents and avatars. MITHOS is a human-centered, socially interactive, and generalizable training system, employing an innovative combination of virtual-immersive environments and social signal analysis to simulate interactive, realistic social interactions. The MITHOS system currently supports classroom conflicts for teacher training but was developed to be adaptable to other professional contexts where socio-emotional skills are essential. It specifically allows the training on self-compassion and the adoption of the other perspective. We present results validating the approach in terms of scenario realism, as well as of a systematic testing of the effects of external avatar similarity on antecedents of self-awareness with behaviour similarity. The chapter contributes to a common methodology of conducting interdisciplinary research for human-centered, socially interactive, and generalizable XR and presents a system designed and implemented to support it.

In this chapter we reported on the first phase of the development, following an iterative approach, which is a standalone result of the project. In the next phase, MITHOS aims to have a fully autonomous and adaptive virtual agent, based on an extension of the ALMA (Gebhard, 2005) computational model of emotion, mood, personality and emotion regulation. This enables it to track the teachers' regulated and unregulated emotions, and the conflict potential during the interaction, to trigger the virtual agent reaction. The emotional modelling is further used to increase implicit and explicit self-awareness by direct feedback. A Geni:OS component is integrated in the system and will analyse the teacher's communicative behaviour and affect display concerning speech, gaze, body movement, and interpersonal distance to provide the input to ALMA. Further evaluations are running, which are testing the effects of the overall automated feedback and evaluate the individual stages. The results currently provide evidence for the MITHOS concept of supporting perspective-taking through the situative mixed-reality self-compassion training.

Funding information

This study is funded by the German Federal Ministry of Education and Research (BMBF)[1] within the funding line »Interactive systems in virtual and real spaces – Innovative technologies for the digital society«.

Literature

Ajzen, Icek (1991). The theory of planned behavior [Theories of Cognitive Self-Regulation]. In: *Organizational Behavior and Human Decision Processes*, vol. 50(2), p. 179–211. https://doi.org/https://doi.org/10.1016/0749-5978(91) 90020-T

Alves da Silva, Claudio et al. (2023). Visual similarity for socially interactive agents that support self-awareness. In: *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*, p. 1–3. https://doi.org/10.1145/3570945.36073

Aspy, David N. & Roebuck, Flora N. (1974). From humane ideas to humane technology and back again many times. In: *Education*, vol. 95(2), Project Innovation Austin LLC, Austin, USA, p. 163–171.

Aspy, David N., Roebuck, Flora N. (1972). An investigation of the relationship between student levels of cognitive functioning and the teacher's classroom behavior. In: *The Journal of Educational Research*, vol. 65(8), Taylor & Francis Ltd., Abingdon, UK, p. 365–368.

Batrinca, Ligia et al. (2013). Cicero – towards a multimodal virtual audience platform for public speaking training. In: *Intelligent virtual agents, Conferene Proceedings 13th International Conference, IVA 2013, Edinburg, UK,* Springer Berlin/Heidelberg, p. 116–128. https://link.springer.com/book/10.1007/978-3-642-40415-3 [access 20.03.2025]

Batson, C. Daniel, Early, Shannon & Salvarani, Giovanni (1997). Perspective taking: Imagining how another feels versus imaging how you would feel. In: *Personality and social psychology bulletin*, vol. 23(7), p. 751–758. https://doi.org/10.1177/0146167297237008

Baumert, Jürgen & Kunter, Mareike et al. (2011). Das Kompetenzmodell von COACTIV. In: *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*, Waxmann Verlag, Münster, p. 29–53.

Bautista, Nazan Uludag, Boone, William J. (2015). Exploring the impact of teachmeTM lab virtual classroom teaching simulation on early childhood education majors' self-efficacy beliefs. In: *Journal of Science Teacher Education*, vol. 26(3), Taylor & Francis Ltd., Abingdon, UK, p. 237–262.

Bhuvaneshwara, Chirag et al. (2023). Mithos-mixed reality interactive teacher training system for conflict situations at school. In: *ISLS Annual Meeting 2023*, Conference Paper, Montreal, Kanada, p. 42. https://2023.isls.org/proceedings/ [access 20.03.2025]

1 At the time of printing: Federal Ministry of Research, Technology and Space (BMFTR)

Chang, Mei-Lin (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. In: *Educational psychology review*, vol. 21, Springer Science+Business Media, Berlin, Heidelberg, p. 193–218.

Choi, Heeryung et al. (2023). Logs or self-reports? Misalignment between behavioral trace data and surveys when modeling learner achievement goal orientation. In: *LAK23: 13th international learning analytics and knowledge conference*, Arlington, USA, p. 11–21. https://doi.org/10.1145/3576050.35760

Connolly, Thomas M. et al. (2012). A systematic literature review of empirical evidence on computer games and serious games. In: *Computers & Education*, vol. 59(2), Elsevier Ltd, Amsterdam, NL, p. 661–686. https://doi.org/https://doi.org/10.1016/j.compedu.2012.03.004

Dale, Di & James, Chris (2015). The importance of affective containment during unwelcome educational change: The curious incident of the deer hut fire. In: *Educational Management Administration and Leadership*, vol. 43(1), Sage Journals, Washington D.C., USA, p. 92–106. https://doi.org/10.1177/1741143213494885

Dalinger, Tara et al. (2020). A mixed reality simulation offers strategic practice for pre-service teachers. In: *Computers & Education*, vol. 144, Elsevier Ltd, Amsterdam, NL, p. 103696. https://doi.org/10.1016/j.compedu.2019.103696

Darnon, Céline et al. (2006). Mastery and performance goals predict epistemic and relational conflict regulation. In: *Journal of educational psychology*, vol. 98(4), American Psychological Association, Washington D.C., USA, p. 766. https://doi.org/10.1037/0022-0663.98.4.766

Dicke, Theresa et al. (2015). Reducing reality shock: The effects of classroom management skills training on beginning teachers. In: *Teaching and teacher education*, vol. 48, p. 1–12. https://doi.org/10.1016/j.tate.2015.01.013

Dieker, Lisa et al. (2016). Mixed reality environments in teacher education: Development and future applications. In: Ludlow, Barbara L. & Collins, Belva C. (ed.), *Online in real time: Using Web 2.0 for distance education in rural special education*, American Council on Rural Special Education, Morgan Town (WV), p. 122–131.

Dieker, Lisa, A. et al. (2015). TLE TeachLivE^{TM}: Using technology to provide quality professional development in rural schools. In: *Rural Special Education Quarterly*, vol. 34, p. 11–16. https://api.semanticscholar.org/CorpusID:146586395

Downs, Edward, Bowman, Nicholas, D. & Banks, Jaime (2019). A polythetic model of player-avatar identification: Synthesizing multiple mechanisms. In: *Psychology of popular media culture*, vol. 8(3), p. 269. https://psycnet.apa.org/doi/10.1037/ppm0000170

Duarte, Joana, Pinto-Gouveia, José, & Cruz, Bárbara (2016). Relationships between nurses' empathy, self-compassion and dimensions of professional quality of life: A cross-sectional study. In: *International Journal of Nursing Studies*, vol. 60, p. 1–11. https://doi.org/10.1016/j.ijnurstu.2016.02.015

Eckstein, Florian, & Rüth, Ulrich (2015). Adventure-based experiential therapy with inpatients in child and adolescent psychiatry: An approach to practicability and evaluation. In: *Journal of Adventure Education & Outdoor Learning*, vol. 15(1), p. 53–63. https://doi.org/10.1080/14729679.2013.844573

Ericsson, K. Anders (2003). Development of elite performance and deliberate practice. In: Janet L. Starkes & K. Anders Ericsson (ed.), *Expert performance in sports: Advances in research on sport expertise*, Human Kinetics, Champaign, p. 49–83.

Fitrianie, Siska et al. (2020). The 19 unifying questionnaire constructs of artificial social agents: An iva community analysis. In: *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, Virtual Event, Scottland, UK, p. 1–8.

Fonagy, Peter, Gergely, Gyorgy & Jurist, Elliot L. (2005). *Affect regulation, mentalization, and the development of the self*, Other Press, New York.

Fonagy, Peter & Bateman, Anthony. W. (2006). Mechanisms of change in mentalization-based treatment of BPD. In: *Journal of clinical psychology*, vol. 62(4), p. 411–430. https://doi: 10.1002/jclp.20241

Frenzel, Anne C., Daniels, Lia, & Burić, Irena (2021). Teacher emotions in the classroom and their implications for students. In: *Educational Psychologist*, vol. 56(4), p. 250–264. https://doi.org/10.1080/00461520.2021.1985501

Gebhard, Patrick (2005). Alma: A layered model of affect. In: *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, Association for Computing Machinery, The Netherlands, Amsterdam, p. 29–36. https://doi.org/10.1145/1082473.1082478. [access 18.08.2023]

Gebhard, Patrick, Mehlmann, Gregor, & Kipp, Michael (2012). Visual scenemaker—a tool for authoring interactive virtual characters. In: *Journal on Multimodal User Interfaces*, vol. 6(1), p. 3–11.

Gebhard Patrick et al. (2018). MARSSI: Model of appraisal, regulation, and social signal interpretation. In: *Proceedings of the 17th International Conference on Autonomous Agents MultiAgent Systems (AAMAS)*, Stockholm, Sweden, p. 497–506.

Gergely, György & Watson, John S. (1996). The social biofeedback model of parental affect-mirroring. In: *The International journal of psycho-analysis*, vol. 77(6), p. 1181–1212.

Gilbert, Paul (2016). *Human nature and suffering*, Routledge, London.

Glock, Sabine et al. (2020). Die Einstellungen von Lehrpersonen gegenüber Schüler*innen ethnischer Minoritäten und Schüler*innen mit sonderpädagogischem Förderbedarf: Ein Forschungsüberblick. In: Glock, Sabine & Kleen, Hannah (ed.), *Stereotype in der Schule*, Springer, Heidelberg, p. 225–279.

Goetz, Thomas et al. (2021). Getting along and feeling good: Reciprocal associations between student-teacher relationship quality and students' emotions. In: *Learning and Instruction*, vol. 71, no. 101349. https://doi.org/10.1016/j.learninstruc.2020.101349

Greenwald, Scott, W. et al. (2017). Technology and applications for collaborative learning in virtual reality. In: *Making a Difference: Prioritizing Equity and Access*, 12th International Conference on Computer Supported Collaborative Learning (CSCL), Philadelphia, USA, International Society of the Learning Sciences, p. 719–726. https://repository.isls.org/handle/1/210. [Zugriff 28.03.2025]

Gross, James J. (2002). Emotion regulation: Affective, cognitive, and social consequences. In: *Psychophysiology*, vol. 39(3), p. 281–291. https://doi.org/https://doi.org/10.1017/S0048577201393198

Gross, James J. et al. (2014). Emotion regulation: Conceptual and empirical foundations. In: Gross, James J. (ed.), *Handbook of emotion regulation*, The Guilford press, New York, p. 3–20.

Gross, James J. (2015). Emotion regulation: Current status and future prospects. In: *Psychological Inquiry*, vol. 26, p. 1–26. https://doi.org/10.1080/1047840X.2014.940781

Gyurak, Anett, Gross, James J., & Etkin, Amit (2011). Explicit and implicit emotion regulation: A dual-process framework. In: *Cognition and emotion*, vol. 25(3), p. 400–412. https://doi.org/10.1080/02699931.2010.544160

Hargreaves, Andy (2000). Mixed emotions: Teachers' perceptions of their interactions with students. In: *Teaching and teacher education*, vol. 16(8), p. 811–826. https://doi.org/10.1016/S0742-051X(00)00028-7

Hartmann, Christian et al. (2023). Imagine & immerse yourself: Does visuospatial imagery moderate learning in virtual reality? In: *Computers & Education*, vol. 207, no. 104909. https://doi.org/10.1016/j.compedu.2023.104909

Hattie, John (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge, London.

Hattie, John. (2009). The black box of tertiary assessment: An impending revolution. In: *Tertiary assessment & higher education student outcomes: Policy, practice & research*, vol. 259, p. 275.

Heppt, Birgit M. (2016). *Verständnis von Bildungssprache bei Kindern mit deutscher und nicht-deutscher Familiensprache*, Dissertation im Fach Psychologie, Humboldt-Universität zu Berlin, Berlin.

Hilpert, Bernhard et al. (2024). Avatar visual similarity for social HCI: Increasing self-awareness. https://arxiv.org/abs/2408.13084 [Zugriff 28.03.2025]

Holt, Jennifer L. & DeVore, Cynthia James. (2005). Culture, gender, organizational role, and styles of conflict resolution: A meta-analysis. In: *International Journal of Intercultural Relations*, vol. 29(2), p. 165–196. https://doi.org/10.1016/j.ijintrel.2005.06.002

Hooi, Rosali & Cho, Hichang (2012). Being immersed: Avatar similarity and self-awareness. In: *Proceedings of the 24th Australian Computer-Human Interaction Conference*, Melbourne, Victoria, Australia, p. 232–240.

Hooi, Rosalie, & Cho, Hichang (2014). Avatar-driven self-disclosure: The virtual me is the actual me. In: *Computers in Human Behavior*, vol. 39, p. 20–28. https://doi.org/10.1016/j.chb.2014.06.019

Keller, Gert (2001). *Konfliktmanagement in der Schule: Moderieren, lösen, vorbeugen*, Kallmeyersche Verlagsbuchhandlung, Wolfenbüttel.

Keller, Melanie M. et al. (2014). Teachers' emotional experiences and exhaustion as predictors of emotional labor in the classroom: An experience sampling study. In: *Frontiers in psychology*, vol. 5, no. 118409. https://doi.org/10.3389/fpsyg.2014.01442

Kervin, Lisa, Ferry, Brian, & Carrington, Lisa (2006). Classism: Preparing tomorrows teachers for classroom reality. In: *Society for Information Technology & Teacher Education International Conference*, Orlando, USA, p. 3204–3211.

Kessler, Klaus & Thomson, Lindsay Anne (2010). The embodied nature of spatial perspective taking: Embodied transformation versus sensorimotor interference. In: *Cognition*, vol. 114(1), p. 72–88. https://doi.org/10.1016/j.cognition.2009.08.015

Krause, Andreas, Dorsemagen, Cosima & Baeriswyl, Sophie (2013). Zur Arbeitssituation von Lehrerinnen und Lehrern: Ein Einstieg in die Lehrerbelastungs- und –gesundheitsforschung, Springer, Heidelberg.

Kunter, Mareike et al. (2013). The development of teachers' professional competence. In: Kunter, Mareike et al. (ed.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project*, p. 63–77, Springer, Heidelberg.

Langer, Wiebke, Bruns, Julia & Erhorn, Jan (2022). Entwicklung und Validierung eines videobasierten Testinstruments zur Erfassung des Noticing mit dem Fokus auf Anerkennungs-

prozesse im inklusiven Sportunterricht. In: *German Journal of Exercise and Sport Research*, vol. 52(3), p. 386–397. https://doi.org/10.1007/s12662-022-00807-0

Lenske, Gerlinde, & Mayr, Johannes (2015). Das Linzer Konzept der Klassenführung (LKK). Grundlagen, Prinzipien und Umsetzung in der Lehrerbildung. In: Arnold, Karl-Heinz et al. (ed.), *Jahrbuch für Allgemeine Didaktik*, Schneider, Hohengehren, p. 71–84.

Lerner, Dieter et al. (2020). An immersive multi-user virtual reality for emergency simulation training: Usability study. In: *JMIR serious games*, vol. 8(3), e18822. https://doi.org/10.2196/18822

Lotz, Miriam, & Lipowsky, Frank (2015). Die Hattie-Studie und ihre Bedeutung für den Unterricht. In: Mehlhorn, Gerlinde, Schöppe, Carola & Schulz, Frank (ed.), *Begabungen entwickeln und Kreativität fördern*, Kreaplus, München, p. 97–136.

Lugrin, Birgit, Pelachaud, Catherine & Traum, David (2021). The handbook on socially interactive agents: 20 years of research on embodied conversational agents, intelligent virtual agents, and social robotics volume 1: Methods, behavior, cognition. Association for Computing Machinery, New York. https://doi.org/10.1145/3563659

Lugrin, Jean-Luc et al. (2016). Breaking bad behaviors: A new tool for learning classroom management using virtual reality. In: *Frontiers in ICT*, vol. 3, p. 26. http://dx.doi.org/10.3389/fict.2016.00026

Mainhard, Tim et al. (2018). Student emotions in class: The relative importance of teachers and their interpersonal relations with students. In: *Learning and instruction*, vol. 53, p. 109–119. https://doi.org/10.1016/j.learninstruc.2017.07.011

Manstead, Anthony S. & Fischer, Agneta H. (2001). Social appraisal: The social world as object of and influence on appraisal processes. In: Scherer, Klaus, R., Schorr, A. & Johnstone, T. (eds.), *Appraisal processes in emotion: Theory, methods, research*, University Press, Oxford, p. 221–232.

Marzano, Robert J. et al. (2005). *A handbook for classroom management that works*. ASCD, Danvers, MA.

Mehrabian, Albert, & Russell, James A (1974). The basic emotional impact of environments. In: *Perceptual and motor skills*, vol. 38(1), p. 283–301. https://doi.org/10.2466/pms.1974.38.1.283

Mertes, S. et al. (2024). *The affecttoolbox: Affect Analysis for Everyone*. https://arxiv.org/abs/2402.15195

Moser, Ulrich, & Von Zeppelin, Ilka (1996). Die Entwicklung des Affektsystems. In: *Psyche*, vol. 50(1), p. 32–84.

Murray, Desiree W., & Rosanbalm, Katie (2017). *Promoting self-regulation in adolescents and young adults: A Practice Brief*. Office of Planning, Research and Evaluation, Administration for Children & Families, US Department of Health and Human Services, 330 C Street SW, Washington, DC 20201. https://www.acf.hhs.gov/opre. [Zugriff 28.03.2025]

Nathanson, Donald L. (2014). Affect theory and the compass of shame. In: Lansky, Melvin, R. & Morrison, Andrew, P. (ed.), *The widening scope of shame,* Routledge, London, p. 339–354.

OPD Working Group (2001). *Operationalized psychodynamic diagnostics: Foundations and manual*. Hogrefe & Hutber, Seattle, Toronto.

Neff, Kristin D., Kirkpatrick, Kristin L. & Rude, Stephanie. S. (2007). Self-compassion and adaptive psychological functioning. *Journal of Research in Personality*, vol. 41(1), p. 139–154. https://doi.org/https://doi.org/10.1016/j.jrp.2006.03.004

Nichols, Joe D. (2004). An exploration of discipline and suspension data. In: *Journal of Negro Education*, p. 408–423. https://www.jstor.org/stable/4129625 [Zugriff 28.03.2025]

Park, Sheenyoung & Tsovaltzi, Dimitra (2022). Implicit and explicit emotion regulation for conflict resolution: Narrative and self-compassion as anti-bullying training. In: *Proceedings of the 16th International Conference of the Learning Sciences-ICLS 2022*, ISLS Annual Meeting 2022, Hiroshima, Japan, p. 187–194. https://2022.isls.org [Zugriff 01.03.2025]

Pertaub, David-Paul, Slater, Mel, & Barker, Chris (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. In: *Presence*, vol. 11(1), p. 68–78.

Pilling, Stephen et al. (2020). Long-term outcomes of psychological interventions on children and young people's mental health: A systematic review and meta-analysis. In: *PloS one*, vol. 15(11), e0236525. https://doi.org/10.1371/journal.pone.0236525

Poeschl, Sandra, & Doering, Nicola (2013). The German VR simulation realism scale-psychometric construction for virtual reality applications with virtual humans. In: *Annual Review of Cybertherapy and Telemedicine*, vol. 11, p. 33–37. https://doi:10.3233/978-1-61499-282-0-33

Praetorius, Anna-Katharina et al. (2018). Generic dimensions of teaching quality: The German framework of three basic dimensions. In: *ZDM*, vol. 50, p. 407–426. https://doi.org/10.1007/s11858-018-0918-4

Prilop, Christopher Neil, Weber, Kira Elena, & Kleinknecht, Marc (2020). Effects of digital video-based feedback environments on pre-service teachers' feedback competence. In: *Computers in Human Behavior*, vol. 102, p. 120–131. https://doi.org/10.1016/j.chb.2019.08.011

Prilop, Christopher Neil, Weber, Kira Elena, & Kleinknecht, Marc (2021). The role of expert feedback in the development of pre- service teachers' professional vision of classroom management in an online blended learning environment. In: *Teaching and Teacher Education*, vol. 99, no. 103276. https://doi.org/10.1016/j.tate.2020.103276

Protzel, Manuela (2015). Wertschätzung und Persönlichkeitsentwicklung in der Grundschule: Welchen Einfluss Lehrerhaltungen auf Kinder haben, Optimus, Göttingen.

Puhl, Thomas, Tsovaltzi, Dimitra & Weinberger, Armin (2015). A Long-Term View on Learning to Argue in Facebook: The Effects of Group Awareness Tools and Argumentation Scripts. In: Lindwall, O. et al. (ed.), *Exploring the Material Conditions of Learning – The Computer Supported Collaborative Learning (CSCL) Conference 2015*, University of Gothenburg, Gothenburg, Sveden, p. 110–117.

Rahim, M. Afzalur (2001). *Managing conflict in organizations*, third edition, Quorum Books, London.

Rattay, Cathrin & Wensing, Rainer (2011). Fünf Eskalationsstufen von Unterrichtsstörungen und mögliche Gegenmaßnahmen. In: Rattay, C. et al. (eds.), *Unterrichtsstörungen souverän meistern – Das Praxisbuch. Profi-Tipps und Materialien aus der Lehrerfortbildung*, Auer, Donauwörth, p. 59–74.

Rocque, Michael, & Paternoster, Raymond (2011). Understanding the antecedents of the »school-to-jail« link: The relationship between race and school discipline. In: *The Journal of Criminal Law and Criminology*, p. 633–665. https://www.jstor.org/stable/23074048 [Zugriff 28.03.2025]

Ryan, Richard M., & Deci, Edward L. (2017). *Self-determination theory: Basic psychological needs in motivation, development*, and wellness, Guilford, London.

Scheeler, Mary Catherine, Ruhl, Kathy L. & McAfee, James K. (2004). Providing performance feedback to teachers: A review. In: *Teacher education and special education*, vol. 27(4), p. 396–407.

Schneeberger, Tanja et al. (2019). Would you follow my instructions if I was not human? Examining obedience towards virtual agents. In: *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Cambridge, United Kingdom, p. 1–7. https://doi.org/10.1109/ACII.2019.8925481

Schneeberger, Tanja et al. (2019). Can social agents elicit shame as humans do? In: *8th International Conference on Affective Computing and Intelligent Interaction (ACII)*, International Conference on Affective Computation and Intelligent Interaction (ACII-2019), Cambridge, United Kingdom, p. 1–6. https://doi.org/10.1109/ACII.2019.8925482

Schwarz, Gerhard (2014). *Konfliktmanagement: Konflikte erkennen, analysieren, lösen*. Springer, Heidelberg.

Seidel, Tina, Blomberg, Geraldine, & Stürmer, Kathlen (2010). »Observer« – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. In: Klieme, Eckhard, Leutner, Detlev & Kenk, Martina (ed.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes,* Beltz, Weinheim, p. 296–306.

Seufert, Christian et al. (2022). Classroom management competency enhancement for student teachers using a fully immersive virtual classroom. In: *Computers & Education*, vol. 179, no. 104410. https://doi.org/10.1016/j.compedu.2021.104410

Sirois, Fuschia M. (2015). A self-regulation resource model of self-compassion and health behavior intentions in emerging adults. In: *Preventive medicine reports*, vol. 2, 218–222. https://doi.org/10.1016/j.pmedr.2015.03.006

Skiba, Russell J. et al. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. In: *The urban review*, vol. 34, p. 317–342.

Slater, Mel (1999). Measuring presence: A response to the Witmer and Singer presence questionnaire. In: *Presence: teleoperators and virtual environments*, vol. 8(5), p. 560–565.

Smith, Patricia L. & Ragan, Tillman J. (1993). *Instructional design*, Wiley, New York.

Tausch, Reinahrd (2008). Personzentriertes Verhalten von Lehrern in Unterricht und Erziehung. In: Schweer, Martin K. W. (ed.), *Lehrer-Schüler-Interaktion. Inhaltsfelder, Forschungsperspektiven und methodische Zugänge,* Verlag für Sozialwissenschaften, Wiesbaden, p. 155–176.

Kumschick, Irina Rosa et al. (2017). Optimierung einer videobasierten Lerngelegenheit im Problem Based Learning Format durch Cognitive Tools. Eine interventionsstudie mit Lehramtsstudierenden. In: *Zeitschrift für Erziehungswissenschaft*, vol. 20(1), p. 93–113. https://doi.10.1007/s11618-017-0728-4

Townsend, Brenda L. (2000). The disproportionate discipline of African American learners: Reducing school suspensions and expulsions. In: *Exceptional children*, vol. 66(3), p. 381–391.

Tsovaltzi, Dimitra et al. (2017). Leveraging social networking sites for knowledge co-construction: Positive effects of argumentation structure, but premature knowledge consolidation after individual preparation. In: *Learning and Instruction*, vol. 52, p. 161–179. https://doi.org/10.1016/j.learninstruc.2017.06.004

Tsovaltzi, Dimitra et al. (2014). Group awareness support and argumentation scripts for individual preparation of arguments in Facebook. In: *Computers & Education*, vol. 76, p. 108–118. https://doi.org/10.1016/j.compedu.2014.03.012

Weber, Kira Elena et al. (2018). Promoting pre-service teachers' professional vision of classroom management during practical school training: Effects of a structured online-and video-

based self-reflection and feedback intervention. In: *Teaching and Teacher Education*, vol. 76, p. 39–49. https://doi.org/10.1016/j.tate.2018.08.008

Stresssensitives, adaptives Mixed-Reality-Sicherheitstraining – NEXT-Reality: Entwicklung einer Lernplattform zum Training der Selbstrettungsmaßnahmen bei einer Notwasserung eines Helikopters unter physischer und psychischer Belastung

Nicola Drüeke, Anna Larissa Rauschelbach, Thomas Jahn, Andreas Probst, Marvin Künsken, Michael Stöber und Steve Dübel

Zusammenfassung

Psychisch und physisch belastende Situationen können mit immersiven Technologien zu Lehrzwecken realitätsnah, aber in einem geschützten Umfeld trainiert werden. Durch unser Mixed-Reality-Sicherheitstraining mit individuell parametrisierbaren Umgebungseinflüssen und Stressoren sowie der zeitgleichen Bewertung von Vitaldaten werden Personen in einem sicheren Umfeld auf die Notwasserung eines Helikopters vorbereitet. Die virtuelle Umgebung wird durch haptische und aktive Elemente erweitert, wodurch eine immersive Trainingsumgebung geschaffen wird. Während des Trainings werden Vitaldaten erfasst, auf Grundlage derer eine Anpassung der Simulation vorgenommen wird. Das Ziel ist die optimale Vorbereitung der Trainees auf die Stressbelastung des Realtrainings »Helicopter Underwater Escape Training« (HUET) und die Erhöhung der Sicherheit und Handlungskompetenz in gefahrengeneigten Arbeitsumfeldern und Notfallsituationen unter Berücksichtigung ethischer und rechtlicher Fragestellungen.

1. Motivation und Ansatz

Bei der Arbeit im Offshore-Windenergie-Bereich ist das operative Fachpersonal verschiedenen Belastungen, Herausforderungen und Risiken ausgesetzt. Daher ist es von besonderer Bedeutung, ein Bewusstsein für Gefahrensituationen und geeignete Notfallstrategien zu schaffen. Um in kritischen Situationen schnell und

sicher handeln zu können, ist eine umfangreiche Sicherheitsausbildung der Fachkräfte wichtig. Bisherige Trainingsmethoden finden vor allem im realen Umfeld statt. Sie sind in der Regel mit hohem Aufwand und Kosten sowie starken psychischen und physischen Belastungen für die teilnehmenden Personen verbunden, welche nicht selten zum Abbruch des Trainings führen.

Ein alternativer Ansatz ist das Training in einer realitätsnahen, aber geschützten Umgebung. Hierfür bietet sich immersive Virtual Reality (VR) an. Um die Trainees angemessen in eine VR-Umgebung einzubinden und den Lernprozess effektiv zu gestalten, ist jedoch ein hoher Grad an Immersion erforderlich. Der alleinige Einsatz einer VR-Brille reicht hierfür nicht aus. Die Erweiterung der virtuellen Realität um haptische Elemente und zusätzliche Aktorik zur Vermittlung vielfältiger Sinneseindrücke erhöht die Realitätsnähe der Anwendung und ermöglicht ein tieferes Eintauchen in das Lernerlebnis. Allerdings ist die kognitive Belastung der Trainees beim Einsatz solcher Technologien sehr hoch. Neben dem Erlernen und Ausführen der zu trainierenden Handlungsabläufe werden die Trainees mit einer realistisch wirkenden kritischen Gefahrensituation konfrontiert und müssen zusätzlich technische Hürden der VR-Technologie wie z. B. Motion Sickness oder komplexe Interaktionsmetaphern überwinden. Der Erhalt der Handlungsfähigkeit in einer kritischen Gefahrensituation, d. h. unter induziertem Stress, gehört zwar zu den Trainingszielen, gleichzeitig kann aber eine Überforderung der Trainees zum Abbruch des Trainings und damit zum Misserfolg führen. Die Abwägung zwischen größtmöglicher Immersion auf der einen Seite und der Belastung der Trainees auf der anderen Seite ist daher eine große Herausforderung, zumal sie von verschiedenen Variablen wie der Erfahrung der Trainees mit VR-Technologien, der Anfälligkeit für Motion Sickness und der individuellen Belastungsgrenze abhängt.

Das Projekt »NEXT-Reality« verfolgt daher einen adaptiven Ansatz, der es erlaubt, den Grad der Immersion und der kognitiven Beanspruchung durch die VR-Technologie an die jeweilige Belastungsgrenze und individuellen Voraussetzungen anzupassen. Dazu werden die Vitaldaten der Teilnehmenden erfasst und daraus die aktuelle Stressreaktion ermittelt. Diese dient als Grundlage, um die Intensität des Szenarios an die Person anzupassen. Dies kann sowohl durch automatische Anpassungen des Systems als auch manuell durch den Trainer erfolgen. Dadurch können die Trainees Handlungssicherheit in der Gefahrensituation erlangen und die Abläufe unterhalb der Belastungsgrenze trainieren. Durch mehrmaliges Durchlaufen der Trainingssituation kann der Immersionsgrad immer wieder angepasst werden, mit der Zielstellung, das Szenario möglichst realitätsnah zu trainieren und so den Trainingserfolg zu erhöhen.

In NEXT-Reality wird hierfür eine innovative Mixed-Reality-Lernplattform (Virtual Reality und weitere Sinneseindrücke, wie Haptik oder Gleichgewicht)

entwickelt, umgesetzt und evaluiert. Der konkrete Anwendungsfall ist das Trainieren einer Unterwasserevakuierung aus einem Helikopter. Beim sogenannten »Helicopter Underwater Escape Training« werden zertifizierte Rettungsszenarien bei der Notwasserung eines Hubschraubers im Wasser simuliert und trainiert. Dazu werden verschiedene Varianten der Notwasserung und die damit verbundenen Selbst- und Teamrettungsfähigkeiten erlernt. Zielgruppe sind hier alle Personen, die mit einem Hubschrauber zu Notfall- oder technischen Einsätzen transportiert werden, sowie Pilot:innen und Besatzungen. Bei der Evakuierung im Ernstfall sind erhebliche physische und psychische Einflüsse zu berücksichtigen, z. B. Situation kopfüber unter Wasser, Atemnot, Orientierungsprobleme und Panikreaktionen.

Ziel ist es, kritische Arbeitssituationen mithilfe von Mixed Reality (MR) realitätsnah zu trainieren. Dies soll das praktische reale Training im Wasser zunächst nicht ersetzen, jedoch als Vorbereitung dienen und somit das Training optimal ergänzen. Eine Gefahr für Leib und Leben besteht in den virtuellen Trainingsszenarien nicht. Darüber hinaus kann das MR-Training als unterstützendes Verfahren zur Eignungsfeststellung eingesetzt werden, um die Abbrecherquote im praktischen Teil der Ausbildung zu reduzieren. Auf der Lernplattform können sich die Trainees bereits im Vorfeld auf die Stressoren der Trainingssituation vorbereiten. Das Verhalten in komplexen Trainingsszenarien kann in MR-gestützten Szenarien wiederholt und in Feedbackgesprächen reflektiert werden. Individuelle Daten der Trainees bezüglich der Stresseinflüsse und der Handlungsfähigkeit des zu beherrschenden Rettungsablaufs im MR-Szenario werden gesammelt und für weitere Analysen zur Verbesserung der Trainingsmethodik und der VR-Darstellung genutzt. Ethische, soziale, arbeitsrechtliche und datenschutzrechtliche Aspekte sind hierbei besonders zu beachten.

2. Stand der Technik

Der Ausbildung qualifizierten Fachpersonals kommt in Zeiten des Fachkräftemangels eine immer größere Bedeutung zu. Die Anwendung von Virtual Reality (VR) zu Ausbildungszwecken hat dabei aufgrund der technischen Fortschritte in diesem Bereich in den letzten Jahren stark zugenommen. VR bietet eine immersive Erfahrung, die es den Trainees ermöglicht, realistische Szenen zu erleben, ohne das reale Umfeld vor Ort zu haben. Dadurch entstehen gegenüber dem herkömmlichen Training weniger Aufwand und Kosten.

MR ergänzt die VR um weitere Sinneseindrücke, z. B. Haptik und reale Bewegung. Diese Real-Life-Elemente (RL-Elemente) führen zu einer noch realistischeren Darstellung. Auch typische Probleme der VR, wie Motion Sickness, kön-

nen so verringert werden (Lucas et al., 2020: 15–22). Die realitätsnahe Erfahrung führt jedoch auch zu Belastungen der Trainees, welche Überforderung und einen Trainingsabbruch nach sich ziehen können. Im Gegensatz hierzu führt ein ausbleibendes Stressempfinden zu einer fehlenden Einbindung in das Training, was sich ebenfalls negativ auf den Trainingserfolg auswirken kann. Beim Training von Gefahrensituationen werden die Trainees so nicht ausreichend auf das Training unter realen Bedingungen und letztendlich den Ernstfall vorbereitet.

Um dem entgegenzuwirken, können adaptive Techniken eingesetzt werden (Jones & Dechmerowski, 2016: 23–33; Finseth, 2025: 5–20). Adaptive Systeme reagieren dynamisch auf die individuellen Bedürfnisse der Trainees, indem sie die Intensität der Szenarien und den Immersionsgrad in Echtzeit anpassen. Eine Anpassung an das Stresserleben der Trainees wird vorgenommen, um sicherzustellen, dass die Belastung angemessen ist. Dadurch werden Überlastung oder Unterforderung vermieden. Das Stresslevel kann über verschiedene Sensorik festgestellt werden, beispielsweise durch das Tracking des Verhaltens oder der Augenbewegung (Gaggioli et al., 2014: 114–120; Hirt et al., 2020: 5977–5989; Stoeve et al., 2022: 19:1–19:17). Als effektive Indikatoren haben sich physiologische Parameter wie Herzfrequenz und Hautleitwert erwiesen (Acerbi et al., 2017: 31–50; Rahma et al., 2022: 155–162; Narciso et al., 2024: 14). Auf diese Weise können Vitaldaten zur Steuerung einer Trainingsszene eingesetzt werden (Finseth et al., 2021: 1541–1545).

Zur Bestimmung des Stresses auf Basis von Vitalparametern kommen verschiedene Methoden des Maschinellen Lernens (ML) infrage (Sánchez Reolid et al., 2020: 1–29; Hosseini et al., 2022: 2619–2623). Dabei stellen sich Support Vector Machines (SVM) als ein geeignetes ML-Verfahren zur binären Klassifikation des Stresses heraus (Sánchez Reolid et al., 2020: 1–29). In diesem Projekt wurde eine adaptive MR-Umgebung entwickelt, welche sich auf Basis einer kontinuierlichen Vitaldatenmessung an das Stressniveau der Trainees anpasst. Aufgrund der Verwendung verschiedener VR- und RL-Komponenten sowie der Erhebung sensibler medizinischer Daten wurden in NEXT-Reality rechtliche und ethische Fragestellungen besonders berücksichtigt.

3. Rechtliche und ethische Fragestellungen

Rechtliche Fragen im Projekt NEXT-Reality können sich zum einen aus der zivilrechtlichen Haftung für Personen- und Sachschäden ergeben, die bei der Nutzung der MR-Trainingsumgebung entstehen. Schäden, deren Ersatz zu erörtern ist, können sich sowohl aus Fehlfunktionen der RL- als auch der VR-Komponenten der MR-Trainingsumgebung ergeben. Diese Schäden können

insbesondere bei den Nutzer:innen, in besonderen Fällen auch bei unbeteiligten Dritten, eintreten. Insoweit geht es einerseits um die vertragliche Mängelhaftung der Verkäufer oder Hersteller der MR-Trainingsumgebung bzw. ihrer einzelnen Komponenten, andrerseits um die gesetzliche Schadensersatzhaftung nach den Vorschriften des Bürgerlichen Gesetzbuchs (BGB) über unerlaubte Handlungen (§§ 823 ff. BGB) und nach dem Produkthaftungsgesetz.

Zum anderen sind die mit den Haftungsfragen eng verbundenen Fragen zu erörtern, wie die MR-Trainingsumgebung und ihre Komponenten im Einzelnen ausgestaltet, überwacht und gewartet werden müssen, um eine zivilrechtliche Haftung möglichst zu vermeiden. Zudem sind Fragen der vertraglichen Versicherbarkeit möglicher Schäden sowie des Eingreifens der gesetzlichen Unfallversicherung zu untersuchen. Als eine weitere, wesentliche Frage im Rahmen der Erprobung mit Proband:innen stellte sich die ethische Bewertung dar. Ethische Bedenken können durch mehrere Aspekte hervorgerufen werden: So werden im Rahmen der Untersuchung sensible Daten der Proband:innen erhoben, sowohl in Gestalt der während des simulierten Trainings überprüften körperlichen Messwerte als auch mittels der von Proband:innen im Anschluss an das Training auszufüllenden Fragebögen. Daneben werden die Proband:innen einer psychischen und physischen Belastungssituation ausgesetzt. Schließlich ist ein potenzielles Abhängigkeitsverhältnis der Proband:innen zur Studienleitung nicht ausgeschlossen, da diese im beruflichen Umfeld der Konsortialpartner:innen kontaktiert werden.

Diesen Bedenken wird mit folgenden Maßnahmen begegnet: Die Teilnahme ist freiwillig. Teilnehmen dürfen nur volljährige Personen ohne psychische oder neurologische Vorerkrankungen. Vor Beginn der Teilnahme erfolgt eine umfangreiche Instruktion. Eine Unterbrechung oder ein Abbruch sind jederzeit möglich. Die Versuchsleitung ist bei dem gesamten Vorgang präsent und steht damit nicht nur als Ansprechpartnerin zur Verfügung, sondern kann auch selbst den Vorgang überwachen und, wenn nötig, einschreiten. Daneben erfolgen die Datenspeicherung und Auswertung jeweils anonym. Ausschließlich die Studienleitung hat Zugriff auf die pseudonymisierten Daten, um einem potenziellen Abhängigkeitsverhältnis Rechnung zu tragen.

4. Entwicklung der Trainingsumgebung

Die Entwicklung der Trainingsumgebung zur Vermittlung von Selbstrettungsmaßnahmen hat den Fokus, ein immersives Lernerlebnis zu schaffen. Der Begriff Immersion beschreibt das Phänomen des möglichst umfassenden Eintauchens

eines Probanden oder einer Probandin in die MR-Umgebung mittels Stimulation der Sinnesorgane durch verschiedene Ausgabegeräte (Doerner et al., 2022: 14).

Um einen hohen Grad an Immersion zu erreichen, wurde die Trainingsumgebung nach den technischen Eigenschaften der Ausgabegeräte, auf der Immersion nach Slater und Wilbur (1997: 603–616) basiert, gestaltet. Bei den technischen Eigenschaften der Ausgabegeräte handelt es sich um die Isolierung der Proband:innen von der Umgebung, das Stimulieren mehrerer Sinne, ein vollständig umgebendes Blickfeld und eine möglichst realistische Darstellung (Slater & Wilbur, 1997: 603–616; Doerner et al., 2022: 14). In Anbetracht dieser Eigenschaften wurde eine virtuelle Trainingsumgebung umgesetzt, die die Sinne der Proband:innen auf visuelle, akustische, haptische und vestibuläre Weise stimuliert.

Um eine realistische und immersive VR-Umgebung für das Training umzusetzen, sind eine Vielzahl von technischen Details und Implementierungsschritten notwendig. Diese umfassen die Wahl der richtigen Soft- und Hardwarelösungen, die Modellierung von Objekten und Szenen, das Simulieren von realistischen, physikalischen Effekten und die Steuerung der Immersionsparameter in Echtzeit. Aufgrund ihrer leistungsstarken Grafik- und Simulationsfähigkeiten wurde Unreal Engine (Das leistungsstärkste Werkzeug für 3D-Echtzeit-Entwicklung, 2024) als Entwicklungsplattform gewählt. Die Engine bietet ein breites Spektrum an Funktionen, die speziell für VR-Anwendungen geeignet sind, darunter die Möglichkeit, hochdetaillierte 3D-Modelle in Echtzeit zu rendern, komplexe Physiksimulationen zu erstellen und in Echtzeit auf Nutzereingaben zu reagieren.

Die Simulation verfügt über drei Szenarien: das Boarding, der Rundflug und das Notfallszenario. Das Boardingszenario vermittelt mittels eines virtuellen Trainers die Selbstrettungsmaßnahmen zur Evakuierung aus einem im Wasser havarierten Helikopter. Der virtuelle Trainer ist menschlich gestaltet und macht die Bewegungsabläufe der Selbstrettung vor; dieses Trainingsszenario können die Proband:innen beliebig oft durchlaufen. Die Bewegungsabläufe zur Selbstrettung wurden durch Motion-Capture-Daten animiert. Hierbei wurde die Faceware-Motion-Capture-Technologie verwendet, um die Gesichtsausdrücke des Trainers realistisch zu gestalten.

Auf das Boardingszenario folgt der Rundflug, bei dem der Helikopter über einen Offshore-Windpark fliegt. Die Proband:innen können sich an das Fluggefühl gewöhnen und mit virtuellen und haptischen Elementen interagieren. Daraufhin folgt das Notfallszenario, in dem der Helikopter eine Notwasserung durchführt und kentert. Die Proband:innen sollen die im Boardingszenario erlernten Bewegungsabläufe durchführen und sich selbst aus dem abgestürzten Helikopter retten. Besonders im Notfallszenario wurde die Sicht unter Wasser durch spezielle

Grafikeffekte eingeschränkt. Die Sichtweite wird stark reduziert, um die Bedrohlichkeit und Unsicherheit des Szenarios zu erhöhen.

Die Szenen in der VR-Umgebung bestehen aus einer Vielzahl von Objekten und Umgebungen, die detailliert modelliert und texturiert wurden. Der Helikopter wurde von Grund auf modelliert, einschließlich der Sitze, Gurte, Steuerungsinstrumente und Fenster. Speziell der Fenstermechanismus wurde so modelliert, dass die Proband:innen im Notfallszenario interaktiv das Fenster herausstoßen können.

Das Wasser reagiert dynamisch auf die Bewegung des Helikopters und der Proband:innen, was besonders im Notfallszenario von großer Bedeutung ist. Wir nutzen spezialisierte Algorithmen zur Simulation von großen Wasserflächen, wie etwa einem Ozean, die es uns ermöglichen, die Stärke und Höhe der Wellen anhand natürlicher Parameter, wie Windgeschwindigkeit oder Entfernung zur Küste, realistisch abzubilden. Abhängig von der Windintensität entsteht zudem Schaum auf der Wasseroberfläche. Für den Windpark wurden realistische Modelle von Windturbinen erstellt. Um reale Wetterphänomene wie volumetrische Wolken, Regen, Schnee, Stürme und Blitze zu simulieren, wurden spezielle vorgefertigte Elemente verwendet. Darüber hinaus werden sämtliche Audioeffekte wie Wind, Regen und Donner synchron integriert. Der Einsatz von Audio spielt eine entscheidende Rolle bei der Verstärkung der Immersion und trägt maßgeblich zur Intensivierung des Erlebnisses bei. Zudem sind Wetterbedingungen und Ozeansimulation programmatisch miteinander verknüpft, sodass stärkere Winde automatisch zu höheren Wellen führen, was eine nahtlose und realistische Darstellung der Umgebung gewährleistet.

Um die Immersion weiter zu erhöhen, wurde der Sound-Design-Prozess stark in die Gestaltung des Erlebnisses integriert. Der Helikoptergeräuschpegel und die Umgebungsgeräusche passen sich dynamisch an die Wetterbedingungen und die Fluggeschwindigkeit an. Dies wird durch das Audio-Component-System der Unreal Engine ermöglicht, welches 3D-Sound lokalisiert und in Echtzeit an die Position und Ausrichtung der Proband:innen anpasst.

Die Simulationssoftware kann von den Proband:innen in der VR-Brille Quest 3 von Meta (Meta Quest 3, 2024) erlebt werden. Die Einbindung des Handtrackings der Quest 3 ermöglicht es den Proband:innen, sich in den Szenarien zu orientieren und mit virtuellen Gegenständen zu interagieren. Zur Simulierung der Flugbewegungen des Helikopters wurde der Seatmover (Bewegungsplattform) YAWII von YAWVR (User Manual for Yaw2, 2024) verwendet, der Bewegungen aus der Simulationssoftware in Echtzeit auf drei Achsen (Kippen, Drehen, Rollen) abbilden kann. Gesichert sind die Proband:innen mit einem 4-Punkt Sicherheitsgurt, dieser entspricht dem Gurt, mit dem Insassen eines Helikopters gesichert werden. Die Flugvibrationen werden über die Vibrati-

onsweste von Woojer (VEST 3, 2024) auf die Proband:innen übertragen. Des Weiteren verfügt der Demonstrator (siehe Abbildung 1) über einen Nachbau eines Helikopterfensters, der als haptisches Element die Simulation ergänzt. Im Rahmen der Selbstrettung wird das Helikopterfenster virtuell und als haptisches Element des Demonstrators herausgestoßen. Weitere Rettungsgeräte, wie ein Atemgerät und ein Mundstück, sind als virtuelle sowie haptische Elemente Teil der Trainingsumgebung.



Abbildung 1: Demonstratoraufbau mit Trainee
Quelle: Eigene Aufnahme, © 2025 OffTEC

Die Steuerung der immersiven Parameter wird über das Open-Sound-Control-Protokoll (OSC) realisiert. OSC ermöglicht es, in Echtzeit diverse Parameter der Umgebung anzupassen, darunter: Wettereffekte, Tageszeit, Windgeschwindigkeit, Fluggeschwindigkeit, Steuerung der Rotationsachsen des Demonstrators, Szenarioauswahl, Ausrichtung des VR-Headsets und Anpassung der grafischen Qualität.

Zur Erfassung der subjektiven Eindrücke der Proband:innen wurden verschiedene Fragebögen entwickelt, die vor, während und nach dem VR-Training von den Proband:innen händisch oder virtuell beantwortet werden. Vor Beginn

des VR-Trainings werden mittels eines Basisfragebogens die Vorerfahrungen zum HUET-Training sowie zu VR-Anwendungen im Allgemeinen erfragt. Während des VR-Trainings wird das subjektive Stressempfinden mehrfach durch einen virtuellen Fragebogen erhoben; nach dem Training wird dies ebenfalls in Form eines schriftlichen Fragebogens erfasst. Weitere subjektive Eindrücke wie Immersion, Motion Sickness und Usability werden im Rahmen von drei separaten Fragebögen nach Abschluss des VR-Trainings aufgenommen. Die erfassten subjektiven Eindrücke dienen als Kontrollvariablen sowie als Feedback für das System.

5. Entwicklung der Vitaldatenmessung

Das MR-Übungsszenario wurde als Ergänzung zum Realtraining entwickelt, um die Schwelle für Erstteilnehmende herabzusetzen, die Abbrecherquote zu reduzieren und den Erfolg der Ausbildung zu steigern. Dennoch ist auch die MR durch den hohen Grad der Immersion und die Trainingsanforderungen an die Trainees herausfordernd. Aus diesem Grund ist das System adaptiv gestaltet, sodass es individuell für alle Trainees parametrisierbar ist. Auf diese Weise kann eine wenig belastende Version für eine Person mit hohem Stressempfinden oder, bei keinen MR-Vorerfahrungen, auch zur Gewöhnung an die virtuelle Umgebung genutzt werden, wohingegen, wenn die Bewegungsabläufe und Selbstrettungsmaßnahmen bereits gefestigt sind, das Szenario und die Immersion verschärft werden können.

Zur Personalisierung des Trainings wird das individuelle Stresserleben der Trainees in Echtzeit beurteilt. Dafür werden Vitalparameter erfasst und zur Stressklassifizierung herangezogen. Bei den für die Messung der Vitaldaten verwendeten Sensoren handelt es sich um Shimmer3 GSR+ Units (Shimmer3 GSR+ Unit, 2024; siehe Abbildung 2). Diese zeichnen sowohl ein Photoplethysmogramm (PPG) als auch die elektrodermale Aktivität (EDA, auch Galvanic Skin Response [GSR]) auf. Mithilfe der PPG werden Volumenänderungen im Blutgefäßsystem gemessen, woraus sich beispielsweise die Pulsrate ableiten lässt. Die elektrodermale Aktivität beschreibt die Leitfähigkeit der Haut. Aus diesen Daten werden relevante Features gewonnen, mithilfe derer der Stress automatisiert klassifiziert werden kann.

Zum Training eines Stressklassifikators werden annotierte Ground-Truth-Daten benötigt. Um diese zu gewinnen, wurden Datenerhebungen, bestehend aus einer Entspannungsphase und einer Stressphase, durchgeführt. In der ersten Phase saßen die Proband:innen in Ruhe und hörten für drei Minuten Entspannungsmusik. Sie wurden gebeten, sich zu entspannen. Zur Stressinduk-



Abbildung 2: Shimmer3-GSR+-Unit zur Vitaldatenmessung; Zeige- und Mittelfinger: Hautleitwert, Ringfinger: PPG
Quelle: Eigene Aufnahme, © 2025 OffTEC

tion im zweiten Teil wurde ein Kaltwassertest (Cold Pressor Test [CPT]) (Hines & Brown, 1932: 332–335), angewandt. Dabei wird die rechte Hand so lange wie möglich, aber maximal drei Minuten, in stark heruntergekühltes (ca. 4°C) Wasser gehalten. Während des gesamten Versuches wurden die PPG- und GSR-Signale mithilfe eines Shimmer3-Sensors an der linken Hand erfasst.

Die erhobenen Sensordaten wurden entsprechend der Phasen annotiert und in sich überlappende Zeitfenster geteilt. Pro Fenster wurden sowohl aus dem PPG- als auch aus dem GSR-Signal Features extrahiert und mit diesen der binäre Klassifikator trainiert. Als Machine-Learning-Methode wird eine Support-Vector-Machine (SVM) (Cortes & Vapnik, 1995: 273–297) eingesetzt. Diese besitzt eine große Generalisierungsfähigkeit und ist daher zur Lösung realer Probleme unter Einbeziehung eines kleinen Datensatzes geeignet.

Während der Ausbildung wird die trainierte SVM zur Stresserkennung in Echtzeit genutzt. Dabei werden die Daten von den Shimmer3-Sensoren gestreamt, gepuffert und fensterweise abgearbeitet, sodass die Verzögerung zwischen Auftreten des Stresses und der Reaktion darauf möglichst niedrig gehalten werden kann. Das Ergebnis der Stressklassifikation wird dem Trainer angezeigt, dieser kann daraufhin Änderungen in der VR-Umgebung vornehmen. Eine Automatisierung ist ebenfalls möglich. Es wurden drei Adaptionsstufen vordefiniert. Dies ermöglicht eine schnelle Anpassung an das Stressniveau des Trainees während des Trainings.

Während der Erprobungsphase des Systems wird zusätzlich zum physiologisch bestimmten Stress das subjektive Empfinden der Proband:innen erfasst. Dazu werden im Verlaufe des Trainings zu verschiedenen Zeitpunkten Befragungen durchgeführt. Während des Trainings wird das Befinden des Trainees mehrfach abgefragt. Hierfür wird das Self-Assessment Manikin (SAM) (Bradley & Lang, 1994: 49–59) verwendet. Es findet vor dem Rundflug-, nach dem Rundflug- und nach dem Notfallszenario statt. Die Abfragen sind in die VR eingebunden und können auf diese Weise von den Proband:innen direkt beantwortet werden. Die Antworten werden in maschinenlesbarer Form gespeichert, sodass diese automatisiert ausgewertet werden können. Direkt nach dem VR-Training wird eine abgewandelte Form des NASA Taskload Index (NASA-TLX) (Hart & Staveland, 1988: 139–183) ausgehändigt. Dieser ist detaillierter als SAM und erhebt rückblickend das subjektive Stressempfinden der Proband:innen im Training.

6. Fazit und Ausblick

Im Projekt NEXT-Reality wird der Prototyp einer immersiven Trainingsumgebung für die Schulung von Selbstrettungsmaßnahmen aus einem im Wasser ha-

varierten Helikopter umgesetzt. Für eine virtuelle, aber realitätsnahe Trainingssimulation wurden moderne Technologien wie leistungsstarke Hardware, Motion Capture und physikalische Simulationen eingesetzt. Die Kombination aus visueller, akustischer und haptischer Aktorik sowie die Einbeziehung des Gleichgewichtssinns erlauben eine hohe Immersion und hierdurch ein effektives Training der Notfallsituation.

Um eine Überforderung der Trainees zu vermeiden, wurde ein adaptives System entwickelt, welches verschiedene Belastungsgrade ermöglicht und diese dynamisch durch Änderung der immersiven MR-Parameter anpassen kann. Grundlage hierfür ist die aktuelle Belastungssituation der Trainees. Diese wird anhand von in Echtzeit aufgenommenen Vitaldaten und des hieraus mittels Machine-Learning-Algorithmen bestimmten aktuellen Stresslevels abgeleitet. Einfluss hierauf hat zum einen ungewollter, durch die Plattform induzierter Stress, welcher beispielsweise durch Motion Sickness oder umständliche VR-Interaktionen hervorgerufen wird, und zum anderen Stress, der durch die Notfallsituation und beim Training der Handlungsabläufe entsteht. Ziel der dynamischen Anpassung ist es, die Belastung der Trainees in einem optimalen Bereich zu halten. Konkret wird die Aktorik, welche den Seh-, Hör-, Tast- und Gleichgewichtsinn beeinflusst, in seiner Intensität heruntergeregelt, wenn die Belastung für die Trainees zu hoch wird. So kann beispielsweise die Bewegungsplattform in ihrer Beschleunigung eingeschränkt oder die Sicht- und Wettersituation in der VR-Welt einfacher eingestellt werden.

Momentan findet eine umfangreiche Evaluation statt. Es werden unterschiedliche Frage- und Feedbackbögen eingesetzt, um das subjektive Empfinden der Proband:innen sowie den Lernerfolg festzuhalten und auszuwerten. Bereits während der Entwicklung des Demonstrators wurde das MR-Szenario von Testpersonen mit verschiedenem Kenntnisstand, sowohl in Bezug auf VR-Vorkenntnisse als auch Vertrautheit mit den Selbstrettungsabläufen des Realtrainings, durchlaufen. Das Training wurde von den verschiedenen Gruppen positiv aufgenommen. Insbesondere sieht das Fachpersonal vielversprechende Einsatzmöglichkeiten der MR-Trainingsumgebung als ergänzende Maßnahme zur bisherigen Ausbildung. Eine Erkenntnis aus den Tests ist, dass der sichere Umgang mit und das intuitive Zurechtfinden in der VR-Welt den Trainingserfolg maßgeblich fördern. Basierend auf diesem Feedback wurden Möglichkeiten entwickelt, um den Einstieg in die VR-Umgebung zu erleichtern. So können z. B. die Interaktionselemente in der VR anfangs farbig hervorgehoben werden. Bezüglich der möglichen Interaktion mit den realen Elementen des Demonstrators kam es vereinzelt zu Unsicherheiten bei den Trainees. So fehlte beispielsweise das Bewusstsein, dass das Fenster physisch herausgestoßen werden kann. Hier

wurde das Trainingsdesign entsprechend überarbeitet, seitdem werden die Teilnehmenden explizit auf die Interaktionsmöglichkeiten hingewiesen.

In Zukunft sollen die Möglichkeiten der Integration des MR-Trainings in die praktische Ausbildung sowie die Übertragbarkeit des Konzeptes auf andere Trainingsinhalte, wie z. B. das Verhalten in Notsituationen bei Feuer, überprüft werden. Darüber hinaus soll erforscht werden, wie eine feinere und individuellere Unterteilung der anpassbaren Immersionsstufen erfolgen kann. Für eine robustere Stressklassifizierung wird außerdem die Erweiterung des Systems um zusätzliche Sensorik, wie beispielsweise Eye Tracking, untersucht werden.

Förderhinweis

Das Projekt NEXT-Reality wurde vom Bundesministerium für Bildung und Forschung (BMBF)[1] gefördert.

Literatur

Acerbi, Giorgia (2017). A Wearable System for Stress Detection Through Physiological Data Analysis. In: Fiorini, Laura et al. (Hg.), *Ambient Assisted Living*, Springer International Publishing (Lecture Notes in Electrical Engineering), Basel, S. 31–50. https://doi.org/10.1007/978-3-319-54283-6_3.

Bradley, Margaret M., Lang, Peter J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. In: *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 25(1), Elsevier Ltd., Amsterdam, S. 49–59. https://doi.org/10.1016/0005-7916(94)90063-9.

Cortes, Corinna, Vapnik, Vladimir (1995). Support-vector networks. In: Hendrik Blockeel, *Machine Learning*, Vol. 20(3), IOP Publishing, S. 273–297. https://doi.org/10.1007/BF00994018.

Epic Games (2024). *Unreal Engine. Das leistungsstärkste Werkzeug für 3D-Echtzeit-Entwicklung*, Epic Games, Cary, North Carolina, USA. https://www.unrealengine.com/de/home [Zugriff 07.10.2024]

Doerner, Ralf et al. (Hg.) (2022). *Virtual and Augmented Reality (VR/AR). Foundations and Methods of Extended Realities (XR)*, Springer International Publishing, Basel. https://doi.org/10.1007/978-3-030-79062-2

Finseth, Tor et al. (2021). The effectiveness of adaptive training for stress inoculation in a simulated astronaut task. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications Sage CA, Los Angeles, S. 1541–1545. https://doi.org/10.1177/1071181321651241

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Finseth, Tor et al. (2025). Virtual Reality Adaptive Training for Personalized Stress Inoculation. In: *Human Factors. The Journal of the Human Factors and Ergonomics Society*, Vol. 67(1), Sage Journals, New York, S. 5–20. https://doi.org/10.1177/00187208241241968.

Gaggioli, Andrea (2014). A decision support system for real-time stress detection during virtual reality exposure. In: *Studies in Health Technology and Informatics*, Vol. 196, IOS Press, Amsterdam, NL, S. 114–120. https://doi.org/10.3233/978-1-61499-375-9-114

Hart, Sandra G., Staveland, Lowell E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In: Hancock, P.A., Meshkati, N. (Hg.), *Human Mental Workload*, Amsterdam, NL, S. 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Hines, Edgar Alphonso, Brown, George Elgie (1932). A standard stimulus for measuring vasomotor reactions: its application in the study of hypertension. In: *Proceedings of the Staff Meetings of the Mayo Clinic*, Vol. 7, Rochester, Minnesota, S. 332–335.

Hirt, Christian, Eckard, Marcel & Kunz, Andreas (2020). Stress generation and non-intrusive measurement in virtual environments using eye tracking. In: Andreu-Perez, Javier *Journal of Ambient Intelligence and Humanized Computing*, Vol. 11(12), Springer Nature, Berlin, Heidelberg, S. 5977–5989. https://doi.org/10.1007/s12652-020-01845-y

Hosseini, Elahe et al. (2022). A Low Cost EDA-based Stress Detection Using Machine Learning. In: 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas: IEEE, S. 2619–2623. https://doi.org/10.1109/BIBM55620.2022.9995093.

Jones, David & Dechmerowski, Sara (2016). Measuring Stress in an Augmented Training Environment: Approaches and Applications. In: Schmorrow, Dylan D., Fidopiastis, Cali M. (Hg.), *Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience*, 10th International Conference, AC 2016, Toronto, Canada, Springer International Publishing, Basel, S. 23–33. https://doi.org/10.1007/978-3-319-39952-2_3

Lucas, Guillaume et al. (2020). A simulation sickness study on a driving simulator equipped with a vibration platform. In: *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 68, Elsevier Ltd., Amsterdam, S. 15–22. https://doi.org/10.1016/j.trf.2019.11.011

Meta (2024). *Meta Quest 3*. Meta Platforms Ireland Ltd., Dublin. https://www.meta.com/de/quest/quest-3/ [Zugriff 07.10.2024]

Narciso, David et al. (2024). Assessing the perceptual equivalence of a firefighting training exercise across virtual and real environments. In: *Virtual Reality*, Vol. 28(1), Springer Nature, Berlin, Heidelberg, S. 14. https://doi.org/10.1007/s10055-023-00917-6

Rahma, Osmalina Nur et al. (2022). Electrodermal Activity for Measuring Cognitive and Emotional Stress Level. In: *Journal of Medical Signals and Sensors*, Vol. 12(2), Medical Image and Signal Processing Research Center, Wolters Kluwer N.V. Alphen aan den Rihn, S. 155–162. https://doi.org/10.4103/jmss.JMSS_78_20

Sánchez Reolid, Roberto et al. (2020). Machine Learning for Stress Detection from Electrodermal Activity: A Scoping Review. In: *Computer Science Medicine*. https://doi.org/10.20944/preprints202011.0043.v1

Shimmer Discovery in Motion (2024). *Shimmer3 GSR+ Unit*. Shimmer Wearable Sensor Technology. https://shimmersensing.com/product/shimmer3-gsr-unit/ [Zugriff 07.10.2024]

Slater, Mel, Wilbur, Sylvia (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. In: *Teleoperators and Vir-*

tual Environments, Vol. 6(6), MIT Press, Cambridge, S. 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Stoeve, Maik et al. (2022). Eye Tracking-Based Stress Classification of Athletes in Virtual Reality. In: *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, Vol. 5(2) No. 19, Association for Computing Machinery, New York, S. 1–17. https://doi.org/10.1145/3530796

YAW VR (2024). *User Manual for Yaw2*. Yaw VR Ltd., Tatabanya, Ungarn. https://www.yawvr.com/help-yaw2 [04.10.2024]

Woojer (2024). *VEST 3*. Epic Games Inc., Cary, USA. https://www.woojer.com/products/vest-3 [07.10.2024]

Understanding Neurodiverse Cognitive-Affective Functioning in a Multimodal Technology-Mediated Framework: Ethical Guidelines and Human-Centered Design Implications – UFO

Laura Bareiß, Laura Wansitler, Alina Schwarten and Maria Wirzberger

Abstract

The perspective of neurodiversity emphasizes societal benefits from recognizing and developing the strengths of cognitive differences in perception, information processing, and subsequent actions, particularly in the workplace. With the goal of promoting a mutual awareness and understanding for both neurotypical and neurodivergent patterns of cognitive-affective functioning, the chapter describes a novel social technology that builds cognitive and socio-emotional competences. While combining multimodal wearable technologies for sensing, interaction, and feedback, the framework employs a part-task training approach to build core cognitive abilities, which subsequently transfer to realistic application scenarios. Aligned with vested evidence on user-centered design and established bioethical principles of non-maleficence, beneficence, respect for autonomy, and justice, fundamental requirements and guidelines for responsibly developing assistive technology for neurodiverse target groups are derived.

1. Boarding the UFO journey: Understanding and facilitating neurodiverse cognitive-affective functioning with a multimodal user-centered social technology

Neurodiversity is not a deficit but an asset – a claim that is fundamental to the project *Use of neUrophysiological interFaces and tactile-assisted virtual reality to promote*

occupational inclusiOn, for short, *UFO*. Tying in with the vision of the related funding scheme »Interactive systems in virtual and real spaces – Innovative technologies for the digital society«, established by the German Federal Ministry of Education and Research, the project explores novel ways of extended tactical perception that aim to transfer cognitive-affective user states into tangible experiences. The resulting virtual training program centers around a protected UFO in outer space, employing both the building of foundational skills in executive functions and avenues to real-world transfer. The UFO training approach thereby connects two important goals: On the one hand, it is designed to empower individuals exerting perceptual characteristics beyond the cognitive majority as, for instance, evident in autistic individuals. On the other hand, it strives to sensitize the broader society for the diverse variety of cognitive demands and sensitivities, thus promoting advocacy and facilitating social change.

Taken together, the emerging technology unites essential qualities: Being a multimodal system, it combines visual, auditory, and haptic sensory channels. Following a participatory user-centered design approach, the entire process of technology design, implementation, and evaluation iteratively solicits relevant stakeholder perspectives. As a social technology, the system follows the purpose of improving interpersonal communication qualities and raising awareness. By putting the intended target groups at the center of conceptual and technical development, technological advances emerging from the project hold the unique position of fostering inclusion by promoting understanding and social acceptance for cognitive diversity at the workplace, enhancing important potentials for individual and societal growth.

The chapter will start by introducing evidence of how incorporating human cognitive diversity brought forward core achievements of modern societies, followed by a conceptual introduction of the paradigm of neurodiversity, its origins in the neurodiversity movement as well as implications for leveraging emerging potentials in the workplace. After summarizing the core rationale and essential features of the technological advances developed in the UFO project, building on the state of research in the multimedia learning and cognitive training domains, core design guidelines for user-centered social technologies like the one developed are derived. Finally, the chapter provides reflections on the ethical, legal, and social implications of implementing innovative user-centered technological advances for social purposes and bringing them to use, thereby emphasizing an incremental approach of large-scale technology adoption. The final conclusion summarizes the scope of the chapter and takes a visionary look at future societies valuing cognitive diversity.

2. Neurodiverse origins in human evolution: Early gains from including different perspectives

Looking back into the early history of humanity reveals that societal progress only developed due to human diversity: Taking autistic thinking as an example, Spikins (2009: 181–191) emphasizes how modern human behavior emerged from incorporating different minds into society, thereby leveraging specified individual talents and assigning dedicated roles. Building on archaeological evidence, she concluded that characteristics attributed to neurodivergent rather than neurotypical patterns of thought and action, such as an enhanced focus on detail, concern for rules, and understanding of physical and mechanical systems, benefitted innovations in technology tailored to specific purposes, standardized rules for tool use, and enhanced tool efficiency, complexity, and precision (e. g., Bar-Yosef, 2002: 379–381; Mellars, 2005: 20–25). Likewise, unique memory capacities, a desire for isolation, and an understanding of natural systems could have provided the foundations for more efficient exploitation of resources, abilities to creatively deal with environmental changes, and geographical population expansion, triggered by mapping unfamiliar lands due to exploration (e.g., Finlayson, 2004: 94–128; Mellars, 2006: 933 f.). Venturing off to new shores and fearlessly exploring possibilities beyond the existing realm has indeed provided the foundations for economic, cultural, and culinary wealth – achievements still valued in modern days.

Although evolutionary benefits provide vested explanations for the emergence of cognitive differences, many neurodivergent individuals face a negative public view on exerted cognitive and behavioral characteristics (Botha et al., 2022: 428 f.). Upon closer inspection, rather than putting stronger emphasis on arising strengths (Wright et al., 2020: 2–4), medical models and related diagnostic criteria focus on deficits underneath functional impairments in daily life activities (American Psychiatric Association, DSM-5 Task Force, 2013: 31 f.; Kapp et al., 2013: 59 f.). Consequently, individuals cognitively deviating from the majority of society, often referred to as neurominorities, meet a lack of understanding and neglect of individual needs in daily life contexts, struggle to thrive on their differences, and experience a feeling of not belonging (e. g., Kapp et al., 2013: 59). Such dynamics are salient especially in the workplace (e.g., Humphrey & Lewis, 2008: 42 f.; Huws & Jones, 2008: 104–106; Portway & Johnson, 2005: 81 f.). Furthermore, the rather biased public opinion is reinforced by stereotypes presented in the media and by cultural perceptions and misconceptions about neurodivergent individuals (Wright et al., 2020: 3).

3. The strengths of cognitive differences: Neurodiversity at work

By neglecting the fact that, despite shared similarities, each individual operates on a diverse set of cognitive patterns, society faces a loss of unique problem-solving perspectives and creative potential relevant to the modern economy. Consequently, the urgent demand to counteract existing societal patterns of excluding different minds arises (Spikins, 2009: 194 f.; Wright et al., 2020: 4). To inspire society to rethink the potential of cognitive differences, the neurodiversity movement advocates for fundamental changes of perspective: Society should refrain from merely recognizing deficits attributed to neurominorities but perceive differences in favor of society as untapped, essentially valuable benefits, especially in work-related contexts (Doyle, 2020: 109).

Neurodiversity as a paradigm is defined and discussed from multiple perspectives, such as neuroscience, psychology, sociology, anthropology, and education. Conceptually, neurodiversity considers cognitive differences to be naturally occurring variations in perception, information processing, and resulting actions (Armstrong, 2010: 3; Doyle, 2020: 108; LeFevre-Levy et al., 2023: 1). This view has been driven by bottom-up initiatives of neurominority interest groups, in particular, individuals labeled as autistic who refused to be stigmatized as *disabled*, unlike the rather dominating medical perspective, but opted to be neutrally described as *different* (Armstrong, 2010: 8; Ortega, 2009: 425; Wright et al., 2020: 3 f.). Indeed, autistic people may face challenges related to time management, concentration, focusing, multitasking, productivity, or social connection and communication with above average frequency (Doyle, 2020: 116; Howlin et al., 2005: 545; Tonizzi et al., 2022: 3 f.; Wehman et al., 2016: 62). However, unique strengths such as enhanced memory abilities (Meilleur et al., 2015: 1359), strong attention to detail (Frith & Happé, 2005: R789; Happé, 2018: 280; LeFevre-Levy et al., 2023: 7), or innovative thinking (Doyle, 2020: 135) contribute skills of essence to precise and efficient working outcomes.

The neurodiversity paradigm became a popular concept in the working environment (Doyle, 2020: 109), even though it faces certain challenges. More precisely, besides lacking adequate reception and consistent definitions, it provides little guidance for implementing concrete support measures at work. Nevertheless, the resulting change of perspectives is reflected, among others, in the promotion of targeted inclusion programs by leading companies such as SAP or Microsoft, which have already reformed their hiring processes to recruit neurodivergent individuals (Austin & Pisano, 2017: 2). Especially skills such as analytical thinking, complex problem-solving, strong attention to detail, innovation, and creativity are crucial for the modern workforce, hence, individuals possessing en-

hanced scopes in these qualities may be particularly in demand (LeFevre-Levy et al., 2023: 6; Walkowiak, 2021: 3).

4. A user-aligned social technology to support cognitive-affective competences

Tying in with the change of perspectives proposed by the neurodiversity paradigm, efforts in research and practice should focus on demands and challenges emerging from the everyday working lives of neurominorities, to derive best practices for professional inclusion (Doyle, 2020: 110 f.). Thereby, benefits arise from dedicated training programs and assistive technology, which hold value in both building defined and potentially lacking skills as well as mutual awareness for naturally occurring individual differences (Boyd et al., 2018: 105 f.; Motti, 2019: 1; Walkowiak, 2021: 1 f.). Leveraging related potentials, a user-aligned technology-mediated framework is introduced, dedicated to supporting both neurodivergent and neurotypical adults in building vital cognitive-affective competences for navigating daily life requirements in both private and professional spheres. Thereby, two avenues of potential use cases arise, related to both supporting neurodivergent individuals in navigating the neurotypical world and building awareness in neurotypical individuals for the manifold challenges emerging from neurodivergent cognitive configurations.

The core idea of the novel social technology relates to individually training and subsequently applying fundamental executive functions such as inhibition or sustained attention, which essentially contribute to individual quality of life, job success, and physical and mental health (Diamond, 2013: 135 f.), particularly due to their facilitating role in emotion regulation (Kalantari et al., 2021: 1 f.). Tying in with Rapp et al. (2018: 49 f.), the safe and distraction-free virtual world of a UFO in outer space provides separate training tasks for core cognitive mechanisms building on standard cognitive task paradigms. Figure 1 displays both the UFO deck with a view into the space environment with an Earth-like planet and an exemplary cognitive task interface.

Taking inhibition as an example, which is defined as the ability to cancel or suppress irrelevant behavior, thoughts, and/or affective reactions while focusing on a relevant selected task (Diamond, 2013: 137; Raud et al., 2020: 1), meta-analyses emphasize the practical benefit of dedicated training programs, in particular, for neurodivergent individuals (Geurts et al., 2014: 407–418). Consequently, we leverage two vested cognitive task paradigms targeting inhibitory response behavior, involving a simple reaction task or choice reaction task as a primary re-



Figure 1: Virtual reality environment developed in the UFO project. Users arrive on the main deck (a) that provides a view into outer space and serves as a starting point for navigation to control and training facilities. Cognitive training tasks (b) are presented via dedicated interfaces in separate rooms. In the displayed exemplary training task, the grey hexagonal tiles will always light up in one of three possible colors in succession. Depending on the color displayed, an action should either be performed or suppressed.
Source: The images were kindly provided by the auticon GmbH, who took on the main responsibility for developing the virtual reality environment.

sponse (Raud et al., 2020: 1). Figure 2 shows an exemplary task procedure, requiring participants to choose either to react to a presented stimulus (GO) or not (NO-GO), depending on the presented stimulus and corresponding task-related rules.

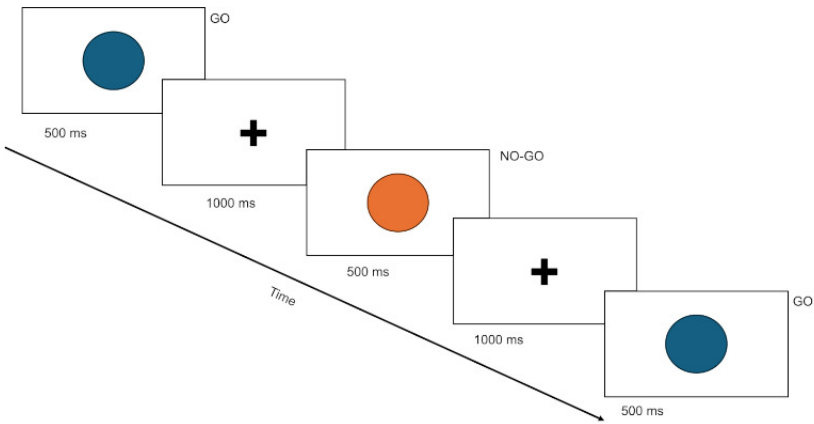


Figure 2: General schematic task procedure for an inhibition task. Screens showing blue circles require participants' response (GO), whereas screens showing orange circles demand response inhibition (NO-GO).
Source: The image was designed by the authors.

In general, the Stop Signal Task (SST) paradigm (Logan et al., 1984: 276 f.) is characterized by participants receiving a GO stimulus first, indicating they should execute a required reaction. After a short period of delay, on rare occasions a STOP signal might follow, which requires participants to stop an already

initiated response (Aron, 2011: e55; Isherwood et al., 2021: 246; Raud et al., 2020: 1; see Figure 2). For the Go/No-Go Task (GNGT) paradigm (Drewe 1975: 8 f.) participants are required to respond quickly to all stimuli (GO) except a predefined NO-GO stimulus response (Aron, 2011: e55 f.; Isherwood et al., 2021: 246; Raud et al., 2020: 1). Each task covers a different inhibitory component: Whereas the SST assesses the ability to cancel an initiated action, the GNGT targets the ability to refrain from starting an action at all (Raud et al., 2020: 1; Schachar et al., 2007: 230). Therefore, enhancing a cognitive task paradigm such as SST with emotional valence is particularly suitable for addressing the prevalence of distracting emotional stimuli and subsequent regulatory efforts (Yerys et al., 2013: 537–545).

The chosen conceptual embedding of a virtual training environment allows users to build confidence and competence in a protected setting before applying their skills in actual work scenarios (Bradley & Newbutt, 2018: 108; Didehbani et al., 2016: 703; Schmidt et al., 2024: 1819). Hence, with the goal of preparing for real-world transfer, following an individually determined period of engagement with the cognitive training tasks, users can beam themselves to an Earth-like planet and engage in simulated everyday tasks such as succeeding in a job interview or managing to schedule and receive medical support. When we consider an exemplary job interview, managing to shield oneself against potentially distracting events or impulses constitutes a crucial prerequisite for being able to adequately demonstrate acquired skills and inherent potentials. Hence, building and honing these skills through training holds practical value in daily life.

To provide the foundation for sustainable skill acquisition, a closed-loop system leverages combined visual and haptic feedback. It employs neurophysiological signals and behavioral response patterns related to the included training tasks such as SST or GNGT. More precisely, while users engage with these cognitive tasks in a virtual reality (VR) environment, the training system uses functional near-infrared spectroscopy (fNIRS) to record brain signals. These signals relate to experiencing task-induced mental workload, which might be a potential trigger for negative affective experiences (Kalantari et al., 2021: 1; Schmitz-Hübsch et al., 2024: 2083 f.). Building on these measurements, the system derives individual user states and converts them into tangible sensory perceptions to facilitate users' awareness of seemingly unconnected affective reactions and their actual situational embedding. For this purpose, users wear a T-shirt equipped with vibration pads that send haptic feedback in the form of vibration patterns. As these patterns vary in intensity with regard to amplitude and frequency, they can communicate the strength of experienced discomfort. The resulting sensory augmentation supports users with understanding their own cognitive-affective states and acting accordingly in the long term.

Developing human-centered technology like the one described can only happen by synthesizing related research across different domains and incorporating the derived design implications. These particularly relate to (1) accounting for individual differences in perception, information processing, and behavioral responses, (2) structured part-task training and segmented information presentation, (3) sustained engagement by embedded gamified elements, and (4) skill transfer to build prerequisites to cope with daily life challenges, which will be discussed subsequently. Following the motto »Nothing about us without us« (Charlton, 1998: 3), which inseparably relates to the neurodiversity movement, a core cross-linking effort relates to implementing a participatory design approach (Spinuzzi, 2005: 169–171). The inherent criteria of quality of life for workers, collaborative development, and iterative process design could be represented by incorporating a consequent co-design approach, embedding both neurodivergent and neurotypical perspectives in conceptual and technical development. While continuously soliciting diverse expertise and critical feedback on the training tasks and potential application scenarios, prerequisites for successful real-world skill transfer and awareness building are maintained.

1. *Accounting for individual differences in perception, information processing, and behavioral response*. Cognitive diversity is characterized by individual sensitivities related to perceiving internal and external stimuli, encoding and processing this information, and consequently engaging in behavioral reactions (Donnellan et al., 2013: 1; Walkowiak, 2021: 3). Applied to technology-facilitated skill building, the necessity for highly adaptive training approaches able to account for such diverse needs can be established. In general, adaptive training is a method that uses technology to personalize a user's experience based on individual needs, skills, and progress (Shute & Zapata-Rivera, 2012: 5–8). It has been demonstrated to promote users' motivation, acceptance, and identification (Streicher & Smeddinck, 2016: 332). Evidence further indicates that adaptive interaction patterns inherent in virtual environments can exert a strong positive influence on cognitive development, for instance, related to strategies of mindfulness or an awareness of individual wellbeing (Jain et al., 2023: 243 f.). Especially in contexts with above-average heterogeneity in cognitive backgrounds and learning preferences, due to adaptive training, each user has the opportunity to receive the most appropriate content in the most suitable form to benefit from optimal learning gains (Capuano & Caballé, 2020: 97 f.; Müller et al., 2020: 273; Streicher & Smeddinck, 2016: 345). Thereby, features such as the order, pace, or difficulty of the presented contents as well as design elements in the user interface can be customized. Related system capabilities are characterized by enhanced

potentials to dynamically respond to users' cognitive states and adjust the environment accordingly. Characteristically, virtual environments put emphasis on presenting information predominantly in the visual modality, potentially resulting in sensory information overload in this channel. Tying in with the multiplicity of processing channels, proposed among others by the Cognitive Theory of Multimedia Learning (CTML) (Mayer & Moreno, 2003: 43 f.) and vested memory models (e. g., Baddeley, 2000: 417–423; Paivio, 1969: 241–263), unfavorable dynamics could be avoided when employing both visual and auditory channels of information processing. Involving sensory channels beyond this scope, particularly related to tactile experiences, offers broader potentials for adaptive information presentation related to diverse sensory needs, which should even enlarge individual benefits for skill building (Motti, 2019: 4–6).

2. *Structured part-task training and segmented information presentation*. Research has demonstrated that part-task training, allowing learners to train and master individual skill components before integrating them into a larger skill set, can substantially enhance skill acquisition (Frerejean et al., 2019: 515 f.; Wickens et al., 2013: 462 f.). To achieve this, structured part-task training uses segmented information presentation as a way of delivering information step by step (Wightman & Lintern, 1985: 270 f.). The segmenting effect proposed in multimedia learning research suggests that people learn more effectively when multimedia instructions are presented in meaningful and coherent segments. These segments should be user-paced, allowing for better understanding compared to continuously presented information units (Mayer & Pilegard, 2014: 317; Rey et al., 2019: 414). As a vested method to manage task-inherent cognitive demands, this approach provides the opportunity to focus on one segment at a time, preventing overload and improving comprehension (Liu, 2024: 10 f.; Rey et al., 2019: 415). Consequently, it exerts a positive effect on task-related retention (Liu, 2024: 9 f.) and skill transfer (McDermott et al., 2012: 2505 f.; Wightman & Lintern, 1985: 273 f.) accompanied by faster learning times (McDermott et al., 2012: 2505 f.; Rey et al., 2019: 413). Part-task training allows for controlled, repetitive practice, which reinforces learning and boosts individual confidence. Existing research indicates that neurodivergent individuals, in particular, benefit from opportunities for repeated practice of specific skills before moving on to more complex tasks (Klorfeld-Auslender et al., 2022: 3206 f.).
3. *Sustained engagement by embedded gamified elements*. Leveraging benefits of gamification, described as applying game elements in non-game settings (Deterding et al., 2011: 10), inherent mechanics such as simulations, interactive challenges, and feedback loops can be used to engage users in focused,

repetitive practice. By isolating specific competences within a gamified context, users can incrementally build proficiency while maintaining motivation and enjoyment, consequently optimizing their performance outcomes (Li et al., 2023: 13; Sailer & Homner, 2020: 106 f.). This can happen because gamification satisfies important psychological needs related to intrinsic motivation (Ryan & Deci, 2002: 10 f.; Sailer et al., 2017: 374). In terms of gamified environments dedicated to particularly support neurominorities, most contributions emphasize the substantial benefits from using game mechanics and dynamics to facilitate knowledge and skill acquisition (López-Bouzas & Del Moral-Pérez, 2023: 8 f.; Malinverni et al., 2017: 540 f.). Additionally, these environments can foster self-awareness, self-regulation, motivation, and empathy (Derks et al., 2022: 8 f.; López-Bouzas & Del Moral-Pérez, 2023: 9). Further evidence could show enhancement related to users' interpersonal relationships, social skills (Alcorn et al., 2011: 12 f.; Wendt et al., 2020: 425 f.), and emotional recognition (Pires et al., 2022: 9 f.). More precisely, in a study conducted by Grund et al. (2020), gamification elements such as progress indicators, collectable rewards, and explicit praise were employed to encourage users' self-motivation while working on the respective task. However, beneficial gamified support for skill building can only emerge by stakeholder inclusion throughout the design process (Grund et al., 2020: 493 f.). Effective gamification relies on collaboration between educators, game developers, and researchers, to ensure that diverse needs are adequately addressed (Jadán-Guerrero et al., 2023: 461; Motti, 2019: 9 f.).

4. *Skill transfer to build prerequisites to cope with daily life challenges*. Embedding training situations in realistic application scenarios supports the transfer of acquired competences to users' daily lives (Barnett & Ceci, 2002: 623; Smid et al., 2020: 532–535). It is therefore important to ensure representing familiar elements from prior knowledge when designing novel opportunities for skill building, which has been demonstrated to enhance cognitive connections and deepen comprehension. Shedding light on the underlying mechanisms, users can draw parallels between known and unknown concepts, scaffold their understanding, and subsequently benefit from eased efforts related to internalizing new information (Barnett & Ceci, 2002: 632 f.; Jackson et al., 2019: 473 f.). This approach is founded upon several educational and psychological theories such as constructivism (Mattar, 2018: 204–211) or situated learning (Li et al., 2020: 30), which emphasize the significance of context and realism in skill acquisition (Choi & Lee, 2021: 8 f.; Dreimane, 2020: 16; Mattar, 2018: 204 f.). Virtual Reality Learning Environments (VRLE's) (Conrad et al., 2024: 4), in particular, open sophisticated avenues to provide authentic experiences for building daily-life skills. Several studies have demonstrated

that transferring acquired skills from a virtual environment to a real-world situation is possible and can improve situational coping as well as resulting performance (Cooper et al., 2021: 14 f.; Ganier et al., 2014: 831 f.). Virtual learning environments can thereby support individuals across a wide range of sensory and processing demands by providing flexibility and diverse learning materials (Li et al., 2020: 30), fostering communication and social skills (Boyd et al., 2018: 109 f.; Parsons & Cobb, 2011: 357 f.; Wang & Reid, 2011: 13 f.), cognitive skills (Boyd et al., 2018: 110), and motor skills (Levac et al., 2019: 9 f.). Ultimately, the related skill building can enhance well-being and reduce negative affective responses such as anxiety or frustration (Zheng et al., 2024: 878–880).

5. Addressing ethical challenges through human-centered design

Even if desired, an early anticipation of the full scope of potential societal consequences of new technologies, such as the one developed, is limited. Hence, the introduction of new technologies into society should be considered as an experience accompanied by potentially occurring unanticipated social consequences (van de Poel, 2016: 669). Following an incremental approach, a novel social technology is gradually and experimentally introduced while emerging social effects are monitored. In such a piecemeal social engineering approach (Popper, 1945: 138 f.), obtained observations are used to both improve the technology itself and ensure an efficient and suitable process of its adoption into society.

In that context, van de Poel (2016: 675–678) argues that established bioethical principles of non-maleficence, beneficence, respect for autonomy, and justice provide a crucial framework of general imperative moral values. To regulate adherence to these principles, the author further introduces the principle of responsibility.

- The principle of *non-maleficence* emphasizes obligations to do no harm, including obligations to minimize risks, or take precautions against possible risks or harms from emerging technologies (van de Poel, 2016: 678–681).
- The principle of *beneficence* underlines obligations to do good, including obligations to take away existing harm, or to prevent harm or risks that do not originate in the emerging technology, to produce more good than harm, to create or increase benefits (van de Poel, 2016: 681 f.).
- The principle of *respect for autonomy* highlights obligations relating to protecting and guaranteeing autonomy, including an autonomous choice of individ-

uals and groups when interacting with an emerging technology (van de Poel, 2016: 682 f.).
- The principle of *justice* stresses obligations of an emerging technology relating to issues of distributive justice, to special protection of vulnerable groups, to avoiding exploitation, but also to producing justice (just procedures) (van de Poel, 2016: 682 f.).
- The principle of *responsibility* specifies who is obligated to fulfill or uphold the previously listed moral commitments (van de Poel, 2016: 677 f.).

The described bioethical principles relate to obligations outlined in established psychological and clinical frameworks for human experimental research (American Psychological Association, 2017: 3 f.; World Medical Association, 2013: 2191–2194). Consequently, they establish crucial ethical, legal, and social implications (ELSI) for evaluating experimental human-centered technology. Linking these implications with the previously introduced emerging technology, potential challenges and associated countermeasures imply including a broad set of inter- and transdisciplinary stakeholder perspectives throughout the process of technology development to achieve accurate solutions aligned with users' needs (Motti, 2019: 6 f.).

With regards to *non-maleficence*, a novel technology should be comprehensively tested before its introduction into society. To this end, already in the pre-development planning stages, it is of essence that both user researchers and technical developers engage in joint efforts to account for iteratively conducted controlled user studies. By such means, it is possible to explore potential multidimensional risks in protected settings prior to upscaling implementation efforts. Nevertheless, there remains a risk of unforeseen hazards that have not yet been identified. Considering the collection and processing of personal data, which cannot be avoided at all costs, even while standards defined in the General Data Protection Regulation (GDPR) are maintained (Hoofnagle et al., 2019: 65–69), risks emerging from inadequate data handling persist. Potential approaches toward mitigating such risks relate to processing all data in a pseudonymized manner and storing it securely on access-limited, password protected servers. Non-maleficence further implies that an experiment must immediately be terminated if it causes harm. While controlled iterative user testing has widely demonstrated to provide valuable insights (Göransson et al., 2003: 111; Gulliksen et al., 2003: 401 f.), related methodological inventories have commonly been developed and tested for typical users. Hence, faced with a sample with a beyond average level of cognitive diversity, the risk of lacking adequate recognition of harm across individuals beyond the cognitive majority needs to be considered. Consequently, technology adoption may be halted too late if unexpected and undetected harm occurs. At this instance,

existing efforts at the European level can be considered helpful, as they propose concrete guidelines for evaluating sensitivities of intelligent technology for diversity, non-discrimination, and fairness (European Commission: Directorate-General for Education, Youth, Sport, and Culture, 2022: 8–11). Among others, these aspects constitute key requirements for a responsible human-technology partnership (High-Level Expert Group, 2019: 12 f.).

To adhere to the principle of *beneficence*, flexibility and the avoidance of technological lock-in, a restriction of multifold technological options by focusing on a more controllable single technology (Collingridge, 1980: 19), form crucial prerequisites. As concerns highly specialized, complex technology tailored to adaptively train cognitive-affective skills, this risk is highly prevalent. According to Wildavsky (1988: 9–15), when introducing novel technology that comes with potentially unexpected risks, to a certain extent, reliance on trusting in individuals' resilience to cope with such risks is required. However, resilience as a highly dynamic individual characteristic differs between and even within individuals over time (Rutter, 2012: 341). Therefore, solely relying on participants' resilience to manage unexpected harm comes with disproportionate challenges. Ultimately, any potential harm caused by a novel technology needs to be closely monitored and weighed against expected societal benefits, implying at least some justification only in cases when the advantages clearly exceed the disadvantages (van de Poel, 2016: 670). With the previously introduced exemplary emerging technology, which is dedicated to contributing to equal occupational opportunities, social awareness, and mutual understanding, clearly assigned advantages seem at hand. Nevertheless, any emerging negative consequences need to be closely monitored from the early stages on to immediately scale back if needed. Responsibilities related to preparing, conducting, monitoring, evaluating, adjusting, and ending the process of technology development thereby need to be distributed and communicated clearly among stakeholders (i.e., scientific researchers, developers in companies, support and training personnel).

To ensure *respect for autonomy*, collective agreement on the outlined steps constitutes an important prerequisite, which usually implies approval by an independent committee with democratic legitimation. In the context of academic research, local university ethics committees usually serve as first and central points of contact for this matter. Alternatively, interdisciplinary professional or specialist ethics committees can be consulted, such as the Ethics Association of German Business (EVW), the Central Ethics Committee (ZEKO) of the German Medical Association, or the Ethics Committee (EK) of the German Psychological Society. However, solely relying on collective agreement with regard to emerging experimental technology poses an unneglectable risk: Individual needs could fall victim to majority interests (van de Poel, 2016: 682). Hence, participants themselves

must be informed about every step of the process and the respective collected data, modalities of storage and access, as well as options for deletion. Tying in with both the key requirements of human agency and oversight as well as privacy and data governance proposed for intelligent interactive technologies (European Commission: Directorate-General for Education, Youth, Sport, and Culture, 2022: 21; High-Level Expert Group, 2019: 14 f.), individual consent from each participant is required, thereby ensuring they have the option to withdraw from participation at any time. Simultaneously, participants should have the ability to influence the conditions, execution, monitoring, evaluation, adaptation, and termination of their interaction with the evaluated technology.

In the interest of *justice*, according to van de Poel (2016: 682 f.), potentially vulnerable groups should either not be included in experimental technology development, be provided with additional protection, or must particularly benefit from the technology. With regards to the previously introduced exemplary technology, potentially more vulnerable neurodivergent groups are intended to benefit from the emerging assistive system. This implies sensitivity towards individual needs diverging from standard use cases, which can be facilitated by trained personnel and adherence to established guidelines (e. g., European Commission: Directorate-General for Education, Youth, Sport, and Culture, 2022: 11). Considering that the previously described exemplary emerging technology holds value for both neurodivergent and neurotypical individuals, thereby promoting awareness, understanding, and inclusion of neurodiverse potentials across various societal stakeholders, it is crucial to ensure that no single group bears the entire burden of the project without benefiting from it. Hence, a broad spectrum of perspectives should have the opportunity to iteratively contribute during the development stages, which needs to be ensured by all involved parties. When evaluating technological artifacts, experimenters must continuously adapt to each participant and adjust the respective conditions accordingly. Here, a core challenge emerges from establishing an adequate balance between accommodating individual needs and maintaining standardized procedures.

In summary, numerous actions need to be taken in the process of developing innovative social technology to implement an incremental approach for evaluation according to vested bioethical principles (van de Poel, 2016: 669 f.). Many of the considerations discussed can be applied to other research and development projects and thus serve as recommendations for ethically sound human-centered technical assistance.

6. Conclusion

In many instances, evolution has proven the benefits of cognitive differences as an inherent part of society. Consequently, the neurodiversity paradigm aims to reduce stigma as well as pathologization of weaknesses through arguing that society could benefit from including the strengths and unique experiences of neurodivergent individuals, exerting cognitive characteristics beyond the societal majority. The related, neurominority-driven movement claims equal rights for all humans, emphasizes the appreciation of diversity in human cognition, and advocates against structures of exclusion.

Technological advances can crucially support individuals in building essential and potentially lacking skills required to manage everyday tasks – independent of whether we consider neurotypical or neurodivergent target groups. Thereby, due to broad cognitive variety, no »prototypical« user can be taken as an independent default for all emerging design decisions. Rather, technology requires participatory development, highly personalized and individualized training situations, multisensory feedback modalities, engaging and well-structured design elements, as well as transferability of the acquired skills to daily life contexts. By implementing these recommendations, emerging assistive systems are more likely to be accepted and applicable for a broad set of use cases.

Nevertheless, the process of designing, implementing, and adopting novel social technology constitutes an experimental endeavor with unforeseen challenges along the way. In consequence, a safe and responsible upscaling of such efforts implies adherence to the vested bioethical principles of non-maleficence, beneficence, autonomy, and justice. Shared responsibilities across the inter- and transdisciplinary stakeholders involved are necessary to broadly account for the variety of perspectives involved. In this vein, research can pave the way for a more inclusive working culture that strives to meet an important goal: Honing existing potential instead of losing it.

Acknowledgements

We thank our project partners at the University of Stuttgart, Fraunhofer Institute for Industrial Engineering, auticon GmbH, NIRx GmbH, and Sensovo GmbH for the joint development and implementation of the UFO project. We further thank Valerie Holzwarth for her support during the preparation of the manuscript.

Funding information

The project UFO is funded by the German Federal Ministry of Education and Research (BMBF) [1] *within the funding line »Interactive systems in virtual and real spaces – Innovative technologies for the digital society« (grant number 16SV8722).*

Literature

Alcorn, Alyssa et al. (2011). Social Communication Between Virtual Characters and Children with Autism, In: Biswas, Gautam, et al. (eds.), *Artificial Intelligence in Education: AIED 2011*, Lecture Notes in Computer Science, vol. 6738, Springer, Berlin/Heidelberg Springer, p. 7–14. https://link.springer.com/chapter/10.1007/978-3-642-21869-9_4 [01.03.2025]

American Psychiatric Association, DSM-5 Task Force. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5™ (5th ed)*, American Psychiatric Publishing, Inc., Washingdon D.C. https://doi.org/10.1176/appi.books.9780890425596

American Psychological Association (2017). *Ethical Principles and Code of Conduct for Psychologists*, American Psychological Association, Washingdon D.C. http://www.apa.org/ethics/code/index.aspx [01.03.2025]

Armstrong, Thomas (2010). Neurodiversity: Discovering the Extraordinary Gifts of Autism, ADHD, Dyslexia, and Other Brain Differences, Da Capo Press, Boston.

Aron, Adam R. (2011). From Reactive to Proactive and Selective Control: Developing a Richer Model for Stopping Inappropriate Responses. In: Krystal, John H. (ed.), *Biological Psychiatry*, vol. 69, no. 12, Yale University School of Medicine, VA Connecticut Healthcare System, West Haven, p. e55–e68. https://doi.org/10.1016/j.biopsych.2010.07.024

Austin, Robert D. & Pisano, Gary P. (2017). Neurodiversity as a Competitive Advantage. In: *Harvard Business Review*, vol. 95, no. 3, Harvard Business Publishing, Brighton, p. 96–103. https://www.hbs.edu/faculty/Pages/item.aspx?num=52624

Baddeley, Alan (2000). The Episodic Buffer: A New Component of Working Memory? In: *Trends in Cognitive Sciences*, vol. 4, no. 11, p. 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Barnett, Susan M. & Ceci, Stephen J. (2002). When and Where Do We Apply What We Learn? A Taxonomy for Far Transfer. In: *Psychological Bulletin*, vol. 128, no. 4, p. 612–637. https://doi.org/10.1037/0033-2909.128.4.612

Bar-Yosef, Ofer (2002). The Upper Paleolithic Revolution. In: *Annual Review of Anthropology*, vol. 31, no. 1, Stanford Universtiy Press, Redwood City, p. 363–393. https://doi.org/10.1146/annurev.anthro.31.040402.085416

Botha, Monique, Dibb, Bridget & Frost, David M. (2022). »Autism is Me«: An Investigation of How Autistic Individuals Make Sense of Autism and Stigma. In: *Disability & Society*, vol. 37, no. 3, p. 427–453. https://doi.org/10.1080/09687599.2020.1822782

1 At the time of printing: Federal Ministry of Research, Technology and Space (BMFTR)

Boyd, LouAnne E. et al. (2018). Leveling the Playing Field: Supporting Neurodiversity via Virtual Realities. In: *Technology and Innovation*, vol. 20, no. 1–2, p. 105–116. https://doi.org/10.21300/20.1-2.2018.105

Bradley, Ryan & Newbutt, Nigel (2018). Autism and Virtual Reality Head-Mounted Displays: A State-of-the-Art Systematic Review. In: *Journal of Enabling Technologies*, vol. 12, no. 3, p. 101–113. https://doi.org/10.1108/JET-01-2018-0004

Capuano, Nicola & Caballé, Santi (2020). Adaptive Learning Technologies. In: *AI Magazine*, vol. 41, no. 2, p. 96–98. https://doi.org/10.1609/aimag.v41i2.5317

Charlton, James I. (1998). *Nothing About Us Without Us: Disability Oppression and Empowerment*, University of California Press, Berkeley.

Choi, Seoyoon E. & Lee, Sun (2021). Applying Constructivism in Neurodiverse Classrooms. In: *Review of Educational Theory*, vol. 4, vo. 4, p. 6–11. https://doi.org/10.30564/ret.v4i4.3876

Collingridge, David (1980). *The Social Control of Technology*, Frances Pinter, London.

Conrad, Matthias, Kablitz, David & Schumann, Stephan (2024). Learning Effectiveness of Immersive Virtual Reality in Education and Training: A Systematic Review of Findings. In: *Computers & Education: X Reality*, vol. 4, e100053. https://doi.org/10.1016/j.cexr.2024.100053

Cooper, Natalia et al. (2021). Transfer of Training – Virtual Reality Training with Augmented Multisensory Cues Improves User Experience during Training and Task Performance in the Real World. In: *PLOS ONE*, vol. 16, no. 3, p. 1–22. https://doi.org/10.1371/journal.pone.0248225

Derks, Suzanne, Willemen, Agnes M. & Sterkenburg, Paula S. (2022). Improving Adaptive and Cognitive Skills of Children with an Intellectual Disability and/or Autism Spectrum Disorder: Meta-Analysis of Randomized Controlled Trials on the Effects of Serious Games. In: *International Journal of Child-Computer Interaction*, vol. 33, e100488. https://doi.org/10.1016/j.ijcci.2022.100488

Deterding, Sebastian et al. (2011). From Game Design Elements to Gamefulness: Defining ›Gamification‹. In: MindTrek '11: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, Association for Computing Machinery, New York, p. 9–15. https://doi.org/10.1145/2181037.2181040

Diamond, Adele (2013). Executive Functions. In: *Annual Review of Psychology*, vol. 64, no. 1, p. 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Didehbani, Nyaz et al. (2016). Virtual Reality Social Cognition Training for Children with High Functioning Autism. In: *Computers in Human Behavior*, vol. 62, p. 703–711. https://doi.org/10.1016/j.chb.2016.04.033

Donnellan, Anne M., Hill, David A. & Leary, Martha R. (2013). Rethinking Autism: Implications of Sensory and Movement Differences for Understanding and Support. In: *Frontiers in Integrative Neuroscience*, vol. 6, p. 124. https://doi.org/10.3389/fnint.2012.00124

Doyle, Nancy (2020). Neurodiversity at Work: A Biopsychosocial Model and the Impact on Working Adults. In: *British Medical Bulletin*, vol. 135, no. 1, p. 108–125. https://doi.org/10.1093/bmb/ldaa021

Dreimane, Lana F. (2020). Virtual Reality Learning Experience Evaluation Tool for Instructional Designers and Educators. In: Daniela, Linda (ed.), *New Perspectives on Virtual and Augmented Reality: Finding New Ways to Teach in a Transformed Learning Environment*, Routledge, London, p. 3–21. https://doi.org/10.4324/9781003001874

Drewe, Elaine A. (1975). Go – No Go Learning After Frontal Lobe Lesions in Humans. In: *Cortex*, vol. 11, no. 1, p. 8–16. https://doi.org/10.1016/S0010-9452(75)80015-3

European Commission: Directorate-General for Education, Youth, Sport, and Culture (2022). *Ethical Guidelines on the Use of Artificial Intelligence (AI) and Data in Teaching and Learning for Educators*. Publications Office of the European Union. https://data.europa.eu/doi/10.2766/153756

Finlayson, Clive (2004). *Neanderthals and Modern Humans: An Ecological and Evolutionary Perspective*, Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511542374

Frerejean, Jimmy et al. (2019). Designing Instruction for Complex Learning: 4C/ID in Higher Education. In: *European Journal of Education*, vol. 54, no. 4, p. 513–524. https://doi.org/10.1111/ejed.12363

Frith, Uta & Happé, Francesca (2005). Autism Spectrum Disorder. In: *Current Biology*, vol. 15, no. 19, p. R786–R790. https://doi.org/10.1016/j.cub.2005.09.033

Ganier, Franck, Hoareau, Charlotte & Tisseau, Jacques (2014). Evaluation of Procedural Learning Transfer from a Virtual Environment to a Real Situation: A Case Study on Tank Maintenance Training. In: *Ergonomics*, vol. 57, no. 6, p. 828–843. https://doi.org/10.1080/00140139.2014.899628

Geurts, Hilde M., et al. (2014). Prepotent Response Inhibition and Interference Control in Autism Spectrum Disorders: Two Meta-Analyses. In: *Autism Research*, vol. 7, no. 4, p. 407–420. https://doi.org/10.1002/aur.1369

Göransson, Birger, Gulliksen, Johan & Boivie, Ingela (2003). The Usability Design Process – Integrating User-Centered Systems Design in the Software Development Process. In: *Software Process Improvement and Practice*, vol. 8, no. 2, p. 111–131. https://doi.org/10.1002/spip.174

Grund, Jonas et al. (2020). A Gamified and Adaptive Learning System for Neurodivergent Workers in Electronic Assembling Tasks. In: *Proceedings of Mensch und Computer 2020 (MuC '20)*, New York: Association for Computing Machinery, p. 491–494. https://doi.org/10.1145/3404983.3410420

Gulliksen, Jan et al. (2003). Key Principles for User-Centred Systems Design. In: *Behaviour & Information Technology*, vol. 22, no. 6, p. 397–409. https://doi.org/10.1080/01449290310001624329

Happé, Francesca (2018). Why Are Savant Skills and Special Talents Associated with Autism? In: *World Psychiatry*, vol. 17, no. 3, p. 280–281. https://doi.org/10.1002/wps.20552

High-Level Expert Group on Artificial Intelligence (2019). *Ethics Guidelines for Trustworthy AI*, 6. European Commission.

Hoofnagle, Chris J., Van Der Sloot, Bart & Borgesius, Frederik Z. (2019). The European Union General Data Protection Regulation: What It Is and What It Means. In: *Information & Communications Technology Law*, vol. 28, no. 1, p. 65–98. https://doi.org/10.1080/13600834.2019.1573501

Howlin, Patricia, Alcock, Jennifer & Burkin, Catherine (2005). An 8-Year Follow-Up of a Specialist Supported Employment Service for High-Ability Adults with Autism or Asperger Syndrome. In: *Autism*, vol. 9, no. 5, p. 533–549. https://doi.org/10.1177/1362361305057871

Humphrey, Neil & Lewis, Sarah (2008). ›Make Me Normal‹: The Views and Experiences of Pupils on the Autistic Spectrum in Mainstream Secondary Schools. In: *Autism*, vol. 12, no. 1, p. 23–46. https://doi.org/10.1177/1362361307085267

Huws, Jaci C. & Jones, Robert S. P. (2008). Diagnosis, Disclosure, and Having Autism: An Interpretative Phenomenological Analysis of the Perceptions of Young People with Autism. In:

Journal of Intellectual and Developmental Disability, vol. 33, no. 2, p. 99–107. https://doi.org/10.1080/13668250802010394

Isherwood, Scott J. S. et al. (2021). Cortical and Subcortical Contributions to Interference Resolution and Inhibition – An fMRI ALE Meta-Analysis. In: *Neuroscience & Biobehavioral Reviews*, vol. 129, p. 245–260. https://doi.org/10.1016/j.neubiorev.2021.07.021

Jackson, Denise, Fleming, Jenny & Rowe, Anna (2019). Enabling the Transfer of Skills and Knowledge Across Classroom and Work Contexts. In: *Vocations and Learning*, vol. 12, no. 3, p. 459–478. https://doi.org/10.1007/S12186-019-09224-1

Jadán-Guerrero, Janio et al. (2023). Gamification in Inclusive Education for Children with Disabilities: Global Trends and Approaches – A Bibliometric Review. In: Gervasi, Osvaldo et al. (eds.), *Computational Science and Its Applications – ICCSA 2023 Workshops*, Lecture Notes in Computer Science, vol. 14104, Cham: Springer, p. 461–477. https://doi.org/10.1007/978-3-031-37105-9_31

Jain, Abhinanadan, Maes, Pattie & Sra, Misha (2023). Adaptive Virtual Neuroarchitecture. In: Simeone, Angelo et al. (eds.), *Everyday Virtual and Augmented Reality*, Cham: Springer, p. 227–249. https://doi.org/10.1007/978-3-031-05804-2_9

Kalantari, Niloofar et al. (2021). Emotion Regulation for Neurodiversity through Wearable Technology. In: *9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, IEEE Los Alamitos/CA. https://doi.org/10.1109/ACII52823.2021.9597458

Kapp, Steven K. et al. (2013). Deficit, Difference, or Both? Autism and Neurodiversity. In: *Developmental Psychology*, vol. 49, no. 1, p. 59–71. https://doi.org/10.1037/a0028353

Klorfeld-Auslender, Shira et al. (2022). A Distinct Route for Efficient Learning and Generalization in Autism. In: *Current Biology*, vol. 32, no. 14, p. 3203–3209. https://doi.org/10.1016/j.cub.2022.05.059

LeFevre-Levy, Rose et al. (2023). Neurodiversity in the Workplace: Considering Neuroatypicality as a Form of Diversity. In: *Industrial and Organizational Psychology*, vol. 16, no. 1, p. 1–19. https://doi.org/10.1017/iop.2022.86

Levac, Danielle E., Huber, Meghan E. & Sternad, Dagmar (2019). Learning and Transfer of Complex Motor Skills in Virtual Reality: A Perspective Review. In: *Journal of NeuroEngineering and Rehabilitation*, vol. 16, 121, p. 1–15. https://doi.org/10.1186/s12984-019-0587-8

Li, Minzi, Ma, Siyu & Shi, Yuyang (2023). Examining the Effectiveness of Gamification as a Tool Promoting Teaching and Learning in Educational Settings: A Meta-Analysis. In: *Frontiers in Psychology*, vol. 14, e1253549. https://doi.org/10.3389/fpsyg.2023.1253549

Li, Ping et al. (2020). Virtual Reality for Student Learning: Understanding Individual Differences. In: *Human Behaviour and Brain*, vol. 1, no. 1, p. 28–36. https://doi.org/10.37716/HBAB.2020010105

Liu, Dongyang (2024). The Effects of Segmentation on Cognitive Load, Vocabulary Learning and Retention, and Reading Comprehension in a Multimedia Learning Environment. In: *BMC Psychology*, vol. 12, no. 1. https://doi.org/10.1186/s40359-023-01489-5

Logan, Gordon D., Cowan, William B. & Davis, Kevin A. (1984). On the Ability to Inhibit Simple and Choice Reaction Time Responses: A Model and a Method. In: *Journal of Experimental Psychology: Human Perception and Performance*, vol. 10, no. 2, p. 276–291. https://doi.org/10.1037/0096-1523.10.2.276

López-Bouzas, Nerea & Del Moral-Pérez, Esther M. (2023). Gamified Environments and Serious Games for Students with Autistic Spectrum Disorder: Review of Research. In: *Review Journal of Autism and Developmental Disorders*, p. 1–13. https://doi.org/10.1007/s40489-023-00381-7

Malinverni, Laura et al. (2017). An Inclusive Design Approach for Developing Video Games for Children with Autism Spectrum Disorder. In: *Computers in Human Behavior*, vol. 71, p. 535–549. https://doi.org/10.1016/j.chb.2016.01.018

Mattar, Joao (2018). Constructivism and Connectivism in Education Technology: Active, Situated, Authentic, Experiential, and Anchored Learning. In: *Revista Iberoamericana de Educación a Distancia*, vol. 21, no. 2, p. 201–217. https://doi.org/http://dx.doi.org/10.5944/ried.21.2.20055

Mayer, Richard E. & Moreno, Roxana (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. In: *Educational Psychologist*, vol. 38, no. 1, p. 43–52. https://doi.org/10.1207/S15326985EP3801_6

Mayer, Richard E. & Pilegard, Celeste (2014). Principles for Managing Essential Processing in Multimedia Learning: Segmenting, Pre-Training, and Modality Principles. In: Mayer, Richard E. (ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed.), Cambridge University Press, Cambridge, p. 316–344. https://doi.org/10.1017/CBO9781139547369.016

McDermott, Patricia L., Carolan, Thomas & Wickens, Christopher D. (2012). Part Task Training Methods in Simulated and Realistic Tasks. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, no. 1, p. 2502–2506. https://doi.org/10.1177/1071181312561509

Meilleur, Andrée-Anne A., Jelenic, Patricia & Mottron, Laurent (2015). Prevalence of Clinically and Empirically Defined Talents and Strengths in Autism. In: *Journal of Autism and Developmental Disorders*, vol. 45, p. 1354–1367. https://doi.org/10.1007/s10803-014-2296-2

Mellars, Paul (2005). The Impossible Coincidence. A Single-Species Model for the Origins of Modern Human Behavior in Europe. In: *Evolutionary Anthropology: Issues, News, and Reviews*, vol. 14, no. 1, p. 12–27. https://doi.org/10.1002/evan.20037

Mellars, Paul (2006). A New Radio-Carbon Revolution and the Dispersal of Modern Humans in Eurasia. In: *Nature*, vol. 439, no. 7079, p. 931–935. https://doi.org/10.1038/nature04521

Motti, Vivian G. (2019). Designing Emerging Technologies for and with Neurodiverse Users. In: *Proceedings of the 37th ACM International Conference on the Design of Communication (SIGDOC '19)*, Association for Computing Machinery, no. 11, p. 1–10. https://doi.org/10.1145/3328020.3353946

Müller, Ina, Moebert, Tobias & Lucke, Ulrike (2020). Lessons Learned from Designing Adaptive Training Systems: An Ethical Perspective. In: Pinkwart, Niels & Liu, Shaofei (eds.), *Artificial Intelligence Supported Educational Technologies, Advances in Analytics for Learning and Teaching*, Springer, Cham, p. 273–290. https://doi.org/10.1007/978-3-030-41099-5_16

Ortega, Francisco (2009). The Cerebral Subject and the Challenge of Neurodiversity. In: *BioSocieties*, vol. 4, no. 4, p. 425–445. https://doi.org/10.1017/S1745855209990287

Paivio, Allan (1969). Mental Imagery in Associative Learning and Memory. In: *Psychological Review*, vol. 76, no. 3, p. 241–263. https://doi.org/10.1037/h0027272

Parsons, Sarah & Cobb, Sue (2011). State-of-the-Art of Virtual Reality Technologies for Children on the Autism Spectrum. In: *European Journal of Special Needs Education*, vol. 26, no. 3, p. 355–366. https://doi.org/10.1080/08856257.2011.593831

Pires, Gabriel et al. (2022). A New Error-Monitoring Brain–Computer Interface Based on Reinforcement Learning for People with Autism Spectrum Disorders. In: *Journal of Neural Engineering*, vol. 19, no. 6, p. 1–15. https://doi.org/10.1088/1741-2552/aca798

Popper, Karl R. (1945). *The Open Society and Its Enemies*, Routledge, London.

Portway, Suzannah M. & Johnson, Barbara (2005). Do You Know I Have Asperger's Syndrome? Risks of a Non-obvious Disability. In: *Health, Risk & Society*, vol. 7, no. 1, p. 73–83. https://doi.org/10.1080/09500830500042086

Rapp, Amon et al. (2018). Designing Technology for Spatial Needs: Routines, Control and Social Competences of People with Autism. In: *International Journal of Human-Computer Studies*, vol. 120, p. 49–65. https://doi.org/10.1016/j.ijhcs.2018.07.005

Raud, Liisa et al. (2020). Differences in Unity: The Go/no-go and Stop Signal Tasks Rely on Different Mechanisms. In: *NeuroImage*, vol. 210, e116582. https://doi.org/10.1016/j.neuroimage.2020.116582

Rey, Günter D. et al. (2019). A Meta-analysis of the Segmenting Effect. In: *Educational Psychology Review*, vol. 31, no. 2, p. 389–419. https://doi.org/10.1007/s10648-018-9456-4

Rutter, Michael (2012). Resilience as a dynamic concept. In: *Development and Psychopathology*, vol. 24, no. 2, p. 335–344. https://doi.org/10.1017/S0954579412000028

Ryan, Richard M. & Deci, Edward L. (2002). Overview of Self-Determination Theory: An Organismic Dialectical Perspective. In: Ryan, Richard M. & Deci, Edward L. (eds.), *Handbook of Self-Determination Research*, University of Rochester Press, Rochester, p. 3–33.

Sailer, Michael et al. (2017). How Gamification Motivates: An Experimental Study of the Effects of Specific Game Design Elements on Psychological Need Satisfaction. In: *Computers in Human Behavior*, vol. 69, p. 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Sailer, Michael & Homner, Lisa (2020). The Gamification of Learning: A Meta-Analysis. In: *Educational Psychology Review*, vol. 32, no. 1, p. 77–112. https://doi.org/10.1007/s10648-019-09498-w

Schachar, Russell et al. (2007). Restraint and Cancellation: Multiple Inhibition Deficits in Attention Deficit Hyperactivity Disorder. In: *Journal of Abnormal Child Psychology*, vol. 35, p. 229–238. https://doi.org/10.1007/s10802-006-9075-2

Schmidt, Matthew et al. (2024). Toward a Strengths-Based Model for Designing Virtual Reality Learning Experiences for Autistic Users. In: *Autism*, vol. 28, no. 7, p. 1809–1827. https://doi.org/10.1177/13623613231208579

Schmitz-Hübsch, Alina et al. (2024). Towards Enhanced Performance: An Integrated Framework of Emotional Valence, Arousal, and Task Demand. In: *Ergonomics*, vol. 67, no. 12, p. 1–14. https://doi.org/10.1080/00140139.2024.2370440

Shute, Valerie J. & Zapata-Rivera, Diego (2012). Adaptive Educational Systems. In: Durlach, Paula J. & Lesgold, Alan M. (eds.), *Adaptive Technologies for Training and Education*, Cambridge University Press, Cambridge, p. 7–27. https://doi.org/10.1017/CBO9781139049580.004

Smid, Claire R., Karbach, Julia & Steinbeis, Nikolaus (2020). Toward a Science of Effective Cognitive Training. In: *Current Directions in Psychological Science*, vol. 29, no. 6, p. 531–537. https://doi.org/10.1177/0963721420951599

Spikins, Penny (2009). Autism, the Integration of ›Difference‹ and the Origins of Modern Human Behaviour. In: *Cambridge Archaeological Journal*, vol. 19, no. 2, p. 179–201. https://doi.org/10.1017/S0959774309000262

Spinuzzi, Clay (2005). The Methodology of Participatory Design. In: *Technical Communication*, vol. 52, no. 2, p. 163–174.

Streicher, Alexander & Smeddinck, Jan D. (2016). Personalized and Adaptive Serious Games. In: Dörner, Ralf et al. (eds.), *Entertainment Computing and Serious Games*, *Lecture Notes in Computer Science*, vol. 9970, Springer, Cham, p. 332–377. https://doi.org/10.1007/978-3-319-46152-6_14

Sweller, John, van Merriënboer, Jeroen J. G. & Paas, Fred (2019). Cognitive Architecture and Instructional Design: 20 Years Later. In: *Educational Psychology Review*, vol. 31, p. 261–292. https://doi.org/10.1007/s10648-019-09465-5

Tonizzi, Irene, Giofrè, David & Usai, Maria C. (2022). Inhibitory Control in Autism Spectrum Disorders: Meta-Analyses on Indirect and Direct Measures. In: *Journal of Autism and Developmental Disorders*, vol. 52, p. 4949–4965. https://doi.org/10.1007/s10803-021-05353-6

van de Poel, Ibo (2016). An Ethical Framework for Evaluating Experimental Technology. In: *Science and Engineering Ethics*, vol. 22, no. 3, p. 667–686. https://doi.org/10.1007/s11948-015-9724-3

Walkowiak, Emmanuelle (2021). Neurodiversity of the Workforce and Digital Transformation: The Case of Inclusion of Autistic Workers at the Workplace. In: *Technological Forecasting & Social Change*, vol. 168, e120739. https://doi.org/10.1016/j.techfore.2021.120739

Wang, Michelle & Reid, Denise (2011). Virtual Reality in Pediatric Neurorehabilitation: Attention Deficit Hyperactivity Disorder, Autism, and Cerebral Palsy. In: *Neuroepidemiology*, vol. 36, no. 1, p. 2–18. https://doi.org/10.1159/000320847

Wehman, Paul et al. (2016). Employment for Adults with Autism Spectrum Disorders: A Retrospective Review of a Customized Employment Approach. In: *Research in Developmental Disabilities*, vol. 53–54, p. 61–72. https://doi.org/10.1016/j.ridd.2016.01.015

Wendt, Oliver et al. (2020). Optimized User Experience Design for Augmentative and Alternative Communication via Mobile Technology: Using Gamification to Enhance Access and Learning for Users with Severe Autism. In: Stephanidis, Constantine et al. (eds.), *HCI International 2020 – Late Breaking Papers: Universal Access and Inclusive Design*, vol. 12426, Springer, Cham, p. 295–306. https://doi.org/10.1007/978-3-030-60149-2_32

Wickens, Christopher D. et al. (2013). Effectiveness of Part-Task Training and Increasing-Difficulty Training Strategies. In: *Human Factors*, vol. 55, no. 2, p. 461–470. https://doi.org/10.1177/0018720812451994

Wightman, Dennis C. & Lintern, Gavan (1985). Part-Task Training for Tracking and Manual Control. In: *Human Factors*, vol. 27, no. 3, p. 267–284. https://doi.org/10.1177/001872088502700304

Wildavsky, Aaron B. (1988). *Searching for Safety*, Transatlantic, Publisher, New Brunswick/London.

World Medical Association (2013). World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. In: *JAMA*, vol. 310, no. 20, p. 2191–2194. https://doi.org/10.1001/jama.2013.281053

Wright, Barry, Spikins, Penny & Pearson, Hannah (2020). Should Autism Conditions Be Characterised in a More Positive Way in Our Modern World? In: *Medicina*, vol. 56, no. 5, 233, p. 1–12. https://doi.org/10.3390/medicina56050233

Yerys, Benjamin E. et al. (2013). Separate Components of Emotional Go/No-Go Performance Relate to Autism versus Attention Symptoms in Children with Autism. In: *Neuropsychology*, vol. 27, no. 5, p. 537–545. https://doi.org/10.1037/a0033615

Zheng, Yujia et al. (2024). P-4.34: Virtual Reality Content Generation for Anxiety and Stress Management: Current Status, Challenges, and Future Developments. In: *Digest of Technical Papers*, vol. 55, p. 878–881. https://doi.org/10.1002/sdtp.17227

Was ist Neurodiversität und was bedeutet das für die VR-Entwicklung? – UFO

Marc Ristau[1]

Zusammenfassung

Das UFO-Projekt »Einsatz virtueller Realität (VR) zur Förderung beruflicher Inklusion« hat sich zum Ziel gesetzt, die berufliche Teilhabe insbesondere für neurodivergente Menschen insgesamt zu erhöhen. Eine virtuelle Umgebung bietet einen sicheren Trainingsraum, der die Bedürfnisse der Zielgruppe passgenau unterstützen kann. Das VR-Projekt trägt einerseits zur gesellschaftlichen Sensibilisierung für das Thema Neurodiversität insgesamt bei und fördert andererseits Empathie und gegenseitiges Verständnis im gelebten Arbeitsalltag von und mit neurodivergenten Menschen. Die technische Innovation des Projekts liegt darin, eine »taktile Sinneserweiterung« zu schaffen und entlang der Herausforderungen und Empfindungen von und mit neurodivergenten Menschen erlebbar zu machen. Das Besondere an diesem Projekt ist, dass die Zielgruppe – neurodivergente Menschen – direkt in das partizipative Entwicklungsprojekt integriert wurde – nicht nur als adressierte Zielgruppe, sondern auch – und das ist wichtig hervorzuheben – als neurodivergentes Entwicklungsteam.

Der vorliegende Artikel spiegelt die gesammelten Erfahrungen und Recherche-Ergebnisse innerhalb des Projektverlaufs aus der Perspektive eines neurodivergenten Mitarbeitenden wider. Eine ebenso ungewohnte wie innovative (in Teilen sogar fluide) Sichtweise auf ein Projekt mit sehr hohem Innovationsindex – gerade durch die Einbeziehung der Zielgruppenperspektive in allen Phasen des konzeptionellen und technischen Entwicklungsprozesses. Das Herausarbeiten der neurodivergenten Perspektive ist explizit beabsichtigt und auch zwingend notwendig, um Partizipation und Softwaregestaltung konsequent zusammenzudenken. Eine Guideline zum Mitnehmen und ein Muss für zukünftige Entwicklungsprojekte im VR-Sektor.

---

1 Unter freundlicher Mitwirkung von Niko Lang, auticon Deutschland GmbH

1. Einleitung

Neurodiversität beschreibt die natürliche Vielfalt menschlicher Gehirne, ihrer Neurologie und ihrer funktionalen Ausprägungen. Abweichungen von der quantitativen Norm werden dabei besonders berücksichtigt und als Neurodivergenzen bezeichnet. Dazu zählen bekannte Phänomene wie Autismus, ADHS (Aufmerksamkeitsdefizit-/Hyperaktivitätssyndrom), Dyslexie und Dyspraxie, um die für VR am relevantesten zu nennen, sowie auch weitere, wie Hochbegabung oder auch Tourette. Obwohl diese in der digitalen Welt überproportional vertreten sind und geradezu omnipräsent erscheinen, machen sie insgesamt nur etwa 15 % bis 20 % der Bevölkerung aus (vgl. Goldberg, 2023: 172 ff.). Das Paradigma der Neurodiversität überwindet die historisch bedingte Pathologisierung und versteht sich als sozialer Gegenentwurf zu vielfältig erlebter Stigmatisierung und Diskriminierung neurodivergenter Lebenswirklichkeiten. Gerade im autistischen Teil der Bevölkerung sind zusätzlich auftretende Merkmale von weiteren Neurodivergenzen wie ADHS, Dyslexie und Dyspraxie weit verbreitet. Diese treten aber auch unabhängig voneinander auf. Viele sind sich ihrer neurodivergenten Merkmale mangels Aufklärung gar nicht bewusst. Dementsprechend ist es sinnvoll, VR-Erlebnisse für einen möglichst breiten Bevölkerungsanteil inklusiv zu gestalten.

Die sich daraus ergebende Vielfalt birgt sowohl Potenziale als auch Herausforderungen, die es bei der Konzeption von VR-Anwendungen zu berücksichtigen gilt – allerdings nicht ausschließlich, um neurodivergente Menschen zielgruppengerecht anzusprechen. Es ist grundsätzlich vorteilhaft, benutzungsfreundliche und anpassbare Bedienoberflächen zu gestalten. Dazu gehören eine klare und konsistente Navigation sowie die Berücksichtigung sensorischer Besonderheiten, z. B. bei der Gestaltung von VR-Welten und -Interaktionen. Durch inklusive und partizipative Gestaltungsansätze können Softwareprodukte und VR-Erlebnisse geschaffen werden, die die Bedürfnisse neurodivergenter Menschen so berücksichtigen, dass diese ihre Stärken spielerisch entwickeln und das Nutzungserlebnis insgesamt genießen können.

2. Nutzendenzentrierung bedeutet: Partizipatives Design

2.1 Die Einbeziehung neurodivergenter Personen

Bei der Gestaltung neuroinklusiver User Experiences kann es von Vorteil sein, neurodivergente Nutzende während des gesamten Designprozesses aktiv mit einzubeziehen. Ihre Einblicke sind unverzichtbar, um spezifische Bedürfnisse

und Vorlieben zu verstehen. Durch direkte Zusammenarbeit können Designschaffende besser einschätzen, welche Elemente in einer VR-Umgebung hilfreich oder hinderlich sein können. Dies führt zu Lösungen, die den Anforderungen einer diversen Nutzendenschaft umfangreicher entsprechen, als dies in homogenen Teams möglich wäre.

Auch in der Forschung wird der partizipative Ansatz verfolgt. »In der partizipativen Forschung stehen die Menschen, die an ihr teilhaben, im Mittelpunkt – ihre Perspektiven, ihre Lernprozesse und ihre individuelle und kollektive (Selbst-)Befähigung. Partizipative Forschung ist damit nie ein rein akademisches Unterfangen, sondern immer ein Gemeinschaftsprojekt mit nichtwissenschaftlichen, gesellschaftlichen Akteuren.« (von Unger, 2014: 2)

Insbesondere die *direkte Mitarbeit im Projekt* kann von unschätzbarem Wert sein, denn eine divergente Perspektive im Team wirft häufig bereits Fragen auf, die sonst nicht gestellt würden. Idealerweise sind alle Teammitglieder reflektiert über mögliche Neurodivergenzen und werden darin unterstützt, divergenten Perspektiven kontinuierlich Ausdruck zu verleihen, damit diese von der neurotypischen Mehrheit nicht als Einzelmeinungen abgetan werden. Dies sollte keine einmalige Gelegenheit sein, sondern Teil des Feedbackprozesses werden. Gerade autistische Nutzende haben einerseits ein gutes Auge für Details, andererseits häufig eine eher langfristig angelegte Verarbeitung von Information, insofern bietet es sich an, diese in alle Testzyklen von Anfang bis zum Ende miteinzubeziehen.

Eine weitere Methode der Einbindung stellen *Co-Design Workshops* dar: Die Durchführung von Co-Design-Workshops bietet eine Plattform, auf der eine Vielzahl neurodivergenter Individuen direkt Ideen einbringen und Feedback geben kann. In diesen Sitzungen arbeiten Designschaffende und Nutzende gemeinsam an der Entwicklung und Verbesserung der VR-Umgebung. Dies stellt sicher, dass das Endprodukt nicht nur den Erwartungen entspricht, sondern auch die Nutzungserfahrung optimiert. Durch diesen kollaborativen Ansatz wird der Designprozess inklusiv gestaltet und fördert Innovationen, die aus der Vielfalt der Perspektiven entstehen. Hier bietet es sich an, bereits Vorarbeiten in verschiedenen Varianten, idealerweise modular, präsentieren zu können.

Für viele neurodivergente, insbesondere autistische Personen ist es von großer Bedeutung, sich gut vorbereiten zu können, das gilt auch für Co-Design-Workshops. Eine ausführliche Vorbereitung ermöglicht es, sich mental und emotional auf die bevorstehenden Interaktionen und Aufgaben einzustellen. Autistische Menschen profitieren besonders von klaren Informationen über den Ablauf, die Ziele und die Erwartungen des Workshops. Das geht häufig so weit, dass sie sich für erwartete Situationen Formulierungen überlegen, die möglicherweise hilfreich sind, oder auch Dinge, die sie auf jeden Fall sagen wollen. Dies

hilft, Unsicherheiten zu reduzieren, was wiederum ermutigt, aktiv und effektiv am kreativen Prozess teilzunehmen. Durch die Bereitstellung von detaillierten Vorabinformationen können mögliche Stressfaktoren minimiert und ein Umfeld geschaffen werden, in dem sie ihre wertvollen Perspektiven und Ideen optimal einbringen können. Auch hier sollte darauf geachtet werden, dass die Annäherung an neurodivergente Nutzungsperspektiven ein kontinuierlicher Prozess ist, der beim ersten Ansatz mit an Sicherheit grenzender Wahrscheinlichkeit nicht abgeschlossen sein wird.

2.2 Kontinuierliches Feedback von neurodivergenten Personen via Community

Das Einholen von kontinuierlichem Feedback von neurodivergenten Personen über Communities und Plattformen wie Discord oder Reddit ist von großem Wert für die Entwicklung inklusiver und nutzendenzentrierter Lösungen. Diese bieten den Nutzenden eine zugängliche und vertraute Umgebung, in der sie ihre Gedanken, Erfahrungen und Anregungen teilen können, wann immer es ihnen danach ist. Die Schaffung separater Kanäle für neurodivergente Perspektiven kann daher lohnenswert sein, um einen geschützten Rahmen zu schaffen, der es erleichtert, spezifische Bedürfnisse und Vorlieben intuitiv zu kommunizieren, ohne Unverständnis bei neurotypischen Nutzenden hervorzurufen. Statt als kontinuierlicher Chat, der schnell unübersichtlich wird, sollten Diskussionen in Threads/Themenbäumen strukturiert und übersichtlich geführt werden, was die aktive Teilnahme erleichtert und sicherstellt, dass wertvolles Feedback nicht übersehen wird. Solch kontinuierliche Interaktion ermöglicht es, direkt auf die Anliegen der Nutzenden einzugehen und fördert eine Zusammenarbeit, die von gegenseitiger Inspiration geprägt ist und so zu besseren und inklusiveren Ergebnissen führt.

2.3 Zusammenarbeit mit Fachleuten und Sachverständigen

Die Konsultation von Fachleuten, wie beispielsweise aus den Bereichen Ergotherapie, Psychologie und Pädagogik, die auf Neurodiversität spezialisiert sind, kann einen signifikanten Beitrag zum Design leisten. Ihre Expertise ermöglicht wertvolle Einblicke in spezifische Bedürfnisse und effektive Strategien zur Unterstützung neurodivergenter Nutzenden. Es ist von entscheidender Bedeutung, auf dem aktuellen Stand der Forschung zu bleiben, um aktuelle Erkenntnisse über Neurodivergenz und Technologienutzung zu berücksichtigen. Durch kontinuierliches Lernen und Anpassen können Gestaltende sicherstellen, dass ihre

VR-Umgebungen den neuesten Standards entsprechen und die bestmögliche Benutzungserfahrung bieten. Ein herausragendes Beispiel in diesem Kontext ist auticon, ein Unternehmen, das sich aktiv an Forschungsprojekten beteiligt. Bei auticon sind ausschließlich autistische Mitarbeitende als IT-Consultants angestellt (vgl. Abramova et al., 2025). Viele dieser Mitarbeitenden zeigen eine hohe Motivation, sich an der Konzeption und Entwicklung von VR-Projekten zu beteiligen und als Tester und Testerinnen zur Verfügung zu stehen. Die Einbindung autistischer Fachpersonen in die Entwicklungsprozesse ermöglicht es Entwickelnden, von den Erfahrungen und Besonderheiten autistischer Wahrnehmung zu profitieren. Dies resultiert in Produkten und Dienstleistungen, die nicht nur besser auf die Bedürfnisse neurodivergenter Nutzenden abgestimmt sind, sondern auch eine höhere Benutzungsfreundlichkeit insgesamt und damit auch Qualität aufweisen. Durch die Berücksichtigung und Anwendung dieser Methoden und Best Practices können VR-Umgebungen geschaffen werden, die wirklich inklusiv sind. Ein fundamentales Verständnis und eine Wertschätzung der Vielfalt neurodivergenter Erfahrungen bilden die Grundlage für die Gestaltung von Erlebnissen, die für alle zugänglich und bereichernd sind.

2.4 Sensorische Anpassungsoptionen

Die Bereitstellung von sensorischen Anpassungsoptionen in VR-Umgebungen ist entscheidend, um den unterschiedlichen Empfindlichkeiten neurodivergenter Personen gerecht zu werden. *Anpassbare sensorische Einstellungen* ermöglichen es Nutzenden, Helligkeit, Kontrast, Lautstärke und Bewegungsempfindlichkeit nach ihren individuellen Bedürfnissen zu modifizieren. Bereits durch solche unscheinbar wirkenden Einstellungen können Nutzende das Erlebnis so gestalten, dass es ihren Komfortlevels entspricht und Überreizung vermieden wird. Beispielsweise kann jemand mit hoher Lichtempfindlichkeit die Bildschirmhelligkeit reduzieren, während eine Person, die empfindlich auf laute oder intensive Geräusche reagiert, die Lautstärke anpassen oder bestimmte Audiosignale deaktivieren kann. Dies kann beim Testen berücksichtigt werden, indem Szenenkontraste einerseits für herabgestellte Helligkeit herausgearbeitet werden und andererseits bei hoher Helligkeit nicht zur Reizüberflutung durch zu hohe Sättigung führen. Ähnlich verhält es sich mit der Audio-Kanalabstimmung oder auch so etwas Simplem wie der Tonqualität. Rauschen, Knacksen, klirrende Geräusche oder Ähnliches, aber auch inkongruentes räumliches Audio kann zwar nicht sofort, aber bei anhaltender Exposition zu Überforderung führen. Was darüber hinaus einen Effekt haben kann, sind realistische Materialeigenschaften und insgesamt die Umgebungen so zu gestalten, dass diese auch in Audio

und Haptik Kongruenz herstellen, z. B. wenn man mit der virtuellen Hand eine Oberfläche streift, einen entsprechenden Ton abzuspielen und den Controller mit einem entsprechenden Muster vibrieren zu lassen. Das fördert nicht nur die Immersion, die hergestellte Kongruenz ist für jede Art von Neurologie, ob neurodivergent oder nicht, energiesparend, sodass damit eine längere Nutzungsdauer ermöglicht wird.

Darüber hinaus bietet es sich an, *starke Reize optional zu minimieren*, um sensorische Überlastung zu verhindern. Dies kann bedeuten, starke visuelle oder auditive Stimuli, die für einige Nutzende überwältigend sein könnten (wie Explosionen, Klirren, schrille Geräusche), optional zu machen oder zu dämpfen. Der Einsatz von beruhigenden Farben und Klängen kann dazu beitragen, eine angenehme und stressfreie Umgebung zu schaffen. Indem Designschaffende bewusst Elemente auswählen, die entspannend wirken, fördern sie ein intensiveres und angenehmeres VR-Erlebnis für neurodivergente Personen. Dies kann sich nicht nur im »Flair« oder auch »Vibe« einzelner Assets ausdrücken, sondern insbesondere auch in harmonischer Formsprache der Weltgestaltung.

Es lohnt sich außerdem, sich an etablierten Designsprachen, Konzepten und Frameworks mindestens zu orientieren, da diese Wiedererkennung und Anknüpfungspunkte bieten und auch deren Fallstricke teilweise bereits bekannt sind und diskutiert werden können. Beispielsweise werden Memphis Design oder auch Pop Art eher als reizüberflutend wahrgenommen, wohingegen Minimalismus und Art Deco häufig als beruhigend empfunden werden.

2.5 Haptisches Feedback

Haptisches Feedback ist ein essenzieller Bestandteil von VR-Erfahrungen, da es durch physische Rückmeldungen wie Vibrationen oder virtuell dargestelltem physischen Widerstand das Erlebnis signifikant realistischer gestaltet. Für neurodivergente Personen kann die Wahrnehmung von haptischem Feedback jedoch variieren. Während einige Personen von der zusätzlichen sensorischen Information profitieren und diese ihnen hilft, sich in der virtuellen Umgebung zurechtzufinden, können andere durch unerwartete oder intensive haptische Signale überfordert werden. Aus diesem Grund sollten VR-Anwendungen die Möglichkeit bieten, haptisches Feedback individuell anzupassen, sodass Nutzende die Intensität regulieren oder bestimmte haptische Effekte deaktivieren können, um ein angenehmes und kontrolliertes Erlebnis zu gewährleisten.

3. Häufige Mechanismen in Social VR

3.1 Audio- und haptische Bubbles

In einigen sozialen VR-Umgebungen werden sogenannte »Bubbles« eingesetzt, die visuelle, akustische als auch haptische Reize umgeben und auf diese Weise konfigurierbare persönliche Bereiche schaffen. In diesen Bereichen können Nutzende ihre sensorischen Erfahrungen nach ihren Präferenzen steuern. Die Bubbles reduzieren oder eliminieren wahlweise Geräusche, haptische Interaktionen sowie die Sichtbarkeit außerhalb eines bestimmten Radius um die Nutzenden.

Für neurodivergente Personen, die häufig empfindlicher auf sensorische Reize reagieren, bieten solche Bubbles die Möglichkeit, Überstimulation zu vermeiden, indem ein Fokusbereich geschaffen und alles dahinter ausgeblendet werden kann. Die Reduktion von Intensität und Reichweite von auditiven und visuellen oder auch haptischen Reizen ermöglicht einen Grad an Immersion, der individuellen Komfortbedürfnissen entspricht und gleichzeitig soziale Interaktionen über einen längeren Zeitraum ermüdungsärmer ermöglicht.

3.2 »Earmuffs«

Ähnlich funktionieren »Earmuffs«, virtuelle Ohrenschützer, die flexibel auf- oder abgesetzt und konfiguriert werden können, um Hintergrundgeräusche oder unerwünschte auditive Reize oder Frequenzen zu reduzieren oder ganz auszublenden. Für neurodivergente Nutzende, die empfindlich auf Lärm reagieren oder schnell von auditiven Stimuli überreizt werden, kann diese Funktion erheblich zum Wohlbefinden beitragen. Nutzende können sich auf Interaktionen konzentrieren, ohne von Umgebungsgeräuschen überwältigt oder abgelenkt zu werden. Autistische Menschen neigen dazu, alles im Vordergrund zu hören, was es schwierig macht, sich in lauten echten wie virtuellen Umgebungen auf ein Gespräch zu konzentrieren. Earmuffs bieten eine Möglichkeit zur individuellen Kontrolle der akustischen Umgebung und ermöglichen so die Teilhabe an Gesprächen in lauter Umgebung. Dies trägt zu einer inklusiveren und angenehmeren sozialen Erfahrung in der virtuellen Realität bei.

3.3 Iglu-Spawning als Safer Space gegen Überwältigung

Ähnlich wie die Bubbles dient das Iglu-Spawning als effektiver Kompensationsmechanismus insbesondere für neurodivergente Nutzende. Dabei wird eine igluähnliche Struktur um sich herum »gespawned« (in der Welt instanziiert), um einen persönlichen Rückzugsort oder Schutzraum innerhalb der virtuellen Welt zu schaffen. Dieses virtuelle Iglu dient dem Schutz vor übermäßigen sensorischen Reizen oder unerwünschten sozialen Interaktionen. Für autistische und neurodivergente Menschen kann dies von Vorteil sein, da es ihnen erlaubt, die Intensität von Eindrücken zu regulieren und ein Gefühl der Sicherheit zu erlangen. Ein signifikanter Aspekt ist die Möglichkeit, andere Nutzende selektiv in das Iglu einzulassen. Dadurch können autistische Personen selbst bestimmen, mit wem sie interagieren möchten, was die Kontrolle über soziale Kontakte erhöht und positive Interaktionen fördert. Die Schaffung eines persönlichen Safer Space ermöglicht es den Nutzenden, ihre Umgebung besser zu kontrollieren und somit ein angenehmeres und weniger stressiges VR-Erlebnis zu genießen.

4. Inklusive Gestaltung in VR

4.1 Synästhesie und Phantom Sense in VR

Synästhesie ist eine neurologische Besonderheit, bei der die Stimulation eines Sinnes gleichzeitig eine Reaktion in einem anderen Sinn auslöst, z. B. das Sehen von Farben beim Hören von Musik. In der VR können solche synästhetischen Erfahrungen verstärkt oder gezielt angesprochen werden. Bei neurodivergenten Personen kommt das nicht selten vor und ist auch intensiver, deshalb bietet VR die Möglichkeit, personalisierte und einzigartige Sinneserlebnisse zu schaffen, um die Wahrnehmungswelt um eindrucksvolle Selbsterfahrungen zu bereichern. Allerdings besteht auch das Risiko einer sensorischen Überladung, wenn zu viele Reize gleichzeitig angeflutet werden. Optionen sollten es daher anbieten, sensorische Eindrücke zu filtern und anzupassen, sodass alle Nutzende die Intensität und Kombination der Sinneseindrücke steuern können.

Phantom Sense ist ebenfalls eine Art von Synästhesie und bezeichnet das Phänomen, bei dem Nutzende in der VR-Umgebung sensorische Empfindungen erleben, die nicht durch physische Reize ausgelöst werden. Dies kann beispielsweise das Gefühl von Wärme, Kälte oder Berührung sein, obwohl kein entsprechendes tatsächliches sensorisches Signal anliegt. Für neurodivergente Individuen kann Phantom Sense besonders intensiv sein und sowohl positive als auch negative Auswirkungen haben. Einerseits kann es das Eintauchen in die virtuelle

Welt vertiefen, andererseits kann es zu Verwirrung oder Unbehagen führen, wenn die Grenzen zwischen realen und virtuellen Empfindungen verschwimmen. Dies kann aber auch als Chance begriffen werden, wohlige Wärme als angenehm und flimmernde Hitze oder klirrende Kälte gezielt als unangenehm einzusetzen, um die Immersion einer Szenerie zu erhöhen. Es ist daher wichtig, dies vorab abzufragen, darüber aufzuklären und die VR-Erfahrungen so zu gestalten, dass auf diese Sensibilität Rücksicht genommen wird und gegebenenfalls Optionen zur Anpassung oder Deaktivierung entsprechender Effekte angeboten werden.

4.2 Vorhersehbarkeit und Konsistenz

Ein weiterer Schlüssel zum erfolgreichen Design inklusiver VR-Umgebungen ist die *Konsistenz der Designelemente*. Durch die Aufrechterhaltung eines einheitlichen Navigationssystems, konsistenter Interface-Elemente und wiederkehrender Interaktionen wird die kognitive Belastung für alle Nutzende reduziert. Konsistenz hilft auch beim Weltenbau dabei, Verwirrung zu vermeiden, und ermöglicht es den Nutzenden, sich leichter in der Umgebung zurechtzufinden. Wenn beispielsweise Menüs und Schaltflächen stets an denselben Positionen bleiben und gleich funktionieren, erzeugt das ein Gefühl von Kontrolle, Sicherheit und Wohlbehagen. Wiederholungen sind starke Elemente jeder neurodivergenten Wahrnehmung, und da diese ein Gefühl von Behaglichkeit erzeugen, kann damit ganz bewusst umgegangen werden. Insbesondere für einen gewollten Lern- oder Aha-Effekt oder auch bloß zur auffrischenden Erinnerung an vergangene Spielstunden darf sich gezielte Wiederholung auch gerne in der Storyline widerspiegeln.

Zusätzlich sind *klare Hinweise und auch konsistentes haptisches Feedback* essenziell. Durch den Einsatz von eindeutigen haptischen, visuellen und auditiven Indikatoren für Aktionen und Reaktionen können Nutzende besser antizipieren, was gerade oder als Nächstes passiert. Auch das schafft ein Gefühl der Kontrolle und Vorhersehbarkeit, was besonders für Personen mit Angstzuständen oder solchen, die Schwierigkeiten mit unerwarteten Ereignissen haben, hilfreich sein kann. Klare Anweisungen und sofortiges Feedback nach einer Aktion stärken das Vertrauen der Nutzenden in ihre eigenen Fähigkeiten innerhalb der VR-Umgebung und darüber hinaus.

4.2.1 Realismus in den Materialeigenschaften

Die Berücksichtigung realistischer Materialeigenschaften in VR-Umgebungen, einschließlich authentischer Audio- und haptischer Feedbacks, ist entscheidend,

um *Kongruenz* zwischen den verschiedenen Sinneseindrücken herzustellen. Für neurodivergente Personen kann diese Kongruenz besonders hilfreich sein, da sie dazu beiträgt, die kognitive Belastung zu reduzieren. Wenn visuelle, auditive und haptische Reize stimmig sind und den Erwartungen entsprechen, muss das Gehirn weniger Energie aufwenden, um Unstimmigkeiten zu verarbeiten. Dies wirkt sich als besonders *energiesparend auf die neurodivergente Neurologie* aus und ermöglicht ein angenehmeres und weniger anstrengendes Erlebnis. Durch die Schaffung einer konsistenten Sinneswelt können VR-Erfahrungen für neurodivergente Nutzende zugänglicher und ansprechender gestaltet werden, was Interaktion und Beteiligung fördern.

4.2.2 Klarheit und intuitive Benutzungsoberflächen

Die Gestaltung *klarer und intuitiver Benutzungsoberflächen* ist von zentraler Bedeutung, um die Zugänglichkeit zu verbessern. *Vereinfachte Navigation* setzt voraus, Menüs und Steuerungen so zu gestalten, dass sie leicht verständlich und zugänglich sind. Unübersichtlichkeit und unnötige Komplexität sollten vermieden werden, um Überforderung zu verhindern. Ein minimalistisches Design mit klaren Strukturen unterstützt die Nutzenden dabei, sich auf das Wesentliche zu konzentrieren.

Der Einsatz von *visuellen Hilfsmitteln und Symbolen* kann ebenfalls die Verständlichkeit erhöhen. Die Verwendung bekannter Symbolik fördert auch hier die Wiedererkennung, die Navigation wird erleichtert und sogar Sprachbarrieren können reduziert werden. Piktogramme und Icons helfen dabei, Funktionen schnell zu erkennen und zu nutzen, ohne dass die Beschriftung gelesen werden muss, was insbesondere Nutzenden mit Leseschwierigkeiten oder Konzentrationsproblemen zugutekommt.

4.3 Bewegungskontrolle und Autonomie

Die Einstellbarkeit der *Bewegungssteuerung* ist ein weiterer wichtiger Aspekt in Bezug auf *Autonomie*. Insbesondere erfahrene Nutzende fühlen sich häufig von zu langsamer Fluid-Locomotion oder Joystick-Drehung unangenehm eingeschränkt, was je nach Situation ein großer Frust- und Nervfaktor sein kann. Anfänger:innen hingegen profieren von Teleportation, Snap-Turn und Vignetten.

Die *Auswahl verschiedener Interaktionsmodi* bietet zusätzliche Flexibilität. Indem verschiedene Methoden zur Interaktion mit der Umgebung angeboten werden, wie Sprachbefehle, Gestensteuerung oder die Konfiguration traditioneller Einga-

begeräte wie typische VR-Controller, können Nutzende die für sie in der jeweiligen Situation angenehmste Option wählen.

Geschwindigkeitsoptionen bezüglich der Wiedergabe von Audio- und Videoinhalten ermöglichen es Nutzenden, das Tempo des Erlebnisses selbst zu bestimmen, Abschnitte zu pausieren, zu überspringen oder erneut abzuspielen, was besonders hilfreich ist, um Informationen im eigenen Tempo zu verarbeiten. Dies unterstützt Lernprozesse und gibt den Nutzenden die Freiheit, die zeitliche Komponente eines Erlebnisses an ihre individuellen Bedürfnisse anzupassen.

4.3.1 Anpassbare Avatare und Repräsentation

Die optionale *Personalisierung von Avataren* ermöglicht es Nutzenden, ihre Identität in der VR-Umgebung widerzuspiegeln. Diese können das Komfortgefühl und die Identifikation mit dem Erlebnis erhöhen. Wenn Nutzende sich selbst oder Aspekte ihrer Persönlichkeit im Avatar wiedererkennen, fördert dies das Engagement und die Verbundenheit mit der Umgebung.

Repräsentation zählt: Die Einbindung vielfältiger Charaktere und Erzählungen, die mit neurodivergenten Personen resonieren, trägt auch ganz allgemein zur Inklusivität bei. Durch Geschichten und Figuren, die ihre Erfahrungen widerspiegeln, fühlen sich Nutzende gesehen und verstanden. Dies kann nicht nur positive Auswirkungen auf das Gesamterlebnis haben, sondern auch auf das Selbstwertgefühl.

Wenn die Möglichkeit besteht und die Umgebung dafür geeignet ist, sorgt entsprechende Symbolik aus der Neurodiversity-Community (wie das Regenbogen-Unendlichkeitszeichen oder Referenzen zu neurodivergenten Charakteren, Organisationen, Kanälen usw.) zusätzlich für Wiedererkennungswert und obendrein für ein nicht zu vernachlässigendes *Willkommensgefühl*.

4.3.2 Pro und Kontra bei der Verwendung von Avatar-Baukastensystemen

Die Verwendung eines Avatar-Baukastens, z. B. von Meta oder Readyplayer.me, bietet gleich mehrere Vorteile bei der Avatar-Gestaltung. Ein wesentlicher Vorteil ist die *Benutzungsfreundlichkeit* dieser Systeme. Sie ermöglichen den Nutzenden, ohne tiefergehende technische Kenntnisse schnell und intuitiv eigene Avatare zu erstellen. Dies erleichtert auch auf Entwicklungsseite die Integration und reduziert auf beiden Seiten den Aufwand für individuelle Gestaltung. Zudem sorgen solche Baukästen für eine *hohe Kompatibilität und Konsistenz* über verschiedene Plattformen hinweg, was die Integration in unterschiedliche VR-Umgebungen erleichtert und ein einheitliches Nutzungserlebnis fördert. Darüber hinaus enthalten solche Baukästen oft bereits *Barrierefreiheitsfunktionen*, die speziell darauf

ausgelegt sind, inklusiv zu sein und die Bedürfnisse einer breiten Nutzendenbasis, einschließlich neurodivergenter Personen, zu berücksichtigen.

Es gibt jedoch auch Nachteile bei der Verwendung vorgefertigter Avatar-Baukästen. Die *Anpassungsmöglichkeiten* können eingeschränkt sein, was die individuelle Ausdrucksfähigkeit der Nutzenden begrenzt. Dies kann besonders dann problematisch sein, wenn spezifische kulturelle Merkmale, persönliche Vorlieben oder einzigartige Identitätsaspekte dargestellt werden sollen, die nicht im Baukasten enthalten sind. Zudem begibt man sich in *Abhängigkeit zu diesen Plattformen*, was Risiken hinsichtlich Lizenzierung, Datenschutz und langfristiger Verfügbarkeit mit sich bringen kann. Schließlich kann die Nutzung solcher Baukästen die *Kreativität einschränken*, da Nutzende und Entwickelnde gleichermaßen auf vordefinierte Optionen beschränkt sind und weniger Möglichkeiten haben, innovative, stimmige oder gar unkonventionelle Avatare zu erstellen. In Bezug auf neurodivergente Nutzende überwiegt jedoch der Vorteil, allein der Konformität mit Gewohnheiten gerade auch wegen der identitätsstiftenden Wirkung von Avataren.

4.3.3 Pro und Kontra der individuellen Größenanpassung in VR

Die Möglichkeit, in VR die eigene Größe selbst zu bestimmen, bietet Nutzenden einerseits ein hohes Maß an Personalisierung und Freiheit in der Gestaltung ihres virtuellen Selbst. Dies fördert ein tieferes Eintauchen in die virtuelle Welt, da sich Nutzende so darstellen können, wie sie sich fühlen oder gesehen werden möchten. Für manche bedeutet dies, sich größer oder kleiner zu machen, um bestimmte Emotionen oder Persönlichkeitsaspekte auszudrücken. Darüber hinaus kann die Größenanpassung die Zugänglichkeit erhöhen, indem sie es Personen mit körperlichen Einschränkungen ermöglicht, Objekte leichter zu erreichen oder sich besser in der Umgebung zu orientieren. Durch das Experimentieren mit unterschiedlichen Größen können Nutzende zudem neue Perspektiven entdecken und ein vielfältigeres Spielerlebnis genießen, was die Kreativität anregt und das Engagement in der VR-Umgebung steigert.

Trotz der Vorteile kann die freie Bestimmung der Körpergröße in VR auch Nachteile mit sich bringen. In Mehrspieler-Umgebungen kann das zu Ungleichgewichten führen, da Nutzende durch Vergrößerung oder Verkleinerung unfaire Vorteile erlangen könnten, etwa durch leichteres Erreichen von Objekten oder schwerere Sichtbarkeit für andere. Dies kann das Spielerlebnis beeinträchtigen und Frustrationen hervorrufen. Zudem kann die Veränderung der eigenen Größe bei einigen Personen Desorientierung oder gar Motion Sickness verursachen, insbesondere wenn die Anpassungen abrupt erfolgen. Technisch gesehen erfordert die Implementierung einer Größenkonfiguration zusätzlichen Entwicklungsaufwand, um sicherzustellen, dass alle Interaktionen und Umgebungen

korrekt auf unterschiedliche Größen reagieren. Schließlich kann eine unzureichende Gestaltung der Größenanpassung das immersive Erlebnis stören, wenn Proportionen unnatürlich wirken oder Skalierungsprobleme auftreten, was die Glaubwürdigkeit der virtuellen Welt mindert. Auf der anderen Seite gibt es das Phänomen der körperdysmorphen Wahrnehmung, bei der man sich als kleiner oder größer wahrnimmt, als man tatsächlich ist, und eine Berücksichtigung dessen bei der Zusammenstellung eines Avatars hilft möglicherweise sogar, sich damit auseinanderzusetzen oder gar anzufreunden.

4.3.4 Vermeidung von potenziellen Triggern

Um das Wohlbefinden der Nutzenden nicht zu gefährden, ist es wichtig, *potenzielle Auslöser von Angstzuständen oder epileptischen Anfällen zu vermeiden*. Gerade in Bezug auf Neurodivergenz ist dies wichtig, denn ca. jede dritte autistische Person entwickelt im Laufe des Lebens eine Epilepsie.

Inhaltswarnungen informieren Nutzende deshalb über Elemente wie blinkende Lichter, laute Geräusche oder intensive Sequenzen. Dies gibt ihnen die Möglichkeit, sich darauf vorzubereiten, entsprechende Anpassungen in den Einstellungen vorzunehmen oder die Sequenz zu überspringen. Transparenz über die Inhalte fördert außerdem das Vertrauen und Sicherheitsgefühl der Nutzenden.

Auch die Abschaltbarkeit von z. B. Partikel-Effekten kann den Komfort erhöhen, einerseits durch die Abwesenheit der gegebenenfalls chaotischen und überfordernden Partikel-Wolke, andererseits durch gegebenenfalls dadurch vermiedene Performance-Engpässe, die zu Rucklern oder reduziertem Detailgrad und damit auch zu mentaler Beanspruchung führen können, weil die Kongruenz darunter leidet.

Opt-Out-Funktionen ermöglichen es den Nutzenden allgemein, bestimmte z. B. als gewalttätig gekennzeichnete Inhalte zu überspringen oder zu deaktivieren, die Unbehagen oder Stress auslösen könnten. Indem man ihnen die Kontrolle gibt, unangenehme Elemente zu vermeiden, wird das Erlebnis auf die individuelle Belastbarkeit angepasst und das Risiko negativer emotionaler oder kognitiver Reaktionen minimiert. Dies ist besonders wichtig für Personen mit sensorischen Empfindlichkeiten oder traumatischen Erfahrungen.

4.3.5 Hilfe beim Zurechtfinden

Die Bereitstellung von *Tutorials und Anleitungen* hilft Nutzenden, sich mit einer VR-Applikation vertraut zu machen. Schritt-für-Schritt-Einführungen oder Hilfebereiche unterstützen beim Verständnis von Navigation und Interaktion. Dies

ist besonders hilfreich für Neulinge oder diejenigen, die allgemein mit Technologie weniger vertraut sind.

Kontextuelle Hilfe bietet Unterstützung genau dann, wenn sie benötigt wird, ohne das Gesamterlebnis zu unterbrechen. Wenn Nutzende auf Schwierigkeiten stoßen, können sie schnell auf Hilfsressourcen zugreifen, was Frustration reduziert und die Lernkurve verkürzt. Dies fördert eine selbstbestimmte Nutzung und erhöht die Zufriedenheit.

4.4 Adaption von Accessibility-Standards

Es bietet sich an, die Anpassung bestehender Zugänglichkeitsstandards an die virtuelle Realität mitzudenken, um ein inklusives Erlebnis für alle Nutzenden zu gewährleisten. Es empfiehlt sich, etablierte Richtlinien wie die Web Content Accessibility Guidelines (WCAG) als Ausgangspunkt zu nutzen und sie für die spezifischen Anforderungen von VR anzupassen. Beispielsweise können Grundsätze zur Lesbarkeit und zum Farbkontrast von Text auf 3D-Oberflächen übertragen werden.

In einer VR-Anwendung sollte der Text in der virtuellen Umgebung so gestaltet sein, dass er stets gut lesbar bleibt, unabhängig von der Hintergrundfarbe oder der Beleuchtungssituation. Indem wir bestehende Standards erweitern und neue Kriterien für VR-spezifische Aspekte wie räumliches Audio und 3D-Navigation entwickeln, schaffen wir konsistente und zugängliche Erlebnisse für eine vielfältige Nutzendenschaft, einschließlich neurodivergenter Personen.

4.4.1 Standardisierungsbestrebungen

Die Konsultation mit Normungsgremien wie dem World Wide Web Consortium (W3C) kann helfen, Zugänglichkeitsstandards VR-spezifisch zu adaptieren. Einheitliche Standards fördern die Zugänglichkeit branchenweit und erleichtern es Nutzenden, sich in verschiedenen Anwendungen zurechtzufinden.

4.4.2 Visuell

Um die visuelle Zugänglichkeit in VR zu verbessern, sollten Texte klar dargestellt und in der Größe skalierbar sein. Nutzenden sollte die Möglichkeit gegeben werden, Aspekte wie Textgröße, Schriftart, Zeilenabstand und Kontrast individuell anzupassen. Eine VR-Anwendung könnte z. B. eine Einstellungsoption bieten, mit der Nutzende den Kontrast erhöhen können, um Texte und Objekte deutlicher zu erkennen. Zusätzlich können Pfeile oder subtile Indikatoren im Sicht-

feld eingeblendet werden, um die Aufmerksamkeit auf wichtige Elemente zu lenken, anstatt das Blickfeld automatisch auszurichten, was schnell verwirren kann. Durch solche Anpassungen wird das Erlebnis für Menschen mit visuellen Beeinträchtigungen oder neurodivergenten Wahrnehmungen verbessert, da sich die Umgebung nach ihren Bedürfnissen ausrichtet, und nicht umgekehrt.

4.4.3 Auditiv

Für Nutzende mit Hörbeeinträchtigungen ist es wichtig, dass auditive Informationen auch visuell verfügbar sind. Die Implementierung von Untertiteln für Dialoge und Tonhinweise, die im Blickfeld platziert sind, ohne die Sicht zu behindern, ist hierbei essenziell. In einer VR-Erfahrung sollten wichtige Geräusche wie Warnsignale durch visuelle Effekte oder Vibrationen des Controllers komplementär ergänzt werden. Die Einbindung von Gebärdensprache nutzende Avatare für non-player characters (NPCs) sollte ebenfalls erwogen werden. Solche Maßnahmen stellen sicher, dass alle Nutzende, unabhängig von ihrer Hörfähigkeit, die volle Erfahrung der Anwendung genießen können. Für neurodivergente Personen kann die multisensorische Darstellung von Informationen zudem die Verarbeitung erleichtern.

4.4.4 Motorisch

VR-Erlebnisse sollten verschiedene Eingabemethoden unterstützen, um Nutzenden mit eingeschränkter Mobilität gerecht zu werden. Die Unterstützung von Eye-Tracking, Face-Pointing, Sprachsteuerung oder vereinfachter Controller-Bedienung, z. B. mit nur einem Controller, kann hier einen großen Mehrwert bieten. Beim Berühren von Buttons einen Ton abzuspielen oder eine Vibration auszulösen, hilft bei intuitivem Bedienen von UI-Elementen (User Interface). Außerdem sollte z. B. bei der Trigger-Taste oder Grip-Taste die Position berücksichtigt werden, die bei Beginn des Tastendrucks registriert wird, d. h. nicht bei vollem Durchdrücken, weil die Position durch die Bewegung des Drückens »verrutschen« kann. Eine Interaktion kann darüber hinaus die Option bieten, Gestensteuerungen durch einfache Kopfbewegungen oder Sprachbefehle zu ersetzen. Zudem kann das Erlebnis so gestaltet werden, dass es vollständig im Sitzen erlebt werden kann. Durch die Anpassung der Interaktionsmechanismen wird die Teilnahme für Personen mit motorischen Einschränkungen oder neurodivergenten Bewegungsmustern ermöglicht und erleichtert. Auch virtuelle Beine und insgesamt ein Full-Body-Avatar, idealerweise mit Position Estimation, tragen als virtuelle Repräsentation sowohl zum Immersionsgefühl als auch für Nutzende mit

eingeschränkter Gehmobilität zu einem Stück Freiheitserleben bei, diesbezüglich kann z. B. auch ein Liegemodus implementiert werden.

4.4.5 Reduzierte kognitive Belastungen bei der Nutzung von VR

Um kognitive Belastung zu reduzieren, sollten VR-Anwendungen eine einfache Navigation durch Menüs und Welten sowie klare Navigationsziele bieten. Die Beschränkung auf maximal zwei Navigationsebenen erleichtert die Orientierung. Konsistente UI-Elemente helfen Nutzenden, die Funktionalität besser einzuschätzen. Ein VR-Lernprogramm sollte z. B. immer die gleichen Symbole für bestimmte Aktionen verwenden und geführte Tutorials mit Schritt-für-Schritt-Anweisungen anbieten. Solche Anpassungen unterstützen insbesondere neurodivergente Personen, die von klaren Strukturen und Vorhersehbarkeit profitieren, und fördern ein stressfreies Erlebnis.

4.4.6 Personalisierung in Nutzendenprofilen für individualisierte VR-Erfahrungen

Die Möglichkeit, individuelle Einstellungen in Profilen zu speichern, erleichtert die Nutzung für verschiedene Personen. Familienmitglieder können jeweils eigene Profile mit ihren bevorzugten Zugänglichkeitseinstellungen anlegen. Personalisierung erhöht die Nutzendenzufriedenheit und ermöglicht es insbesondere neurodivergenten Personen, die Erfahrung nach individuellen Bedürfnissen anzupassen.

4.4.7 Sicherheitsmerkmale in VR

Sicherheitsfunktionen wie ein Pausieren bei Überschreiten der virtuellen Begrenzungen sind essenziell, um Unfälle zu vermeiden. Auch virtuelle Notausgänge oder Notausknöpfe können dazu beitragen, ein Gefühl der Sicherheit zu vermitteln. Wenn Nutzende sich der physischen Grenze des Spielbereichs nähern, sollte die Anwendung nicht nur automatisch pausieren, sondern auch den Wiedereinstieg abfedern, z. B. indem es bei zeitkritischen Interaktionen eine Transitionspause mit Countdown gibt, damit man sich mental auf den Wiedereinstieg vorbereiten kann. Solche Funktionen schützen nicht nur vor physischen Gefahren, sondern bieten auch neurodivergenten Personen ein kontrollierbareres und damit stressfreieres Erlebnis.

5. Fazit

Die kontinuierliche Anpassung von VR-Erlebnissen an neurodivergente Nutzende sowie die Adaption von Zugänglichkeitsstandards an die virtuelle Realität sind Prozesse, die in hohem Maße interdisziplinäre Zusammenarbeit und Innovationsgeist erfordern. Die Priorisierung von Barrierefreiheit in der Design- und Entwicklungsphase ermöglicht die Schaffung von VR-Erlebnissen, die nicht nur zugänglicher, sondern auch genießbarer sind. Die Nutzung offener Standards sowie die aktive Beteiligung an ihrer Weiterentwicklung sind dabei von entscheidender Bedeutung, um Inklusion zu fördern, die Benutzungsfreundlichkeit zu erhöhen und die Flexibilität in der Nutzungsperspektive zu erhöhen. Von diesen Maßnahmen profitieren nicht nur neurodivergente Nutzende, sondern alle Beteiligten. Die Berücksichtigung der Bedürfnisse von neurodivergenten Nutzenden ist daher von entscheidender Bedeutung, um eine vielfältigere und empathischere digitale Zukunft zu gestalten.

Förderhinweis

Das Projekt UFO wurde vom Bundesministerium für Bildung und Forschung (BMBF)[2] gefördert.

Literatur

Abramova Olga, Recker Jan, Schemm Ursula und Barwitzki Lukas-Daniel (2025). Inclusion of Autistic IT Workforce in Action: An Auticon Approach. In: Information Systems Journal, Early View, John Wiley & Sons, Hoboken. First published online 08 February 2025. https://doi.org/10.1111/isj.12588

Goldberg Hagar (2023). Unraveling Neurodiversity: Insights from Neuroscientific Perspectives. In: *Encyclopedia 2023*, Vol. 13, MDPI, Basel, S. 972–980. https://doi.org/10.3390/encyclopedia3030070

von Unger, Hella (2014). *Partizipative Forschung. Einführung in die Forschungspraxis*, Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-01290-8.

2 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Virtual Professors in Mixed Reality – VoluProf

Sergio Tejeda Pastor, Jangwoo Son, Johann-Christian Põder, Andy Räder, Juliane Wegner, Armin Keller, Tanja Bäumel, Bernd Kiefer, Michael Dietrich, Torsten Schimmer, Atilla Azgin, Benjamin Körner, Sven Wischnowsky, Thomas Buchholz, Mustafa Tevfik Lafci, Birgit Nierula, Peter Eisert, Anna Hilsmann, Sebastian Bosse and Cornelius Hellge

Abstract

The VoluProf platform uses mixed reality (MR) to integrate volumetric avatars of educators into virtual classrooms, enhancing engagement and accessibility. With real-time avatar animation, speech synthesis, and scalable cloud rendering, it supports interactive and personalized learning across devices. Emphasizing ethical design and user inclusivity, VoluProf complements traditional education by fostering flexibility and enriched experiences, setting a foundation for future immersive educational tools.

1. Introduction

VoluProf is transforming online education by addressing the key challenges of traditional e-learning methods, such as limited interactivity, isolation, and accessibility barriers. Static video and audio formats often fail to engage students effectively, while the lack of personal interaction can make learning experiences feel disconnected. Moreover, standard platforms struggle to adapt to diverse user needs, leaving many students underserved. VoluProf introduces an immersive mixed reality (MR) experience, where students interact with volumetric avatars of professors integrated directly into their physical spaces (cf. Figure 1). These avatars replicate human expressions and gestures, creating a dynamic and engaging learning environment. Using MR headsets, students can seamlessly blend physical and virtual worlds, enhancing the depth and quality of their educational experience. The platform is built on a client-server architecture, where the server handles computationally intensive tasks like rendering and voice processing,

while the client manages real-time integration of avatars into students' environments. This split-rendering approach ensures accessibility, even on lower-end devices, by optimizing resource usage between server and client. Advanced technology underpins VoluProf's innovation. Volumetric animation, text-to-speech systems, and chatbots allow for lifelike interactions between students and virtual professors. These features foster real-time dialogue, enabling a personalized and engaging learning process. Ethical considerations are central to VoluProf's design. Privacy and data control are priorizised, ensuring transparent handling of user data. Inclusivity is another key focus, with avatars designed to reflect diverse cultural backgrounds and adaptive features that accommodate various learning styles and physical abilities. By creating an immersive, personalized, and accessible educational experience, VoluProf addresses the challenges of traditional online learning. It bridges the gap between physical and virtual education, offering an inclusive solution that sets new standards for engagement, adaptability, and accessibility in modern learning environments.

The project contributes to the Federal Ministry of Education and Research (BMBF)[1] announcement 2912 »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« (cf. BMBF, 2020). Central to this initiative is the BMBF's Human-Technology Interaction (HTI) research programme, which emphasizes interactive systems for knowledge and learning.

This article is structured as follows: In Section 2 we give an overview of related work. Section 3 describes the scenarios and section 4 gives a detailed overview of the developed technologies. In section 5, we discuss the user studies, ethical and legal aspects and conclude in section 6 with an overview of the overall system. In section 7 we conclude with a summary.

2. Related work

2.1 Educational technology

There have been many MR-based educational programs for decades. These have been used in several very different application areas such as mathematics (cf. Nikolakis, 2004: 93–102), teacher training (cf. Ke et al., 2016: 212–220) or mechatronics (cf. Mueller, 2003). The use of virtual reality in education is not new either (cf. Bricken, 1991: 178–184) and has led to demonstrable successes (cf. Kaminska,

1 At the time of printing: Federal Ministry of Research, Technology and Space (BMFTR)



Figure 1: VoluProf scenario
Source: Own illustration. Copyright: © iStock.com/TheItern/Ljupco/klikk edited by Fraunhofer HHI

2019: 318) – however, the use of MR technology for rich multimodal interaction with photo- and audio-realistic avatars of real teachers, as envisaged in this work, has not yet been explored.

An important aspect of the interaction with the avatar is speech and gesture-based discourse. There are systems that support the gestures of avatars in educational technologies (cf. Ke et al., 2016: 212–220), but these avatars are not designed to be realistic. There are also several learning bot systems that enable educational discourse between users and virtual teachers. For example, several universities in Germany and around the world use chatbots for various aspects of university information. These include the chatbots Clara (MLU Halle), Hermine Winf (TU Darmstadt), the chatbot at the University of Vienna, Sydbot (University of Sydney) and Pounce (Georgia State University). These chatbots support navigation on the institutions' websites and in some cases allow free text input. However, they do not yet allow longer natural language interactions and do not relate to courses. Clarizia (2018) and Azcona (2019) have each developed a chatbot that supports learners in computer science and programming courses according to their competencies. Azcona et al. shows that 87.1% of the answers and advice given by the chatbot were correct. Le et al. investigated how a chatbot can take the lead in group discussions and showed that the instructions and explanations were rated as useful (cf. Le & Huse, 2016: 69–78).

2.2 Volumetric Video and Animation

Techniques for creating realistic and interactive virtual humans classically utilize parametric models for geometry, reflection and lighting for human bodies (cf. Theobalt, 2007: 663–674) and faces (cf. Garrido, 2016: 1–15). They allow arbitrary animations, where the body movements are controlled by an underlying skeleton, while the facial expressions are described by a set of blendshapes (predefined basic shapes). However, a considerable amount of manual modelling is required to obtain high quality models, and true photorealism is still difficult to achieve: The results often appear fake (uncanny valley) and the parametric models are often unable to describe fine details such as loose clothing or complex reflections. In addition, subtle details such as the eyes or mouth are difficult to model. In recent years, advances in deep learning (cf. Yamaguchi, 2018: 1–14) have been used to fill in missing parts such as eyes (cf. Paier, 2020: 359–369) or to learn parameters for reflectance models (cf. Li, 2018: 1–11).

On the other hand, the generation of high-quality, photorealistic 3D models from video data captured with multi-camera systems is an active field of research (cf. Collet, 2015: 1–13; Ebner, 2017: 151–157; Guo, 2019: 1–19; Schreer, 2019:4310-4314; Zollhöfer, 2014: 1–12). In recent years, state-of-the-art volumetric acquisition and processing systems (cf. Collet, 2015: 1–13; Ebner, 2017: 151–157) as well as companies and start-ups such as 8i, Omnivor, Volucap, Microsoft and Metastage, have become established. These systems can generate high-quality photorealistic 3D video data for AR/VR systems. This data or the resulting virtual people can then be viewed from any angle, but this is limited to pre-recorded situations and lacks interactivity. To create an animatable model that resembles a real person, parametric models are usually adapted to the real data in terms of shape and body pose and then used for animation. However, these models have the same disadvantages as described above. A more recent trend is therefore to combine data from the real world with parametric models for animations: Recently presented methods combine Computer Graphic elements (parametric models) for animation with real-world data for realism (e. g. volumetric video, video-based textures) (cf. Paier, 2017: 784–797; Fechteler, 2019: 91–109) and allow the manipulation of the photorealistic real-world data itself (cf. Hilsmann, 2020: 350–358; Paier, 2020: 359–369).

2.3 Streaming of volumetric videos

The fifth generation of mobile networks, 5G, promises significant performance improvements over the current LTE standard, particularly in terms of data trans-

mission capacity. This advancement is especially beneficial for data-intensive audiovisual streaming services, which already account for approximately 70 % of global internet traffic—a figure expected to grow further. Beyond supporting existing services, 5G also enables innovative applications like those targeted by VoluProf, which require not only high bandwidth but also ultra-low response times and latencies.

Common visualization formats for volumetric data are polygon meshes and point clouds. The streaming of volumetric videos in such 3D data formats requires enormous bit rates (cf. Schwarz, 2018: 133–148). Some recent approaches in the literature present various optimizations for efficient streaming of point clouds, especially using adaptive HTTP streaming approaches (cf. van der Hooft, 2019: 2405–2413). Even if streamable bit rates can be achieved using such approaches and emerging 5G networks, the limited computing power of current mobile devices such as head-mounted displays is not sufficient to decode and render complex volumetric videos in real time (cf. Qian, 2019: 135–140). Furthermore, efficient hardware implementations of mesh/point cloud decoders are not yet available. Software-based decoding can be prohibitively expensive in terms of battery consumption and may also not be able to fulfill real-time rendering requirements (cf. Gül, 2020: 27–33).

One approach to significantly reduce the computational load on the mobile device is to offload the processing to a powerful cloud server and stream the result to the mobile device over a network (remote rendering) (cf. Shi, 2015: 1–29). Although remote rendering has previously been used in various application areas such as cloud gaming (cf. Shea, 2013: 16–21), its application for streaming volumetric videos poses new challenges due to the increased interactivity and freedom of movement of users.

2.4 Speech synthesis

In the field of speech synthesis, the unit selection approach was ›state of the art‹ until a few years ago. In this approach, the spoken language is dynamically pieced together from a large number of recordings of a speaker when the text is entered using sound characteristics (cf. Taylor, 2009: 474–516). While this approach can already achieve a naturalness that is almost indistinguishable from human speech with sufficient data material, this naturalness is often limited to the content domain for which the recordings are available.

In recent years, however, deep learning approaches have opened new possibilities for speech synthesis. These methods rely on end-to-end ›translations‹ from text to audio. This involves transferring a text into a mole spectrogram, which in

turn is converted into an audio signal with the help of a vocoder (cf. Wang, 2017: 4006–4010). This type of machine learning also makes it possible to influence the prosody and voice sound via additional embeddings when training the text-to-melspectrogram model.

3. Scenarios, user interaction and requirements

In the scope of VoluProf we wanted to answer some research questions related to the use of the new technology of photorealistic avatars like:

- Is the acceptance of the learning lessons higher if the avatar is known to the students?
- Does the acceptance of the avatar depend on the study subject?
- Do students see an advantage in using the new technology?
- Are genders having an impact on the learning process of the students?

To be able to answer as many of these questions as possible the consortium came up with two scenarios, which are briefly described in the following sections. In general, two avatars with photorealistic appearances and lifelike original voices were to be created and used during the project. One of the avatars would be female, the other male. Both avatars would wear neutral clothing suitable for both summer and winter.

A typical lecture situation, frontal teaching, was selected as the teaching format. In terms of content, one lecture from the humanities and one from the natural sciences should be used. Both scenarios were designed as an extension to face-to-face teaching.

3.1 Scenario 1

Table 1 summarizes the elements of scenario 1.

During the learning phase, slides, video and audio recordings and internet sources (websites) are used. The lecturer gives a lecture of about 10 minutes and questions about the lecture are only allowed after the lecture. The text template is provided by the real lecturer and questions/answers are generated from it. This can either be done manually by people or by using question/answer generation software (section 0). In the case of unknown questions, the script is searched for relevant keywords and the relevant text modules are repeated. If the question is still not answered, curated alternative learning materials are displayed. In this

Subject area	**Humanities**
Avatar/tutor	Female, known to students as a real person, possibly with physical and/or linguistic limitations
Teaching method	Lecture, frontal teaching followed by questions
Teaching materials	Slides, whiteboard, video and audio clips, online research, questions/test sheets to evaluate understanding
Learning phase	Lecture, questions/discussion with the lecturer after the lecture
Consultation phase	General questions about the lecture (overall, paragraph by paragraph, sentence by sentence), about the teaching materials

Table 1: Description of scenario 1
Source: Own representation

scenario we assume an ideal learning situation, i. e. with no spatial or social distractions.

In the subsequent consultation phase, the student could ask questions about the process, organization and content of the provided text template. If there are no suitable answers to the questions, the student is referred to the office hours of the known lecturer. A course evaluation form should then be completed.

3.2 Scenario 2

Table 2 summarizes the elements of scenario 1.

Subject area	**Natural science**
Lecturer	Male, not known to students as a real person, possibly physical or language restrictions
Teaching form	Lecture, frontal teaching, incl. experiment (video if necessary)
Teaching materials	Slides, whiteboard/chalkboard, video and audio clips, possibly model for experiment (e. g. digital molecular construction kit)
Learning phase	Lecture and discussion with lecturer, demonstration of experiment, questions can be asked at any time
Consultation phase	Questions about the lecture (complete, paragraph by paragraph, sentence by sentence), questions about the experiment (repetition of experiment)

Table 2: Description of scenario 2
Source: Own representation

The second scenario is similar to the first, but the avatar should be male, and the students can ask questions about the lecture at any time. In addition to slides, video and audio clips, other interactive elements such as an experimental model (e. g. a molecular kit) should be used where possible. Another difference from scenario 1 is that the learning situation may not be optimal, e. g. there may be a second person in the room, noise or interruptions, e. g. the doorbell ringing.

Due to the agile nature of research projects, not all requirements necessary to build the scenarios were met. The most serious hurdle encountered was the lack of video footage to create a male avatar, so Scenario 2 also had to be implemented with a female avatar. As we had two female avatars, we were able to stick to the requirement that one avatar was unknown to the students. Based on the two scenarios described above, the following ethical and privacy related points have been identified:

- The VoluProf application must be easily accessible and as barrier-free as possible, and information about the data collected must be transparently communicated to the users. Only the user data that is necessary for providing the functionality should be collected.
- Trust of users in the application should be increased by giving them control over individual components. The avatar should have a high degree of presence but still interact with the user as an equal, especially by praising them, to promote a positive feeling of interaction.
- Users do not always need to be attentive or active and the application should reflect the diversity of our society, e. g. through avatars with different body types, including disabilities.

The technical requirements of the VoluProf system were identified as adaptation to the user's technical equipment and the available internet connection. In addition, the functional scope of the VoluProf application depends on the user's permissions to use personal data. On the content side, the scripts for the lectures and additional materials such as curated bibliographies, administrative information or link lists are required.

4. System and technical modules

The design and implementation of the VoluProf platform involves the integration of a variety of tools and developments to achieve an interactive system that enables real-time multi-level interaction between a user and a virtual professor in an XR environment.

A general overview of the platform can be seen in figure 2. The architecture follows the client-server model, and it has been designed to be accessible in a remote manner.

The server, hosted on a cloud environment, carries out the most intensive data-management and graphic-processing tasks. A central element, named Controller, acts as scheduler and dispatcher of the different input and output data flows. The rest of the server modules are connected to the controller and

take care of one of the steps of the pipeline. This way, there are audio processing elements named »Speech 2 Text« and »Text 2 Speech« engines, in charge of the transcription and generation of audio respectively, and a »Chatbot« that holds the state of the playback of the lecture and reacts to the commands and questions emitted by the user. On top of this, the »Avatar Animation« module generates the 3D representation of the professor (mesh and texture) for each frame of the playback and the split rendering element oversees the rendering and streaming of the final video and audio.

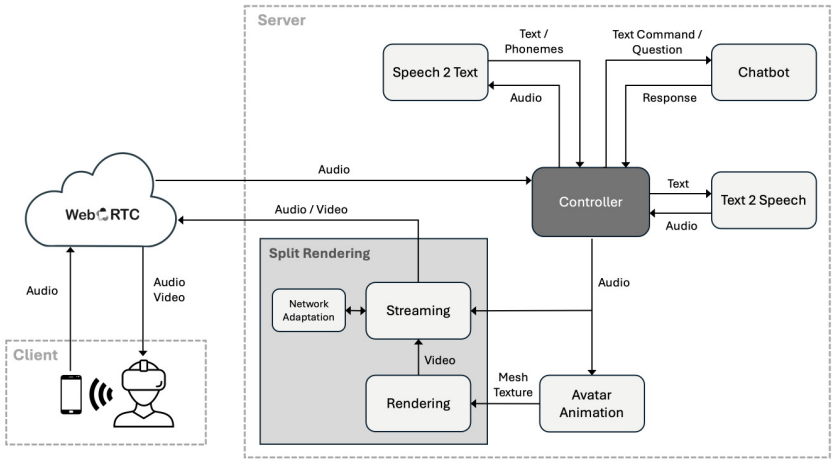


Figure 2: General architecture of the VoluProf platform
Source: Own illustration

The client application, on the other hand, is used to play the video and audio streams sent by the server and the inclusion of these representations in an XR environment, where the user can see the professor completely integrated in his/her actual surroundings. At the same time, this application allows the user to interact with the avatar, both by sending voice commands and questions to which the professor will react and by modifying the position of the avatar within the scene. This client application has been implemented for a variety of devices, with the intention of enabling access to the platform to as many users as possible. In this manner, mixed reality glasses, such as the Lenovo ThinkReality a3 (cf. Lenovo, 2020) model or VR headsets like the Meta Quest 3 (cf. Meta, 2020) are supported by the VoluProf platform. Apart from these XR-specific options, common devices such as Android phones and tablets can also be connected to the server.

4.1 Avatar Recording

The creation of a highly realistic avatar of a professor begins with an extensive and meticulous recording process conducted by Volucap. Utilizing cutting-edge volumetric capture technology, the professor is recorded in a controlled studio environment from 360 degrees, ensuring that every aspect of their appearance, expressions, and movements is captured with high fidelity.

4.1.1 Capture setup and equipment

The recording process employs a high-resolution camera setup consisting of:

– 38 Cameras at 65 Megapixels Each: This setup provides an extensive amount of detail, capturing the subject from multiple angles to create a comprehensive visual dataset with precise reconstruction of fine details.

4.1.2 Recording sessions

During the recording sessions, the professor performs:

– Three Different 10-Minute Talks: These sessions capture the professor's natural speech patterns, facial expressions, and gestural nuances, providing a rich dataset for creating an authentic avatar.
– A Series of Base Motions: Additional recordings of fundamental movements such as arm gestures, leg positions, and head rotations are captured. These base motions are essential for developing an animatable model that can perform a wide range of actions beyond the recorded sequences.

4.1.3 From raw data to high-resolution meshes

The initial raw data consists of multi-view video footage from all cameras, capturing the professor from every possible angle. This data is processed to generate high-resolution 3D meshes for each frame, resulting in temporal 3D mesh sequences. Each mesh comprises several million polygons, representing an incredibly detailed reconstruction of the professor's appearance and movements.

4.1.4 Compression and optimization process

To make these high-resolution meshes usable and deliver them in an optimized format for further use, a multi-stage compression and optimization process is applied:

1. Polygon Reduction: The original meshes with millions of polygons are computationally intensive and unsuitable for direct use. The meshes are therefore reduced to 380,000 polygons. This significant reduction is crucial for ensuring that the avatar can be rendered efficiently while maintaining a high level of visual fidelity.
2. Texture Optimization: The high-resolution textures derived from the camera footage are optimized to an 8k texture to ensure they retain sufficient detail for realistic rendering. This process preserves the essential visual characteristics of the professor while ensuring that the textures are suitable for further processing in the avatar pipeline.

4.1.5 Pipeline Summary

The entire process from capture to the final rendered avatar can be summarized in the following stages:

1. Capture: High-resolution volumetric data is captured using the multi-camera setup. The professor's performances and base motions are recorded in meticulous detail.
2. Initial Processing: The multi-view videos are processed to generate high-resolution 3D meshes for each frame, creating detailed temporal mesh sequences.
3. Compression and Optimization: Meshes are simplified, and textures are optimized to meet the performance requirements of the target platform. Advanced algorithms ensure that this reduction in data does not significantly impact visual quality.
4. Final Rendering: The optimized meshes and textures are assembled into an OBJ sequence, ready for further use. The use of standardized formats ensures compatibility with various rendering engines and platforms.

The resulting high-resolution meshes provide a rich dataset for AI training purposes. In the next step the data is utilized to develop advanced models for animation, gesture recognition, and interactive experiences. The detailed capture of the professor's unique expressions and movements allows for the creation of avatars that not only look like the professor but also behave authentically.

4.2 Avatar Modelling and Animation

As mentioned in section 4.1, the professor was captured in a volumetric studio from 360 degrees. The recorded multi-view videos were then processed into temporal 3D mesh sequences (cf. Eisert, 2022: 289–326). We captured three different

talk sequences of 10 minutes each as well as several base motions (arm, leg, head rotation etc.). Our goal was to not only replay the sequences but to learn an animatable model from this data as exemplarily illustrated in Figure 3.

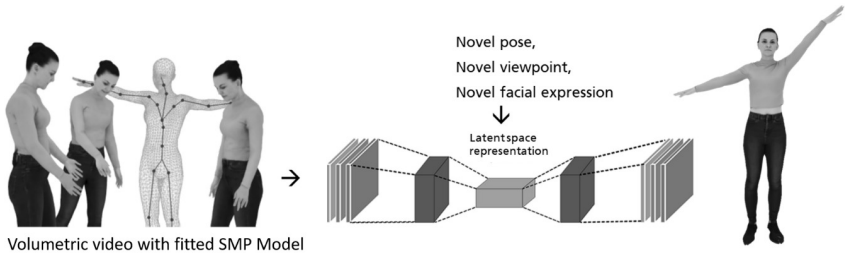


Figure 3: We learn an animatable human model avatar from real captured volumetric video data.
Source: Own illustration

4.2.1 Body modelling and animation

A volumetric video sequence typically consists of independent meshes for each frame, which complicates tasks such as modification, blending, animation, or data learning. To address this, we first temporally register the meshes of the volumetric video stream using the key-frame based method (cf. Zimmer, 2022: 123–139), yielding short sequences of temporally consistent topology. We then fit a parametric, kinematically animatable human body model (e. g., SMPL-X [cf. Pavlakos, 2019: 10975–10985]) to the temporally registered mesh sequences.

Model fitting and semantic enrichment: different methods exist to estimate temporally consistent model parameters to temporal mesh sequences, such as a modified smplify-X (cf. Zimmer, 2022: 123–139), PIXIE (cf. Feng, 2021: 792–804) and EasyMocap (cf. Shuai, 2022: 1–10). We processed the data with a variant of EasyMocap, which we extended with an additional loss term representing the average 3D distance between the fitted SMPL template mesh surface and the surface of the scan from the dynamic mesh sequence. In the optimization process, we ensured that the SMPL shape parameters for each identity were shared for all recorded frames (cf. Knoll, 2024: 404–413). The tracked template mesh, aligned with the dynamic mesh sequence, provides semantic labels and maintains a consistent topology, establishing correspondences over time. Additionally, it can be animated using an underlying skeleton, allowing motion transfer to the unstructured mesh sequence.

Kinematic template-based animation: by enriching the 3D mesh stream with semantic data through fitting a human body model with an underlying skeleton, it becomes possible to kinematically animate the volumetric video. This allows for moderately modifying or correcting body poses while preserving the high level of detail in the meshes and textures which were originally captured in the studio. This is accomplished by animating the template model using the skeleton input data and transferring the surface motion from the rigged template to the unstructured volumetric sequence. For each mesh vertex, its position relative to the nearest triangle of the template mesh is calculated. This relative position remains fixed, essentially »gluing« the mesh vertex to the template mesh triangle while maintaining a constant distance and orientation. As a result, the animation of the model can be directly applied to the volumetric video frame. This approach yields plausible results only for slight changes of the pose as pose-dependent geometry or appearance changes are not modeled. We use this approach to maintain eye contact with the user in the AR/VR scenario, enhancing the sense of real conversation with the virtual human. To achieve this correction, the user's position relative to the virtual human must be obtained, usually from headset tracking data. Using this information, the joints in the neck of the fitted template model can be adjusted to align the actor's face with the user's point of view. To prevent geometric artifacts or unnatural poses, joint rotation offsets can be limited to a specific range around the original pose, or a human body prior can be enforced.

Learning pose-dependent geometry and appearance: for more sophisticated body animation, we learn pose dependent appearance and geometry from the dynamic mesh sequences. Statistical body models like SMPL-X (cf. Pavlakos, 2019: 10975–10985) provide valuable prior knowledge that can be utilized to constrain the dimensionality of the search space, allowing for more efficient and focused learning while defining pose-dependency. Instead of learning absolute pose-dependent geometry directly, the difference between the observed geometry and the fitted SMPL model is learned (cf. Morgenstern, 2024: 2644–2650). This enables encoding of both pose-dependent appearance and geometry within the consistent UV space of the SMPL model. For this purpose, we further move the surface of the SMPL mesh to match the mesh's surface. The captured textures are then mapped onto the SMPL mesh displacement maps between the SMPL mesh surface and the scan data are calculated. After the network is trained, it can be used to generate an animatable virtual human in new poses.

Audio-driven gesture synthesis: to develop an approach for audio-driven animation, we developed a generative model designed to synthesize human motion from a

limited set of training data (cf. Moreno, 2024: 1–10). Our framework enables conditional generation and blending across different temporal scales. By leveraging skeletal convolution layers and a multi-scale architecture, the model effectively captures human motion dynamics. The model includes a combination of generative and adversarial networks, along with specialized embedding modules that generate motions at distinct frame rates while allowing control over motion content and detail.

4.2.2 Face modelling and animation

We reconstruct a personalized parametric head model that captures the 3D geometry, head motion, facial expressions, and appearance of the person being recorded (cf. Paier, 2020: 359–369, Paier 2021: 52–63; Paier, 2023). Our hybrid head model merges the benefits of lightweight, robust model-based representations with the realism of neural rendering techniques, enabling real-time animation, facial expression editing, and highly realistic rendering.

This model is learned from multi-view data collected using a multi-view video rig and a microphone to capture the speech signal. For animation control, we utilize a latent parameter vector from an autoencoder, which compresses the captured facial performance into a sequence of low-dimensional parameter vectors. This latent vector can then be driven by input data from text, speech, or video through a corresponding mapping network. To represent the speech content, we automatically annotate the captured voice signal with time-aligned phonemes using forced alignment. We train a neural animation model along with a speech style encoder using the annotated speech signals and corresponding animation parameters. This process separates the style and content of the captured performance. After training, the neural animation model can be conditioned with a low-dimensional style/emotion vector, enabling the synthesis of realistic and emotionally expressive talking head videos. For added realism, we train a neural rendering model that enhances rasterized images of the head model, allowing us to generate photorealistic videos of the animated head. Moreover, automatically learned foreground/background segmentation masks allow for simple integration of rendered head images with new backgrounds as well as with 3D/VR applications.

4.3 Split rendering

Streaming volumetric objects through a network in real time presents several technical hurdles, the main one being the magnitude of the data load to be

sent, especially when handling realistic representations of humans, where the complexity of the geometry is significant. On top of this, the already mentioned amount of data varies from object to object, making it difficult to assess what the exact requirements for the network must be.

To overcome this, a Split Rendering mechanism has been included in the VoluProf platform. It addresses the main challenge encountered when trying to show complex volumetric objects on a portable device: the rendering capabilities. Split Rendering offloads the most intensive part of the rendering process to a cloud server, where currently available hardware encoders can handle this task. Then, only fine-tuning steps of the rendering, such as background removal, are left to the client.

An element of the Split Rendering library present in the server named Object Pool handles the position, rotation and scale of the objects (avatar, blackboard) and the virtual cameras. This will serve as a central repository that can be used at any moment to obtain updated information. Using data channels such as WebSocket and the one provided by WebRTC (cf. Webrtc, 2024), the client will continuously inform the server about the viewpoint (position and orientation) of the mobile device/XR glasses and the objects within the 3D scene. The server will then apply these changes to the scene.

This outcome is a single image, or an image formed as a combination of two camera renders, placed side by side, in the stereoscopic case. The background is monochromatic. The rendered image is then encoded in a Gstreamer (cf. Gstreamer, 2024) pipeline and sent through the network using WebRTC.

When the client receives the image, the second part of the rendering process takes places. This time, the client application removes the monochromatic background (whose specific color is known by both server and client) and provides an image where the avatar is integrated in the real scene. Once the consumer user can see and hear the avatar, certain interaction functionalities, such as dragging and dropping the avatar, become available. As mentioned, after each one of these actions, the new parameters (position, rotation, scale) of the avatar are sent to the server, so it can act accordingly.

4.4 Transport

For seamless user interaction with the cloud server which performs the actual rendering for a particular viewpoint, the rendered image and audio need to be transmitted in real time with minimal latency. Modern 5G and Wifi networks provide the necessary technology basis. Furthermore, to ensure broad compatibility, the transmission protocol should adhere to established standards and be eas-

ily distributable. Accordingly, we exploit the Low Latency, Low Loss, and Scalable Throughput (L4S) (cf. Briscoe, 2023) congestion control mechanism, ensuring its implementation is compatible with WebRTC. Here, L4S, defined as a standard by the Internet Engineering Task Force (IETF), aims to ensure stable streaming by minimizing queue delay and jitter, and WebRTC is specifically designed to prioritize low latency in real-time communication protocols.

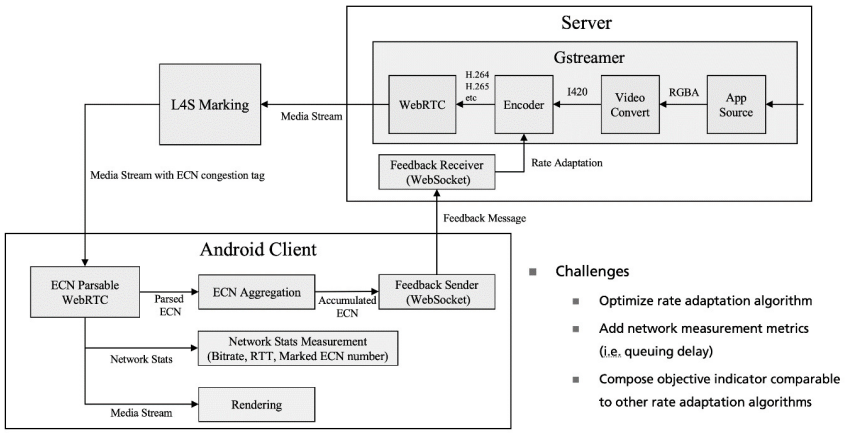


Figure 4: Media stream architecture
Source: Own illustration

Figure 4 illustrates the process of transmitting rendered video and audio via WebRTC, integrated with L4S. The media stream processed by the Gstreamer pipeline on the server is transmitted to the client through WebRTC over the 5G network which supports Explicit Congestion Notification (ECN) marking for L4S. More concretely, L4S utilizes a two-bit ECN field in the Internet Protocol (IP) header to signal queuing delay information (cf. Briscoe, 2023), and in practice, the scheduler in the 5G network performs the ECN marking proportionally to the queuing delay. In (cf. Son, 2024: 46–52) we have designed an algorithm that minimizes variation in latency with detailed results given in (cf. Son, 2024: 474–516).

4.5 Audio synthesis

The previously prepared content, the VoluProf avatar, is going to lecture about any topic. For that, a highly flexible method of converting Text to Speech (TTS)

is needed that enables the conversion of any kind of textual input into spoken audio.

The TTS component consists of two models: one model based on Tacotron (cf. Wang, 2017: 4006–4010; Shen, 2020) to generate a mel-scale spectrogram from a phone-sequence and a vocoder model (cf. Yang, 2021: 492–498), which uses the mel-scale spectrogram to generate the audio signal. To give the students a fully immersive experience and to improve learning success we aim to show the volumetric avatar not only as a photo-realistic version of the original teacher but also equip him/her with a TTS voice version of the original person.

4.5.1 Challenges of data quality and quantity

The previous approach for training such TTS models required hours of audio data to generate robust models. In the case of VoluProf the challenge arose that it is hardly acceptable for teachers to do recordings for several hours for about 5 or more days. This would not be feasible for teachers on top of their other duties in university or school. Nevertheless, the training approach so far does not create sufficient results with an amount of audio data that would be reasonable for teachers to record between their classes. Apart from the amount of data itself, the quality of the resulting TTS voice is influenced by the quality (e. g. pronunciation, recording quality etc.) of the training dataset. An hour of unsupervised home recordings will most likely not meet the quality criteria.

However, supervised audio recordings are time consuming and less likely to fit into usually busy schedules. For VoluProf, we tried an approach that consisted of two steps. First, we recorded the voice of a lecturer at the University of Rostock for 3–4 hours in a supervised audio recording session. Secondly, we augmented this data to increase the number of training samples.

4.5.2 Overcoming data scarcity

To tackle the problem of requiring a large dataset to create a cloned voice, speaker embeddings were introduced. These are generated by a speaker encoder model, which is trained by a large dataset of short audio samples of varying speakers. The model learns the differences of these speakers and generates speaker embeddings. In production, the speaker encoder gets a sample of the avatar and returns an embedding of the particular voice. This embedding can be introduced in an extra layer of the Tacotron model, which generates the Mel spectrogram. The resulting spectrogram is expected to resemble the voice of the speaker for the avatar.

4.6 Interactive speech interface

The purpose of the interactive speech interface is twofold: firstly, it provides a command-and-control interface, allowing the user to perform actions like restarting previous sessions, selecting lectures or navigating during a lecture. Secondly, it allows the student to pose questions in natural language and also get the answer as output spoken by the avatar, if such an answer is available. The runtime system of the interface builds upon the results of an offline process that extracts metadata, questions, and synopses from presentation data (optionally with notes) or lecture scripts. The offline processing pipeline, which is sketched in figure 5, takes presentations with optional text comments and lecture scripts to produce the following data:

- Presentation and text data for the automatic presenter.
- Metadata for the controller and interactive component, e. g., titles, synopses, chapter lists etc.
- Additional question/answer pairs for the Q/A component, which are then manually checked.

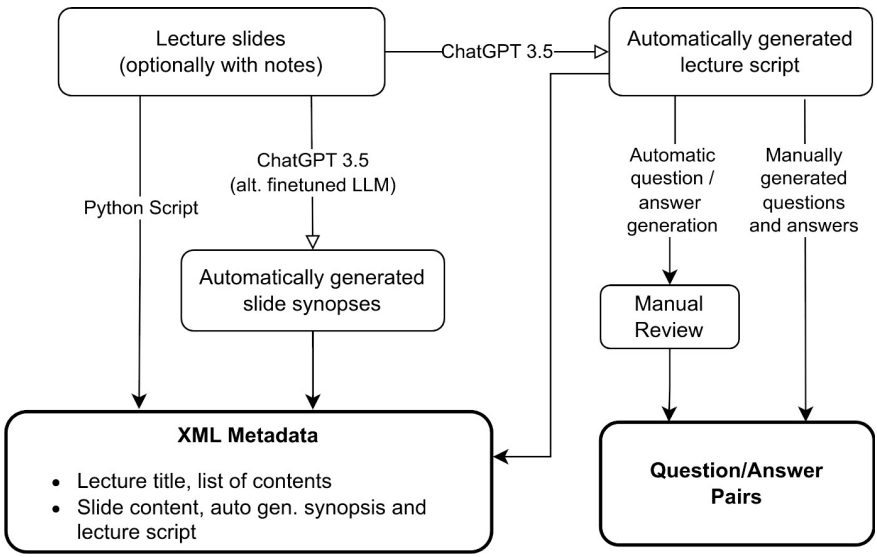


Figure 5: Runtime system architecture
Source: Own illustration

The architecture of the runtime part gets its input as text transcriptions from the automatic speech recognition provided by the project partner Aristech, which

also provides the text-to-speech functionality. To identify if the incoming text is a command, it is analyzed using a natural language understanding (NLU) model. For this, a corpus of possible utterances for the available commands was created and processed with the rasa NLU framework (cf. Rasa, 2024) to train a model. Rasa was used since it can cope with relatively little training data and the resulting model is very compact and needs very little time for inference.
The dialogue processing is implemented with the VOnDA framework (cf. Vonda, 2024), which was designed for asynchronous social interaction with embodied agents, such as robots or avatars. VOnDA is a hybrid rule-based / statistical system that was inspired by the Information State / Update paradigm. It organizes the communication with the main controller and takes care of handling the dialogue state.
The command interface understands around 24 different commands, where about half of them affect the look and feel of the system, like audio volume, subtitle size or showing or hiding an info board or the lecture list, and the other half are navigation commands like starting, stopping or pausing the lecture or moving between chapters with absolute or relative references (»go to previous chapter«). In case the NLU does suggest that the incoming string is a question or the probability for all command categories is too low, the string is passed to the question answering module to see if the input is possibly a question concerning the lecture.

The question answering module consists of two parts: One with hand curated question/answer pairs from the content provider, and one which uses extractive question answering from lecture text data to find answers which are not in the curated set. In the project, we deliberately do not use internet sources to find possible answers because the correctness of these answers cannot be verified at run time, and we prefer to refer the student back to the human professor rather than giving wrong information.

4.7 Integrated system and demonstrator

The integration of the different technical modules described in this section has brought multiple challenges, especially related to the desired real-time performance of the VoluProf platform. Several iterations of the demonstrator were required to refine aspects such as the latency on the transmission pipeline, the proper synchronization between the animation of the avatar and the audio or the smoothness of the interaction via voice questions and commands. However, a fully functional demonstrator has been successfully integrated, with all the capa-

bilities described in this document. Figure 6 shows screenshots of the VoluProf platform using two of the available avatars.

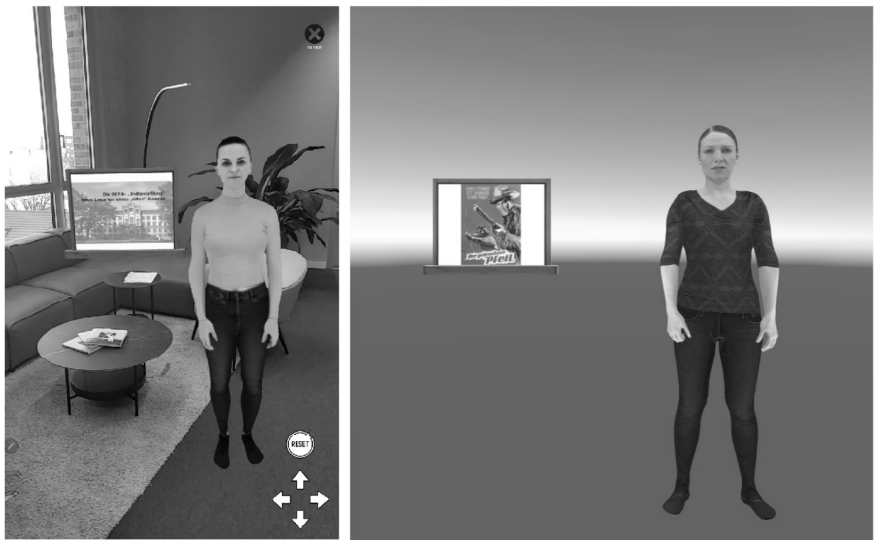


Figure 6: Screenshots of the integrated VoluProf platform demonstrator
Source: Own illustration

The demonstrator allows the user to log into the system by selecting a specific account and lecture. This way, the chatbot keeps the current state of the playback, which can be altered via voice commands. At any time, the user can stop the lecture and formulate a question that will be answered. Aspects such as the presence of subtitles or the size of the blackboard can also be controlled. In addition to this, the user can enjoy an immersive experience by receiving audio and video of the virtual elements that are integrated in the room, while being able to move around and see the objects in 3D. These objects can also be moved and relocated in the scene. The professor will maintain eye contact with the user to enhance the experience. The demonstrator can be used with a variety of devices such as phones, tablets and XR/VR glasses.

5. User studies, ethical and legal discussion

5.1 Ethical and legal discussion

The special feature of ethical and user-centered design in the VoluProf project is the symbiotic integration of ELSI/ethical aspects and user acceptance (UA) dimensions. This was made possible by an interdisciplinary collaboration between ethics and communication and media sciences at the University of Rostock. Our approach drew on both state-of-the-art Value Sensitive Design (VSD) and Human Centered Design (HCD) principles, enabling a discursive and participatory mixed-methods approach to technology development, involving different stakeholders such as engineers, ethics and communication experts and potential users (Friedman & Hendry, 2019; Norman, 2023). In our work packages, we used a mixed-methods approach to integrate ELSI/ethical considerations and insights from human-machine interaction into the development process. This included literature review, conceptual analysis, expert workshops, participant observation and empirical user studies.

We employed a value-anticipatory and exploratory methodology, beginning with predictive scenarios to evaluate the technology and its potential future applications. This informed the development of a checklist based on scientific literature and ethics guidelines that incorporates ELSI considerations and UA factors, highlighting key ethical, social and legal issues and facilitating discourse on ELSI and UA. For the VoluProf project, we identified the four following categories: (1) acceptance (usefulness, necessity, and trust); (2) control (data sovereignty, privacy); (3) diversity (gender, accessibility, and social equity); (4) connectivity (alignment with the higher education context, collaborative learning, enhancement of learning and communication).

In the second step, we aimed to empirically validate, specify and expand the ELSI and UA checklist and to explore ELSI and UA issues through stakeholder workshops and user studies. We conducted a mixed-method study involving various stakeholders, including three distinct tracks:

1. First, we used a discourse-ethical, participatory panel discussion, known as a MEESTAR workshop (Model for the ethical evaluation of socio-technical arrangements; see Manzeschke et al., 2013). Using predictive scenarios, this ELSI discourse format allows for the exploration of different dimensions of value and user acceptability, as well as different levels (individual, organisational, societal, technical) and risk assessment categories (harmless, sensitive, problematic and unacceptable). Our distinctive approach to the MEESTAR included the use of our own value and UA categories and the explicit inclusion

of user acceptability concerns in the methodology. The ELSI and UA aspects can be in tension but are mutually enriching and need to be carefully balanced. In addition, we carried out a validation study (via questionnaires) to measure the effectiveness of the MEESTAR method.

Our empirical-inductive analyses of MEESTAR showed the importance of the following evaluation categories: »interaction« (quality and authenticity of human-machine interaction); »digital lecturer/teaching person« (authenticity of voice and facial expressions; acceptance by students; control by the lecturer over his/her digital twin); »application and system design« (privacy and security; degree of autonomy of AI-based functions; inclusive design; technical feasibility); »data management« (data sovereignty, privacy, transparency, data rights); »teaching methods and patterns« (teaching methods, quality of teaching materials, power relations, impact on learning outcomes, flexible learning environments); and »framework conditions« (equity and diversity, technical requirements, institutional ›fit‹).

2. The qualitative (and, to a lesser extent, quantitative) user studies (group discussions with students, and individual qualitative interviews with lecturers) included semi-structured questionnaire with open-ended questions, facilitating more nuanced and unstructured responses to explore ELSI as well as UA issues. The first user study was based on predictive scenarios, while subsequent studies also involved a demonstrator experience of the technological application. For these semi-structured group discussions, we employed both our ELSI and UA checklist as well as the results of the MEESTAR workshop and literature reviews. The demonstrator experience enabled participants from the potential user group to gain firsthand knowledge of the application's actual experience, moving beyond purely imaginative predictive scenarios.
3. In a third, separate track, we validated and discussed the ELSI and user acceptability checklist, the predictive scenarios, and the risks and benefits of the VoluProf technology with expert perspectives from the Advisory Board workshop.

   Experts from the fields of applied ethics, educational sciences, AI, and extended reality emphasized the importance of gradually and exploratively testing the application with users, incorporating a feedback loop to enhance its quality. They also stressed the need for transparent data management and clarity regarding AI-generated teaching inputs or materials. Furthermore, attention was drawn to the quality of social interaction between the digital lecturer and students, ensuring a respectful distance is maintained and being mindful of factors such as the size and voice of the digital teacher. The importance of adaptive or personalized user options, such as adjustable speed and subtitles, was also highlighted.

The results of all these methodological steps were incorporated into the ELSI and UA guidelines, which were updated and fine tuned throughout the project, and which serve as our »soft« Key Performance Indicators (KPIs) for technology development. The intermediate results and KPIs were made available to the VoluProf project team both in regular project meetings and in ongoing communication with partners developing different technical aspects of our VoluProf technology.

5.2 User Studies

From a methodological point of view, the most central tool in our Ethical, Legal and Social Implications (ELSI) and User Acceptance (UA) research was to conduct user studies. The empirical user studies focus on the expectations and needs of the potential user groups in relation to the project and their use of the VoluProf application. The underlying data material for the (qualitative and, to a lesser extent, quantitative) analysis consists of individual interviews with the lecturers (teachers) and group discussions with the students (learners). The content of the interviews was reduced, paraphrased and the core statements listed in a category system (categories 1–6) (Mayring 2014). The results provide an overview of the identified expectations and needs of the user groups.

5.2.1 Interaction

The digital human twin of a lecturer should establish a close relationship with the students, take their individual needs and social norms seriously and encourage a culture of interacting and discussing as equals. At the same time, the digital lecturer should moderate social norms without replacing the real teacher. He or she should be able to show emotions, such as humour, a good or bad mood or other feelings, and at the same time respect the students' emotions. The digital lecturer should actively involve the students, make eye contact, give affirmation and respond to requests to speak in order to increase the students' motivation to learn.

5.2.2 Data management and law

The security and privacy of users must be guaranteed. The uncontrolled dissemination of educational content should be restricted. The involvement of external service providers or institutions in the development and licensing of the application should be made transparent. The legal responsibility for the statements of the digital lecturer and the use of teaching material must be clearly regulated. It

should be ensured that the answers of the digital lecturer are correct and comparable with the real person's answers. Incorrect information can lead to disadvantages for students (e. g. in exams).

5.2.3 Application and system design

The application should use different forms of teaching and learning. The volume should be adapted to the room size. The user's range of movement should be restricted as little as possible (see section 5.2.8 about Proxemics). Changes and additions in the application should only be made with the needs of the users in mind and the consent of the lecturer. The application must be understandable for all users and easy to use or barrier-free (e. g. physical or cognitive limitations). Information on accessibility should be provided within the application and questions answered in order to reduce inhibitions.

5.2.4 Teaching and learning setting

The digital lecturer should enable a personalized learning experience in which users can set their own learning times and speeds. Students should be actively involved and allowed to contribute their own ideas. Content should only be provided and updated by qualified teaching staff. The application should be scientifically evaluated on a regular basis.

5.2.5 Digital lecturer

The relationship and hierarchy between teachers and learners should be emphasized in order to create the conditions for authority, respect and trust in the teaching-learning setting. The acceptance of the lecturer is of great importance, as users may test the limits of the application. Real lecturers who use the application should be actively involved in the onboarding process. Characteristics and mannerisms of the lecturers should be adopted, including language, voice, gestures, facial expressions and physique. Realistic speech and movement skills, especially in the upper body have priority. A complete audiovisual match with the real lecturer is not necessary. Realistic clothing and the adoption of features (e. g. glasses, hair) of the real person increase the authenticity of the digital human twin and thus the acceptance of the users.

5.2.6 Framework conditions and social responsibility

The effective use of the application in the teaching-learning process not only requires an existing technical infrastructure but also technical and media skills among teachers and learners. Unrestricted access to technological equipment is just as important as support in using the application. The amount of work involved in using the application should not exceed that of classroom teaching.

5.2.7 Reflection and discussion

One of the challenges in the interdisciplinary and participatory collaboration in technology development and the iterative research involving ELSI and UA is finding a common language in order to utilize the different technical, engineering, humanities and social science perspectives. In addition, the dependence of empirical research on technological development proved to be challenging, as ELSI and UA research must always react to changes in technological development in the event of technical delays or changes in the application. Predictive scenarios can partially compensate for this shortcoming, but they cannot fully replace firsthand knowledge of the actual experience of the application.

A clear strength of our empirically based approach was the early involvement in the design process. This made it possible (despite the challenges mentioned above) to reflect on value change and potential negative impacts early in the development of the VoluProf application and to involve different stakeholders – such as developers, experts and (potential) users – in the critical analysis and design process of the VoluProf technology.

Our early-stage ELSI and UA research can also be challenged by the fact that, despite predictive scenarios, it does not fully capture the uncertainties and unknowns related to the widespread use of VoluProf technology in society (such as its impact on the education system, or on the concept of »education«). While this may be true, it is also too late to begin addressing ELSI and UA concerns once the technology is already widespread. Our mixed-methods study made it possible, in an empirically grounded way, to facilitate a critical discourse on the kind of impact we want our technology to have (before it produces its predicted – or unpredicted – effects).

The user studies particularly emphasized the authenticity of the interaction with the digital human twin of a lecturer, focusing on relationship, emotions, natural dialogue, respect, and active participation. The digital lecturer should offer a realistic user experience, though there was no expectation of a perfect audiovisual match for authenticity. At the same time, there was critical awareness of poten-

tial negative effects, such as social isolation and concerns about data security and privacy.

The studies also stressed the need to ensure the quality and transparency of the teaching materials. The system needs to be user-friendly, adaptable to different environments and address issues of diversity and accessibility. It was expected to support personalised learning and encourage active student participation. Adequate technical infrastructure and support should ensure ease of use without increasing the workload compared to traditional teaching and learning.

In conclusion, there was a high level of acceptance of this new technology among the various user groups of the application. Both students and lecturers saw advantages in the temporal and spatial flexibility of the application and the improved accessibility for restricted user groups. However, all respondents pointed out that the application is not intended to replace classroom teaching, but merely to supplement it, e. g. for individual exam preparation or to relieve the burden on the teacher or the university (space issues).

These results reflect important concerns that inform both our technology design in the VoluProf project and facilitate an ongoing ELSI and UA discourse for emerging mixed reality and AI technologies in higher education. However, these findings need to be balanced and synthesized with insights from stakeholder workshops, the scientific advisory board, and the academic literature. Achieving a »reflective equilibrium« (Rawls, 1971; Daniels, 1996) between the different »voices« in our mixed method ELSI and UA approach allows for acceptable coherence across various sources of considerations and information. This, in turn, supports the creation of a well-balanced ELSI and UA checklist/guideline that informs the design of VoluProf technology.

5.2.8 Proxemics

In human interaction, interpersonal distance (IPS) functions as a significant nonverbal cue, reflecting the nature and closeness of social relationships (cf. Hall, 1966). This concept is central to the study of proxemics, which provides a framework for understanding how spatial behaviour influences and is influenced by social interactions. Closer interpersonal distances are typically observed between individuals with intimate or close relationships, whereas greater distances are maintained between strangers. The violation of personal space boundaries can lead to feelings of arousal or discomfort (cf. Hayduk, 1978).

Given the flexibility to position avatars within virtual environments, this research investigates whether the spatial placement of a VoluProf can be optimized to enhance students' attention levels. Building upon previous research demonstrating similar discomfort and arousal responses to IPS violations in virtual and

real-world settings using subjective reports and objective measures of electrodermal activity (EDA) (cf. Tootell, 2021), we designed a study that additionally utilizes electroencephalography (EEG) to identify neural correlates associated with changes in IPS. Participants were approached by an avatar from an initial distance of 8 m. The design manipulated two factors: personal space violation (present or absent) and the avatar's facial expression (angry or neutral). The experiment encompassed 60 trials, presented in 4 blocks, and each trial consisted of an approach that varied based on the combination of the two experimental factors. To control for habituation effects, we measured the subjective personal space boundary before each block and adjusted the presented distance accordingly. 64-channel EEG data were recorded during all trials. After artifact removal via independent component analysis (ICA), our analysis focused on the relationship between alpha oscillations, maximized via spatio-spectral decomposition (SSD) (cf. Nikulin, 2011: 1528–1535), and avatar distance using Source Power Comodulation (SPoC) analysis (cf. Dähne, 2014: 111–122).

Increased EDA was observed when the virtual avatar entered the participant's personal space, indicating heightened arousal. EEG analysis, focusing on the relationship between alpha oscillations and avatar distance, revealed responses within parietal brain regions when the avatar intruded into their personal space. These regions are associated with attentional processes, suggesting that IPS intrusion may increase attentional engagement.

Based on these findings, an open question is to investigate the influence of the IPS between VoluProf and Student on learning outcomes and susceptibility to distraction. Concentration has been shown to shield against distraction due to reduced background processing and a more robust attentional focus (cf. Sörqvist, 2015: 267–272). To test whether the positioning of the VoluProf, i. e. the IPS between VoluProf and student, affects students' concentration during the encoding of the learning content, one could position the VoluProf at three different distances and measure the student's susceptibility to distraction. In such an experiment, the VoluProf will be positioned at three distances relative to the participant: within their subjective personal space, slightly outside their personal space, and at a further distance. At each distance, participants follow the lecture of the VoluProf while distractors such as a ringing phone or a text message sound are incorporated into the VR-environment. Participants gaze needs to be captured to measure attentional shifts towards these distractors. Additionally, participants' comprehension of the lecture needs to be assessed. Taken together these future studies would allow a better understanding of spatial placement of the professor in a virtual classroom scenario.

6. Summary

In conclusion, the VoluProf platform represents an approach to online education by integrating mixed reality (MR) technologies with advanced volumetric avatars. By creating immersive and interactive learning environments, VoluProf addresses key challenges in modern education, including student engagement, interactivity, and accessibility. Through its sophisticated client-server architecture, split rendering, and real-time avatar animation, the platform ensures that students can access high-quality, personalized learning experiences from anywhere. Moreover, VoluProf places significant emphasis on ethical considerations, inclusivity, and user-centered design, ensuring that the platform meets diverse learning needs while maintaining privacy and data control. As education continues to evolve in the digital age, VoluProf has the potential to foster more meaningful connections between students and virtual educators.

Funding information

This Project VoluProf is funded by the German Federal Ministry of Education and Research (BMBF) [2] within the funding line »Interactive systems in virtual and real spaces – Innovative technologies for the digital society«.

Literature

Azcona, David et al. (2019). CoderBot: AI chatbot to support adaptive feedback for programming courses. In: *Proceedings of 20th Internationale Conference Artificial Intelligence in Education 2019*, Chicago, vol. 11625, Springer, Cham. https://doi.org/10.1007/978-3-030-23207-8

Balakrishnan, Guha et al. (2018). Synthesizing Images of Humans in Unseen Poses. In: *2018 IEEE /CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, p. 8340–8348. https://doi.org/10.48550/arXiv.1804.07739

BMBF (2020). Richtlinie zur Föderung von Forschung und Entwicklung auf dem Gebiet »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft«, Bundesanzeiger, Bundesministerium für Bildung und Forschung, Berlin. https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2020/04/2912_bekanntmachung.html

Bricken, Meredith (1991). Virtual reality learning environments: Potentials and challenges.In: *ACM SIGGRAPH Computer Graphics*, vol. 25(3), p. 178–184. https://doi.org/10.1145/126640.126657

2 At the time of printing: Federal Ministry of Research, Technology and Space (BMFTR)

Briscoe, Bob et al. (2023). Low Latency, Low Loss, and Scalable Throughput (L4S) Internet Service: Architecture. In: *IETF Internet Egineering Task Force*, RFC 9330. https://doi.org/10.17487/RFC9330

Clarizia, Fabio et al. (2018). Chatbot: An education support system for students. In: Castiglione, Arcangelo et al. (eds.), *Cyberspace safety and security*, vol. 11161, p. 291–302. https://doi.org/10.1007/978-3-030-01689-0_23

Collet, Alvaro et al. 2015). High-quality Streamable Free-viewpoint Video. In: *ACM Transactions on Graphics*, vol. 34(4), p. 1–13. https://doi.org/10.1145/2766945

Dähne, Sven et al. (2014). SPoC: A novel framework for relating the amplitude of neuronal oscillations to behaviorally relevant parameters. In: *NeuroImage*, vol. 86, p. 111–122. https://doi.org/10.1016/j.neuroimage.2013.07.079

Ebner, Thomas et al. (2017). Multi-view reconstruction of dynamic real-world objects and their integration in augmented and virtual reality applications. In: *Journal of the Society for Information Display*, vol. 25(3), p. 151–157. https://doi.org/10.1002/jsid.538

Eisert, Peter et al. (2022). Volumetric Video – Acquisition, Interaction, Streaming and Rendering. In: Valenzise, Guiseppe et al. (eds.), *Immersive Video Technologies*, Academic Press, p. 289–326. https://doi.org/10.1016/B978-0-32-391755-1.00017-1

Fechteler, Philipp, Hilsmann, Anna & Eisert, Peter (2019). Markerless Multiview Motion Capture with 3D Shape Model Adaptation. In: *Computer Graphics Forum*, vol. 38(6), p. 91–109. https://doi.org/10.1111/cgf.13608

Feng, Yao et al. (2021). Collaborative Regression of Expressive Bodies using Moderation. In: *Proceedings of the 2021 International Conference on 3D Vision (3DV)*, London, p. 792–804. https://doi.org/10.1109/3DV53792.2021.00088.

Garrido, Pablo et al. (2016). Reconstruction of personalized 3D face rigs from monocular video. In*: ACM Transaction on Graphics,* vol. 35(3), no. 28, p. 1–15. https://doi.org/10.1145/2890493

GStreamer (2024). *GStreamer Open Source Multimedia Network*. https://gstreamer.freedesktop.org [20.03.2025]

Gül, Serhan et al. (2020). Low-latency cloud-based volumetric video streaming using head motion prediction. In: *Proceedings of the 30th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, NOSSDAV '20, Istanbul, p. 27–33. https://doi.org/10.1145/3386290.3396933

Guo, Kaiwen et al. (2019). The relightables: volumetric performance capture of humans with realistic relighting. In: *ACM Transaction of Graphics*, vol. 38(6), no. 217, p. 1–19. https://doi.org/10.1145/3355089.3356571

Hall, Edward Twitchell (1966). *The hidden dimension*, Doubleday, New York.

Hayduk, L. A. (1978). Personal space: An evaluative and orienting overview. *Psychological Bulletin*, vol. 85(1), p. 117–134. https://org/doi/10.1037/0033-2909.85.1.117

Hilsmann, Anna et al. (2020). Going beyond Free Viewpoint: Creating Animatable Volumetric Video of Human Performances. In: *IET Computer Vision*, vol. 14(6), p. 350–358. https://doi.org/10.1049/iet-cvi.2019.0786

van der Hooft, Jeroen et al. (2019). Towards 6DoF HTTP Adaptive Streaming Through Point Cloud Compression. In: *Proceedings of the 27th ACM International Conference on Multimedia MM '19*, p. 2405–2413. https://doi.org/10.1145/3343031.3350917

Kamińska, Dorota (2019). Virtual reality and its applications in education: Survey. In: *Information 2019*, vol. 10(10), 318, p. 1–20. https://doi.org/10.3390/info10100318

Ke, Fengfeng, Lee, Sungwoong & Xu, Xinhao (2016). Teaching training in a mixed-reality integrated learning environment. In: *Computers in Human Behavior*, vol. 62, p. 212–220. https://doi.org/10.1016/j.chb.2016.03.094

Knoll, Paul et al. (2024). Animating NeRFs from Texture Space: A Framework for Pose-Dependent Rendering of Human Performances. In: *Proceedings of the 19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP)*, Rome, p. 404–413. https://doi.org/10.5220/0012373800003660

Le, Nguyen-Thinh & Huse, Nico (2016): Evaluation of the Formal Models for the Socratic Method. In: *ITS 2016: Proceedings of the 13th Int. Conf. Intelligent Tutoring Systems*, vol. 9684, p. 69–78. http://dx.doi.org/10.1007/978-3-319-39583-8_7

Lenovo (2020). *ThinkReality A3 Smart Glasses,* Lenovo Ltd., Honkong. https://www.lenovo.com/us/en/thinkrealitya3/ [20.03.2025]

Li, Zhengqin et al. (2018). Learning to Reconstruct Shape and Spatially-varying Reflectance from a Single Image. In: *ACM Transactions on Graphics*, vol. 37(6), p. 1–11. https://doi.org/10.1145/3272127.3275055

Meta (2020). *Meta Quest 3 VR-Headset.* Meta Inc., Menlo Park. https://www.meta.com/de/en/quest/quest-3/ [20.03.2025]

Moreno-Villamarin, David Eduardo Hilsmann, Anna & Eisert, Peter (2024). Multi-Resolution Generative Modeling of Human Motion from Limited Data. In: *CVPM '24: The 21st ACM SIGGRAPH European Conference on Visual Media Production*, no. 7, p. 1–10. https://doi.org/10.1145/3697294.3697309

Morgenstern, Wieland et al. (2024). Animatable virtual humans: Learning pose-dependent human representations in UV space for interactive performance synthesis. In: *IEEE Transactions on Visualization and Computer Graphics*, vol. 30(5), p. 2644–2650. https://doi.org/10.1109/TVCG.2024.3372117

Mueller, Dieter & Martins Ferreira, José Manuel (2003). MARVEL: A mixed-reality learning environment for vocational training in mechatronics. In: *Proceedings of the Technology Enhanced Learning International Conference (TEL 03)*. https://repositorio-aberto.up.pt/bitstream/10216/84622/2/56048.pdf [Zugriff 20.03.2025]

Nikolakis, Georgios et al. (2004). A mixed reality learning environment for geometry education. In: *Proceedings of the Hellenic Conference on Artificial Intelligence*, Samos, p. 93–102. http://dx.doi.org/10.1007/978-3-540-24674-9_11

Nikulin, Vadim V., Nolte, Guido & Curio, Gabriel (2011). A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition. In: *NeuroImage*, vol. 55(4), p. 1528–1535. https://doi.org/10.1016/j.neuroimage.2011.01.057

Paier, Wolfgang et al. (2017). A Hybrid Approach for Facial Performance Analysis and Editing. In: *IEEE Transactions on Circuits and Systems for Video Technology*, vol.27, no.4, p. 784–797. https://doi.org/10.1109/TCSVT.2016.2610078

Paier, Wolfgang, Hilsmann, Anna & Eisert, Peter (2020). Interactive facial animation with deep neural networks. In: *IET The Institute of Engineering and Technology, Computer Vision,* vol. 14(6), p. 359–369. https://doi.org/10.1049/iet-cvi.2019.0790

Paier, Wolfgang, Hilsmann, Anna & Eisert, Peter (2021). Example-Based Facial Animation of Virtual Reality Avatars using Auto-Regressive Neural Networks. In: *IEEE Computer Graphics and Applications*, p. 52–63. https://doi.org/10.1109/MCG.2021.3068035

Paier, Wolfgang, Hilsmann, Anna & Eisert, Peter (2023). Unsupervised learning of style-aware facial animation from real acting performances. In: *Graphical Models, Science Direct*, vol. 129, no. C, 101199, p. 1–16. https://doi.org/10.1016/j.gmod.2023.101199

Pavlakos, Georgios et al. (2019). Expressive body capture: 3D hands, face, and body from a single image. In: *Proceedings of the IEEE /CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 10975–10985. https://doi.ieeecomputersociety.org/10.1109/CVPR.2019.01123

Qian, Feng et al. (2019. Toward Practical Volumetric Video Streaming on Commodity Smartphones. In: *Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications (HotMobile '19)*, Santa Cruz CA USA, p. 135–140. https://doi.org/10.1145/3301293.3302358

RASA 2024. *Rasa Platform*. Rasa Technologies Inc., Berlin. https://www.rasa.com. [20.03.2025]

Schreer, Oliver et al. (2019). Capture and 3d video processing of volumetric video. In: *2019 IEEE International Conference on Image Processing (ICIP)*, Taipei, p. 4310–4314. https://doi.org/10.1109/ICIP.2019.8803576

Schwarz, Sebastian et al. (2018). Emerging MPEG standards for point cloud compression. In: *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, p. 133–148. https://doi.org/10.1109/JETCAS.2018.2885981

Shea, Ryan et al. (2013). Cloud gaming: architecture and performance. In: *IEEE network*, vol. 27, no. 4, p. 16–21. https://doi.org/10.1109/MNET.2013.6574660

Shen, Jonathan et al. (2020). Non-attentive Tacotron: Robust and controllable neural TTS synthesis including unsupervised duration modeling. In: *arXiv preprint arXiv:2010.04301*. https://doi.org/10.48550/arXiv.2010.04301

Shi, Shu & Hsu, Cheng-Hsin (2015). A survey of interactive remote rendering systems. In: *ACM Computing Surveys (CSUR)*, 47(4), no. 57, p. 1–29. https://doi.org/10.1145/2719921

Shuai, Quing et al. (2022). Novel View Synthesis of Human Interactions from Sparse Multi-view Videos. In: *ACM SIGGRAPH 2022 Conference Proceedings (SIGGRAPH '22)*, Association for Computing Machinery, New York, no. 57, p. 1–10. https://doi.org/10.1145/3528233.3530704

Son, Jangwoo et al. (2024). Split Rendering with L4S Over 5G for Latency Critical Interactive XR Applications. In: *IEEE Communications Magazine*, vol. 62(8), p. 46–52. https://doi.org/10.1109/MCOM.006.2300800

Son, Jangwoo et al. (2024). Adaptable L4S Congestion Control for Cloud-Based Real-Time Streaming Over 5G. In: *IEEE Open Journal of Signal Processing*, vol. 5, p. 841–849. https://doi.org/10.1109/OJSP.2024.3405719

Taylor, Paul (2009). *Text-to-Speech Synthesis*, Cambridge University Press, Campbridge, p. 474–516.

Sörqvist, Patrik & Marsh, John E. (2015). How concentration shields against distraction. In: *Current Directions in Psychological Science*, vol. 24(4), p. 267–272. http://dx.doi.org/10.1177/0963721415577356

Theobalt, Christian et al. (2007). Seeing people in different light: Joint shape, motion, and reflectance capture. In: *IEEE Transactions on Visualization and Computer Graphics*, vol. 13(4), p. 663–674. https://doi.org/10.1109/TVCG.2007.1006

Tootell, Roger et al. (2021). Psychological and physiological evidence for an initial ›rough sketch‹ calculation of personal space. In: *Scientific Reports*, vol. 11, no. 20960. https://doi.org/10.1038/s41598-021-99578-1

Vonda GitHub (2024. *VOnDA, a Versatile Ontology based Dialogue management Architecture,* GitHub Inc., San Francisco. https://github.com/bkiefer/vonda [20.03.2025]

Webrtc (2024). *Echtzeitkommunikation im Web.* Web Real-Time Communication, https://www.webrtc.org. [20.03.2025]

Wang, Yuxuan et al. (2017). Tacotron: Towards End-to-End Speech Synthesis. In: *2017 Interspeech*, p. 4006–4010. http://dx.doi.org/10.21437/Interspeech.2017-1452

Yamaguchi, Shugo et al. (2018). High-fidelity facial reflectance and geometry inference from an unconstrained image. In: *ACM Transactions on Graphics*, vol. 37(4), no. 162, p. 1–14. https://doi.org/10.1145/3197517.3201364

Yang, Geng et al. (2021). Multi-band MelGAN: Faster waveform generation for high-quality text-to-speech. In: *Proceedings of the 2021 IEEE Spoken Language Technology Workshop (SLT)*, p. 492–498. https://doi.org/10.1109/SLT48900.2021.9383551

Zimmer, Alexandra et al. (2022). Imposing temporal consistency on deep monocular body shape and pose estimation. In: *Computational Visual Media*, vol. 9(1), p. 123–139. https://doi.org/10.1007/s41095-022-0272-x

III.
Transformative Technologien im Öffentlichen Raum

Beam me up, Goethe! Gemeinsam im hybriden Museum – Goethe-Live-3D

Annika Schlimm, Kristina Johannes, Dirk Wintergrün, Sophia Gröschke und Alexander Kulik

Zusammenfassung

Das vom Bundesministerium für Bildung und Forschung (BMBF)[1] geförderte Verbundprojekt »Goethe-Live-3D« erforscht innovative Wege der digitalen Vermittlung von Kulturerbe am Beispiel des Goethe-Wohnhauses in Weimar. Es kombiniert virtuelle und analoge Elemente zu einem hybriden Museumserlebnis, das neue Dimensionen in der musealen Präsentation eröffnet. Besucher:innen können Goethes Wohnhaus mithilfe einer digitalen Rekonstruktion ortsunabhängig und barrierefrei erkunden. Die Anwendung erlaubt es, dem musealen Ort anders gegenüberzutreten und mit ihm zu interagieren. Technische Innovationen wie Multi-User-Virtual-Reality (VR), immersive 3D-Modelle und soziale Mixed-Reality-Erfahrungen ermöglichen nicht nur individuelle, sondern auch gemeinschaftliche Museumsbesuche. Das Projekt setzt auf interaktive Vermittlungskonzepte, die klassische Museumsangebote um neue, medienadäquate Ansätze erweitern. Erste Evaluierungen zeigen ein großes Potenzial für die nachhaltige Transformation des Museumsbesuchs, indem soziale, inklusive und spielerische Dimensionen gefördert werden. Goethe-Live-3D ist ein Modellprojekt für die digitale Transformation im Kulturbereich und bietet einen dynamischen Rahmen für innovative hybride Museumsformate.

1. Einleitung – Gemeinsam im hybriden Museum

Die Projektpartner:innen des Goethe-Live-3D-Projekts arbeiteten gemeinsam an der Entwicklung, Realisierung und Evaluation eines hybriden Museums. Im Rahmen des BMBF-Förderschwerpunktes »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« ermöglicht

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Goethe-Live-3D den virtuellen Besuch von Goethes Wohnhaus in Weimar und erlaubt Besucher:innen ortsunabhängig einen gemeinsamen und interaktiven Museumsbesuch. Darüber hinaus bildet das Projekt nun einen dynamischen Rahmen, anhand dessen eine Vielzahl von Inhalten mittels innovativer Vermittlungsformen dargestellt werden kann.

Im Folgenden werden in diesem Beitrag zunächst die Voraussetzungen und Herausforderungen des Museums im digitalen Zeitalter erläutert, auf denen das Projekt fußt. Der zweite Teil widmet sich der technischen Umsetzung, bevor es um die möglichen Vermittlungsweisen und zu vermittelnden Inhalte geht. Den Abschluss bilden dann die das Projekt begleitenden Evaluationen, gefolgt von einem kurzen Fazit.

Die Bewahrung und Vermittlung von materiellem und immateriellem Kulturerbe ist die traditionelle Aufgabe von Museen, Bibliotheken und Archiven. Im digitalen Zeitalter müssen sie ihre Rolle neu definieren. Ausgelöst durch die Notwendigkeit der räumlichen Distanzierung während der COVID-19-Pandemie rückt die Bedeutung digitaler Museumsangebote seitdem verstärkt in den Fokus. Dies betrifft sowohl die öffentliche Wahrnehmung als auch die finanzielle Förderung. Das Projekt Goethe-Live-3D hat es sich zur Aufgabe gemacht, die Grenzen des traditionellen Museums zu überschreiten und ein hybrides Erlebnis zu schaffen, das virtuelle und analoge Welten miteinander verbindet.

Die museale Aufbereitung historischer Orte und Objekte erfordert in der Regel die Festlegung auf einen spezifischen Zustand, dessen statische Inszenierung durch digitale Medien durchbrochen werden kann. Das Wohnhaus von Johann Wolfgang von Goethe in Weimar bildet als bedeutendes Denkmal und als Ort historischer Schichtung und Veränderlichkeit ein ideales Experimentierfeld für ein immersives, digitales Angebot. Letzteres erlaubt einerseits, den Ort von überall zu besichtigen, und andererseits, verschiedene Zeitschichten und museale Inszenierungen zu entdecken und damit die scheinbare Unveränderlichkeit des Denkmals infrage zu stellen. Die digitale Rekonstruktion des Hauses bietet die Möglichkeit, das Museumserlebnis neu zu gestalten und die Erwartungen der Besucher:innen zu verändern. Historische Gebäude werden mittels Verknüpfung von digitalen und analogen Strategien zu Experimentalräumen eines gesellschaftlichen Austauschs über Kulturgeschichte.

Im BMBF-Verbundprojekt »Goethe-Live-3D« erforschen die Klassik Stiftung Weimar, die Bauhaus-Universität Weimar, die Universität Hamburg sowie die ArcTron 3D GmbH, die Curvature Games GmbH und die Consensive GmbH die weitreichenden Möglichkeiten einer digitalen Rekonstruktion historischer Gebäude am Beispiel von Goethes Wohnhaus in Weimar. Die digitale 3D-Rekonstruktion bildet hier die Grundlage für neue Erlebnisse und die interaktive Auseinandersetzung mit dem Haus und seinem Mobiliar. Besucher:innen können

Goethes Wohnhaus barrierefrei und trotz unterschiedlicher Standorte gemeinsam besuchen (vgl. Abb. 1). Darüber hinaus sollen digitale Vermittlungsstrategien und didaktische Konzepte neue museale Lernräume eröffnen. Das Projekt stellt sich die Frage, wie es gelingen kann, ein tatsächlich hybrides Museum zu erschaffen, das nicht nur reale und digitale Ausstellungen nebeneinanderstellt, sondern diese intelligent verbindet, um sie so für lokale und nichtlokale Besuchende gemeinsam und auch gleichzeitig erkundbar zu machen.

Für die Klassik Stiftung Weimar hat sich aus dem Projekt heraus ein multimediales Konzept entwickelt, auf dessen Grundlage in Zukunft museale Angebote entwickelt werden. Der Fokus liegt daher im Folgenden auf den Projektteilen, die in Weimar zur Verstetigung vorgesehen sind.



Abbildung 1: Einblick in die Zukunft der digitalen Museumsführung?
Quelle: Klassik Stiftung Weimar/Bauhaus-Universität Weimar, Social-XR-Anwendung: Consensive GmbH, 3D-Modell des Goethe-Wohnhauses: ArcTron 3D Vermessungstechnik & Softwareentwicklung GmbH

2. Technische Umsetzung

Im Projekt entstanden verschiedene Demonstratoren der beteiligten Entwicklungspartner:innen mit jeweils unterschiedlichem Fokus. Neben den interaktiven Möglichkeiten sozialer gemischter Realität, die von allen Partnern auf unterschiedliche Weise ausgelotet wurden, widmete sich Consensive auch praktischen Fragen des musealen Einsatzes der Technologie. Dazu gehören zunächst

ein möglichst unkomplizierter Einsatz an wechselnden Orten und die Reduktion technischer Anforderungen. Darüber hinaus wurden aber auch die Voraussetzungen für eine langfristige Wartung und Aktualität der Inhalte sowie die Übertragbarkeit auf vergleichbare Anwendungsfälle bearbeitet.

Unter Verwendung von Vorarbeiten und parallelen Entwicklungen der jungen Firma entstand ein Technologiedemonstrator, der die Präsentation der umfangreichen und hochdetaillierten Inhalte auf unterschiedlicher, vergleichsweise leistungsarmer mobiler Hardware sowie im Browser ermöglicht (https://www.vr4more.com/demo/goethe/). Unabhängig von der gewählten Plattform können die jeweiligen Softwareinstanzen miteinander vernetzt werden, um gemeinsame Erlebnisse in der virtuellen Umgebung zu ermöglichen – vor Ort und auch ortsübergreifend.

Im Falle des Einsatzes mehrerer XR-Geräte im Museum oder z. B. auch in einer Schule werden die Avatare weiterer Beteiligter an ihren korrekten räumlichen Positionen dargestellt. 2D-Darstellungen auf großen Bildschirmen und/oder Tablets ermöglichen es, den Aktivitäten in der virtuellen Umgebung auch ohne VR-Brille zu folgen. Im Headset erleichtert die visuelle Überlagerung von realer und virtueller Umgebung den Einstieg in die Interaktion mit den digitalen Modellen.

Die gemeinsam erlebten Räume können in gemischter Realität erlaufen, einzelne Objekte mit den Händen gegriffen und neu platziert werden. Die gemeinsame Navigation über den erlaufbaren Raum hinaus wird durch dedizierte Wechsel der virtuellen Räume bzw. verschiedener »Points of Interest« erreicht. Diese Sprünge von Szene zu Szene können von Guides, d. h. geschultem Personal, vor Ort explizit ausgelöst oder auch als Tour automatisiert werden. Bei Bedarf können Guides die Gruppe auch spontan an Orte einladen, die für ihre Erzählungen und den Austausch mit Besucher:innen relevant sind.

An 2D-Bildschirmen (Flatscreens) können die Transitionen zwischen Points of Interest mit animierten Kamerapfaden ergänzt werden, um die räumlichen Beziehungen zu illustrieren. Bei Nutzer:innen immersiver Displays (Deepscreens) würden solche virtuellen Kameraflüge Unwohlsein auslösen (Cybersickness). Hier ist es sinnvoller, die räumlichen Beziehungen im Miniaturmodell darzustellen. Animierte Touren ermöglichen Kurator:innen, Inhalte nach Bedarf jederzeit neu zu verknüpfen. Mit Tonspuren unterlegt, können Besucher:innen lineare Filme erleben, die jedoch die Option bieten, jederzeit unterbrochen und verändert zu werden. Vor allem aber laden XR-Geräte vor Ort ein, in die jeweils dargestellten Räume und Orte »einzusteigen« und diese immersiv und gemeinsam zu begehen.

Hybride soziale Erlebnisse entstehen durch die Möglichkeit, Freund:innen und Verwandte bei ihrem Rundgang durch das reale Haus in der digitalen Kopie

zu begleiten. Dies ermöglicht die Teilhabe über Ortsgrenzen hinweg, ist aber auch vor Ort nützlich. Wenn Gruppen das Museum besuchen, sind nicht immer alle in der Lage, die Treppen des alten Hauses zu steigen und durch seine engen Gänge zu manövrieren. Statt im Foyer zu warten, können sie die Tour durch das Haus am Bildschirm oder in gemischter Realität erleben, wobei sie der Tour ihrer Familie oder ihrer Freund:innen folgen. Letztere müssen dafür ihre Position im realen Museum teilen – sporadisch, z. B. durch Einscannen von QR-Codes für den Abruf von Zusatzinformationen, kontinuierlich, z. B. durch 3D-Tracking einer AR-Software oder eine Kombination dessen. Die Forschung der Bauhaus-Universität Weimar zeigt, wie gut das funktionieren kann (Schott et al., 2023: 3 f.). Mit den Möglichkeiten, Points of Interest durch QR-Codes aufzurufen und AR-Tracking auf mobilen Geräten, bietet auch der Demonstrator von Consensive die technischen Grundlagen dafür.

Nach Abschluss des Projektes entwickelt Consensive den Demonstrator Schritt für Schritt zu einem robusten Produkt für den Einsatz in Museen, auf Messen, in Schulen und anderen Bildungseinrichtungen weiter. Dies beinhaltet einerseits die weitere Vereinfachung der Bedienung, aber auch die permanente Weiterentwicklung wesentlicher Grundlagentechnologien, die wir im Folgenden kurz vorstellen.

Eine wesentliche Herausforderung digitaler Museumsexponate ist, dass einerseits die darzustellenden Modelle oft groß und detailliert sind und andererseits möglichst vielen Interessierten eine kosteneffiziente Zugänglichkeit ermöglicht werden soll. Museale Anwendungen von XR-Technologien spitzen diesen Zielkonflikt zu, wofür das Goethe-Wohnhaus als hybrides Museum ein sehr gutes Beispiel ist.

Für interaktive Anwendungen werden photogrammetrische Rekonstruktionen historischer Orte und Objekte oft manuell optimiert und nachmodelliert – ein nicht zu unterschätzender Aufwand, den die Kolleg:innen von ArcTron mit viel Erfahrung und Sachverstand für das Projekt geleistet haben. Trotzdem besteht das resultierende Modell letztlich aus etwa 290 Einzelobjekten mit insgesamt etwa 12 Millionen Dreiecken und etwa 100 4K-Texturen (ca. 16 Megapixel pro Textur). Während leistungsfähige Grafik-PCs diese Datenmenge flüssig darstellen können, genügt die Rechenleistung und auch der Grafikspeicher heutiger XR-Geräte nicht, um den gesamten Datensatz zu laden und mit den nativen Bildwiederholraten der Geräte (72–120 Hz) flüssig darzustellen. Natürlich werden zukünftige Geräte mehr Leistung zur Verfügung stellen, aber im gleichen Zuge werden auch die Anforderungen an Bildwiederholraten, Darstellungsqualität, Detaillierungsgrad und Informationsmenge steigen. Zudem verbraucht jeder Rechenschritt Energie, und die Akkus der Geräte sollen lange laufen.

Unter dem Namen VR4more-Date entwickelt Consensive daher eigene Grundlagentechnologien für ausgabesensitive Grafikverarbeitung. Dabei gilt es, nur so viele Modelldaten zu verarbeiten, wie für eine aktuelle Perspektive tatsächlich benötigt werden. Modellbereiche, die ohnehin verdeckt, also unsichtbar sind, werden ausgelassen. Der Detailgrad der sichtbaren Bereiche wird ihrer Darstellungsgröße auf dem resultierenden Bild angepasst. Die Modelldaten können von unterschiedlichen Quellen geladen werden – typischerweise von Online-Servern. Sobald bestimmte Bereiche nicht mehr sichtbar sind, werden die entsprechenden Geometrie- und Texturdaten im Arbeitsspeicher freigegeben, um Platz für andere, nun sichtbare Modellbereiche zu schaffen. Dabei bleiben sie jedoch auf lokalen Datenträgern im »Cache« – für schnellere Verfügbarkeit beim nächsten Bedarf und minimiertem Datentransfer mit dem externen Datenserver. Da Softwarefunktionen und Inhalte separat verwaltet werden, können die angezeigten 3D-Modelle und deren inhaltliche Ergänzungen jederzeit verändert und ergänzt werden.

Eine weitere Herausforderung bildet das generelle Softwaredesign und die Infrastruktur für verteilte Anwendungen, also soziale Interaktion in virtuellen Umgebungen. Verschiedene Softwareinstanzen auf unterschiedlichen Geräten an unterschiedlichen Orten müssen dafür permanent Zustandsänderungen der Anwendung und Szenerie miteinander synchronisieren. Dafür ist ein gemeinsam erreichbarer Netzwerkserver notwendig. Digitale museale Anwendungen sollten nun aber vor Ort im Museum genauso funktionieren wie auf Wanderausstellungen, in Schulen, auf Messen oder auch ad hoc bei öffentlichen Veranstaltungen im Freien.

In manchen dieser Kontexte ist Internet vorhanden, teilweise nur mit bestimmten Einschränkungen, in anderen Fällen wiederum gar nicht. Mit der eigenen Netzwerksoftware CommonGround erreicht Consensive diesbezüglich eine hohe Flexibilität. Die gleiche Anwendung kann ohne Änderungen ihrer Konfiguration heute über Online-Server und morgen in einem dafür konfigurierten lokalen Netzwerk über einen lokalen Server synchronisiert werden. CommonGround unterstützt nicht nur die Synchronisation der gleichen Anwendung auf unterschiedlichen Clients, sondern erleichtert auch den Echtzeit-Datenaustausch mit gänzlich anderen Anwendungen und Webservices. Dadurch wird umfassende Anschlussfähigkeit gewährleistet.

3. Vermittlungs- und Inhaltskonzepte

Anschließend an die Ausführungen zur technischen Basis der Anwendung sollen im Folgenden kurz bereits umgesetzte sowie mögliche Verwendungskontex-

te aufgezeigt werden, die sich auf der Grundlage einer digitalen Rekonstruktion ergeben. Dies betrifft zum einen die Art und Weise der Vermittlung, zum anderen die Inhalte, Geschichten und Ideen, die mit Goethe-Live-3D erzählt werden können. Dabei steht besonders im Vordergrund, das virtuelle Wohnhaus nicht als bloße »Kopie des Originals« zu verstehen, sondern die spezifischen dynamischen Möglichkeiten zu beleuchten, die die Umgebung bietet.

Nachdem der offizielle Projektabschluss im Oktober 2024 in Weimar gemeinsam mit allen Projektpartnern begangen wurde, werden die im Rahmen des Projekts entwickelten Formate nun im Kontext des Goethe-Nationalmuseums in Weimar weitergeführt, gezielt für die Vermittlungsarbeit genutzt und weiterentwickelt. Besonders im Hinblick auf die anstehende Schließung des Wohnhauses im Zuge umfassender Sanierungsarbeiten, die ab 2026/27 geplant sind, bietet das Angebot die wertvolle Möglichkeit, den Besucher:innen weiterhin einen Einblick in das Wohnhaus zu ermöglichen sowie die Vermittlungsarbeit in dieser herausfordernden Phase mithilfe neuer Formate aufrechtzuerhalten und zu stärken.

3.1 Vermittlungsmöglichkeiten

Museale Vermittlungskonzepte für eine Multi-User-VR-Anwendung stecken noch in den Kinderschuhen. Im Museumsbetrieb sind Online-Repräsentationen meist nur separat gedacht, obwohl sie ideale Voraussetzungen bieten, Vor- und Nachbereitungen sowie den individuellen Museumsbesuch selbst miteinander zu verknüpfen (Kuflik et al., 2014: 20 f., Petrelli et al., 2017: 283). Virtuelle Ausstellungen sind typischerweise immer verfügbar. Ohne konkreten Anlass für deren Besuch werden andere Aktivitäten daher oft priorisiert. Virtuelle Museen sind bisher auch kaum für gemeinsame Freizeitgestaltung geeignet. Die verwendeten Schnittstellen isolieren die Besucher:innen oft sogar von ihrem direkten sozialen Umfeld. Auch Angebote auf der Basis von virtueller und erweiterter Realität (VR/AR) in Museen stehen den Besucher:innen meist als individuelle Erlebnisse zur Verfügung. Das gemeinsame Erlebnis ist jedoch einer der wichtigsten Gründe für den Besuch von Museen (Debenedetti, 2003: 56).

Eine VR-Anwendung, wie in diesem Projekt entwickelt, die den Museumsbesuch als gemeinschaftliches Event in den Vordergrund rückt und Möglichkeiten zur kooperativen Kommunikation zwischen Besucher:innen vor Ort und im virtuellen Raum bietet, ist im musealen Kontext noch weitestgehend einzigartig und stellt grundsätzlich neue Anforderungen an die Art und Weise, wie Inhalte vermittelt werden. Anders als bei 360-Grad-Rundgängen geht es nicht um die bloße Wiedergabe des Museums als unveränderlichem Ort. Das 3D-

Modell erlaubt Interaktionen und Veränderungen. Dabei besteht das Risiko, dass das Vermittlungspotenzial hinter der VR-Anwendung als innovativem technischem Spielzeug zurücktritt, obwohl sicher auch solche Angebote dabei helfen können, neue Zielgruppen für das Museum zu gewinnen und das Museum als Innovationsort zu präsentieren.

Es stellt sich aufgrund der vielfältigen Nutzungs- und Entwicklungsmöglichkeiten die Herausforderung, sich auf bestimmte Szenarien, Zielgruppen und Kommunikationsmittel festzulegen und nicht beliebig zu werden. Jedoch sollte die Anwendung nicht auf nur ein Nutzungsszenario beschränkt sein, sondern unterschiedliche Perspektiven auf den musealen Raum ermöglichen. Das Erfolgskonzept besteht in einem dynamischen Rahmen, der zukünftig im Hinblick auf verschiedene Vermittlungsziele angepasst werden kann.

Museale XR-Anwendungen ermöglichen eine schier unendliche Vielfalt interaktiver Nutzungsszenarien, die wiederum unterschiedliche Vermittlungskonzepte ermöglichen oder erfordern. Diese werden dann im engen Dialog mit den Geschichten und Inhalten, die vermittelt werden sollen, entwickelt.

Mit den innovativen, stereoskopischen Multi-User-3D-Projektionen der Bauhaus-Universität Weimar können beispielsweise mehrere Personen dreidimensionale Räume gemeinsam erleben, ohne dass ihre Sicht auf die direkte Realität wesentlich eingeschränkt wird. Die eingesetzten »Shutter-Brillen« unterscheiden sich kaum von Sonnenbrillen (vgl. Abb.2). Aufgrund des hohen technischen Aufwands und der hohen Rechner- sowie Energieleistung kann dieser Prototyp derzeit nur in den Laboren der Bauhaus-Universität Weimar erprobt werden. Im Zuge eines wissenschaftlichen Begleitprojekts ist jedoch eine weitere Installation beim Virtual Dimension Center (VDC) Fellbach geplant. Das holografische Display erlaubt als Highlight einerseits, innovative Zukunftstechnologien bei der Entwicklung von Vermittlungs- und Inhaltskonzepten mitzudenken, und bildet gleichzeitig den Anlass, die Notwendigkeit flexiblerer Technologien zu reflektieren.

Eine andere Möglichkeit einer Multi-User-VR-Anwendung wird zurzeit bereits im Goethe-Nationalmuseum getestet. Dort können sich Besucher:innengruppen mit VR-Brillen gemeinsam durch das virtuelle Wohnhaus bewegen. An mehreren Tagen der Woche kann das Angebot kostenfrei ausprobiert werden. Dabei sehen sich die Personen als Avatare in der virtuellen Welt in den gleichen Relationen im Raum, sodass Handgesten und Interaktionen miteinander ermöglicht werden. Im virtuellen Raum können die Besucher:innen Gegenstände verrücken oder Statuen in die Hand nehmen. Zeitgleich erlaubt eine Projektion des Geschehens auch Zuschauer:innen ohne Brille die Teilnahme (vgl. Abb.3). In einem nächsten Schritt können sich externe Besucher:innen, die sich nicht vor



Abbildung 2: Einsatz der Shutter-Brillen an der Bauhaus-Universität Weimar
Quelle: Consensive GmbH/Bauhaus-Universität Weimar, 3D-Modell des Goethe-Wohnhauses: ArcTron 3D Vermessungstechnik & Softwareentwicklung GmbH

Ort befinden, dazuschalten. Das Wohnhaus wird so von überall und barrierefrei erlebbar.

Darüber hinaus gibt es im 3D-Modell die Möglichkeit von Live-Broadcasts oder aufgezeichneten Führungen. Nutzer:innen können sich von überall aus über VR oder ihren eigenen Desktop einwählen und an einer Führung durch das Haus teilnehmen. So könnten beispielsweise Kustod:innen anhand von hochaufgelösten 3D-Rekonstruktionen von Grafiken oder Majoliken etwas über Goethes umfangreiche Sammlung erzählen. Außerhalb eines Führungskontextes kann Goethe-Live-3D zur Unterstützung von Vorträgen als Präsentationsumgebung eine Alternative zu statischen Formaten wie PowerPoint-Präsentationen bieten.

Die vorhandenen Vermittlungskonzepte, die bereits umgesetzt wurden, orientieren sich noch stark an klassischen musealen Vermittlungsangeboten. Der Schwerpunkt der weiteren Arbeit liegt nun auf neuen medienadäquaten Vermittlungsszenarien und neuen Formen des Storytellings. Hier sind Elemente des szenischen Spiels, Gamification-Ansätze wie Escape-Rooms oder das Kuratieren eigener Ausstellungen durch die Besucher:innen zurzeit technisch umsetzbar. Workshops mit Studierenden verschiedener Disziplinen haben hier bereits

Potenziale aufgezeigt (https://blog.klassik-stiftung.de/immersives-erleben-des-goethe-wohnhaus/).

An der Universität Hamburg wurden im Rahmen des Projekts bereits Interaktionen mit historischen Charakteren im digitalen Raum getestet, die weitere Entwicklung erfordern: Die Möglichkeiten reichen hier von Avatar-Repräsentationen, bei denen Besucher:innen in verschiedene (historische) Rollen schlüpfen können, bis hin zu KI-Agent:innen, die auf der Grundlage dedizierter Sprachmodelle mit Besucher:innen kommunizieren. Dabei gilt es immer auch darauf zu achten, wie Akteur:innen im digitalen Raum angenommen werden und welche Repräsentationen bei den Besucher:innen eher Unbehagen hervorrufen (Lehmann, 2021: 24). Zu erproben sind noch andere Anwendungsfälle, die jenseits der besucher:innenorientierten Museumsfunktion denkbar sind, etwa die Kommunikation der Forschung im Museum, die Nutzung der Umgebung als virtueller Forschungsraum bis hin zur Anwendung im Rahmen des Kulturgutschutzes.

Goethe-Live-3D kann Forschung neu vermitteln, ermöglicht eine innovative Art der Wissenschaftskommunikation und kann aufgrund der technologischen Flexibilität immer wieder neu angepasst werden. Ein wichtiger Aspekt des Projekts ist es, technische Möglichkeiten auf der Grundlage von konkreten Inhalten zu erproben und zugleich, abgeleitet aus diesen, Anforderungen an neue Entwicklungslinien zu formulieren. So entstehen im Wechselspiel von Technik, Vermittlung und Inhalten neue museale Angebote.

3.2 Inhalte

Aufbauend auf den technischen Entwicklungen und parallel zu den Vermittlungsszenarien entstehen die Inhalte, die Goethe-Live-3D erzählen kann. Inhaltliche Konzepte knüpfen vor allem an die eingangs bereits erwähnte vermeintliche Unveränderlichkeit des musealen Orts an, indem sie Alternativen aufzeigen. Der digitale Museumsraum kann keine exakte Reproduktion des realen Raums sein. Dafür bietet er alternative Möglichkeiten, historische Zusammenhänge zu veranschaulichen und neue narrative Ebenen zu erschließen: So kann etwa Goethes Arbeitszimmer, das normalerweise für Besucher:innen unzugänglich ist, in der virtuellen Realität betreten und erforscht werden. Der digitale Raum ermöglicht es, mit verschiedenen Zeitebenen zu spielen und das Haus zur Goethezeit sowie in späteren Epochen zu erleben. Auch die Betrachtung von Details wie Goethes Farbenlehre oder der musealen Überformung bestimmter Räume wird durch die immersive Erfahrung vertieft.

Einen Überblick bietet ein Modell des gesamten Hauses, das für die Besucher:innen wie ein Puppenhaus erscheint. Die Struktur des Hauses, die dem



Abbildung 3: Besuch im virtuellen Brückenzimmer
Quelle: Klassik Stiftung Weimar, Thomas Müller, Social-XR-Anwendung: Consensive GmbH, 3D-Modell des Goethe-Wohnhauses: ArcTron 3D Vermessungstechnik & Softwareentwicklung GmbH

Blick der Besuchenden im Regelfall verborgen bleibt, wird hier sichtbar. Von Goethe selbst oder später unternommene Umbauten können anhand dieses Modells gezeigt werden. Zudem erlaubt es eine vereinfachte Lokalisierung des eigenen Standpunkts und der Räume im Haus.

Ein inhaltlicher Schwerpunkt sind die »Metamorphosen«, die den Fokus auf die Veränderlichkeit und die Inszenierung der musealen Räume legen und verschiedene Formen der Veränderungen in ausgewählten Räumen sichtbar machen sollen. Der an Goethes botanische Untersuchungen angelehnte Titel zeigt Metamorphosen der Einrichtung, der Nutzung, aber auch der Stimmung mithilfe von Beleuchtung im Majolikazimmer, im Brückenzimmer und in Goethes für Besuchende vor Ort heute unzugänglichem Arbeitszimmer.

Tageszeitabhängige Wechsel in der Beleuchtung des Raumes werden im Brückenzimmer simuliert. Goethe selbst verglich diesen Raum aufgrund seiner besonderen architektonischen Funktion als Verbindung zwischen Vorder- und Hinterhaus mit einem römischen Lararium, das den Schutzgöttern des Hauses bzw. der Familie gewidmet war. Der Raum hat eine symbolische Zentralstellung innerhalb des Hauses. Hier können Besucher:innen die Inszenierung des Raumes nachvollziehen, indem sie das Zimmer einerseits zu verschiedenen Tageszeiten wahrnehmen und andererseits durch Neuanordnung der Objekte verändern kön-

nen. Heute bevölkern sieben Statuen den Raum, zu Goethes Lebzeiten stapelten sich die Plastiken jedoch bis unter die Decke. Ein Eindruck, der – auch wenn sich die genaue Bestückung nicht mehr rekonstruieren lässt – im virtuellen Wohnhaus vermittelt werden kann.

Im sogenannten Majolikazimmer, das als Teil der musealen Inszenierung nur einen Bruchteil der ursprünglich dort ausgestellten Majoliken enthält, werden verschiedene Zeitschichten rekonstruiert. Der Raum wurde zunächst als Schlafzimmer genutzt und erst ab 1826 zum Sammlungszimmer – in jeweils unterschiedlicher Ausgestaltung. Indem diese Stadien sichtbar gemacht werden, wird die Musealisierung des Goethe-Wohnhauses für die Besucher:innen nachvollziehbar.

Das Arbeitszimmer, das heute nur von der Türschwelle aus eingesehen werden kann, ist das Herzstück des Hauses. Hier scheint alles im Zuge der verschiedenen Musealisierungsphasen nahezu so erhalten geblieben zu sein, wie es am 22. März 1832 vorlag, als Goethe im benachbarten Schlafzimmer starb. Das Arbeitszimmer fasst alle wesentlichen Bereiche von Goethes Schaffen zusammen – von der Literatur über die Naturwissenschaften bis hin zu seinen grafischen Sammlungen und Zeichnungen. Es handelt sich um einen Raum, den Goethe mehr als 40 Jahre lang nutzte und immer wieder an seine Bedürfnisse anpasste, umgestaltete und veränderte. Für Besucher:innen ist dieser private Arbeitsbereich und Rückzugsort Goethes heutzutage nicht zugänglich. Es sind keine Objektbeschilderungen vorhanden, sodass in der jetzigen Ausgangssituation nur mittels Audioguide Wissen zu Raum und Zeitschichten vermittelt werden kann. Hier knüpft bereits der »Goethe-Apparat« an, der es Besucher:innen erlaubt, das Arbeitszimmer mittels einer 3D-Rekonstruktion des Raumes zu erkunden (https://www.klassik-stiftung.de/ihr-besuch/ausstellung/goethe-apparat/). Auch innerhalb von Goethe-Live-3D kann das Arbeitszimmer nun betreten und als Vermittlungs- und Interaktionsraum erschlossen werden.

Ein weiterer Raum, anhand dessen die museale Inszenierung deutlich gemacht und das Bild des »authentischen« Ortes aufgebrochen werden kann, ist das zum Frauenplan gelegene Urbinozimmer. Dieses wurde 1945 durch einen Bombentreffer gemeinsam mit einem großen Teil des Dachgeschosses fast vollständig zerstört. Anhand von 15 Fotografien, die den Zerstörungsgrad sichtbar machen, erstellte ArcTron eine digitale Rekonstruktion des Zustands von 1945 – kurz bevor das Haus zu Goethes 200. Geburtstag 1949 in Rekordzeit wieder aufgebaut wurde (vgl. Abb. 4). Auch wenn die Möbel sich zum Zeitpunkt des Bombeneinschlags bereits nicht mehr im Haus befanden und so größtenteils im Original ausgestellt werden können, zeigt die Rekonstruktion eindrücklich, dass die vermeintliche Authentizität des Denkmals selbst eine Art der Inszenierung ist. Goethe-Live-3D macht bauliche Zeitschichten sichtbar.



Abbildung 4: 3D-Rekonstruktion der Kriegszerstörung
Quelle: ArcTron 3D Vermessungstechnik & Softwareentwicklung GmbH

4. Evaluierung und Praxistests

Ein wesentliches übergreifendes Ziel des Projektes ist es, die entwickelten neuartigen immersiven Besuchserfahrungen begleitend zu evaluieren. Die experimentelle Installation im Veranstaltungsraum des Goethe-Nationalmuseums bietet dabei die perfekte Umgebung für eine Evaluation der Usability, Alltagstauglichkeit und Nutzerakzeptanz im laufenden Praxisbetrieb. Durch den User-Centered-Design-Ansatz und iterativen Entwicklungsprozess fließen die Beobachtungen regelmäßig in die Prototypen ein. Das direkte Feedback der Besucher:innen und Nutzer:innen ist essenziell für die Weiterentwicklung von Goethe-Live-3D. Zu diesem Zweck kann das virtuelle Wohnhaus seit August 2023 mehrmals pro Woche ausprobiert werden. Über 1.500 Personen unterschiedlichster Zielgruppen nahmen dieses Angebot bisher an und waren bei verschiedenen Versionen und Entwicklungsstufen dabei.

Hinzu kommen verschiedene Veranstaltungen, an denen immer wieder neue Funktionalitäten und Vermittlungsmöglichkeiten eingeführt und getestet werden können. So konnte im August 2024 die erste virtuelle Führung einer Kuratorin durch das digitale Wohnhaus stattfinden, mit Besucher:innen, die sich aus ganz Europa zugeschaltet hatten. Ein weiterer Anlass war Goethes 275. Geburtstag, an dem neben dem regulären Flanierpublikum auch mit zwei Gruppen Möglichkeiten für Rundgänge im digitalen Wohnhaus getestet wurden. Ebenfalls sind auch aufgrund der räumlichen Nähe der verschiedenen Koopera-

tionspartner:innen insbesondere technische Änderungen häufig sehr kurzfristig umsetz- und anpassbar, die auch spontane Pop-Up-Installationen im Stadtraum erlauben. Diese Pop-Up-Version, in der ein ausgewählter Raum betreten werden kann, bildete wiederum ein Highlight der Langen Nacht der Wissenschaften, die in Weimar am 15. November 2024 stattfand (vgl. Abb. 5).



Abbildung 5: Pop-Up-Version bei der Langen Nacht der Wissenschaften 2024 in Weimar
Quelle: Klassik Stiftung Weimar, Fotothek, Social-XR-Anwendung: Consensive GmbH, 3D-Modell des Goethe-Wohnhauses: ArcTron 3D Vermessungstechnik & Softwareentwicklung GmbH

Dahinter steht die Überzeugung, dass aktive Museumsforschung im laufenden Betrieb eines Museums eine Aufgabe ist, für die sich der Aufwand lohnt. Dazu zählt sowohl die professionelle Betreuung und Einweisung in die Technik als auch das Weitertragen und Ernstnehmen von Bedenken und Kritik. Nur so kann langfristig ein inklusiveres und attraktiveres Museumsangebot für alle entstehen, insbesondere wenn es sich um ein Angebot handelt, das tendenziell mit einer höheren Hemmschwelle verbunden ist.

Die Evaluierungen des letzten Jahres bestätigen grundsätzlich Überlegungen der Konzeptionsphase. So gab ein Großteil der Teilnehmer:innen einer ersten nicht repräsentativen Online-Befragung an, dass das gemeinsame Erlebnis auch beim digitalen Museumsbesuch eine wichtige Rolle spielt. Im direkten Gespräch und in teilnehmenden Beobachtungen zeigt sich, dass die Frage nach dem Mehrwert einer vermeintlichen virtuellen Museumskopie bei den Besucher:in-

nen präsent ist und durchaus kritisch hinterfragt wird. Dieser Umstand bietet gleichzeitig die Chance, die Vorteile bereits bestehender Nutzungsszenarien zu vermitteln, denn insbesondere die Barrierefreiheit von Goethe-Live-3D stellt für Personen mit Mobilitätseinschränkungen einen großen Mehrwert dar. Auch dass ein Besuch des digitalen Wohnhauses nicht an den realen Ort gebunden ist, sondern theoretisch von überall auf der Welt erfolgen kann und sich so Menschen von verschiedenen Orten aus im Haus treffen können, bietet insbesondere im Hinblick auf die anstehende Renovierung des Hauses großes Potenzial. Ein Mehrwert, der auch jetzt schon anerkannt und wertgeschätzt wird, ist die Möglichkeit, Teile des Hauses zu erkunden, die Besucher:innen im Normalfall nicht zugänglich sind. Allseits beliebt ist ein Ausflug aufs Dach, seit Neuestem kann zusätzlich Goethes Arbeitszimmer besichtigt werden.

Was die Weiterentwicklung von interaktiven Vermittlungsformaten betrifft, zeigt sich, dass die Nutzer:innen ganz unterschiedliche Wünsche und Vorstellungen davon haben, was das digitale Museumserlebnis bieten soll. Grundlegend ist in jedem Fall das Bedürfnis nach Interaktion und Information. Mehr als ein Drittel der Rückmeldungen bezieht sich darauf, im Wohnhaus aktiv zu werden, Objekte anzufassen, Schubladen zu öffnen und Möbel zu verrücken. Auffällig ist dabei, dass nahezu unabhängig von der Altersgruppe der Testenden die erste Aktion von Gruppen darin besteht, Chaos anzurichten: Türen mit Statuen zu verbarrikadieren oder Figuren in Wänden zu platzieren. Hier ist also bereits ein Bewusstsein dafür vorhanden, was das digitale Wohnhaus im Vergleich zum realen musealen Ort anbieten kann und soll. Ebenfalls groß ist die Nachfrage nach Informationsvermittlung. Während im realen Wohnhaus ein Audioguide zur Verfügung steht, ist das digitale bisher noch auf die Expertise der Durchführenden angewiesen. Hier knüpfen die inhaltlichen und vermittelnden Konzepte an. Das Gleiche gilt für den Wunsch nach Gamification-Ansätzen, einem virtuellen Goethe oder einem belebteren, »alltagsnäheren« Wohnhaus.

5. Fazit

Das Verbundprojekt Goethe-Live-3D ist ein Katalysator und Demonstrator für die digitale Vermittlung kulturellen Erbes. Vor allem aber muss es als dynamischer Baukasten verstanden werden, auf dessen Grundlage eine Vielzahl von Ansätzen und Formaten realisierbar wird. Die Flexibilität des Angebots fußt dabei zum einen maßgeblich auf der technischen Umsetzung, die vielfältige Nutzungsszenarien erst ermöglicht, ist aber zum anderen bereits im Virtuellen als Möglichkeitsraum angelegt. Ist das klassische Museum oder Denkmal immer noch ein statischer, auratischer Ort, können im digitalen Museum unzugängliche Berei-

che geöffnet, spielerische Interaktionen gefördert und museale Inszenierungen sowie historische Veränderungen sichtbar gemacht werden. Dies schafft sowohl neue Perspektiven für Besucher:innen vor Ort als auch für ein weltweites Publikum. Gleichzeitig zeigt das Projekt auch die Herausforderungen solcher Anwendungen auf, wie die Balance zwischen technischer Faszination und nachhaltigem Vermittlungsziel.

Die technische Innovation eines Multi-User-VR-Erlebnisses erweitert den Zugang zu kulturhistorischen Inhalten überdies, indem auch die soziale Dimension des Museumsbesuchs gestärkt wird. Die Möglichkeit, gemeinsam – physisch oder virtuell – interaktive Erlebnisse zu teilen, hebt das Projekt von klassischen VR-Angeboten ab und setzt neue Maßstäbe für Multi-User-VR-Anwendungen im Kulturbereich.

Erste Evaluierungen des letzten Jahres bestätigen das Potenzial von Goethe-Live-3D, eine barrierefreie, immersive und gemeinschaftliche Form der Kulturvermittlung zu bieten. Gleichzeitig sind sie wesentlich für die stetige Entwicklung des Projekts. Die Rückmeldungen der Besucher:innen zu technischen Problemen sowie zu Inhalten oder Vermittlungsideen hilft dabei, neue Ansätze zu entwickeln – besonders für interaktive Inhalte, spielerische Elemente oder »goethzeitliche« Inszenierungen.

Goethe-Live-3D zeigt, dass XR-Angebote im Museum, die über eine bloße Reproduktion der Ausstellung oder des Raumes hinausgehen, ein kreativer Raum für neue Narrative, inklusive Vermittlungsansätzen und innovativer Forschung, sein können. Mit einer zukunftsorientierten, dynamischen Weiterentwicklung kann es als Modell für die digitale Transformation der Museen gelten.

Förderhinweis:

Das Projekt Goethe-Live-3D wurde vom Bundesministerium für Bildung und Forschung (BMBF)[2] gefördert.

Literatur

Consensive (2024). *VR4more-Date, Goethe-Live-3D*Kulik, Alexander, Consensive GmbH, Weimar, VR4more-Date, Goethe-Live-3D. https://www.vr4more.com/demo/goethe/ [Zugriff 07.01.2025]

2 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Debenedetti, Stéphane (2003). Investigating the Role of Companions in the Art Museum Experience. In: *International Journal of Arts Management*, Vol. 5(3), HEC Montreal, S. 52–63.

Klassik Stiftung Weimar (2024). *Goethe-Apparat.* https://www.klassik-stiftung.de/ihr-besuch/ausstellung/goethe-apparat/ [Zugriff 07.01.2025]

Kuflik, Tsvi et al. (2014). An Integrative Framework for Extending the Boundaries of the Museum Visit Experience: Linking the Pre, During and Post Visit Phases. In: *Information Technology & Tourism*, Vol. 15(1), Springer-Verlag GmbH, Heidelberg, S. 17–47.

Lehmann, Stephanie, Ruf, Esther & Misoch, Sabina (2021). Emotions and attitudes of older adults toward robots of different appearances and in different situations. In: *International Conference on Information and Communication Technologies for Ageing Well and E-Health*, Prague, Springer Interantional Publishing, Cham, S. 21–43. https://doi.org/10.1007/978-3-030-70807-8

Petrelli, Daniela et al. (2017). Tangible data souvenirs as a bridge between a physical museum visit and online digital experience. In: *Personal and Ubiquitous Computing*, Vol. 21(2), Springer-Verlag, London, S. 281–295.

Schlimm, Annika & Zander, Paula (2024). Unveiling Spaces: Immersives Erleben des Goethe Wohnhaus. In: *Blog – Klassik Stiftung Weimar*. https://blog.klassik-stiftung.de/immersives-erleben-des-goethe-wohnhaus/ [Zugriff 07.01.2025]

Schott, Ephraim et al. (2023). UniteXR: Joint Exploration of a Real-World Museum and its Digital Twin. In: *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology (VRST '23)*, Association for Computing Machinery, New York, S. 1–10.

Skalierbare hybride Avatar-Agenten-Technologie für alltagstaugliche soziale Interaktion in XR – HiAvA

Christian Merz, Jonathan Tschanter, Niklas Krome, Nicolas Wagner, Mario Botsch, Stefan Kopp, Marc Erich Latoschik und Carolin Wienrich

Zusammenfassung

Soziale Extended Reality (XR) ist ein vielversprechendes Werkzeug für die Remote-Interaktion, steht jedoch vor Herausforderungen, die es erschweren, die feinen Nuancen realer Kommunikation vollständig abzubilden. Diese Arbeit beleuchtet die Herausforderungen sozialer XR und konzentriert sich dabei insbesondere auf Aspekte wie Skalierbarkeit, Komfort und Realismus für einen breiten Zugang zu dieser Technologie. Der Ansatz umfasst: i) die Entwicklung und Evaluation einer skalierbaren XR-Anwendung, ii) den Einsatz von Techniken des maschinellen Lernens für die Generierung von Gesten für Benutzer:innen mit weniger immersiven Geräten und iii) die Generierung sozial-expressiver Avatare für eine realistischere Repräsentation. In diesem Artikel geht es um Erkenntnisse und bewährte Praktiken für das Design effektiver sozialer XR-Anwendungen auf Basis verschiedener wissenschaftlicher Veröffentlichungen im Rahmen des »HiAvA«-Projektes.

1. Einleitung

In der heutigen sich schnell entwickelnden technologischen Landschaft wächst die Nachfrage nach immersiven und interaktiven virtuellen Umgebungen exponentiell. Das ist insbesondere in Bereichen wie sozialer Kommunikation, Bildung, Gesundheitswesen, Unterhaltung und Zusammenarbeit relevant. Virtuelle Realität (VR), Mixed Reality (MR) und Augmented Reality (AR), zusammengefasst als Extended Reality (XR), stehen an der Spitze dieser Transformation und bieten neue Möglichkeiten, wie Menschen miteinander und mit digitalen Inhalten interagieren können. Die vollständige Nutzung dieses Potenzials wird jedoch oft durch den aktuellen Stand der Avatar- und Agentensysteme eingeschränkt (Magnenat-Thalmann & Thalmann, 2005). Die Avatar-Repräsentation ist elementar

für die visuelle Darstellung und Interaktion von Nutzer:innen in virtuellen Räumen (de Borst & de Gelder, 2015; Fiedler et al., 2023; Latoschik et al., 2017; Smith & Neff, 2018). Diese Einschränkungen umfassen eine mangelnde Skalierbarkeit, eine unzureichende Reproduktionsqualität und einen geringen Komfort für die Nutzer:innen, was die Integration von XR-Technologien in den Alltag verhindert. Es gibt zwar beeindruckende Fortschritte im Bereich XR, jedoch besteht nach wie vor eine erhebliche Lücke zwischen dem derzeit Verfügbaren und dem, was notwendig ist, um vollständig immersive und sozial bedeutsame virtuelle Erlebnisse zu ermöglichen (Merz et al., 2024a).

Das Projekt HiAvA – Skalierbare hybride Avatar-Agenten-Technologien für alltagstaugliche soziale Interaktionen in XR – zielt darauf ab, diese kritischen Lücken zu schließen, indem es innovative, skalierbare und benutzerfreundliche Avatar-Technologien entwickelt. Insbesondere strebt HiAvA an, die technologischen Barrieren zu überwinden, die die effektive Nutzung verkörperter Avatare in sozialen XR-Umgebungen einschränken. Das Ziel ist es, den Benutzer:innen die Möglichkeit zu geben in Echtzeit über personalisierte und lebensechte digitale Repräsentationen miteinander zu interagieren.

Ein Hauptmotiv für diese Forschung liegt in der wachsenden Bedeutung von sozialer XR – einer Form von VR, die die Interaktion zwischen Nutzer:innen in gemeinsamen virtuellen Räumen ermöglicht. Soziale XR ist eine der vielversprechendsten Anwendungen von XR-Technologien und ermöglicht es Nutzer:innen, sich in komplexen und bedeutungsvollen Interaktionen über Entfernungen hinweg zu vernetzen. Diese Kommunikation geht weit über die derzeit verfügbaren videobasierten Kommunikationswerkzeuge hinaus. Dies umfasst nicht nur die verbale Kommunikation, sondern auch die gesamte Bandbreite an nonverbalen Signalen, die für eine natürliche menschliche Interaktion entscheidend sind, wie Gesichtsausdrücke, Gesten und Körpersprache (Buschmeier & Kopp, 2018; Oertl et al., 2016). Durch die Möglichkeit der verkörperten Interaktion kann Kommunikationsverhalten entstehen, welches dem Verhalten in der realen Welt ähnelt (de Borst & de Gelder, 2015; Smith & Neff, 2018). Bestehende Forschung fokussiert sich auf VR-Interaktionen; aktuelle XR-Plattformen versagen jedoch noch häufig darin, nuancierte Interaktionen effektiv zu vermitteln, insbesondere in Bezug auf die Avatar-Treue (Cao et al., 2016; Garrido et al., 2016), die Echtzeit-Reaktionsfähigkeit und die Zugänglichkeit durch verschiedene Geräte (Latoschik et al., 2019; Yassien et al., 2020).

Der weltweite Umstieg auf digitale Kommunikation hat die Grenzen bestehender Systeme offengelegt und die Bedeutung von Technologien hervorgehoben, die immersive und nahtlose Interaktionen ermöglichen, unabhängig von der physischen Nähe. HiAvA baut auf diesen Erkenntnissen auf, indem es eine soziale XR-Plattform entwickelte, die nicht nur immersiv und hochqualitativ ist,

sondern auch breit zugänglich, sodass Nutzer:innen über verschiedene Geräte wie Mobiltelefone, Tablets, Desktop-Computer und High-End-VR-Headsets teilnehmen können. Diese Skalierbarkeit stellt sicher, dass soziale XR nicht mehr auf Benutzer:innen mit spezieller Hardware beschränkt bleibt und so der Zugang zu immersiven virtuellen Erfahrungen inklusiv wird.

Das technologische Herzstück des HiAvA-Projekts besteht in der Entwicklung hybrider Avatar-Agenten-Systeme, die fortschrittliches Körper- und Gesichtstracking mit KI-gestützter Verhaltensmodellierung integrieren. Diese Avatare sind in der Lage, realistische Bewegungen, personalisierte Gesichtsausdrücke und natürliche Interaktionssignale darzustellen, um ein hohes Maß an Immersion und Glaubwürdigkeit zu gewährleisten. Das Projekt nutzt innovative KI-Methoden, um das Verhalten von Avataren vorherzusagen und zu simulieren, wenn Sensordaten begrenzt oder nicht verfügbar sind. Ziel ist es, die Skalierbarkeit und Benutzerfreundlichkeit des Systems über verschiedene Plattformen hinweg zu verbessern, um wie in Abbildung 1 dargestellt Nutzer:innen mit verschiedenen Endgeräten die Teilhabe zu ermöglichen.



Abbildung 1: Nutzer:innen mit verschiedenen Endgeräten. Nutzerin A hat Zugang mit einem Tablet, Nutzer B nimmt über einen Desktop-PC teil und Nutzerin C mit einen VR-Headset. Alle Nutzer:innen sind trotz unterschiedlicher Zugangsgeräte und verfügbarer Sensordaten in einer sozialen VR-Umgebung durch Avatare repräsentiert. Sie können so gleichberechtigt miteinander in VR teilhaben und interagieren.
Quelle: Projekteigene Darstellung

Darüber hinaus ist HiAvA motiviert durch die Erkenntnis, dass Avatare mehr sind als nur digitale Stellvertreter der Benutzer:innen. Sie sind mächtige Werkzeuge für Kommunikation, Identitätsausdruck und soziale Interaktion in virtuellen Welten. Realistische, ausdrucksstarke Avatare können das Gefühl der Präsenz und der sozialen Präsenz der Nutzer:innen erhöhen und virtuelle Interaktionen realer wirken lassen (Bente et al., 2008; Waltemate et al., 2018). Dies ist besonders wichtig in Kontexten wie der Zusammenarbeit, dem Fernunterricht und therapeutischen Anwendungen. Bei diesen Anwendungen kann die Qualität der Interaktion direkte Auswirkungen auf die Ergebnisse haben. HiAvA zielt daher darauf ab, die Grenzen der Avatar-Realität zu erweitern, indem es personalisierte Gesichtsausdrücke und Körpersprache integriert, sodass Avatare natürlicher und effektiver kommunizieren können, auch wenn weniger immersive Endgeräte in der Interaktion beteiligt sind.

Neben den technischen Fortschritten legt HiAvA großen Wert auf nutzerzentriertes Design und Evaluation. Dabei werden technische Innovationen alleine nicht ausreichen, sondern es ist notwendig, ein tiefes Verständnis dafür zu bilden, wie die Nutzer:innen mit diesen Systemen interagieren und von ihnen profitieren. Daher hat HiAvA einen iterativen, nutzerzentrierten Entwicklungsprozess verfolgt, der Feedback aus umfassenden Usability-Studien einbezieht, um ein intuitives, komfortables und fluentes Nutzererlebnis sicherzustellen. Dabei werden auch wichtige Bedenken hinsichtlich des Datenschutzes, der Sicherheit und der Identität der Nutzer:innen in virtuellen Räumen berücksichtigt, die zunehmend an Bedeutung gewinnen, da Avatare nicht nur für Unterhaltung, sondern auch in beruflichen und persönlichen Kontexten genutzt werden. Darüber hinaus wurden umfassende empirische Wirksamkeitsstudien durchgeführt, die die Auswirkung der Systemkomponenten auf verschiedene Aspekte der sozialen Interaktion untersuchten.

Zusammenfassend stellt HiAvA einen bedeutenden Fortschritt im Bereich der XR-Technologien dar, da es den dringenden Bedarf an skalierbaren, realistischen und sozial ausdrucksstarken Avatar-Systemen adressiert. Durch die Überwindung der Einschränkungen bestehender Avatar- und Agententechnologien zielt dieses Projekt darauf ab, soziale XR zu einem praktikablen und zugänglichen Werkzeug für alltägliche Interaktionen zu machen und die Lücke zwischen der physischen und der virtuellen Welt zu schließen.

2. Entwicklung und Forschung

Im Folgenden geht es um wissenschaftliche Arbeiten, die bei internationalen Konferenzen oder Journalen veröffentlicht wurden und zu der Erreichung des

Gesamtziels beitragen und neue Erkenntnisse für den aktuellen Forschungsstand geschaffen haben. Dabei untergliedert sich die Forschung hinsichtlich 2.1) Entwicklung und Nutzerstudien für soziale XR, 2.2) Gestengenerierung für soziale XR und 2.3) Optimierung der Gesichtsanimationen von sozial expressiven Avataren.

2.1 Soziale XR-Plattform – Entwicklung und Nutzerevaluationen

Dieser Abschnitt beschreibt wissenschaftliche Arbeiten, die eine soziale XR-Plattform betreffen, wie diese entwickelt wurde und welche Nutzerstudien durchgeführt wurden, um das Verständnis von sozialen Interaktionen von Nutzenden mit verschiedenen Endgeräten zu erforschen. Abbildung 2 zeigt eine Kollaboration von Nutzer:innen mit VR-Headset und Desktop-PC und wie die jeweilige Perspektive in der virtuellen Welt ist. Dies ist besonders relevant für das allgemeine Projektvorhaben, denn die Nutzerevaluationen können nur mit einer anpassbaren sozialen XR-Plattform durchgeführt werden, um neue Erkenntnisse für die Gestaltung von inklusiven und geräteübergreifenden sozialen XR-Plattformen zu gewinnen. Das ermöglicht es zukünftigen Anwendungen und Applikationen, die auf diesem Paradigma aufbauen, Plattformen mit einer hohen Qualität zu entwickeln.

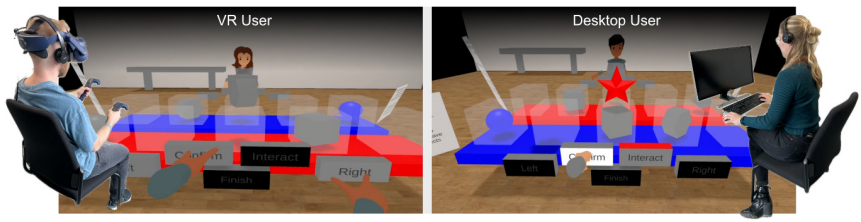


Abbildung 2: VR- und Desktop-Nutzer:innen während einer kollaborativen Aufgabe, die in verschiedenen Studien verwendet wurde.
Quelle: Merz et al., 2024b: 1

2.1.1 Prozessoren in einer Pipeline für die Zerteilung von Charakteranimationen

Die Arbeit »Pipelining Processors for Decomposing Character Animation« (Merz et al., 2024b) beschäftigt sich mit der Entwicklung einer modularen Pipeline-Architektur für die Animation von humanoiden Charakteren. Ziel ist es, häufige Schritte bei der Verarbeitung von Charakteranimationen, wie das Kopieren von

Bewegungsdaten, die Anwendung von Inverse Kinematics (IK) oder das Skalieren von Charakteren, in dedizierte, leicht anpassbare Prozessoren zu unterteilen. Diese modulare Herangehensweise ermöglicht eine flexible Anpassung und Erweiterung der Animationen in Echtzeitsystemen, wie in Anwendungen für virtuelle Realität (VR) oder Extended Reality (XR).

Implementierung
Die Entwicklung basiert auf einer Pipeline-Architektur mit Prozessoren, die in einer an den Anwendungsfall angepassten Reihenfolge ausgeführt werden, um die Animation eines humanoiden Charakters zu verändern. Jeder Prozessor führt eine spezifische Aufgabe aus, wie das Kopieren von Bewegungsdaten, das Skalieren des Charakters oder die Anpassung der Gliedmaßen durch eine IK. Die Architektur ermöglicht es Entwickler:innen, Animationen dynamisch zu bearbeiten und zu erweitern, ohne dass komplexe Systemneukonfigurationen erforderlich sind. Die Implementierung basiert auf Unity und beinhaltet verschiedene grundlegende Prozessoren, die zeigen, wie das System in realen Anwendungen verwendet werden kann.

Ergebnis
Das System bietet eine flexible, modulare und skalierbare Architektur, die es Entwickler:innen erlaubt, humanoide Charaktere realistisch zu animieren. Es reduziert die Komplexität der Animation von humanoiden Charakteren erheblich und bietet gleichzeitig Anpassungsmöglichkeiten. Die Möglichkeit, Prozessoren in Echtzeit hinzuzufügen, zu entfernen oder neu anzuordnen, ermöglicht eine flexible Entwicklung von interaktiven Systemen. Die Implementierung wird als Open-Source-Lösung zur Verfügung gestellt und legt damit den Grundstein für zukünftige Erweiterungen, wie den Einsatz von KI oder die Animation nichthumanoider Charaktere.

2.1.2 *Universeller Zugang für soziale XR über Geräte hinweg: Der Einfluss der Immersion auf die Erfahrung in der asymmetrischen virtuellen Zusammenarbeit*

In der Arbeit »Universal Access for Social XR Across Devices: The Impact of Immersion on the Experience in Asymmetric Virtual Collaboration« (Merz et al., 2024a) untersuchen wir, wie unterschiedliche Eingabe-/Ausgabegeräte und dementsprechend verschiedene Immersionsstufen die Benutzererfahrung in asymmetrischen sozialen XR-Umgebungen beeinflussen. Das primäre Forschungsinteresse liegt darin, den Zugang zu sozialen XR-Umgebungen zu ermöglichen und inklusiver zu gestalten, insbesondere in Szenarien, in denen

nicht alle Nutzer:innen die Möglichkeit haben, hochentwickelte XR-Hardware wie Head-Mounted Displays (HMDs) oder VR-Controller zu nutzen. Dies kann aufgrund von Kosten, Zugänglichkeit oder technischen Einschränkungen der Fall sein. Die Arbeit befasst sich daher mit der Frage, wie solche technischen Unterschiede die Zusammenarbeit zwischen Nutzer:innen mit unterschiedlichen Immersionsstufen beeinflussen. Konkret untersuchen wir, wie die Nutzung von verschiedenen Gerätekonfigurationen—vom Desktop mit Maussteuerung bis hin zu immersiven VR-Setups—die Selbstwahrnehmung (z. B. Präsenz und Verkörperung), die Wahrnehmung anderer (soziale Präsenz und Co-Präsenz) sowie die Wahrnehmung und Bewältigung einer gemeinsamer Aufgaben beeinflusst. Das Ziel ist es, zu prüfen, ob und inwiefern Nutzer:innen mit geringerer Immersion oder ihre Interaktionspartner:innen durch die Verwendung weniger immersiver Geräte in ihrer Fähigkeit zur Zusammenarbeit eingeschränkt sind und ob es möglich ist, XR-Umgebungen zu entwickeln, die eine qualitativ hochwertige Interaktion auch dann ermöglichen, wenn die technologische Ausstattung der Nutzer:innen asymmetrisch ist.

Evaluation

Zur Evaluierung der Forschungsfragen wurde eine Studie mit 87 Teilnehmer:innen durchgeführt, die in Paaren (Dyaden) zusammenarbeiteten, um in einer virtuellen Umgebung eine Sortieraufgabe zu lösen. Diese Aufgabe verlangte von den Teilnehmer:innen, verschiedene Objekte in einem virtuellen Raum zu sortieren. Die Studie umfasste drei unterschiedliche Gerätekonfigurationen: (i) einen Desktop-Bildschirm mit Maussteuerung, (ii) einen Desktop-Bildschirm mit VR-Controllern und (iii) ein vollständig immersives Setup mit HMDs und VR-Controllern. Diese Konfigurationen repräsentieren unterschiedliche Immersionsstufen, wobei die Desktop-Maus-Konfiguration die niedrigste und das HMD-Setup die höchste Immersion bietet. Die Studie war so angelegt, dass jeweils zwei Nutzer:innen mit verschiedenen Gerätekonfigurationen in einer Zusammenarbeit agierten, um zu simulieren, wie eine Person mit einem hochimmersiven Gerät und die andere mit einem weniger immersiven Gerät zusammenarbeiten im Vergleich zu zwei Personen mit hochimmersiven Endgeräten. Die Studie untersucht dabei verschiedene Aspekte der Benutzererfahrung, wie beispielsweise das Gefühl der Präsenz und die Verkörperung. Darüber hinaus wurde die Wahrnehmung des Gegenübers evaluiert mit Fragebögen, die die Co-Präsenz (das Gefühl, aktiv mit den anderen zu interagieren) und soziale Präsenz (das Gefühl, gemeinsam mit anderen in der virtuellen Welt anwesend zu sein) als auch die Plausibilität des Gegenübers betreffen. Zudem wurde die Aufgabenbelastung, die Benutzerfreundlichkeit und der Spaß an der Aufgabe

erfasst, um zu verstehen, wie sich unterschiedliche Immersionsstufen auf die Zufriedenheit und Leistung der Nutzer:innen auswirken. Die Studie erlaubt tiefere Einblicke in die Auswirkungen der asymmetrischen Konfigurationen auf die Zusammenarbeit in sozialen XR-Umgebungen.

Ergebnis

Die Ergebnisse der Studie zeigen signifikante Unterschiede in der Selbstwahrnehmung, abhängig von der Immersionsstufe des verwendeten Geräts. Teilnehmer:innen, die weniger immersive Geräte wie Desktop-Bildschirme mit Maussteuerung nutzten, berichteten von einem deutlich geringeren Gefühl der räumlichen Präsenz und Verkörperung im Vergleich zu denen, die HMDs verwendeten, welche eine größere visuelle und sensorische Kongruenz bieten. Die höhere Immersion ermöglichte den Nutzer:innen ein stärkeres Gefühl des »Dabeiseins« in der virtuellen Umgebung sowie eine intensivere Identifikation mit ihrem virtuellen Körper (Avatar). Dies war insbesondere bei Nutzer:innen mit VR-Controllern und HMDs der Fall, die ein höheres Maß an visuomotorischer Synchronität erfuhren, was zu einer stärkeren Wahrnehmung der Körperkontrolle führte. Trotz dieser signifikanten Unterschiede in der Selbstwahrnehmung stellte sich überraschenderweise heraus, dass die Wahrnehmung der anderen Teilnehmer:innen und die Qualität der sozialen Interaktion kaum von der eigenen Immersion oder der des Gegenübers beeinflusst wurden. Sowohl die soziale Präsenz als auch die Plausibilität des Gegenübers und die Co-Präsenz waren bei allen Teilnehmer:innen auf einem vergleichbaren Niveau, unabhängig davon, ob sie ein hochimmersives oder ein weniger immersives Gerät nutzten. Dies deutet darauf hin, dass asymmetrische Zusammenarbeit auch bei unterschiedlichen Immersionsstufen möglich ist, ohne dass die Qualität der sozialen Interaktion erheblich leidet. Darüber hinaus wurden auch die Wahrnehmung der Aufgabe, die Benutzerfreundlichkeit und der Spaß an der Aufgabe durch die unterschiedlichen Immersionsniveaus nicht wesentlich beeinträchtigt. Dies ist ein wichtiger Befund, da er darauf hinweist, dass Nutzer:innen mit weniger immersiven Geräten nicht zwangsläufig benachteiligt sind, wenn es um die Erledigung gemeinsamer Aufgaben in sozialen XR-Umgebungen geht. Insgesamt unterstützen diese Ergebnisse die Argumentation, dass soziale XR-Umgebungen so gestaltet werden können, dass sie eine effektive Zusammenarbeit über verschiedene technische Plattformen hinweg ermöglichen, was die Zugänglichkeit und Inklusion in solchen Systemen erheblich verbessert. Diese Forschung leistet somit einen wichtigen Beitrag zum Verständnis, wie immersive und nichtimmersive Technologien erfolgreich integriert werden können, um universellen

Zugang zu sozialen XR-Umgebungen zu bieten, ohne die Zufriedenheit von Nutzenden oder die Interaktionsqualität zu beeinträchtigen.

2.1.3 Spielt die Stimme eine Rolle? Die Auswirkung von verbaler Kommunikation und Asymmetrie auf die Erfahrung von kollaborativen sozialen XR-Umgebungen

In der Arbeit mit dem Titel »Does Voice Matter? The Effect of Verbal Communication and Asymmetry on the Experience of Collaborative Social XR« (Merz et al., 2024c) wurden die Auswirkungen von asymmetrischen Gerätekonfigurationen und verbaler Kommunikation auf die Benutzererfahrung in kollaborativen sozialen XR-Umgebungen untersucht. Da soziale XR zunehmend für die digitale Zusammenarbeit an Bedeutung gewinnt, untersucht diese Studie, wie unterschiedliche Geräte, insbesondere HMDs und Desktop-Setups, die Selbstwahrnehmung, die Wahrnehmung anderer und die Aufgabenleistung der Benutzer:innen beeinflussen. Diese Fragestellung ist besonders relevant im Kontext sozialer XR, in der Teilnehmer:innen möglicherweise mit Geräten interagieren, die sehr unterschiedliche Immersionsgrade bieten. Dabei liegt der Fokus darauf, wie diese Geräte-Asymmetrie zentrale Interaktionsdynamiken wie das Gefühl der Präsenz, der sozialen Präsenz und die Wahrnehmung der Aufgabe beeinflusst, was entscheidend für die digitale Zusammenarbeit ist, bei der Benutzer:innen auf unterschiedliche Technologien angewiesen sind. Die Studie beantwortet folgende Forschungsfrage:

Führt die Verwendung von symmetrischen und asymmetrischen Geräte-Setups zu einem stärkeren Einfluss auf das Nutzererlebnis und die Qualität der Interaktion, wenn soziale Signale (d. h. Sprache) hinzugefügt werden?

Das Ziel ist es, zu verstehen, wie immersive Technologien ohne Qualitätsverlust mit weniger immersiven Plattformen integriert werden können, um eine hohe Interaktionsqualität und effektive Zusammenarbeit zu gewährleisten, insbesondere wenn verbale Kommunikation vorhanden ist.

Evaluation

Um die Forschungsfrage zu evaluieren, wurde eine soziale XR-Anwendung entwickelt und eine Studie mit 52 Teilnehmer:innen durchgeführt. Diese arbeiteten paarweise (in Dyaden) zusammen, um in einer virtuellen Umgebung eine Sortieraufgabe zu lösen. Die Studie wurde in symmetrische und asymmetrische Gerätekonfigurationen unterteilt, in der die beiden Teilnehmenden miteinander reden konnten: Eine Gruppe bestand aus Teilnehmer:innen, die beide HMDs mit VR-Controllern verwendeten, während in der asymmetrischen Konfiguration eine Teilnehmer:in ein HMD und die andere Person einen Desktop mit Maussteue-

rung benutzte. Die Daten für die Bedingungen ohne Sprache stammten aus einer vorherigen Veröffentlichung. Die zentralen unabhängigen Variablen waren die Immersionsstufen der Geräte und die Anwesenheit oder Abwesenheit verbaler Kommunikation. Die Studie zielte darauf ab, drei Hauptaspekte der Benutzererfahrung zu messen: Selbstwahrnehmung (z. B. Präsenz und Verkörperung), Fremdwahrnehmung (z. B. soziale Präsenz und Co-Präsenz) und Aufgabenwahrnehmung (z. B. Arbeitsbelastung und Benutzerfreundlichkeit). Eine Reihe validierter Fragebögen und Skalen wurde verwendet, um diese Dimensionen zu bewerten. Die Daten wurden mittels statistischer Methoden, einschließlich ANOVA-Tests, analysiert, um festzustellen, wie diese Faktoren die Erfahrungen der Teilnehmer:innen unter den verschiedenen Bedingungen beeinflussten.

Ergebnis

Die Ergebnisse der Studie ergaben mehrere signifikante Erkenntnisse. Erstens verbesserten höhere Immersionsstufen, wie sie durch HMDs erreicht werden, die Selbstwahrnehmung erheblich, insbesondere in Bezug auf Präsenz und Verkörperung. Teilnehmer:innen, die HMDs nutzten, berichteten von einem stärkeren Gefühl des »Dabeiseins« in der virtuellen Umgebung, und dieser Effekt wurde durch die Möglichkeit verbaler Kommunikation weiter verstärkt. Allerdings hatte die verbale Kommunikation keinen Einfluss auf das Gefühl der Verkörperung, das stärker mit der durch das Gerät bereitgestellten visuomotorischen Synchronität verbunden war. In Bezug auf die Fremdwahrnehmung wurde die soziale Präsenz signifikant verbessert, wenn verbale Kommunikation bei hochimmersiven Geräten möglich war. Die Co-Präsenz (das bloße Bewusstsein für die Anwesenheit einer anderen Person) wurde nicht signifikant von der Gerätekonfiguration oder der verbalen Kommunikation beeinflusst, während die soziale Präsenz (die Qualität der Interaktion) sehr sensibel auf diese Faktoren reagierte. Interessanterweise blieb die Interaktionsqualität für Benutzer:innen mit immersiven Geräten hoch, selbst in asymmetrischen Kollaborationen, bei denen die Partner:innen ein weniger immersives Desktop-Setup verwendeten. Die Wahrnehmung der Aufgabe hingegen wurde weder durch die Gerätekonfiguration noch durch die verbale Kommunikation wesentlich beeinflusst, was darauf hindeutet, dass die Aufgabe selbst — das Sortieren von Objekten in einer virtuellen Umgebung — nicht anspruchsvoll genug war, um Unterschiede in der Leistung zwischen den Bedingungen deutlich zu machen. Insgesamt lässt sich daraus schließen, dass verbale Kommunikation ein wesentlicher Faktor zur Verbesserung der sozialen Präsenz ist und dass Personen in XR-Umgebungen mit asymmetrischen Geräte-Setups erfolgreich integrieren können, ohne die Qualität der Zusammenarbeit zu beeinträchtigen, wenn sie hochimmersive Geräte-Setups verwenden. Die Ergebnisse

deuten jedoch auch darauf hin, dass Entwickler:innen von sozialen XR-Anwendungen das Gleichgewicht zwischen Geräte-Immersion und der Reichhaltigkeit sozialer Hinweise sorgfältig abwägen müssen, um eine effektive Zusammenarbeit über unterschiedliche Immersionsniveaus hinweg zu gewährleisten.

2.2 Sprachbasierte Gestengenerierung für soziale XR

Sprachbasierte Gestengenerierung bietet große Potenziale für Nutzer:innen mit weniger immersiven Endgeräten, die keine Bewegungserfassung von Nutzer:innen erlauben. Diese Generierung ermöglicht die Teilnahme mit verkörperten Avataren in sozialer XR mit weniger immersiven Endgeräten, um die Bewegungen der Nutzer:innen zu synthetisieren. In den folgenden Abschnitten geht es nun darum, wie diese Generierung umgesetzt und evaluiert wurde. Abbildung 3 zeigt wie die Generierung verwendet werden kann, um eine Interaktion zwischen Smartphone und VR Nutzer:innen zu verbessern.

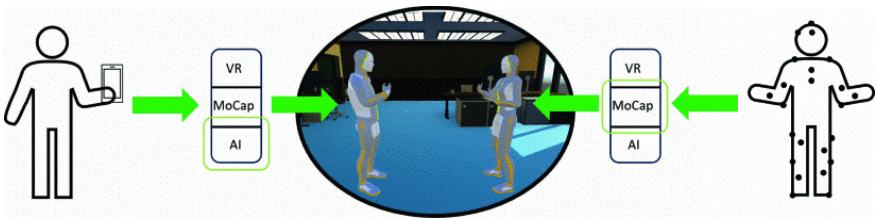


Abbildung 3: Gestengenerierung als Möglichkeit der Avatar-Kontrolle für soziale XR
Quelle: Krome & Kopp, 2023: 2

2.2.1 Co-Speech-Gestengenerierung in Echtzeit für Online-Interaktionen in sozialer XR

Die Arbeit »Towards Real-time Co-Speech Gesture Generation in Online Interaction in Social XR« (Krome & Kopp, 2023) dient der Exploration echtzeitfähiger Modelle zur Generierung sprachbegleitenden Verhaltens sowie der Implementierung einer prototypischen Pipeline zur Integration solcher Systeme in eine soziale XR-Umgebung. Nonverbales Verhalten stellt parallel zu sprachlicher Kommunikation einen wichtigen Stützpfeiler der Verständigung zwischen Menschen dar. Die Übermittlung nonverbalen Verhaltens innerhalb einer virtuellen Umgebung ist allerdings nicht trivial. Um eine vollständig verkörperte Avatar-medierte Unterhaltung zu gewährleisten, wird typischerweise mit aufwendigen und teuren Motion-Tracking-Systemen gearbeitet. Leider haben nicht alle potenziel-

len Nutzer:innen einer sozialen VR-Plattform Zugriff auf solche Hardware, weshalb Alternativen geschaffen werden müssen.

Eine mögliche Alternative ist die Generierung nonverbalen Verhaltens auf Basis der eingehenden Sprache der Nutzer:innen. Existierende Systeme sind bereits in der Lage ein Sprachsignal zu verarbeiten und daraus eine Bewegungssequenz abzuleiten, welche vor allem zum Rhythmus, in Teilen aber auch zum Inhalt der dazugehörigen Sprache passt. Diese Systeme erwarten typischerweise eine Audiospur in Form einer zuvor aufgenommenen Audiodatei als Eingabe und generieren die Bewegungssequenz in einem dafür typischen Dateiformat, welches anschließend mit beliebigen Programmen (Blender, Unity, Unreal etc.) visualisiert werden kann. In diesem Anwendungsfall soll die Generierung der Bewegungen auf einem Echtzeit-Sprachsignal basieren, welches die Nutzer:innen über ihr Mikrofon einsprechen. Die Bewegung soll daraufhin kontinuierlich und möglichst latenzfrei generiert und visualisiert werden.

Die zeitkritische Verarbeitung des Sprachsignals führt zu einer starken Reduktion der zur Verfügung stehenden Informationen, auf deren Basis eine passende Bewegung erstellt werden kann. Der Ansatz zur Gewährleistung einer schnellen und kontinuierlichen Generierung war hier die Einteilung des Sprachsignals in einzelne Abschnitte. Diese Abschnitte werden individuell verarbeitet und nachträglich zu einer zusammenhängenden Bewegung verbunden. Das zugrundeliegende Modell bekommt jeweils eine wenige Sekunden lange Audiosequenz sowie die zuletzt generierte Pose der vorangehenden Bewegungssequenz. Auf Basis dieser Informationen wird die Folgesequenz generiert. Die Bewegungssequenzen werden einzeln in eine virtuelle Umgebung übertragen, wo sie verbunden und auf einem humanoiden Avatar abgespielt werden.

Die Qualität der resultierenden Bewegungssequenz und die Relation zur dazugehörigen Sprache hängen von der Reichhaltigkeit der Eingabesequenz ab. Je mehr Informationen über den Sprachrhythmus sowie den semantischen Inhalt jedes einzelnen Abschnitts extrahiert werden können, desto bessere Bewegungen kann das Modell generieren. Das Eingabesignal in längere Abschnitte zu unterteilen, wirkt sich somit positiv auf die Qualität des nonverbalen Verhaltens des Avatars aus. Umgekehrt proportional beeinflusst die Länge der Abschnitte allerdings die Latenz des Systems, da zur Generierung langer Sequenzen zuvor auf ein ausreichend langes Eingabesignal gewartet werden muss. Nicht zu warten und somit eine Asynchronität zwischen verbalem und nonverbalem Verhalten zu erzeugen, würde zu einer verminderten Gesamtqualität führen. Ein Zurückhalten der Sprache, bis die dazugehörige Bewegung generiert wurde, gewährleistet eine synchrone Übertragung. Zu große Verzögerungen zwischen Spracheingabe und Sprachübertragung an das Gegenüber würden aber in einer interaktiven Anwendung zu Brüchen in der Kommunikation führen. Die optimale Länge der einzel-

nen Eingabesequenzen war somit ein entscheidender Parameter für ein erfolgreiches System.

Die Forschungsfrage lautete also: Können moderne Machine-Learning-Methoden zur Generierung nonverbalen Verhaltens (hier vor allem Gesten) auf Basis eines generativen Modells verwendet werden, um fehlende Bewegungsdaten in Avatar-mediierter Online-Kommunikation zu ersetzen und wie viel Latenz erzeugt die Generierung einer Bewegungssequenz von ausreichender Qualität?

Evaluation

Die Evaluation der Echtzeit-Gestengenerierungs-Pipeline erfolgte anhand der typischen Maße zur Erfassung der Qualität sprachbegleitender Gesten. Es wurden Beispielsequenzen mit verschiedenen Parametern erzeugt und in einer Online-Studie bewertet. Gefragt wurde u. a., ob die auf einem Avatar angezeigten und mit dem dazugehörigen Sprachsignal synchronisierten Bewegungen menschlich wirken und ob sie zum Gesagten passen. Die verschiedenen Sequenzen wurden in Echtzeit generiert, indem das Eingangssignal in entweder 0,5, 1, 2 oder 5 Sekunden lange Abschnitte unterteilt wurde. Verglichen wurde zudem mit einer Sequenz ohne Echtzeitgenerierung sowie mit der echten, zur Audiospur gehörenden Bewegungssequenz, die mittels Motion Capture aufgenommen wurde.

Ergebnis

Die Ergebnisse konnten bestätigen, dass generative Modelle zur Erzeugung sprachbegleitenden Verhaltens verwendet werden können und diese von Beobachter:innen als hinreichend menschlich und passend empfunden werden. Außerdem wurde ein beispielhaftes Modell in eine Echtzeit-Visualisierungs-Pipeline eingebaut, ohne die Qualität der erzeugten Bewegungen grundsätzlich einzuschränken. Die Qualität des erzeugten Verhaltens nimmt aber deutlich ab, wenn die Abschnitte, in die das eingehende Audiosignal unterteilt wurde, zu klein wurden. Ab einer Abschnittslänge von 2 Sekunden oder weniger wurde die Qualität der Gesten signifikant schlechter bewertet als ohne Echtzeiteinschränkungen. Vor allem die Relation zum Inhalt der Sprache litt deutlich. Eine Latenz von 2 oder mehr Sekunden wäre in einem interaktiven System viel zu hoch, weshalb weitere Arbeiten an der Pipeline geplant wurden.

2.2.2 *Sprachgesteuerte Gestenerzeugung mit minimaler Latenz für kontinuierliche Interaktion in sozialer XR*

Die Arbeit »Minimal Latency Speech-Driven Gesture Generation for Continuous Interaction in Social XR« (Krome & Kopp, 2024) stellt eine Fortführung der Ar-

beit »Towards Real-time Co-speech Gesture Generation in Online Interaction in Social XR« dar und führte zu einer maßgeblichen Optimierung des bestehenden Systems. Die vorangehenden Arbeiten haben gezeigt, dass generative Modelle zur Erzeugung nonverbalen Verhaltens auf Basis eines in Echtzeit eingehenden Audiosignals qualitativ hochwertige Bewegungssequenzen erzeugen können. Dies macht sie zu einer potenziellen Lösung, um fehlende Bewegungsdaten von Teilnehmer:innen sozialer virtueller Räume auszugleichen, wenn diese keine Motion-Tracking-Hardware besitzen, und ihnen somit eine vollständig verkörperte Erscheinung zu ermöglichen. Es hat sich aber auch gezeigt, dass die Qualität der synthetisierten Bewegungen und die durch die Generierung erzeugte Latenz in unterschiedlicher Weise mit der Länge der Abschnitte zusammenhängen, in welche ein kontinuierliches Audiosignal aus einem Mikrofon zur Analyse unterteilt werden muss. Je länger die individuell analysierten Abschnitte sind, desto mehr Informationen können daraus extrahiert werden, was die Qualität der Bewegungen verbessert, umso länger muss aber auch vorab gewartet werden, bis die erste Bewegung auf dem Avatar der Nutzer:innen visualisiert werden kann. Die dadurch entstehenden Verzögerungen wirken sich negativ auf die Qualität und Robustheit der vermittelten Kommunikation aus.

Eine Limitation war die individuelle Betrachtung der einzelnen Segmente. Jedes Audiopaket wurde in Isolation betrachtet und lediglich die zuletzt generierte Pose der vorangehenden Sequenz wurde als Startposition der nächsten Sequenz übergeben, um eine kontinuierliche und bruchfreie Bewegung zu erzeugen. Informationen über bereits analysierte Audiosequenzen gingen verloren. Daher war es entscheidend, dass jede einzelne Sequenz von ausreichender Länge war, was eine Mindestlatenz des Systems zur Folge hatte. Durch eine Modifikation des zugrundeliegenden neuronalen Netzes konnten die über die Laufzeit gesammelten Langzeitinformationen in Form eines persistenten »Hidden States« über mehrere Abschnitte erhalten bleiben. Durch diese Hinzunahme und Übertragung von Kontextinformationen sollten Bewegungssequenzen von ausreichender Qualität sogar für wenige Millisekunden eines Audiosignals erzeugt werden können.

Deshalb behandelt dieser Artikel demnach folgende Forschungsfrage: Kann die Latenz einer Pipeline zur Echtzeitgenerierung nonverbalen Verhaltens auf Audiobasis minimiert werden, ohne die Qualität der erzeugten Bewegungen zu beeinflussen?

Evaluation

Zur Evaluation des verbesserten Modells wurden erneut Beispielsequenzen generiert unter Verwendung verschiedener Abschnittslängenparameter. Aufgrund

der Kürze der Abschnitte werden sie nunmehr in »Frames« ausgedrückt. Ein Frame entspricht bei einer Abspielfrequenz von 60 Frames pro Sekunde etwa 16 Millisekunden. Verglichen wurden die Einteilung in 20, 10 und 5 Frames lange Abschnitte sowie in einzelne Frames. Außerdem wurde erneut mit einer Bewegungssequenz ohne Echtzeiteinschränkungen verglichen sowie mit der zuvor kürzesten Abschnittslänge des alten Modells von einer halben Sekunde (also 30 Frames). Gemessen wurden dieselben Qualitätsmaße, also die Menschlichkeit der Bewegungen sowie deren Relation zur verbalen Äußerung.

Ergebnis
Die Ergebnisse konnten die Qualität der nichteingeschränkten Bewegungen sowie der Kontrollsequenz aus 0,5-Sekunden-Abschnitten erneut bestätigen. Relativ zu diesen Bewertungen ließ sich zudem zeigen, dass die Verbesserungen zu einer gleichbleibenden Qualität über alle Abschnittslängen hinweg geführt haben, die jeweils auf dem Niveau der uneingeschränkten Sequenz war. Somit konnte die Latenz des Gesamtsystems bis hinunter auf einen einzigen Frame reduziert werden, ohne die Qualität der generierten Gesten einzuschränken. Eine solche Pipeline kann von nun an in soziale XR-Umgebungen integriert und ohne negative Einflüsse auf die Kommunikationsdynamiken als Ersatz für Motion-Tracking-Systeme verwendet werden, um Nutzer:innen ohne entsprechende technische Ausstattung eine vollständig verkörperte Erscheinung im virtuellen Raum zu ermöglichen.

2.3 Physikbasierte Animationen von Gesichtsbewegungen für sozial expressive Avatare

Bisher werden meistens lineare Blendshape-Gewichte verwendet, um Gesichtsanimationen auf virtuellen Avataren darzustellen. Dabei ermöglichen physikbasierte Simulationen eine verbesserte Ausdrucksstärke im Vergleich zu linearen Blendshapes. In den folgenden Abschnitten geht es darum, wie diese auch effizient auf weniger leistungsstarken Endgeräten berechnet werden können, um die Teilhabe von Nutzer:innen mit weniger immersiven – dementsprechend oft leistungsschwächeren – Endgeräten nicht einzuschränken. Dadurch verbessert sich der Realismusgrad der sozial expressiven Avatare und die Verbesserungen erlauben weiterhin, diese auch auf verschiedensten Endgeräten zu verwenden. Abbildung 4 zeigt den datengetriebenen Ansatz um die Informationen auf Haut, Muskeln und Knochen dargestellt werden.

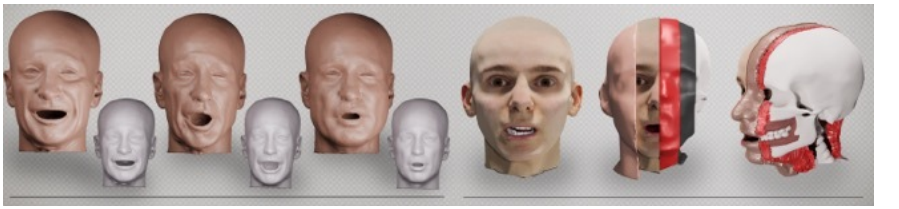


Abbildung 4: Generierung der physikbasierten Gesichtsanimation
Quelle: Wagner et al., 2023: 1

2.3.1 AnaConDaR: Anatomisch eingeschränktes datenadaptives Retargeting von Gesichtern

In der Arbeit mit dem Titel »AnaConDaR: Anatomically-Constrained Data-Adaptive Facial Retargeting« (Wagner et al., 2024a) untersuchen wir, wie physikbasierte Simulationen genutzt werden können, um personalisierte und immersive Gesichtsanimationen zu erzeugen. Oftmals wird das Gesicht einer Person in virtuellen Realitäten animiert, indem abstrakt aufgezeichnete oder in Echtzeit getrackte Gesichtsausdrücke auf ein Oberflächennetz der Person übertragen werden. Eine der beliebtesten und häufigsten Darstellungsformen sind dabei lineare Blendshape-Gewichte. Diese erfassen einen beliebigen Gesichtsausdruck als lineare Kombination von Oberflächennetzen, die die Grundausdrücke einer Person repräsentieren. Die Herausforderung besteht darin, dass der linear kombinierte Ausdruck möglichst realistisch bleibt und sich nicht signifikant vom echten Ausdruck unterscheidet. Gleichzeitig sollte die Anzahl der verwendeten Grundausdrücke gering gehalten werden, da deren Erstellung – sei es durch manuelles Modellieren oder Einscannen – oft sehr zeitaufwendig und kostspielig ist. Obwohl bereits viel vorherige Forschung darauf abzielt, die Ausdrucksstärke von linearen Blendshapes zu verbessern, erfordern authentische Animationen oft noch immer Hunderte von Grundausdrücken. AnaConDaR hingegen nutzt eine nichtlineare Kombination von Blendshapes mithilfe einer physikbasierten Simulation, um die Anzahl der benötigten Grundausdrücke signifikant zu reduzieren.

Implementierung
Genauer gesagt, werden die Oberflächen der Grundausdrücke konsistent in zusammenhängende Einzelteile (Patches) unterteilt. Um einen beliebigen Gesichtsausdruck zu erzeugen, werden diese Patches – anstatt vollständiger Grundausdrücke – mit unterschiedlichen Gewichtsvektoren linear kombiniert. Dieses Vorgehen führt in der Summe zwar zu einer nichtlinearen Kombination der Grundausdrücke, kann jedoch sichtbare Kanten an den Rändern der Patches hervorrufen. Ebenso bleiben andere bekannte Nachteile von linearen Blendshapes be-

stehen. Dazu gehört etwa die Schwierigkeit, Kollisionen der Lippen miteinander oder mit den Zähnen zu verhindern. Es kann auch zu anatomisch unrealistischen Veränderungen des Kopfes kommen, wie z. B. einer Veränderung des Volumens durch die lineare Kombination der Patches. Um diese Probleme zu umgehen und die anatomische Plausibilität zu verbessern, wendet der Algorithmus nach dem Blenden der Patches eine volumetrische, physikbasierte Simulation an. Schlussendlich kann AnaConDaR mit nur zehn Grundausdrücken mindestens genauso authentische Gesichtsanimationen erzeugen wie herkömmliche lineare Verfahren mit Hunderten von Blendshapes.

Ergebnis
Die quantitative Evaluierung von AnaConDaR gestaltet sich schwierig, da die Authentizität von Gesichtsausdrücken subjektiv wahrgenommen wird. Um eine fundierte Bewertung zu ermöglichen, wurde daher eine Nutzerstudie durchgeführt, in der verschiedene bestehende Ansätze mit AnaConDaR verglichen wurden. Diese Studie fand online statt und umfasste 33 Teilnehmer:innen, die unterschiedliche Gesichtsanimationen bewerteten. Im Durchschnitt wurde AnaConDaR als »natürlichere« Methode empfunden und konnte die Studienteilnehmer:innen überzeugen.

2.3.2 SoftDECA: Computereffiziente, physikbasierte Gesichtsanimationen

Mit den Arbeiten »SoftDECA: Computationally Efficient Physics-Based Facial Animations« (Wagner et al., 2023) und »SparseSoftDECA: Efficient High-Resolution Physics-Based Facial Animation from Sparse Landmarks« (Wagner et al., 2024b) untersuchten wir, wie physikbasierte Gesichtsanimationen effizient auch auf leistungsschwachen Systemen ausgeführt werden können. Solche Animationen haben das Potenzial, eine deutlich höhere Qualität und Natürlichkeit zu erzeugen als simplere Verfahren wie beispielsweise lineare Blendshapes, die oft zu starr und unrealistisch wirken. Die verbesserte Qualität physikbasierter Animationen bringt jedoch erhöhte Rechenanforderungen mit sich, was insbesondere bei XR-Systemen schnell zu Leistungsengpässen führen kann. Dies ist einer der Hauptgründe, warum herkömmliche, weniger rechenintensive Animationsverfahren weiterhin weit verbreitet sind. SoftDECA und SparseSoftDECA setzen hier an, um realistische und gleichzeitig performante Gesichtsanimationen auch in ressourcenbeschränkten Umgebungen zu ermöglichen. SparseSoftDECA ist dabei als Erweiterung von SoftDECA konzipiert: Während SoftDECA effiziente physikbasierte Gesichtsanimationen ermöglicht,

erweitert SparseSoftDECA dieses System, indem es auch durch gängige und leicht verfügbare Trackingmethoden gesteuert werden kann.

Implementierung
Beide Systeme basieren auf effizienten neuronalen Netzen, die mit einem umfangreichen Korpus an Trainingsdaten trainiert wurden. Das neuronale Netz von SoftDECA ist so konzipiert, dass es mit bestehenden Standards der Gesichtsanimation interagieren kann, etwa durch das Interface linearer Blendshape-Gewichte. Das Apple ARKit, eine weitverbreitete Plattform für Gesichtsanimationen, ermöglicht somit eine nahtlose Integration. Die Ausgabe von SoftDECA erfolgt in Form von animierten Gesichtsausdrücken, die als Oberflächennetze dargestellt werden. Diese Netze können von den meisten gängigen Rendering-Frameworks problemlos verarbeitet werden, was eine ebenso breite Anwendbarkeit sicherstellt. SparseSoftDECA erweitert das Interface von SoftDECA, indem es populäre Tracking-Frameworks wie Google Mediapipe integriert, welche Gesichtsausdrücke in beliebigen Bildern feststellen können.

Neben dem Design der neuronalen Netze bestand eine zentrale Herausforderung dieser Arbeit darin, geeignete Trainingsdaten zu erstellen. Um die Vielfalt der menschlichen Anatomie und der verschiedenen Gesichtsausdrücke abzudecken, wurde erstmalig eine Kombination unterschiedlichster Datenquellen herangezogen. Dazu zählen computertomografische Aufnahmen von Schädeln, um anatomische Strukturen genau zu erfassen, sowie Oberflächenaufnahmen von Köpfen und getrackte Dialoge, um eine repräsentative Vielfalt an Kopfformen und dynamischen Gesichtsausdrücken zu generieren. Dank dieser umfassenden Datenbasis sind sowohl SoftDECA als auch SparseSoftDECA in der Lage, in realen Anwendungen leistungsfähig und akkurat zu arbeiten.

Ergebnis
Der methodische Schwerpunkt dieser Arbeit lag auf der quantitativen Evaluierung der entwickelten Systeme. Beide neuronalen Netze zeigten auf den Testdaten nur minimale Fehler, die kleiner als ein Millimeter waren, was auf eine hohe Genauigkeit und Generalisierungsfähigkeit hindeutet. Gleichzeitig zeichnet sich die Effizienz der Systeme dadurch aus, dass sie auch auf veralteten Prozessoren in der Lage sind, Gesichtsanimationen in Echtzeit zu berechnen. Sie erreichen Bildwiederholraten, die für den Einsatz in XR-Umgebungen vollkommen ausreichend sind und gewährleisten so eine reibungslose, immersive Benutzererfahrung.

Darüber hinaus wurde SoftDECA auf der ACM SIGGRAPH Conference on Motion, Interaction and Games 2023 mit einer »Honorable Mention« für die zweit-

beste Einreichung ausgezeichnet, was die Relevanz und Innovation dieser Forschung unterstreicht.

3. Diskussion

Die im HiAvA-Projekt durchgeführten Forschungsarbeiten haben einen signifikanten Beitrag zur Weiterentwicklung von sozialer XR und der Integration hybrider Avatar-Agenten-Technologien für alltagstaugliche soziale Remote-Interaktion geleistet. Die Ergebnisse der Studien verdeutlichen die Relevanz von skalierbaren Avatar-Systemen, die durch KI und maschinelles Lernen unterstützt werden, um realistische soziale Interaktionen in virtuellen Umgebungen zu ermöglichen. Der Einsatz von datengetriebenen Technologien hat es ermöglicht, nonverbale Kommunikationssignale, wie Gesten und Gesichtsausdrücke, in Echtzeit zu generieren und darzustellen, was entscheidend zur Erhöhung der sozialen Präsenz und der Interaktionsqualität beiträgt.

Ein zentrales Thema der Arbeit war die Schaffung einer inklusiven Plattform, die Nutzer:innen mit verschiedensten VR-Headsets und weniger immersiven Geräten Zugang zu qualitativ hochwertigen sozialen XR-Erfahrungen bietet (Kern & Latoschik, 2023). Dies ist insbesondere in Anbetracht der wachsenden Popularität von XR in verschiedenen Anwendungsbereichen wie Bildung, Kommunikation, Arbeit und Unterhaltung von großer Bedeutung. Die Studien zeigten, dass Nutzer:innen mit unterschiedlichen technischen Ausstattungen – von hochimmersiven VR-Headsets bis hin zu herkömmlichen Desktop-Systemen – in der Lage waren, erfolgreich in sozialen XR-Umgebungen zu interagieren. Dies unterstreicht die Bedeutung der Entwicklung skalierbarer Lösungen, die die technischen Barrieren für die Nutzung von sozialer VR verringern und eine breite Adaption ermöglichen. Die Studien bringen Erkenntnisse, in welchen Situationen die Nutzererfahrung von weniger immersiven Nutzer:innen beeinträchtigt wird. Daraus lässt sich ableiten, wie soziale XR-Anwendungen gestaltet werden müssen. Die Ergebnisse der Nutzerstudien haben gezeigt, dass zwar Nutzer:innen mit hochimmersiven Geräten wie VR-Headsets ein stärkeres Gefühl der Präsenz und Verkörperung erlebten, dies jedoch die Qualität der Interaktion mit anderen Teilnehmer:innen, die weniger immersive Geräte verwendeten, nicht wesentlich beeinträchtigte. Dies deutet darauf hin, dass asymmetrische Interaktionen – also die Zusammenarbeit zwischen Nutzer:innen mit unterschiedlicher technischer Ausstattung – möglich sind, ohne dass die soziale Präsenz oder die Kommunikationsqualität erheblich leidet. Diese Erkenntnis ist von großer Bedeutung für die Entwicklung von sozialen XR-Anwendungen, die eine breite Nutzerbasis ansprechen sollen, da sie zeigt, dass hochimmersive Technologien

und weniger immersive Plattformen erfolgreich integriert werden können, um eine qualitativ hochwertige Interaktion zu ermöglichen. Abbildung 5 zeigt wie Nutzer:innen mit verschiedensten Endgeräten (VR-Headset, Desktop-PC, und Smartphone) der virtuellen Welt beitreten können und dort miteinander interagieren können. Die zusammengefassten Studien verwenden verschiedene User-Experience(UX)-Indikatoren, die für die Evaluation für XR-Anwendungen verwendet werden (Latoschik & Wienrich, 2022; Wienrich & Gramlich, 2020; Wienrich et al., 2021).



Abbildung 5: Geräteübergreifende soziale XR
Quelle: Projekteigene Darstellung

Ein weiteres wichtiges Ergebnis der Forschung ist die Erkenntnis, dass Realitätsnähe und Glaubwürdigkeit der Avatare nicht nur von der visuellen Wiedergabetreue abhängen, sondern auch maßgeblich durch die Integration von nonverbalen Verhaltensweisen beeinflusst wird. Die entwickelten Technologien zur

Echtzeitgenerierung von Gesten und Bewegungen zeigten, dass Nutzer:innen die Avatare als wesentlich menschlicher und natürlicher wahrnahmen, wenn diese in der Lage waren, nonverbale Signale wie Gestik und Mimik zu synchronisieren. Besonders bemerkenswert ist, dass die Generierung dieser Bewegungen auch ohne aufwendige Hardware wie Motion-Capture-Systeme möglich war. Dies stellt einen bedeutenden technologischen Fortschritt dar, da es die Einstiegshürden für Nutzer:innen ohne spezielle Möglichkeiten der Bewegungserfassung erheblich senkt. Dennoch gibt es weiterhin Herausforderungen, die es zu adressieren gilt. Eine der größten Hürden bleibt die Latenz in der Generierung und Visualisierung von nonverbalen Verhaltensweisen in Echtzeit. Die durchgeführten Untersuchungen zur Gestengenerierung auf Basis von Echtzeit-Sprachanalysen zeigten, dass es schwierig ist, die optimale Balance zwischen der Länge der analysierten Sprachabschnitte und der Generierung von qualitativ hochwertigen Bewegungen zu finden. Längere Sprachabschnitte liefern detailliertere Informationen und ermöglichen genauere Bewegungen, gehen jedoch mit einer höheren Latenz einher. Kürzere Abschnitte reduzieren die Latenz, können jedoch die Qualität und Kohärenz der generierten Gesten beeinträchtigen. Zukünftige Arbeiten müssen sich daher auf die Optimierung dieser Prozesse konzentrieren, um eine möglichst latenzfreie, aber gleichzeitig realistische Bewegungsdarstellung zu gewährleisten.

Für die realistische Darstellung von Avataren und Agenten wurden physikbasierte und datengetriebene Gesichtsanimationen entwickelt. Durch den Einsatz neuronaler Netzwerke werden physikbasierte Simulationen mit hoher Genauigkeit und reduzierten Rechenkosten ermöglicht, was realistische Animationen auch auf Geräten mit begrenzten Ressourcen zugänglich macht. Ein hybrider Ansatz, der daten- und geometriegetriebene Methoden kombiniert, sorgt für eine individualisierte und anatomisch präzise Übertragung von Gesichtsanimationen, selbst wenn nur wenige Ausgangsdaten zur Verfügung stehen. Dieser Ansatz bewahrt dabei charakteristische Ausdrucksmerkmale des Gesichts und steigert so die Ausdruckskraft und Realitätsnähe der Animationen. Insgesamt ermöglichen die vorgestellten Technologien eine flexible Anwendung in Echtzeitszenarien und bieten Raum für kreative und funktionale Anpassungen, etwa bei der Simulation von Alterung oder in medizinischen Kontexten. Diese Entwicklungen schaffen eine solide Grundlage für neue, dynamische und individualisierbare Animationen auf einer breiten Palette von Hardwareplattformen. Diese Ansätze wurden in einer ersten Nutzerstudie untersucht, die darauf hindeutet, dass sie natürlichere Ergebnisse als herkömmliche Ansätze liefern.

Zukünftige Evaluationen müssen darauf achten, eine noch breitere Personengruppe bei der Evaluation zu integrieren. Zwar sollten bei Evaluationen auch Personen mit unterschiedlichem Hintergrund (Alter, Bildung und Beruf) berücksich-

tigt werden, aber der Großteil der Testpersonen waren Studierende. Eine diversere Stichprobe wird es ermöglichen, die bisherigen Erkenntnisse allgemeingültiger zu verifizieren.

Ein weiterer interessanter Aspekt, der im Rahmen der Forschung identifiziert wurde, ist das Potenzial für die Integration solcher Technologien in Bereiche wie das Bildungswesen, die medizinische Rehabilitation und die berufliche Weiterbildung. Die Fähigkeit, realistische und personalisierte Avatare geräteübergreifend zu verwenden, eröffnet neue Möglichkeiten für die Fernschulung und Zusammenarbeit in immersiven Umgebungen (Wienrich et al., 2018). Soziale XR könnte in der Zukunft eine Schlüsselrolle bei der Überbrückung von Distanzen und der Schaffung interaktiver und ansprechender Lernumgebungen spielen. Insbesondere im Bereich der Gesundheitsfürsorge und Therapie könnte die Nutzung von Avataren zur Simulation sozialer Interaktionen sowie zur Unterstützung von Patient:innen in virtuellen Rehabilitationsumgebungen bedeutende Vorteile bringen. Außerdem können auch andere Bereiche von diesen Erkenntnissen profitieren, zum Beispiel bei der Nutzung von sozialer XR für Gerichtsverhandlungen (Wienrich et al., 2022).

4. Fazit

Das HiAvA-Projekt hat bedeutende Fortschritte in der Entwicklung skalierbarer hybrider Avatar-Agenten-Technologien erzielt, die soziale Interaktionen in XR-Umgebungen realistisch, inklusiv und geräteübergreifend ermöglichen. Durch die Kombination aus KI-gestützter Animation, maschinellem Lernen und nutzerzentriertem Design wurden wichtige Herausforderungen bewältigt, die bislang den breiten Einsatz sozialer XR-Technologien eingeschränkt haben. Ein zentrales Ziel war es, immersive soziale Interaktionen auch für Nutzer:innen mit weniger leistungsstarken Geräten zugänglich zu machen. Dieses Ziel wurde erreicht, wie Studien zur asymmetrischen Zusammenarbeit belegen. Sie zeigen, dass Nutzer:innen mit unterschiedlichen Immersionsstufen erfolgreich interagieren können, ohne dass die Qualität der sozialen Präsenz oder Kommunikation erheblich leidet.

Mit fortschrittlichen KI-Modellen wurde es möglich, nonverbales Verhalten basierend auf Spracheingaben zu erzeugen und Avatare menschlicher und natürlicher wirken zu lassen. Die Implementierung physikbasierter Gesichtsanimationen hat ebenfalls gezeigt, dass realistische und ausdrucksstarke Animationen auch auf leistungsschwachen Geräten möglich sind, wodurch die Barrieren für den Zugang zu sozialen XR-Anwendungen deutlich reduziert wurden. Dies ermöglicht die Echtzeitgenerierung von Gesten und Gesichtsanimationen für ver-

schiedenste Endgeräte. Studien zur Nutzung unterschiedlicher Gerätekonfigurationen und der Rolle verbaler Kommunikation verdeutlichten, dass technische Unterschiede keine unüberwindbaren Hindernisse für erfolgreiche Zusammenarbeit darstellen, wenn die Systeme sorgfältig gestaltet sind.

Gleichzeitig hat HiAvA wichtige Erkenntnisse für zukünftige Anwendungen in Bereichen wie Bildung, Gesundheitswesen und beruflicher Weiterbildung geliefert. Die Fähigkeit, realistische und personalisierte Avatare geräteübergreifend einzusetzen, bietet neue Möglichkeiten für interaktive Lernumgebungen, therapeutische Anwendungen und virtuelle Kollaboration. Trotz der Erfolge bleiben jedoch Herausforderungen bestehen. Insbesondere die Reduktion von Latenzzeiten bei der Bewegungs- und Gestengenerierung sowie die weitere Optimierung der Benutzererfahrung sind zentrale Forschungsfelder. Zudem erfordert die breite Einführung sozialer XR-Technologien eine vertiefte Auseinandersetzung mit ethischen und datenschutzrechtlichen Fragen, um Vertrauen in diese Systeme zu schaffen (Lin et al., 2024; Rack et al., 2023; Rack et al., 2024).

Insgesamt hat das HiAvA-Projekt gezeigt, dass technologische Innovationen die soziale Teilhabe in XR-Umgebungen maßgeblich fördern können. Die entwickelten Lösungen bieten eine solide Grundlage, um soziale XR als alltagstaugliches Werkzeug für Kommunikation, Lernen und Zusammenarbeit zu etablieren und dabei die Kluft zwischen virtuellen und physischen Welten zu überbrücken. Die gewonnenen Erkenntnisse und Technologien markieren einen wichtigen Schritt hin zu einer inklusiveren und immersiveren digitalen Zukunft.

Förderhinweis

Das Projekt HiAvA wurde vom Bundesministerium für Bildung und Forschung (BMBF)[1] gefördert.

Literatur

Bente, Gary et al. (2008). Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-Based Collaborations. In: *Human Communication Research*, Volume 34, Issue 2, Oxford Univerity Press (formerly Wiley-Blackwell), Oxford, S. 287–318. https://doi.org/10.1111/j.1468-2958.2008.00322.x

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

De Gelder, Beatrice, de Borst, Aline W. & Watson, Rebecca (2015). The perception of emotion in body expressions. In: Wiley Interdisciplinary Reviews: *Cognitive Science* Vol. 6.2, Wiley-VCH GmbH, Weinheim, Berlin, S. 149–158. https://doi.org/10.3389/fpsyg.2015.00576

Buschmeier Hendrik, Kopp Stefan (2018). Communicative listener feedback in human–agent interaction: Artificial speakers need to be attentive and adaptive. In: *2018 Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems, Stockholm*, International Foundation for Autonomous Agents, Richland, S. 1213–1221. https://dl.acm.org/doi/proceedings/10.5555/3237383 [Zugriff 20.03.2025]

Cao, Chen et al. (2016). Real-time facial animation with image-based dynamic avatars. In: *ACM Transactions on Graphics*, Vol. 35, No. 4, ACM Inc., New York, https://doi.org/10.1145/2897824.2925873

Fiedler, Marie et al. (2023). Embodiment and personalization for self-identification with virtual humans. In: *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, IEEE VR 2023, Shanghai, S. 799–800. https://doi.org/10.1109/VRW58643.2023.00242

Garrido, Pablo et al. (2016). Reconstruction of personalized 3D face rigs from monocular video. In: *ACM Transactions on Graphics (TOG)*, Vol. 35, No. 3, ACM Inc., New York S. 1–15. https://doi.org/10.1145/289049

Kern, Florian & Latoschik, Marc Erich (2023). Reality Stack I/O: A versatile and modular framework for simplifying and unifying XR applications and research. In: *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, IEEE 2023, Sydney, S. 74–76. https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00023

Krome, Niklas & Kopp, Stefan (2023). Towards real-time co-speech gesture generation in online interaction in social xr. In: *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*, Würzburg, S. 1–8. https://doi.org/10.1145/3570945.3607315

Krome, Niklas & Kopp, Stefan (2024). Minimal Latency Speech-Driven Gesture Generation for Continuous Interaction in Social XR. In: *2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*, IEEE 2024, Los Angeles, S. 236–240. https://doi.org/10.1109/AIxVR59861.2024.00038

Latoschik, Marc Erich et al. (2019). Not alone here?! Scalability and user experience of embodied ambient crowds in distributed social virtual reality. In: *IEEE transactions on visualization and computer graphics*, Vol. 25, No. 5, S. 2134–2144. https://doi.org/10.1109/TVCG.2019.2899250

Latoschik, Marc Erich et al. (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, Association for Computing Machinery, New York, S. 1–10. 2017. https://doi.org/10.1145/3139131.3139156

Latoschik, Marc Erich & Wienrich, Carolin (2022). Congruence and plausibility, not presence: Pivotal conditions for XR experiences and effects, a novel approach. In: *Frontiers in Virtual Reality*, Vol. 3, Frontier Media S.A., Lausanne, S. 694433. https://doi.org/10.3389/frvir.2022.694433

Lin, Jinghuai et al. (2024). Usability, Acceptance, and Trust of Privacy Protection Mechanisms and Identity Management in Social Virtual Reality. In: *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, ISMAR 2024, Bellevue (WA), S. 130–139. https://doi.org/10.1109/ISMAR62088.2024.00027

Magnenat-Thalmann, Nadia & Thalmann, Daniel (2005). Virtual humans: thirty years of research, what next? In: *The Visual Computer*, Vol. 21, Springer-Verlag GmbH, Heidelberg, S. 997–1015. https://doi.org/10.1007/s00371-005-0363-6

Merz, Christian et al. (2024a). Universal Access for Social XR Across Devices: The Impact of Immersion on the Experience in Asymmetric Virtual Collaboration. In: *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, IEEE 2024, Orlando, S. 859–869. https://doi.org/10.1109/VR58804.2024.00105

Merz, Christian et al. (2024b). Pipelining Processors for Decomposing Character Animation. In: *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*, Trier, ACM Inc., New York, S. 1–2. https://doi.org/10.1145/3641825.3689533

Merz, Christian, Wienrich, Carolin & Latoschik, Marc Erich (2024c). Does Voice Matter? The Effect of Verbal Communication and Asymmetry on the Experience of Collaborative Social XR. In: *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Bellevue (WA), S. 1127–1136. https://doi.org/10.1109/ISMAR62088.2024.00129

Oertel, Catharine et al. (2016). Towards building an attentive artificial listener: On the perception of attentiveness in audio-visual feedback tokens. In: *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, Tokyo, ACM Inc., New York, S. 21–28. 2016. https://doi.org/10.1145/2993148.2993188

Rack, Christian et al. (2023). Who is Alyx? A new behavioral biometric dataset for user identification in XR. In: *Frontiers in Virtual Reality*, Vol. 4, Frontier Media Ltd., London, S. 1272234. https://doi.org/10.3389/frvir.2023.1272234

Rack, Christian et al. (2024). Motion Passwords. In: *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*, Trier, ACM Inc., New York, S. 1–11. https://doi.org/10.1145/3641825.3687711

Wagner, Nicolas, Botsch, Mario & Schwanecke, Ulrich (2023). Softdeca: Computationally efficient physics-based facial animations. In: *Proceedings of the 16th ACM SIGGRAPH Conference on Motion, Interaction and Games*, Rennes France, ACM Inc., New York, S. 1–11. https://doi.org/10.1145/3623264.3624439

Wagner, Nicolas, Schwanecke, Ulrich & Botsch, Mario. AnaConDaR: Anatomically-constrained data-adaptive facial retargeting. In: *Computers & Graphics*, Vol. 122, John Wiley & Sons Ltd, Chichester, S. 103988. https://doi.org/10.1016/j.cag.2024.103988

Wagner, Nicolas, Schwanecke Ulrich & Botsch Mario (2024b). SparseSoftDECA—Efficient high-resolution physics-based facial animation from sparse landmarks. In: *Computers & Graphics*, Vol. 119, John Wiley & Sons Ltd, Chichester, S. 103903. https://doi.org/10.1016/j.cag.2024.103903

Waltemate, Thomas et al. (2018). The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. In: *IEEE transactions on visualization and computer graphics*, Vol. 24, No. 4, IEEE, S. 1643–1652. https://doi.org/10.1109/TVCG.2018.2794629

Wienrich, Carolin, Fries, Lennart & Latoschik, Marc Erich (2022). Remote at court: Challenges and solutions of video conferencing in the judicial system. In: *International conference on human-computer interaction*, Online-Tagung, Springer International Publishing, Cham, S. 82–106. https://doi.org/10.1007/978-3-031-05014-5_8

Wienrich, Carolin & Gramlich, Johanna (2020). appRaiseVR–An Evaluation Framework for Immersive Experiences. In: *i-com*, Journal of Interactive Media, Vol. 19, No. 2, de Gruyter Wissenschaftsverlag, Berlin, S. 103–121. https://doi.org/10.1515/icom-2020-0008

Wienrich, Carolin et al. (2021). Spatial presence in mixed realities–considerations about the concept, measures, design, and experiments. In: *Frontiers in Virtual Reality*, Vol. 2, Frontiers Media S.A., Lausanne, S. 694315. https://doi.org/10.3389/frvir.2021.694315

Wienrich, Carolin et al. (2018). Social presence and cooperation in large-scale multi-user virtual reality-the relevance of social interdependence for location-based environments. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, IEEE 2018, Tübingen/Reutlingen, S. 207–214. https://doi.org/10.1109/VR.2018.8446575

Yassien, Amal et al. (2020). A design space for social presence in VR. In: *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences*, Shaping Society, Tallin, ACM Inc., New York, S. 1–12. https://doi.org/10.1145/3419249.3420112

Die Stadt von morgen mitgestalten! Partizipative Stadtplanung mit AR- und VR-Technologien – Chancen und Herausforderungen für die Bürgerbeteiligung – INSPIRER

Jan Stepczynski, Christoph Holtmann, Sebastian Keppler, Thomas Jung, Heike Wiesner und Johann Habakuk Israel

Zusammenfassung

Beteiligung an der Stadtplanung bezieht sich auf die Einbeziehung und Ermächtigung von Bürger:innen in Entscheidungsprozesse der Stadtentwicklung und auf die Überbrückung der Kluft zwischen Entscheidungsträger:innen und Interessengruppen. Digitale Werkzeuge wie Virtual Reality (VR) und Augmented Reality (AR) besitzen ein erhebliches Potenzial zur Verbesserung von Partizipationsprozessen. Dieser Artikel beschreibt den Entwicklungsprozess von digitalen Lösungen im Rahmen des INSPIRER-Projektes, welches die Möglichkeiten von VR- und AR-Technologien nutzt, um eine stärker visuell orientierte Bürgerbeteiligung in urbanen Umgebungen zu ermöglichen. Das Akronym INSPIRER steht für PartizipatIoN in StadtPlanungsprozessen In viRtuEllen und Realen Räumen. Das Projektkonsortium war hierfür besonders transdisziplinär aufgestellt mit Partner:innen aus Wissenschaft, Wirtschaft, Stadtverwaltung sowie Partner:innen für ethische, rechtliche und soziale Implikationen. Während der dreijährigen Projektlaufzeit wurden zwei Anwendungen unter frühzeitiger Einbindung unterschiedlicher Akteur:innen des Stadtplanungsprozesses entwickelt und verbessert. Im Folgenden werden der Entwicklungsprozess skizziert sowie einige der wichtigsten Ergebnisse des Projektes zusammengefasst.

1. Einleitung

In demokratischen Gesellschaften ist Partizipation ein vielschichtiger Begriff, dessen Bedeutung sich im Laufe der Geschichte verändert hat und der nur selten eindeutig definiert werden kann (vgl. Gabriel, 2020: 1 ff.). Der Begriff Partizipa-

tion an sich umfasst Teilhabe, Beteiligung und Mitwirkung an gesellschaftlichen Prozessen und hat sich in Deutschland über die Jahre stark gewandelt. Anfangs war gesellschaftliche Beteiligung hauptsächlich ein Minderheitenphänomen, doch seit den 1960er-Jahren werden immer mehr gesellschaftliche Gruppen in unterschiedlichste Prozesse einbezogen. Seit den 2000er-Jahren erlebte die Partizipation im Zuge der Digitalisierung einen Aufschwung. Allerdings steigt hierdurch auch die Komplexität partizipativer Prozesse, was neue Kompetenzen erfordert (Le Blanc, 2020: 9 ff.). In der politischen Landschaft der Bundesrepublik Deutschland gilt mehr Partizipation inzwischen als allgemein anerkanntes politisches Ziel. Sie basiert auf der aktiven Beteiligung und Zustimmung der Bürger:innen und trägt zur Legitimation der Herrschaft in einem freiheitlich-pluralistischen System bei (vgl. Gabriel, 2020: 1 ff.).

Innerhalb des Kontextes von Stadtplanung bezeichnet der Begriff hauptsächlich das Mitwirken und die Mitbestimmung der Bürger:innen in Planungsprozessen (vgl. Senatsverwaltung für Stadtentwicklung und Umwelt Berlin, 2014: 10). Diese Mitwirkung spielt auch im Zusammenhang mit Digitalisierung eine zunehmend bedeutende Rolle in der Stadtplanung. Eines der wichtigsten Ziele hierbei ist es etwa, sicherzustellen, dass die Entwicklung von Städten den Bedürfnissen und Bestrebungen der Gemeinschaft entspricht (vgl. Aguilar et al., 2021: 1594 ff.; Saßmannshausen et al., 2021: 250 ff.).

Das Projekt für Partizipation in Stadtplanungsprozessen in virtuellen und realen Räumen, kurz INSPIRER, steht als Teil der Projektlinie VAR2 des Bundesministeriums für Bildung und Forschung (BMBF)[1] für die Beteiligung an Stadtplanungsprozessen in virtuellen und realen Räumen. Die Forschenden im INSPIRER-Projekt entwickelten eine Augmented Reality (AR) und eine Virtual Reality (VR) Anwendung, die es Nutzenden ermöglicht, gemeinsam an Stadtplanungsprojekten zu partizipieren. Virtuelle Planungszustände können somit auf immersive und interaktive Weise erlebt werden. Die Interaktion innerhalb der Anwendungen basiert auf intuitiven virtuellen Bedienelementen. Das Ziel des Systems ist es, möglichst breiten Bevölkerungsschichten die Teilnahme an demokratischen Entscheidungsprozessen in der Stadtplanung zu ermöglichen. Der Entwicklungsprozess wurde von einer partizipativen Evaluation begleitet. Hierdurch wurde analysiert, welche Funktionalitäten sowie Aktivitäten sich besonders für die Einbindung unterschiedlicher Nutzendengruppen eignen und ob immersive Lösungen mit umfassenden Interaktionsmöglichkeiten und ortsgenauen 3D-Visualisierungen partizipative Stadtplanungsprozesse unterstützen können. Auch ethische, rechtliche und soziale Aspekte sind wichtige Faktoren (ELSI), die im Entwicklungsprozess berücksichtigt wurden. Für entsprechende

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Fragestellungen war das FrauenComputerZentrumBerlin e. V. (FCZB) als Projektpartner verantwortlich. In diesem Beitrag wird zunächst die Entwicklung der Demonstratoren als VR- sowie als AR-Anwendung beschrieben. Es folgt eine Zusammenfassung des projektbegleitenden Evaluationskonzeptes. Abschließend werden die wichtigsten Projektergebnisse beschrieben und erläutert.

Dieser Beitrag skizziert die Schritte eines vollumfänglich partizipativen Entwicklungsprozesses, einschließlich der Anforderungsanalyse, Demonstratoren-Entwicklung und Auswertung der Ergebnisse. Hervorzuheben ist, dass im Projekt eine unabhängige Lokalisierungsmethode eingesetzt wurde, die auf 3D-Punktwolken basiert (vgl. Alfakhori et. al., 2022: 11 ff.). Durch ihren Einsatz kann die Abhängigkeit von großen Dienstleistern vermieden werden. Der Anspruch, mit dem Projekt digitale Souveränität zu fördern, konnte auf diese Weise eingelöst werden. Diese Methode wurde innerhalb der Projektlaufzeit in einem isolierten Demonstrator getestet.

2. Die partizipative Entwicklung von VR- und AR-Demonstratoren zur Nutzung in Stadtplanungsprozessen

Basierend auf einem partizipativen Prozess entwickelte das Team des Projektes zwei Demonstratoren, um die technischen Möglichkeiten von VR- sowie AR-Technologien in der Stadtplanung für Bürger:innen erfahrbar zu machen und diese untersuchen zu können. Hierbei war es möglich, geplante Bauprojekte virtuell zu erkunden, sich über unterschiedliche Feedbackinteraktionen zu beteiligen und diese zu gestalten oder neue Ideen und Feedback für Stadtprojekte einzubringen.

2.1 VR-Entwicklung

Einer der entwickelten Demonstratoren ist eine VR-Anwendung (siehe Abbildung 1), mit welcher rein virtuelle Umgebungen urbaner Gebiete darstellt werden können. Die Visualisierung der Umgebung wird mithilfe von 3D-Punktwolken-Scans, 3D-Karten und städtischen Karten erstellt, die in 3D-Modelle umgewandelt wurden. Neben dem aktuellen Zustand zeigen die virtuelle Szenen auch die geplanten Projekte in Form eines 3D-Modells. Um das Nutzerengagement zu erhöhen, wurden verschiedene Feedbackmechanismen innerhalb der VR-Erfahrung implementiert. Dazu gehören die Möglichkeit, Bewertungen abzugeben, Textkommentare zu hinterlassen, vorgefertigte virtuelle Objekte einzufügen und

dreidimensionale Freihandskizzen zu erstellen. Durch diese Interaktionen können Nutzer:innen nicht nur die vorgeschlagenen Baumaßnahmen kennenlernen und erkunden, sondern aktiv an deren Gestaltung und Entwicklung teilnehmen.



Abbildung 1: Platzieren eines 3D-Stadtmöbel-Modells im virtuellen Stadtraum mithilfe der INSPIRER-VR-Anwendung.
Quelle: Projekteigene Darstellung

2.2 AR-Entwicklung

Zusätzlich zu dem VR-Demonstrator umfasste das Projekt auch die Entwicklung eines AR-Smartphone-Demonstrators (siehe Abbildung 2). Durch die nahtlose Integration virtueller Bauprojekte in die physische Umgebung ermöglicht es die AR-Erfahrung den Bürger:innen, Planungsstände zu erkunden und sich mit ihnen auseinanderzusetzen, bevor diese verwirklicht werden. Nutzer:innen können Informationen und visuelle Darstellungen von Bauprojekten einsehen, die virtuell an den Orten platziert sind, an denen sie gebaut werden sollen. Alle Interaktionen, die in der VR-App möglich sind, wurden in angepasster Form auch im AR-Smartphone-Demonstrator umgesetzt. So ist es beispielsweise in beiden Demonstratoren umsetzbar, virtuelle Objekte zu platzieren und zu bearbeiten. Die Möglichkeit, Ideen und Feedback in Textform sowie als 3D-Skizzen zu hinterlassen, wurde ebenfalls evaluiert. Die Platzierung virtueller Modelle im urbanen Raum wurde basierend auf Kommentaren der Bürger:innen verbessert. Bei

der technischen Umsetzung der Objektplatzierung wurde unteranderem das AR Foundation Package der Unity Engine verwendet, welches auf verschiedenen Geräten sowohl Google ARCore, Apple ARKit und OpenXR (z.B. für die HoloLens) unterstützt.

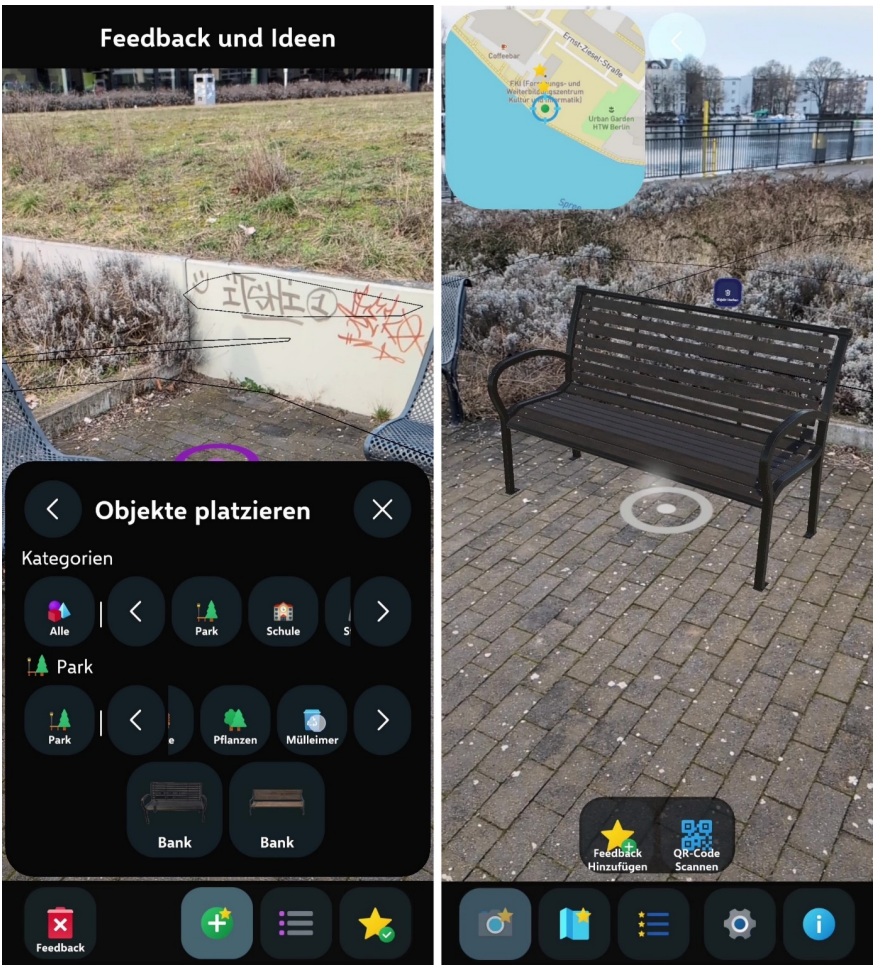


Abbildung 2: AR-Szenen, die eine Benutzeroberfläche für das Platzieren vordefinierter virtueller Objekte (links) und eine platzierte virtuelle Bank (rechts) zeigen.
Quelle: Projekteigene Darstellung

2.3 Realprojekt der Stadt Fellbach zur Veranschaulichung der Demonstratoren

Um die INSPIRER-Demonstratoren in der Praxis analysieren zu können, stand die Stadt Fellbach als Projektpartner des INSPIRER-Projekts zur Verfügung und war auch für die Fallbeispiele zuständig, welche den Bezug zur realen Stadtplanung herstellten. Die Entwicklung der Demonstratoren wurde mithilfe eines Anwendungsszenarios in Fellbach und einer formativen und partizipativen Evaluation durchgeführt. Die assoziierten Partner:innen und verschiedene Bürgergruppen nahmen über den Projektzeitraum kontinuierlich an partizipativen Workshops und Interviews teil, um den Grad der Partizipation und die Akzeptanz digitaler Technologien in städtischen Planungsprozessen zu analysieren.

Eines der wichtigsten partizipativen Ziele des Projektes war es, Bürger:innen so früh wie möglich in den Entwicklungsprozess der Anwendungen einzubeziehen, um neue Erkenntnisse zu gewinnen und die Anwender:innenperspektive in die Entwicklungszyklen zu integrieren. Als reales Szenario wurde für das Projekt eine Kreuzung in Fellbach gewählt, die zu einem Kreisverkehr umgebaut werden soll. In Abbildung 3 ist der geplante Kreisverkehr in Form einer virtuellen Nachbildung in eine virtuelle Repräsentation des Geländes integriert, welche mithilfe von Punktwolken-Scans und 3D-Karten erstellt wurde.



Abbildung 3: Ansichten eines virtuellen 3D-Scans und Punktwolken der Kreuzung in Fellbach mit eingeblendetem 3D-Modell des geplanten Kreisverkehrs
Quelle: Projekteigene Darstellung mit Umgebungsdaten aus Google Maps (oben links) & Projekteigene Darstellung

3. Der Partizipative Evaluationsprozess des INSPIRER-Projekts

Das Evaluierungskonzept des INSPIRER-Projekts verwendete einen Mixed-Methods-Ansatz, um verschiedene Perspektiven auf die Nutzeranforderungen zu sammeln sowie zur qualitativen und quantitativen Validierung der Ergebnisse. Das Feedback wurde während der gesamten Entwicklungsphase der VR- und AR-Apps gesammelt, wobei eine aktive Kommunikation und konstruktive Zusammenarbeit innerhalb des INSPIRER-Projekts durch interne Weitergabe der Ergebnisse der Anforderungsanalyse gewährleistet wurde. Parallel zum Entwicklungsprozess führten die Projektpartner:innen der Hochschule für Wirtschaft und Recht Berlin die Evaluation durch. Diese stellte sicher, dass die erhobenen Anforderungen auf das Gesamtsystem abgestimmt waren und in den Entwicklungsprozess einflossen. Methodisch wurden bei der Evaluation der Demonstratoren sowohl qualitative Einzel- und Gruppeninterviews, teilnehmende Beobachtungen sowie zwei unterschiedliche quantitative Fragebögen verwendet.

Im Rahmen des partizipativen und nutzerzentrierten Prozesses für das INSPIRER-Projekt wurde zunächst eine Bedarfsanalyse und Anforderungserhebung mit elf Stakeholdern im Stadtplanungsprozess durchgeführt. Die Ergebnisse dieser Befragung wurden verwendet, um die ersten Demonstratoren in einem Workshop-Format zu entwickeln. Im Oktober 2022 wurde dann der erste digitale Demonstrator der AR-App mit 16 Bürgerinnen und Bürgern der Stadt Fellbach evaluiert. Dieser Evaluationsprozess bestand aus drei Elementen: einer informativen Einführung in das Projekt und ELSI in der App-basierten Stadtplanung, einem interaktiven Test der Demonstratoren in kleinen Gruppen mit technischer Betreuung sowie einem reflektiven Teil in Form von Interviews und Fragebögen, in denen die Bürger:innen Feedback zum Workshop und zum Entwicklungsstand des Demonstrators geben konnten. Dieser Prozess ermöglichte es den Teilnehmenden, sowohl Kritik als auch konstruktives Feedback zu äußern.

Ergänzt wurden diese Ergebnisse mit teilnehmenden Beobachtungen und Fragebögen, die bei weiteren Workshops und Veranstaltungen nach Testung der Demonstratoren verteilt wurden, sodass bis zum Ende der Projektlaufzeit 78 Fragebögen zur AR-Anwendung und 37 Fragebögen zur VR-Anwendung ausgewertet werden konnten.

4. Ergebnisse der Evaluation

Das folgende Kapitel fasst die wichtigsten Ergebnisse der qualitativen sowie der quantitativen Evaluation zusammen.

4.1 Ergebnisse der qualitativen Evaluation

Die qualitativen Interviews wurden mit der Methode der qualitativen Inhaltsanalyse analysiert, die von Kuckartz und Rädiker entwickelt wurde (vgl. Kuckartz & Rädiker, 2022: 200). Die Kategorien für die Analyse wurden aus der Anforderungserhebung abgeleitet, die im ersten Jahr des Projekts durchgeführt wurde. Der Analyseprozess bestand aus sieben Phasen. In der ersten Phase wurden die Interviews transkribiert und gemäß den Transkriptionsregeln von Kuckartz und Rädiker überarbeitet. In der zweiten Phase wurden die Hauptkategorien entwickelt, entweder aus der Literaturübersicht oder aus den Kategorien der Anforderungserhebung. Anschließend wurden die Transkripte in der dritten Phase nach den Hauptkategorien kodiert. In der vierten Phase wurden auf Grundlage des Interviewmaterials Unterkategorien erstellt, und in der fünften Phase wurde der Text erneut unter Verwendung dieser Unterkategorien kodiert. In der sechsten Phase wurden einfache und komplexe Analysen durchgeführt, die sich auf die Haupt- und Unterkategorien konzentrierten. Dieser Schritt konnte unterschiedliche Formen annehmen. Schließlich wurden die Ergebnisse in Evaluierungsberichten zusammengefasst und der gesamte Prozess ausführlich dokumentiert, um Transparenz und Intersubjektivität, wesentliche Qualitätskriterien qualitativer Forschung, zu gewährleisten. Diese Dokumentation ermöglicht ein vollständiges Verständnis und eine verlässliche Reflexion der Ergebnisse (vgl. Kuckartz & Rädiker, 2022: 132 ff.). Die Ergebnisse der qualitativen Analyse wurden mit den Projektpartner:innen diskutiert, um die darauffolgenden Entwicklungszyklen zu organisieren und die Demonstratoren zu optimieren. Die wichtigsten Erkenntnisse dieser Analyse wurden in sechs Hauptkategorien unterteilt, die im Folgenden ergebnisorientiert zusammengefasst werden:

1. **Inhalt der Anwendung und Schnittstellenaspekte:** Teilnehmer:innen schätzten die Visualisierung für ein besseres Verständnis der Stadtplanung. Der Bezug zu konkreten Stadtplanungsprojekten musste jedoch stärker integriert werden, um Bürger:innen das Gefühl zu geben, dass ihre Beiträge gehört werden. Dies stellt eine Herausforderung aufgrund des komplexen und streng regulierten Stadtplanungsprozesses dar.

2. **Funktionalität der Anwendung:** Eine bessere Kontrolle über angezeigte Objekte wurde besonders in frühen Iterationen der Demonstratoren angesprochen, die Darstellungen in der VR-Umgebung und die Integration zusätzlicher Projektinformationen wurden als wichtig erachtet. Eine Zeichnen-Funktion zur Förderung der Bürgerbeteiligung und Verantwortung wurde integriert und als nützlich sowie interaktionsfördernd empfunden.
3. **Benutzeroberfläche und Benutzererfahrung:** Die Benutzerfreundlichkeit wurde von den Teilnehmenden als wichtiges Kriterium betrachtet. Teilnehmer:innen empfahlen eine verbesserte Zoomfunktion und Anpassung der Sprachoptionen sowie der Textfelder, um inklusiver zu sein und Barrieren für sehbehinderte oder ältere Menschen zu vermeiden. Die Anwendung sollte einfach und benutzerfreundlich sein, um auch technisch unerfahrene Nutzer:innen zu erreichen. Dies führte zu kontinuierlichen Verbesserungen der Anwendung mit geringerer Einarbeitung durch intuitivere Navigation.
4. **Technische Aspekte:** Hier wurden vor allem teilweise lange Zwischenladezeiten in der VR-Anwendung sowie aber auch Probleme in der Objektplatzierung in VR- und AR-Demonstratoren angesprochen.
5. **ELSI (Ethische, rechtliche und soziale Implikationen):** Die Befragten erwarteten die Wahrung des Datenschutzes, ohne dafür Einschränkungen in der Benutzerfreundlichkeit in Kauf nehmen zu müssen. Zur genaueren Adressierung der Zielgruppen wurde außerdem das Evaluationskonzept angepasst, um spezifische Bürgergruppen, wie Frauen und junge Menschen, gezielt zu adressieren.
6. **Beteiligung und Akzeptanz in der Stadtplanung:** Die Visualisierung und Emotionen sind entscheidend für die Beteiligung. Eine interaktive Nutzererfahrung und zusätzliche Informationen fördern das Interesse und die Akzeptanz der Bürger:innen ohne Vorkenntnisse in der Stadtplanung. Zusätzlich sind die in Punkt 3 genannten Aspekte der Benutzerfreundlichkeit ebenfalls wichtige Aspekte für Beteiligung und Akzeptanz.

Die qualitative Evaluation der VR- und AR-App führte zu einigen wesentlichen Schlussfolgerungen für die Verwendung von VR- und AR-Technologien für Partizipationsprozesse in der Stadtplanung:

1. **Positive Bewertung des Workshop-Konzepts:** Die Mehrheit der Teilnehmer:innen bewertete das Workshop-Konzept als sehr positiv. Besonders hervorgehoben wurden die Kombination aus Informationsvermittlung und praktischem Testen der Anwendung sowie die enge Unterstützung durch die Projektpartner:innen während des Workshops.
2. **Beteiligung der Bürgerinnen und Bürger an der Entwicklung wird sehr geschätzt:** Ein unerwarteter positiver Aspekt war, dass diverse Bürger:innen

– gerade unter Berücksichtigung der digitalen Gesellschaft in Deutschland (vgl. Initiative D21, o. J.: 15 ff.) – frühzeitig in Transformationsprozesse einbezogen wurden. Der Ansatz, diverse Bürger:innen in der Testphase zu beteiligen, wurde als sinnvoll erachtet und stieß auf Zustimmung, auch wenn die getesteten Demonstratoren zu Beginn des Projekts noch in einem frühen Entwicklungsstadium waren. Einen Softwareentwicklungsprozess von Anfang an (auch kritisch) zu begleiten, wurde somit als Chance angesehen, eine sehr breite Bürgerperspektive in den Entwicklungsprozess direkt zu implementieren.

3. **Vorschläge zur Verbesserung und Anforderungen:** Die Befragten machten zahlreiche Verbesserungsvorschläge, darunter eine schnellere Entwicklung der App, eine verbesserte Workshop-Organisation, klarere Kommunikation mit der Öffentlichkeit und eine noch stärkere Berücksichtigung von Datenschutzaspekten.
4. **Die technischen Aspekte:** Insbesondere die Benutzerfreundlichkeit der Anwendung wurde diskutiert. Die Anwendung stellte vor allem für Menschen mit Sehbeeinträchtigungen eine Herausforderung dar. Es wurde empfohlen, solche Apps auch auf Tablets verfügbar zu machen und die Möglichkeit zu verbessern, Inhalte zu vergrößern. Diese Anregung wurde in den späteren Demonstratoren auch direkt umgesetzt.
5. **Schnittstellenaspekte und Verbindung zur Stadtverwaltung:** Fragen zur Nutzung der App durch die Stadtverwaltung und zur Integration von Bürgerkommentaren in die Stadtplanung wurden aufgeworfen. Die Befragten äußerten Unsicherheit darüber, ob und wie die Bürgerbeteiligung in den Planungsprozess integriert wird.
6. **Funktionalitätsaspekte:** Die Integration zusätzlicher Informationen über das dargestellte Projekt sowie die Verbesserung der Zeichnen-Funktion wurden als wichtige funktionale Aspekte genannt.
7. **UI/UX-Aspekte (User Interface/User Experience):** Feedback zur Benutzeroberfläche und Benutzererfahrung beinhaltete Vorschläge für eine bessere Anzeige von Interaktionsobjekten, anpassbare Schriftgrößen und eine verbesserte Zeichnen-Funktion. Die Benutzerfreundlichkeit und der Fokus auf wesentliche Funktionen wurden betont.
8. **ELSI (Ethische, rechtliche und soziale Implikationen):** Ethische Bedenken wurden hinsichtlich der Marginalisierung bestimmter Gruppen durch die Technologie geäußert. Datenschutz und rechtliche Aspekte wurden als wichtige Themen für die weitere Entwicklung identifiziert. Soziale Aspekte betrafen hauptsächlich die Zugänglichkeit der App für eine breite Bevölkerung.

4.2 Ergebnisse der quantitativen Evaluation

Im Rahmen des Projekts wurden von den Proband:innen zwei unterschiedliche Fragebögen ausgefüllt, nachdem diese die VR- und AR-Demonstratoren getestet hatten. Hierbei handelte es sich um einen teilstandardisierten projekteigenen Fragebogen zur Einholung von Feedback und zur Bewertung der Anwendung sowie den standardisierten AttrakDiff-Fragebogen zur Bewertung der pragmatischen und hedonischen Qualitäten von Anwendungen, welche sich in vier unterschiedlichen Ausprägungen messen lässt (vgl. Hassenzahl et al., 2008: 79 f.):

- **Pragmatische Qualität (PQ):** Hierbei handelt es sich um die Möglichkeit, mithilfe der INSPIRER-Anwendungen Handlungsziele zu erreichen, indem sie nützliche und benutzbare Funktionen bereitstellen.
- **Hedonische Qualität – Stimulation (HQS):** Beschreibt die Fähigkeit der Anwendungen, das Bedürfnis nach Verbesserung der eigenen Kenntnisse und Fertigkeiten zu befriedigen.
- **Hedonische Qualität – Identität (HQI):** Hierbei werden insbesondere die Fähigkeiten der Anwendungen bewertet, selbstwertdienliche Botschaften zu kommunizieren.
- **Attraktivität (ATT):** Beschreibt die insgesamt Positiv- oder Negativbewertung der Anwendungen.

Die Anzahl der Abstufung bei den Skalen des Fragebogens beträgt 7, wobei höhere Werte bessere Wertungen bedeuten.

Insgesamt konnten 63 AttrakDiff-Fragebögen zur AR-Anwendung und 16 AttrakDiff-Fragebögen zur VR-Anwendung ausgewertet werden. Hierbei ergaben sich die folgenden Mittelwerte:

AttrakDiff-Mittelwerte	VR-Demonstrator	AR-Demonstrator
PQ	4,67	4,74
HQS	6,10	5,26
HQI	5,29	5,02
ATT	5,57	5,48

Tabelle 1: Mittelwerte der AttrakDiff-Dimensionen (Pragmatische Qualität [PQ], Hedonische Qualität – Stimulation [HQS], Hedonische Qualität – Identität [HQI] sowie Attraktivität [ATT]) aus den Evaluationen der VR- und AR-Demonstratoren im Verlauf des Projektes.
Quelle: Eigene Darstellung

Zusammenfassend lässt sich festhalten, dass sowohl die VR- als auch die AR-Anwendung insgesamt positive Bewertungen erhalten haben, wobei es einige Unterschiede in den Ergebnissen der beiden Anwendungen gibt. Die AR-Anwendung erzielte insgesamt mittlere bis gute Werte in allen Bereichen,

mit einer besonders hohen Bewertung in der Attraktivität (ATT: 5,48) und einer soliden Bewertung in der pragmatischen Qualität (PQ: 4,74). Die VR-Anwendung zeigte in Bezug auf die hedonische Qualität – Stimulation (HQS: 6,10) die höchste Bewertung, was auf eine stärkere Wahrnehmung der Anwendung als anregend und lernfördernd hindeutet, während auch die Attraktivität (ATT: 5,57) und die Identität (HQI: 5,29) gut bewertet wurden. Beide Anwendungen wurden als nützlich und funktional wahrgenommen, die VR-Anwendung sticht insbesondere bei dem Merkmal HQS hervor und kann somit besonders dienlich für das Bedürfnis nach Förderung der eigenen Kenntnisse und Fähigkeiten sein. Dies könnte daran liegen, dass der Umgang mit VR-Technologie für viele Proband:innen eine neue Erfahrung war. Allerdings muss man hier hinzufügen, dass die Gruppe der Testenden im Vergleich zur AR-Anwendung durch den Zeitaufwand weitaus kleiner ausfiel und somit die Stichprobengröße einen Effekt auf die Bewertung hat, da Ausreißer stärker ins Gewicht fallen. Insgesamt deuten die Ergebnisse darauf hin, dass beide Technologien ein hohes Potenzial haben, es aber Unterschiede in der Nutzerwahrnehmung und den spezifischen Stärken in den jeweiligen Anwendungsbereichen gibt.

5. Fazit

Die INSPIRER-Anwendungen nutzen VR- und AR-Technologien, um die Bürgerbeteiligung in der Stadtplanung zu fördern. Durch die Einbeziehung der Bürger:innen von Anfang an wird ein inklusiver und transparenter Planungsprozess geschaffen, der auf die tatsächlichen Bedürfnisse der Bürger:innen bei der Stadtplanung und während Beteiligungsprozessen eingeht. Der Entwicklungsprozess ermöglichte es, diverse Perspektiven zu berücksichtigen und ein gemeinsames Verantwortungsgefühl zu fördern. Workshops und Tests mit verschiedenen Bürgergruppen lieferten wertvolles Feedback, das in die Verbesserung und Verfeinerung der Anwendungen einfloss. Die Ergebnisse der qualitativen Evaluation zeigen, dass die Bürger:innen positiv auf den Workshop-Ansatz reagierten und es schätzten, in den Entwicklungsprozess integriert zu werden, selbst wenn die Anwendungen noch nicht final und technisch unausgereift waren. Besondere Bedeutung wurde dabei auf technische Aspekte und ELSI-Fragen (ethische, rechtliche und soziale Implikationen) gelegt, die explizit in Workshops mit Frauen und Jugendlichen thematisiert wurden. Die Akzeptanz und Beteiligung der Teilnehmenden waren insgesamt besonders positiv zu bewerten und können als besondere Chance des Einsatzes von VR- sowie AR-Technologien in der Stadtplanung hervorzuheben. Die Nutzung von VR- sowie AR-Technologien ermöglichte es, die Gegebenheiten vor allem im Hinblick

auf das Realprojekt in Fellbach besser zu visualisieren und die Bereitschaft der Bürger:innen zur aktiven Mitgestaltung und zum Feedback zu erhöhen. Zukünftige partizipative Prozesse können auf diesen Erkenntnissen aufbauen, um die Bürger:innen stärker in die Stadtplanung und die Entwicklung interaktiver digitaler Tools einzubinden. Zusammenfassend lässt sich festhalten, dass sowohl die VR- als auch die AR-Anwendung insgesamt positive Bewertungen erhalten haben. Insgesamt deuten sowohl die Ergebnisse der qualitativen Erhebungen als auch die Ergebnisse der quantitativen Erhebungen darauf hin, dass beide Technologien ein hohes Potenzial vor allem für die Motivation und die Information von Bürger:innen haben, es aber Unterschiede in der Nutzerwahrnehmung und den spezifischen Stärken in den jeweiligen Anwendungsbereichen gibt.

Anhand des Forschungsprojektes INSPIRER wird ersichtlich, dass VR/AR-Technologien neue Wege der Beteiligung eröffnen, die gerade durch ihre einfache Handhabung und die klare Visualisierung von Informationen besonders vielversprechend sind. Eine erfolgreiche Implementierung erfordert jedoch die Zusammenarbeit aller relevanten Akteur:innen und die Integration in den bestehenden rechtlichen Rahmen, was weitere Forschung und Entwicklung zur Integration technologischer Lösungen in praktische Stadtplanungsprozesse notwendig macht.

Förderhinweis

Das Projekt INSPIRER wurde vom Bundesministerium für Bildung und Forschung (BMBF)[2] gefördert.

Literatur

Aguilar, Jose, Díaz, Francisco, Altamiranda, Junior et al. (2021). Metropolis: Emergence in a Serious Game to Enhance the Participation in Smart City Urban Planning. In: Carayannis, Elias G. (Hg.), *Journal of the Knowledge Economy*, Vol. 12, S. 1594–1617. https://doi.org/10.1007/s13132-020-00679-5.

Alfakhori, Muhammad, Dastageeri, Habiburrahman, Schneider, Sven et. al. (2022). Occlusion Screening Using 3D City Models as a Reference Database for Mobile AR-Applications. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. X-4/W3-2022. S. 11–18. 10.5194/isprs-annals-X-4-W3-2022-11-2022.

2 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Fegert, Jonas, Pfeiffer, Jella, Reitzer, Pauline et al. (2021). Ich sehe was, was du auch siehst. Über die Möglichkeiten von Augmented und Virtual Reality für die digitale Beteiligung von Bürger:innen in der Bau- und Stadtplanung. In: Knoll, Matthias (Hg.), *HMD 58*, S. 1180–1195. https://doi.org/10.1365/s40702-021-00772-6.

Gabriel, Oscar (2020). *Partizipation im Wandel*, bpb Bundesnzentrale für politische Bildung, Berlin. https://www.bpb.de/themen/deutsche-einheit/lange-wege-der-deutschen-einheit/308683/partizipation-im-wandel/ [Zugriff 01.03.2025].

Hassenzahl, Marc, Burmester, Michael & Koller, Franz (2008). Der User Experience (UX) auf der Spur: Zum Einsatz von www.attrakdiff.de. In: Brau, Henning et. al. (Hg.), *Usability Professionals 2008: Berichtband des sechsten Workshops des German Chapters der Usability Professionals Asscocation e. V.*, Fraunhofer IRB Verlag, Stuttgart, S. 78–82.

Initiative D21 (Hg.) (2024). *D21-Digital-Index 2023/2024. Jährliches Lagebild zur Digitalen Gesellschaft, Berlin.* https://initiatived21.de/publikationen/d21-digital-index/2023-24. [Zugriff 01.03.2025].

Kuckartz, Udo & Rädiker, Stefan (2022). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung*, Beltz Verlag, Weinheim und Basel.

Le Blanc, David (2020). E-participation: a quick overview of recent qualitative trends. In: *DESA Working Paper*, No. 163, JANUARY 2020. https://doi.org/10.18356/0f898163-en.

Saßmannshausen, Sheree, Radtke, Jörg, Bohn, Nino et al. (2021). Citizen-Centered Design in Urban Planning: How Augmented Reality can be used in Citizen Participation Processes. In: *Proceedings of the 2021 ACM Designing Interactive Systems Conference (DIS '21)*, Association for Computing Machinery, New York, NY, USA, S. 250–265. https://doi.org/10.1145/3461778.3462130.

Senatsverwaltung für Stadtentwicklung und Umwelt Berlin (Hg.) (2014). *Handbuch zur Partizipation*, L.I.S.T. Stadtentwicklungsgesellschaft mbH (Bearb.), Berlin, Kulturbuch-Verlag, Berlin. https://www.berlin.de/raum-fuer-beteiligung/links-downloads/handbuch_partizipation.pdf [Zugriff 01.03.2025].

Den Klimawandel am eigenen Körper erfahren – Klima-ACT!

Maik Helfrich, Barbro Scholz, Sabine Gärtner und Jasmin Röseler

Zusammenfassung

Das Forschungsprojekt »Klima-ACT!« geht der Frage nach, wie eine immersive Erfahrung geschaffen werden kann, um emotionale Betroffenheit und folglich ein positiveres Klimahandeln bei Nutzenden hervorzurufen. Klima-ACT! basiert auf körper- und materialbasierten Ansätzen aus den Bereichen Human-Computer Interaction (HCI) und Game Design, die auf spielerische Motivation und Verbundenheit abzielen. In einem designgetriebenen Prozess wurden explorative Studien und Workshops durchgeführt, um eine Kombination aus Textilien und Virtual Reality (VR) zu erforschen. Textilien sind das Schlüsselmaterial, um eine Brücke zwischen der physischen und der virtuellen Realität zu schlagen und um eine körpernahe Erfahrung zu vermitteln. Das Ergebnis ist ein Prototyp aus interaktiven, innovativen Textilien und einer spielerischen virtuellen Erfahrung, die durch ein Gesamtdesignkonzept miteinander verbunden sind. Vorläufige Ergebnisse bestätigen das Ziel, eine Erfahrung zu schaffen, welche die Nutzenden emotional berührt, Gemeinschaft schafft und dadurch zu positiverem Handeln führen kann.

1. Einleitung

Das bloße Wissen über den Klimawandel und seine Auswirkungen reicht oft nicht aus, um Menschen zu klimafreundlichem Verhalten zu motivieren. Den meisten Menschen ist die Tatsache eines menschengemachten Klimawandels bewusst, dennoch erscheinen ihnen die Folgen zu weit entfernt (Weber, 2006: 116). Vielmehr sind es persönliche emotionale Erlebnisse, die eine intrinsische Motivation zum Handeln schaffen (Grothmann, 2018: 15–18).

Das Forschungsprojekt Klima-ACT! vereint multidisziplinäre Designbereiche – Game Design und Textildesign – in einem Designforschungsprozess, in dem ausgehend von einem gesamtgestalterischen Ansatz (Freyling, 1993: 1–5) explo-

rativ und prototypisch geforscht wird. Beide Bereiche beschäftigen sich mit der Frage, wie man Emotionen durch Materialität und Handlung beeinflussen und durch Semantik und Atmosphäre so gezielt beim Lernen helfen kann (z. B. Döring, 2016: 81–110; Vallgårda, 2014: 577–592; MANND, 2018; Ahn; 2014: 235–245). Eng verknüpft arbeiten sie daran, eine positive Lernumgebung zu schaffen, die den Klimawandel, dessen drastische Folgen und Handlungsoptionen vermittelt.

Das Ziel von Klima-ACT! war es, eine multisensorische immersive Erfahrung zu schaffen, die thematisch stark mit dem Alltag der Nutzenden verbunden ist, den Menschen langfristig die Probleme des Klimawandels näherbringt und diese sowohl haptisch als auch emotional spürbar macht. Basierend auf körper- und materialbezogenen Forschungsansätzen aus der HCI (Hobye et al., 2011: 31–48; Mueller et al., 2020: 207–220; Buso et al., 2022: 8) und aus dem Game Design, die auf Spielfreude, Motivation und Zusammengehörigkeit abzielen (Ahn et al., 2014: 235–245; Bekker et al., 2015: 469–474), wurde untersucht, wie physische und virtuelle Welten durch die Gestaltung der physischen und virtuellen Materialien spielerisch miteinander verbunden werden können. Die Nutzung von innovativen Textilien und die daraus folgende Entwicklung von Multiplayer-Interaktionen, die das Gefühl von Gemeinschaft hervorrufen, waren dabei grundlegend.

Das folgende Kapitel (Verwandte Forschung) skizziert zunächst den multidisziplinären wissenschaftlichen Hintergrund als Grundlage für den Forschungsansatz bei Klima-ACT!. Nach einer kurzen Beschreibung vorausgegangener Studien (Kapitel 3) wird auf den iterativen Designforschungs- und Entwicklungsprozess (Kapitel 4) sowie die prototypische Umsetzung eingegangen (Kapitel 5). In Kapitel 6 wird der finale Showcase vorgestellt, gefolgt von einer Darlegung der Studienergebnisse und letztlich einer Diskussion über eben jene.

2. Verwandte Forschung

Im Folgenden werden die multidisziplinär beteiligten Designbereiche kontextualisiert: körper- und materialbezogene Konzepte aus der HCI und Konzepte aus dem Game Design, die auf Spielfreude, Motivation und Zusammengehörigkeit abzielen.

2.1 Immersion und Präsenz

VR ermöglicht es den Nutzer:innen, in neue Erfahrungen versetzt zu werden und vollständig in diese einzutauchen. »Immersion« und »Präsenz« sind feststehende Begriffe, die für die Beschreibung solcher Erfahrungen herangezogen werden.

Immersion bezieht sich auf die objektive Leistung des Systems, eine virtuelle Welt zu erschaffen: Qualität der Grafik, des Sounds, des haptischen Feedbacks, aber auch der Synchronisation zwischen verschiedenen Sinnen, der Grad an Interaktivität und das Verhalten von virtuellen Dingen (Jareld, 2015: 45–46). Präsenz, so Jared weiter, beschreibt die subjektive psychologische Wahrnehmung der Nutzer: innen in der virtuellen Umgebung. So wird eine virtuelle Welt »besucht« und nicht lediglich wahrgenommen. Immersion kann Präsenz hervorrufen; somit ist die Präsenz vom Grad der Immersion limitiert.

Beide Konzepte sind bei der Erfahrung virtueller Systeme miteinander verzahnt und korrelieren sogar miteinander. Mit Blick auf das übergeordnete Ziel, zu berühren, zu motivieren, das Bewusstsein zu schärfen und Emotionen auszulösen, erscheint es naheliegend, bei Klima-ACT! ein Höchstmaß an Immersion zu schaffen; gleichzeitig sollten Möglichkeiten zur Interaktion mit der Umgebung gegeben sein, um damit das Gefühl der Präsenz bestmöglich zu unterstützen. In Klima-ACT! wird diese Definition von Präsenz und Immersion mit dem Fokus auf den Übergang von der physischen in die virtuelle Welt und umgekehrt erweitert.

2.2 Materialität, Textil und Körper

Textilien und (immersive) Technologien werden kombiniert, um computergestützte Technologie tragbar zu machen oder um sie in einer interaktiven Umgebung einzusetzen. Der »material turn« in der HCI (Robles et al., 2010: 137–144; Wiberg, 2018) beschreibt, wie nicht nur Geräte, sondern auch interaktive Umgebungen geschaffen werden können. Während interaktive Materialien neue Möglichkeiten bieten, wie Technologie aussehen, klingen und sich anfühlen könnte, wird diese dadurch auch stärker in alltägliche Dinge eingebettet.

In vielen Forschungsprojekten fungieren die Textilien jedoch nur als Träger von Technologie und sind nicht Teil des Konzepts der Nutzererfahrung (Hatwell et al., 2017). Gowrishankar u. a. erörtern in ihrem Artikel, dass Textilien als Interaktionsmaterialien mehr und andere materialspezifische Möglichkeiten bieten, wie z. B. das »Ausstopfen, Dehnen, Aufrollen« (Gowrishankar et al., 2017: 233–257). Redström beschreibt, wie Textilien, ein wohlbekanntes Alltagsmaterial, die Art und Weise verändern können, wie Menschen mit Technologie umgehen (Redström, 2010: 9–10): Die Sinneserfahrungen mit Textilien sind körpernah und alltäglich und haben somit das Potenzial, die Nutzenden in ihrer ästhetischen Erfahrung zu beeinflussen. Sie ist niedrigschwellig und kann dadurch Menschen erreichen, die weniger Erfahrung mit Technologie haben.

Das Weben zur Herstellung und Gestaltung von Textilien basiert auf einer der ältesten Techniken der Menschheit (Albers, 2017: 1). Die Werkzeuge und Geräte

haben sich im Laufe der Jahrtausende verändert, nicht aber das Herstellungsprinzip. Meyer zur Capellen beschreibt Gewebe als »textiles Flächengebilde«, das aus mindestens zwei sich rechtwinklig kreuzenden Fadensystemen besteht und auf dem Webstuhl hergestellt wird (Meyer zur Capellen, 2001: 148–149). So kann die ästhetische Erfahrung auf funktionaler und gestalterischer Ebene (Buso et al., 2022; Meiklejohn et al., 2024: 35–49) – zusätzlich in Kombination mit Technologie – vielschichtig gesteigert werden. Die Kombination aus digitaler Steuerung und manuellem Weben am Jacquardmusterwebstuhl »TC2«, wie sie bei Klima-ACT! genutzt wird, ermöglicht große Gestaltungsfreiheit.

Die Flexibilität, Transformation, Performanz und Verspieltheit (engl. »Playfulness«) ist weichem Material inhärent: Unter anderem erforschen Ugur Yavuz et al. den spielerischen Umgang mit Textilien und die Fähigkeit weicher Materialien, neue Interaktionsmöglichkeiten zu bieten (Yavuz et al., 2021: 1–4). In einem von Scholz et al. durchgeführten Workshop wurden Teilnehmende dazu angeleitet, Einzel- und Multi-User-Interaktionen mit Textilien frei zu explorieren, um dies für zukünftige Interaktionsdesignprozesse nutzen zu können (Scholz et al., 2022). Wichtig war dabei auch die Vermittlung des »Play«, angelehnt an das »lenses of play tool«, erarbeitet von Bekker et al., welches die Aktivitäten des Workshops bestimmten (Bekker et al., 2015: 469–474).

In der material- und erlebnisbezogenen HCI-Designforschung wurde in den letzten Jahren u. a. untersucht, wie mit der Berührbarkeit von physischem und der Ungreifbarkeit von digitalem oder computergestütztem Material gestaltet werden kann und was sie für Design und Nutzende bedeutet. Diese Forschung untersucht meist in der physischen Welt und ist häufig nicht mit VR-Anwendungen verbunden. Klima-ACT! bezieht sich deshalb zudem auf die »entanglement«-Theorie (Barad, 2018: 223–239), um die involvierten Materialien (Textilien, Interaktion, VR) in ihrer Transformation zu sehen und darüber hinaus in der Lage zu sein, den miteinander verschränkten technischen, gestalterischen und erfahrungsbezogenen Ebenen zu begegnen.

Konzeptuell beziehen sich Forschende oft auf phänomenologische Ansätze, so auch Vallgårda et al., die diese Ansätze mit materialwissenschaftlichem Vokabular kombinieren, wenn sie interaktive Verbundwerkstoffe (»computational composites«) beschreiben (Vallgårda, 2014: 577–592; Vallgårda & Sokoler 2009: 4147–4152). Sie stellen dabei die Erlebnisebene interaktiver Umgebungen in den Vordergrund und beschreiben, wie die physischen und nichtphysischen Materialien beim Erleben miteinander verwoben sind. Frauenberger beschreibt die Ära des »entanglement« (dt: Verschränkung) in der HCI (Frauenberger, 2020: 1–27), in der sich die Forschung nicht mehr nur allein auf die technischen Vorteile von Materialien und deren Wirkung konzentriert, sondern die Erfahrung interaktiver Materialien im täglichen Leben miteinander verschränkt. Er schlägt

posthumanistische Konzepte als theoretische Ausrichtung vor, um die Komplexität der Verflechtung von Menschen und Technologie zu erforschen. Klima-ACT! knüpft an diesen Ansatz an, um die Vielschichtigkeit der Erfahrung von physischen und nicht physischen Materialien zu untersuchen.

2.3 Multimodal erweiterte VR

Dieses Projekt basiert auf der Forschung im Bereich der Haptik und greifbaren Interaktion. Ein hohes Maß an Propriozeption, d. h. das Gefühl für die Position, Haltung und Bewegung des eigenen Körpers, führt zu einem hohen Maß an Präsenz. Dies kann erreicht werden, wenn die virtuellen Interaktionen so gestaltet sind, dass die zur Durchführung der Interaktion notwendige physische Körperbewegung dem realen Äquivalent entspricht (Jerald, 2015: 270). Jerald kommt auch zu dem Schluss, dass physisches Input-Feedback durch das Berühren passiver Objekte »die Präsenz erhöht, die kognitive Abbildung der Umgebung verbessert und die Trainingsleistung steigert« (Jerald, 2015: 37). Eingabegeräte, bei denen die Hände frei bleiben, z. B. Hand-Tracking mittels Sensoren, verbessern den Propriozeptionssinn, indem sie virtuelle Abbilder der eigenen Hände in VR mit der exakten Position und Pose der realen Hände (aus der Ich-Perspektive) bereitstellen. Dennoch fehlt ihnen in der Regel der Tastsinn, wenn Nutzende virtuelle Objekte ohne physische Präsenz greifen. Die Implementierung von getrackten passiven Objekten – also physische Objekte, deren virtuelles Pendant sich in Lage und Position in VR angleicht – kann diese Nachteile umgehen. Klima-ACT! bezieht sich im Hinblick auf Textilien zudem auf den Begriff des »computional composite« (Vallgårda & Sokoler, 2009: 4147–4152): Hierbei wird greifbares, programmiertes Material durch eine virtuelle Darstellung in VR ergänzt werden.

In der Forschung mit VR-Anwendungen, die haptisches Feedback verwenden, gibt es hauptsächlich zwei Schwerpunkte: So arbeiten Wissenschaftler:innen einerseits mit der Simulation von haptischem Feedback in VR, andererseits mit der somatischen Erfahrung mit Konnotationen aus dem realen Leben. Erstere arbeiten gezielt an technischen Möglichkeiten zur Simulation von Oberflächen oder Materialien. Zum Beispiel wird in Strohmeiers Forschung mit in Sandalen eingebetteten Vibrationsmotoren in Kombination mit VR ein Waldspaziergang simuliert (Strohmeier et al., 2020: 579–593). DieEntwickler:innen des Designstudios MANND machten in ihrer multimodalen VR-Erfahrung die Fast Fashion und ihre Auswirkungen auf Umwelt und Arbeit erlebbar (MANND, 2018): Sie entwickelten einen Parkour aus haptischen Komponenten, die mit den visuellen und auditiven Empfindungen der VR korrespondieren. Unter anderem laufen die Nutzenden durch eine virtuelle Szenerie, in der Menschen im Fluss stehend Textilien färben,

während sie durch ein reales Wasserbecken laufen. Ein anderes Beispiel zeigt den heißen und engen Raum einer Näherei in VR, während in der physischen Welt ein Ventilator heiße Luft in Richtung der Person in der Anwendung bläst.

Folgende Arbeit ist zwar nicht mit (immersiver) Technologie oder VR verbunden, arbeitet jedoch intensiv mit der Gestaltung von somatischen, also auf den Körper bezogenen Erfahrungen: Kargas Ausstellung »Wellbeings« stellt textile Objekte bereit, welche Menschen Gefühle der Sicherheit in klimawandelbedingten Angstmomenten ermöglichen sollen (Karga, 2023). Die Objekte sind von unterschiedlicher Größe; kleinere Objekte können vor dem Gebrauch aufgewärmt werden und größere eine Art hybrides Kuscheltier darstellen, welches sich die Nutzenden um den Körper legen und streicheln können, um sich zu beruhigen. Karga nutzt auch das Prinzip von Gewichtsdecken, um den Menschen dabei zu helfen, sich zu entspannen und ihren Körper besser wahrzunehmen (Eron et al., 2020). Die Gestalt der Objekte ist ein weiterer wichtiger Aspekt: Sie sind verbunden durch eine gemeinsame Designsprache. Die einheitliche Farbgebung und Materialien mit weichen Oberflächen schaffen einen klar definierten Raum und somit eine gemeinsame Basis, um sich auf die physische Interaktion mit den textilen Objekten einzulassen. Die Gestalt der Objekte ist spielerisch, der Ausstellungsraum lädt zur freien Erkundung ein.

2.4 Gemeinsames spielerisches Erlebnis

Huizinga bezeichnet Spiel als »freie, nicht befohlene Handlung« (Huizinga, 1981: 16). Spielende treten dabei in eine Spielwelt ein, die eigenen Regeln und Abläufen unterliegt und sich somit von der realen, ernsten Welt unterscheidet. Ernst wiederum ist nicht gleichbedeutend mit dem Gegenteil von Spiel: Dem Spiel können ernsthafte Züge zugeschrieben werden (Huizinga, 1981: 14). Regeln und Abläufe zeigen sich in zeitgemäßen (Computer-)Spielen meist als Win/Loose-Bedingung, die es zu erfüllen gilt. Durch die bewusste Verwendung des Ausdrucks »spielerisch« verfolgt Klima-ACT!, abgeleitet von Huizinga, den freien Handlungsraum im Spiel, der durch Narrativ, Interaktionsmöglichkeiten und Handlungsablauf definiert wird.

McGonical beschreibt das Potenzial von Spieleszenarien und -mechanismen, die sich um die Rettung der Welt und eventuelle epische Siegesmomente herum entwickeln. Während Spiele die Nutzenden dazu motivieren, eine glaubhafte virtuelle Welt zu verändern und zu beeinflussen (Handlungsspielraum und sinnvolle Interaktionen), stellt sie auch die Frage, wie dieses Potenzial in die reale Welt übertragen werden kann (McGonical, 2011). Sie plädiert für einen spielerischen Bezug realweltlicher Handlungen, um durch eine intensive Immersion,

Zugehörigkeit und Verbundenheit (wie z. B. in Onlinevideospielen) positiv zu einer Handlungsmotivation beizutragen und diesen Beitrag gesamtgesellschaftlich sinnvoll nutzen zu können. Klima-ACT! versucht, diese Verbundenheit u. a. dadurch herzustellen, dass sich die Handlung in Hamburg abspielt. So werden lokale Klimaauswirkungen dargestellt, aber auch die Möglichkeit gegeben, ihnen entgegenzuwirken. Ansässige Menschen sollen so einen emotionalen Bezug zur virtuellen Welt aufbauen können.

Ahn et al. zeigen, dass menschliches Verhalten durch ein interaktives Medium deutlicher positiv beeinflusst werden kann als durch ein passives Medium wie Text oder Video (Ahn et al., 2014: 235–245). Auf der Grundlage von Forschungen im Bereich der Umweltpsychologie folgen sie der Vorgabe, dass »Wissenschaftler:innen Wege finden sollten, um viszerale Reaktionen auf die Gefahr von Umweltproblemen hervorzurufen, vielleicht durch Simulationen konkreter zukünftiger Konsequenzen« (Weber, 2006: 116). In Experimenten konnten sie zeigen, dass immersive virtuelle Umgebungen, in denen die Teilnehmenden aufgefordert wurden, mit einem Controller mit haptischem Feedback aktiv Waldbäume zu fällen, das Umweltverhalten in der realen Welt wirksamer zu beeinflussen scheinen als das Lesen von Beschreibungen oder das Betrachten von Videos der gleichen Aktion.

Multi-User-Interaktionen sind ein weiterer motivierender Faktor: Vrellis et al. haben gezeigt, dass kollaborative virtuelle Lernumgebungen die »soziale Reichhaltigkeit« des Präsenzgefühls fördern, und folgern daraus, dass Nutzende das Eintreten virtueller Aktivitäten dieser Umgebungen dadurch in der Realität für wahrscheinlicher halten (Vrellis et al., 2010: 6–7). Whittle et al. schlagen kooperative Spielabläufe vor, um Nutzende durch gemeinsames Lösen von Herausforderungen darin zu bestärken, diese anzugehen. Unterschiedliche Sichtweisen mehrerer Spielenden sind bei der Lösung von Problemen ebenfalls von Vorteil, während Hinweise einer spielinternen Moderation beruhigend wirken können, wenn sich vorangegangene Lösungsversuche als erfolglos entpuppen (Whittle et al., 2020: 69–71).

Poulsen beschreibt, wie »Play« motivierend und integrativ sein kann, wenn ein breites Spektrum von Nutzenden (in ihrem Fall verschiedene Bewohner:innen einer Stadt) angesprochen wird (Poulsen, 2020: 147–162). Klima-ACT! setzt das Spielerische mit dem Ziel ein, eine vielfältige Gruppe Nutzender emotional zu involvieren, indem spielerisch-freudige Erfahrungen in gelockerter Stimmung (gemeinsame Handlungen, freie Umsetzung auch lustiger Interaktionen mit dem Textil) mit dem ernsten Thema des Klimawandels verbunden werden.

3. Vorausgegangene User Study 1

In der ersten User Study ging es darum, konzeptuell die technischen Komponenten zusammenzubringen. Narrativ und Design waren bereits ausreichend ausgestaltet, um ein zu beforschendes vorläufiges Klima-ACT!-Erlebnis bieten zu können.

3.1 Beschreibung

Das Setup legte den Fokus auf Gamification und war für drei Nutzende konzipiert: eine Person in VR, für die ein technisch funktionierendes Wearable sowie eine haptische Bodenfläche bereitstand, und zwei weitere Personen, die via Desktop-Computer ohne VR-Headset in dieselbe virtuelle Welt eintraten (Abbildung 4). Der virtuelle Ort befand sich in einer Kleingartenanlage im Herzen der Stadt Hamburg und trennte die Nutzenden räumlich voneinander (Abbildung 1), sodass die gestellten Aufgaben nur gemeinsam gelöst werden konnten (Abbildung 2). So mussten aus verschiedenen Medien wie Bild, Ton und Video Informationen zu klimarelevanten Bereichen wie Ernährung, Mobilität und Energie gefunden und für das Lösen von Aufgaben benutzt werden (Abbildung 3). Durch erfolgreiches Lösen wurde das initial extreme Wetter weniger extrem und beispielsweise ein virtuelles Atomkraftwerk verwandelte sich in eine Windkraftanlage.

3.2 Auswertung und Erkenntnisse aus User Study 1

Die entstandenen gemeinsamen Interaktionen, die Kollaboration und damit Kommunikation zwischen den Spielenden konnte positiv gewertet werden. Durch das Multiplayer-Konzept wurden drei sich zunächst unbekannte Personen zu einem Team. Sie begannen nach kürzester Zeit, miteinander zu kommunizieren und zu agieren. In den im Anschluss durchgeführten Gruppeninterviews wurde deutlich, dass ihre Zusammengehörigkeit auch nach der Klima-ACT!-Anwendung erhalten blieb.

Ein weiteres Ergebnis war jedoch auch, dass sich diese Art der Erfahrung zu stark auf die Interaktion im Sinne eines zu gewinnenden Spiels konzentrierte. Die Aufmerksamkeit der Nutzenden war in der Effizienz ihrer Handlungen gebündelt: Sie nahmen weder die sichtbaren Veränderungen in der VR-Umgebung noch die haptischen Komponenten auf dem Boden oder im Wearable wahr. Das multisensorische Erlebnis wurde überlagert vom Ehrgeiz, Aufgaben zu erfüllen.

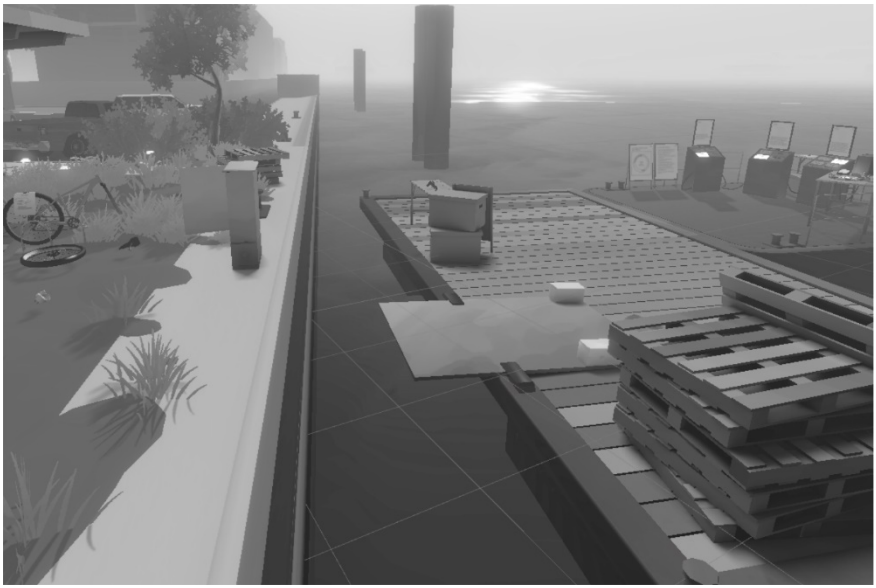


Abbildung 1: Durch Wasser getrennte Spielbereiche
Quelle: Eigene Darstellung

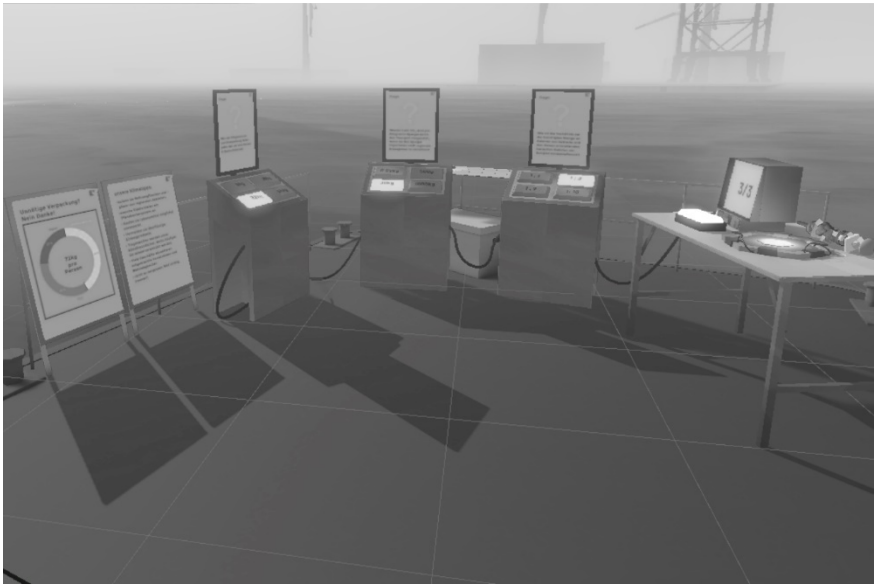


Abbildung 2: Rätselstationen und Informationsposter
Quelle: Eigene Darstellung



Abbildung 3: Sprache und Emojis werden zur Kommunikation genutzt.
Quelle: Eigene Darstellung



Abbildung 4: Studienansicht mit Desktop- und VR-User
Quelle: Eigene Darstellung

Die virtuelle Erfahrung samt Informationen zum Klimawandel sowie zeitgleich die gemeinsamen Interaktionen überforderte die Nutzenden. Die Informationen, wie z. B. Statistiken aus der Klimafolgenforschung oder die tatsächlichen Wetterwechsel, waren für die spielerische Anwendung zu stark vereinfacht und wurden dadurch beliebig. Der Wetterwechsel wurde als »schönes« Wetter (sonnig) und »schlechtes« Wetter (regnerisch und stürmisch) wahrgenommen und nicht kritisch in Bezug auf die durch den menschengemachten Klimawandel entstandenen Extreme verstanden. Insgesamt waren die Textilien und ihre vielschichtigen inhärenten Eigenschaften in den Hintergrund gerückt und auf die Option zur Tragbarkeit von Technologie reduziert. Nach intensiver Reflexion über die Schwerpunkte im Projekt wurde beschlossen, die emotionale Betroffenheit der Nutzenden in den Mittelpunkt zu stellen und den Gamification-Aspekt innerhalb der Interaktionen subtiler einzusetzen. Die klimawissenschaftlichen Informationen wurden im Rahmen einer räumlichen Installation außerhalb des multimodalen VR-Erlebnisses im Realen platziert, um den Input für die Nutzenden kognitiv zu verteilen und mehr Raum für die ästhetische Erfahrung zu bieten. Die Textilien rückten dadurch mehr in den Fokus und konnten über die körperliche Erfahrung die virtuelle und die reale Welt verbinden.

Dies führte zu weiteren Fragestellungen, die im nachfolgenden Prozess untersucht wurden: Wie können Informationen und Erfahrungen miteinander verbunden werden? Wie können die verschiedenen Ebenen des Designs genutzt werden, um die Menschen emotional einzubinden? Und wie kann der starke Moment der Zusammengehörigkeit, der durch die Interaktionen zwischen mehreren Nutzenden entsteht, noch besser genutzt werden?

4. Designgeleiteter Prozess

In einem iterativen Forschungsprozess setzte das Projekt darauf, durch Designforschung und Materialbezug das Gestaltungskonzept und die Textilien in den Vordergrund zu stellen und die Entscheidungen daran zu orientieren. In mehreren explorativen Workshops und User Studies wurden zyklisch verschiedene Designelemente sowohl auf ihre beabsichtigte emotionale und/oder haptische Wirkung auf die Nutzenden als auch auf ihre technische Machbarkeit getestet und weiterentwickelt.

Der Begriff »explorativ« wird im Rahmen des Projekts für einen »research through design«-Ansatz verwendet, der durch gestalterische Iterationen (schnelle Realisierungen von Einzelkomponenten, die wiederum als Forschungsgegenstände dienen) das Design selbst in den Mittelpunkt stellt. Klima-ACT! nutzt diese explorative Herangehensweise vor allem in Bezug auf die sensorischen

multimodalen (also ästhetischen) Erfahrungen von textilen Objekten und Bekleidung sowie auditiver und visueller Atmosphäre in VR. Die Untersuchung des Zusammenspiels dieser Sinneseindrücke, immer in Kombination mit spielerischen Interaktionen und mit Blick auf die Auswirkung des gemeinsamen Erlebens durch das Multi-User-Setting, basiert auf dem Verständnis, dass eine ästhetische Erfahrung stets ganzheitlich zu denken ist. Das bedeutet, die Wechselbeziehung von Körper und Geist, Empfindung und Emotion wird aktiv mit einbezogen (Höök et al., 2021: 1–26; Höök, 2018: 1–29; Shusterman, 1999: 299–313).

In vier explorativen internen Workshops arbeitete das Team, angelehnt u. a. an die Methoden der »Embodied Design Ideation« (Wilde, 2017: 5158–5170), am und mit dem Körper. Den unterschiedlichen Entwicklungszyklen und -strukturen der physischen und der virtuellen Umgebung wurde mit einer gemeinsamen Gestaltungsgrundlage und engem Austausch über die Art Direction in den jeweiligen Entwicklungs- und Gestaltungsschritten begegnet, um somit ihrer Verschränkung gerecht zu werden und dem finalen Prototyp eine Gesamtästhetik zu verleihen.

Die explorativen Workshops in diesem Forschungsabschnitt hatten den Vorteil, die multidisziplinären Perspektiven des Teams in alle Umsetzungsbereiche einbeziehen zu können. Dadurch konnten Workshopergebnisse direkt in Bezug auf Machbarkeit reflektiert werden. Die Resultate sind Rahmenbedingungen für die Realisierung des Prototyps.

4.1 Workshop 1 – Textile Interaktion

Workshop 1 konzentrierte sich auf die spielerische Interaktion mit den Textilien: Eigenschaften und Merkmale der Textilien wurden durch improvisierte Einzel- und Multi-User-Interaktionen spielerisch erforscht. Wichtig war dabei, das gesamte Projektteam für die ästhetische Erfahrung mit Textilien zu sensibilisieren. Insbesondere sollten die interdisziplinär Forschenden die spielerischen Möglichkeiten, die den weichen Materialien inhärent sind, kennenlernen und in eigener Erfahrung erproben, wofür eine Auswahl von gewebten Materialproben zusammengestellt wurde. Als erste Annäherung wurden mit Textilien verschiedene Interaktionen ausprobiert, um dabei zu reflektieren, was es bedeutet, wenn z. B. ein interaktives Material am Körper getragen wird (Abbildung 5). Nach dieser ersten Annäherung wurde zielgerichtet mit der Frage gearbeitet, wie Textilien konkret Emotionen erzeugen und transportieren können und in welcher Form Nutzende in Single- oder Multi-User-Situationen mit Textilien interagieren können, sowohl real als auch in VR.



Abbildung 5: Interaktive Versuche aus der explorativen Materialbegegnung
Quelle: Eigene Fotografie

Die eigene Körpererfahrung mit dem Material Textil setzte eine gemeinsame Grundlage für das ganze Entwicklungsteam. Es wurde ein Repertoire an Interaktionen für einzelne und mehrere Personen entwickelt, um für den weiteren Prozess modular arbeiten zu können. Der Workshop hat die Verschränkung von multisensorischem und emotionalem Erleben (Fühlen) deutlich gemacht: besipielsweise förderte humorvolle gemeinsame Interaktion das Teambuilding. Komplexere textile Strukturen und Faserkombinationen stimulierten den Tastsinn und es wurden Gefühlsstimmungen und Erinnerungen evoziert.

Auf den Einsatz von konventionellen Game Controllern soll verzichtet werden, da sie ein direktes Anfassen von realen Gegenständen unmöglich machen. Stattdessen interagieren die Nutzenden über textile Objekte gemeinsam mit der virtuellen Welt, die mit den eigenen Händen direkt angefasst werden können und gleichermaßen in VR dargestellt werden, um beide Welten multimodal zu verbinden. Somit wurde der Einsatz von Hand-Tracking und dem Tracking der textilen Objekte im Raum notwendig, damit die Lage und Position der virtuellen Pendants mit ihrem realen Ebenbild übereinstimmen.

4.2 Workshop 2 – Narrative Dramaturgie

Workshop 2 diente dazu, eine kohärente Narration für die multimodale Erfahrung zu entwickeln. Insbesondere ging es darum, in welcher Form diese durch textile Objekte unterstützt werden kann. Folgende Fragen wurden dabei in den

Blick genommen: Wie sehr können Nutzende mit der Katastrophe des Klimawandels konfrontiert werden und trotzdem die Klima-ACT!- Erfahrung als positive Erinnerung im Gedächtnis behalten? Wie können Textilien und Objekte gezielt positiv konnotiert eingesetzt werden, um Nutzende zusammenzubringen und um die womöglich negative Atmosphäre und Stimmung der virtuellen Ereignisse positiv auszugleichen? Wie können diese Objekte eine Verbindung von realer und virtueller Welt schaffen? Welche Handlungen und Interaktionen lassen sich mit diesen Objekten durchführen?

Systematisch wurden dramaturgische Abläufe, einzelne Module und ihre Auswirkung auf das Gesamte diskutiert. Hier wurde mit sogenannten Flowcharts gearbeitet, an welchen sich dramaturgische Abläufe übersichtlich darstellen lassen (Abbildung 6).

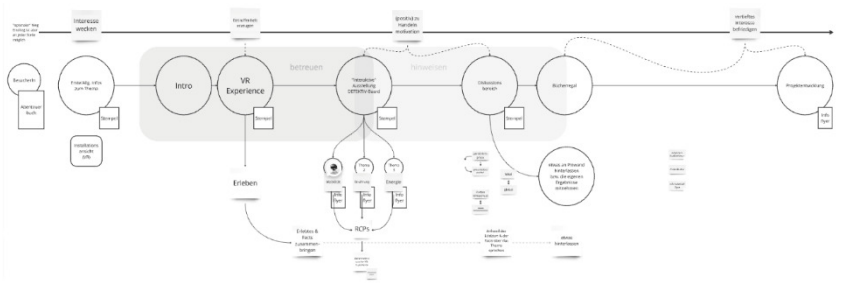


Abbildung 6: Flowchart mit dramaturgischem Ablauf
Quelle: Eigene Darstellung

Zwei Personen sollen sich diesmal gemeinsam in der VR befinden, um das gleiche haptische Feedback zu erhalten und kollaborativ zu interagieren. Dabei sehen sie sich, können miteinander reden und sich physisch und virtuell berühren.

Der virtuelle Schrebergarten in Hamburg soll Ort des Geschehens bleiben. Eine Hollywoodschaukel als Symbol für Komfort soll die Nutzenden einerseits zu Lethargie verleiten. Andererseits sollen Interaktionspunkte (ein Sandkasten, ein Tomatenstrauch, ein Schildkrötengehege, eine Gartenmauer) zum aktiven Handeln bewegen. Ein Verlauf von friedlichen bis extremen Wetterveränderungen formt die allgemeine Atmosphäre. Die VR-Erfahrung soll schließlich in einem dystopischen Szenario der Zerstörung durch extremes Wetter enden.

Um das Teamgefühl zu stärken, sollen die textilen Objekte (Wearable, eine Decke und eine Hollywoodschaukel) durch ihre positiven Konnotationen (wärmend, schützend, tröstlich) zu interaktiven, spielerischen Bezugsobjekten werden. Die

Art der Benutzung (sich mit einer Decke zudecken oder damit etwas vor Wetter schützen) ist bekannt und lässt sich glaubhaft durch eine Narration in die virtuelle Welt einbauen. Diese Narration soll lineare Zeitpunkte aufweisen, innerhalb derer sich die Nutzenden frei für oder gegen Interaktion mit der VR-Welt entscheiden können. Veränderungen, die durch das eigene Handeln an den Interaktionspunkten hervorgerufen werden, sollen den Nutzenden als Erfolgserlebnis multimodal zurückgemeldet werden.

4.3 Workshop 3 – Optisches Tracking

Workshop 3 untersuchte das Tracking von Körper und Textilien. Hierzu mussten Kriterien erarbeitet werden, um die virtuelle Repräsentation der Textilien der realen Haptik und Erscheinung anzugleichen. In Klima-ACT! wird ein optisches VR-System zur Bewegungserkennung eingesetzt, wodurch das Problem entsteht, dass Hände oder Objekte verdeckt und nicht mehr korrekt erkannt werden konnten. Lösungsansätze in beiden relevanten Bereichen, Textilgestaltung und Programmierung, wurden iterativ getestet und evaluiert, wodurch sich beide Bereiche einander annäherten.

Aufgrund der begrenzten Projektlaufzeit wurden Ansätze priorisiert, die als Lösungsweg die technischen Mängel der VR mit einbeziehen: Wann werden Immersion und Präsenz gestört? Wann ist das Erlebnis mit den Textilien so stark, dass technische Fehler verzeihbar sind? So wurden technische Probleme gegen den Nutzen der textilspezifischen Materialität abgewogen (siehe auch Kapitel 5.3).

4.4 Workshop 4 – Gewicht und Wahrnehmung

Workshop 4 diente dazu, die Rolle des Gewichts der Textilien zu bestimmen und gezielt einzusetzen. Getestet wurden flächige und punktuelle Verteilungen verschiedener Gewichtsmengen und Platzierungen sowohl am Wearable, welches direkt am Körper anliegt, als auch an der Decke, die in den Händen gehalten oder über den Körper gelegt werden kann.

Die Beschwerung des Wearables und der Decke soll die Körperpräsenz stärken. Gewichtstaschen an der Decke sollen das Komfortgefühl erhöhen und dadurch die dramaturgische Narration unterstützen (siehe Kapitel 6), zeitgleich von den Nutzenden aber einen erhöhten Kraftaufwand beim Bewegen der Decke erfordern.

5. Prototyping

Basierend auf den Ergebnissen der Workshops wurden alle Bereiche der Umsetzung zusammengeführt. Der hohe gestalterische Anspruch sowohl im Realen als auch im Virtuellen und das Ziel, beides miteinander zu verweben, bestimmten den weiteren Prozess. Das parallele Arbeiten einzelner Bereiche wurde durch stetige Absprache und die gegenseitige Präsentation von Iterationsschritten begleitet.

5.1 Farb- und Formkonzept

Als Grundlage für die gestalterischen Designentscheidungen wurde ein kollaboratives Moodboard erstellt: Gesammelt wurden Gestaltungseinflüsse aus den Bereichen Computergrafik, Modedesign und Textildesign sowie Beispiele für Farbkonzepte und deren Wirkungsweisen, um übereinstimmende Gestaltungsvoraussetzungen zu bestimmen (Abbildung 7). Die zu gestaltende Klima-ACT!-Anwendung wurde in die Phasen 1) Installation in der realen Welt mit textilen Objekten, 2) virtuelle Welt, in der sich die realen Objekte durch ein digitales Ebenbild wiederfinden sowie 3) den Übergang zwischen den beiden Welten aufgeteilt. Bei den Designentscheidungen war ein hohes Maß an ästhetischer Verbindung zwischen diesen Phasen maßgeblich. Das übergeordnete Farbkonzept korrespondiert sowohl in Form eines gewebten Farbverlaufs im Textil als auch in den digitalen Grafiken mit den wechselnden Stimmungen der Dramaturgie.

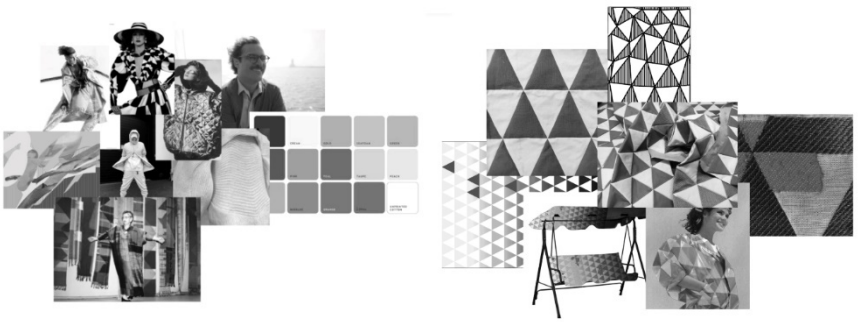


Abbildung 7: Bildsammlungen aus dem kollaborativen Moodboard
Quelle: Eigene Darstellung

5.2 Textiltechnische und -gestalterische Umsetzung

Die textilen Objekte sind komplett auf dem Webstuhl entstanden. Gewebearten wurden so kombiniert, dass auf eine zusätzliche Bearbeitung der Gewebe weitgehend verzichtet werden konnte (Abbildungen 8–10). Die verwendeten Schussfarben und Gewebestrukturen erscheinen hell und farbig und verlaufen zunehmend ins Dunkle und Graue. Der partiell eingesetzte Broché-Schuss (Abbildung 11) in einem leuchtenden Orange wirkt als Signalfarbe und wird auch im Virtuellen als positives Feedback genutzt.

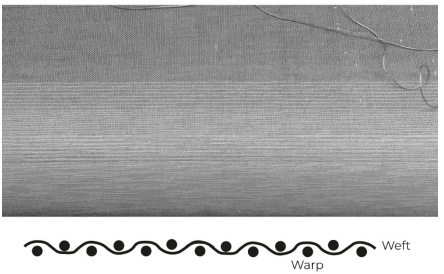


Abbildung 9: Leinwandbindung: Textilgewebe aus mindestens zwei sich rechtwinklig kreuzenden Fäden
Quelle: Eigene Darstellung

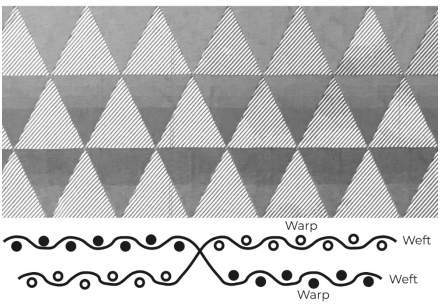


Abbildung 9: Doppelgewebe: Gewebe aus zwei nicht verbundenen Gewebelagen, die in einem Arbeitsgang entstehen. Eine Verbindung kann durch Änderung der Bindung hergestellt werden.
Quelle: Eigene Darstellung

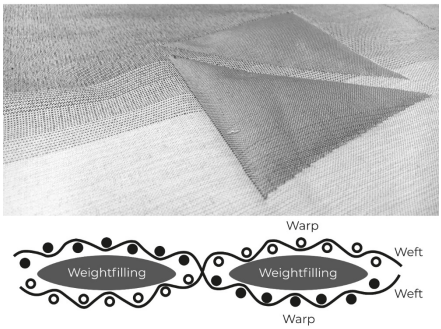


Abbildung 10: Doppelgewebe mit Gewichtseinlagen in den Hohlräumen zwischen den Lagen
Quelle: Eigene Darstellung

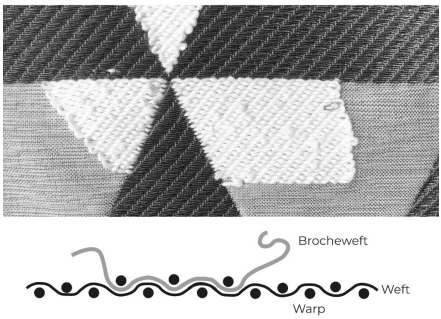


Abbildung 11: Broché: Gewebe mit einem zusätzlichen dekorativen Schussfaden in bestimmten Bereichen
Quelle: Eigene Darstellung

Die Musterstruktur der Dreiecksgitter hat seinen Ursprung in der Darstellung digitaler dreidimensionaler polygonaler Objekte. Die gewebten Icons der realen Decke verändern sich auf dem digitalen Pendant dynamisch und abhängig vom Fortschritt und Zeitpunkt der Erzählung. Die Griffe der Decke und die Halterungen für die optischen Tracker sind gewebetechnisch integriert.

Das Wearables ist mit einer »one size fits all«-Lösung, die unterschiedliche Körpergrößen und somit unterschiedliche Teilnehmende inkludiert, konzipiert. Die Schnittteile wurden mit dem Farbverlauf sowie der Positionierung und Größe der Gewichtstaschen entworfen, für den Webstuhl programmiert und formgewebt (Abbildung 12). Für die Integration der Vibrationsaktuatoren (siehe Kapitel 5.3) und das technische Zubehör diente das Doppelgewebe (Schulter, Rücken und

Vorderseite). Die in Dreiecksform eingewebten Gewichtssäckchen sorgen für den richtigen Sitz des Wearables und drücken die Aktuatoren auf den Körper (Abbildungen 13–14). Sie müssen also nicht, wie bei früheren Iterationen, durch größenverstellbares Design festgeschnürt werden.

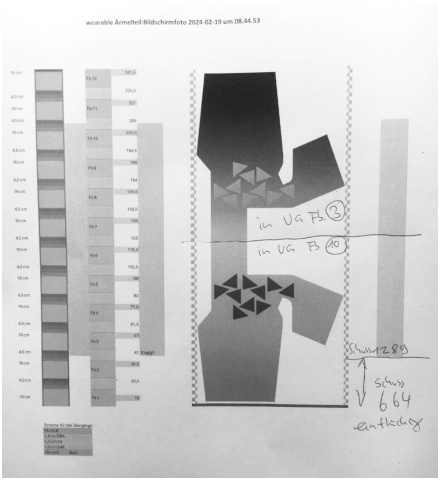


Abbildung 12: Entwurf Schnittteil
Quelle: Eigene Darstellung



Abbildung 13: Dreiecke mit Gewichtseinlage und Aktuator
Quelle: Eigene Fotografie



Abbildung 14: Rückseite des gewebten Schnittteils
Quelle: Eigene Fotografie

5.3 Game Design und Engine-Programmierung

Die Arbeit mit der Unity 3D-Engine bietet alle Möglichkeiten, um eine Umsetzung der in den Workshops definierten Vorgaben der Interaktion, des physikalischen Verhaltens und der grafischen Darstellung in der virtuellen Welt zu ermöglichen.

Die Nutzung einer Decke als textiles Interaktionsobjekt und Ersatz eines Game Controllers folgt den Ergebnissen der Materialbegegnung des Workshops 1 und der Erarbeitung des Narrativs in Workshop 2 (Kapitel 4.1 und 4.2). Ein rechtwinkliges Objekt bietet jedoch auch aus Game-Design-Sicht Vorteile: Wie eine Decke zu benutzen ist, muss in der Anwendung selbst (beispielsweise durch ein »Tutorial«) nicht mehr erklärt werden, vor allem wenn das virtuelle Objekt die gleichen physikalischen Eigenschaften (Bewegung durch Schwerkraft) aufweist wie das reale Ebenbild. Die Größe der Decke sowie das Gewicht legen nahe, dass sie zu zweit benutzt werden soll, wodurch sich wiederum ergibt, wo Griffe positioniert werden müssen, um die angebrachten optischen Tracker nicht mit den Händen zu verdecken (Abbildungen 15–16). Auch lässt sich das Tracking einer rechtwinkligen Decke weniger schwer umsetzen als das einer komplexen geometrischen Form. Letztlich kann durch das eingewebte zusätzliche Gewicht der Decke das Eigengewicht der Tracker kaschiert werden.

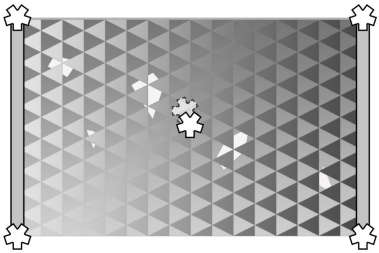


Abbildung 15: Schematischer Aufbau der Decke mit Haltegriffen
Quelle: Eigene Darstellung



Abbildung 16: Vertikale Haltung mit der Decke
Quelle: Eigene Fotografie

Erfolgserlebnisse werden spielerisch-narrativ (z. B. wird eine durch Regen zerstörte Sandburg wieder aufgerichtet oder ein Loch in einer Steinmauer wird geschlossen; siehe Kapitel 6.2), aber auch visuell, auditiv und haptisch zurückgegeben: Die Decke leuchtet in Signalfarben auf (siehe Kapitel 4.2). Die Welt scheint glaubwürdiger, wenn sie auf Handlungen der Nutzenden reagiert.

Die im Wearable verbauten sogenannten »Linear Resonance Actuators« (LRA), also linear schwingende Motoren, simulieren durch ihr Vibrationsverhalten auf dem Körper der Nutzenden auftreffende Regentropfen: Sitzen die Nutzenden unter dem Dach der Hollywoodschaukel oder halten sie sich die Decke über den Kopf, reagieren die LRA entsprechend auf diese Handlung und schalten ab. Hierbei werden die realen Positionen der LRA im Wearable den Positionen der Trigger-Bereiche auf dem Avatar in der Engine angepasst (Abbildung 17).

Die LRA werden im System über eine Audioschnittstelle der Firma Actronika angesprochen. Das einer Lautsprechermembran ähnelnde Verhalten bietet sich

dafür an. Audioaufnahmen von echtem Regen (sog. »field recordings«) konnten wegen zu vieler Nebengeräusche nicht genutzt werden. Deshalb wurden mit einem Synthesizer Sinustöne eines bestimmten Frequenzbereichs mit schneller Ansprech- und Abklingzeit generiert, die zur Nachbearbeitung durch eine Effektkette mit zufällig wechselnden Hall-Einstellungen geleitet wurden (Abbildung 18). Die dadurch entstandenen Audioloops klingen randomisiert genug, sodass keine sich wiederholenden Muster erkennbar sind.

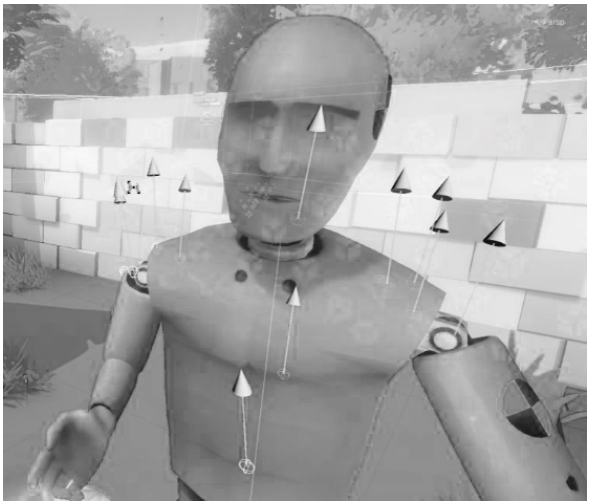


Abbildung 17: In-Engine-Screenshot mit Avatar. Pfeilsymbole repräsentieren die Positionen der Triggerpunkte auf dem Körper.
Quelle: Eigene Darstellung

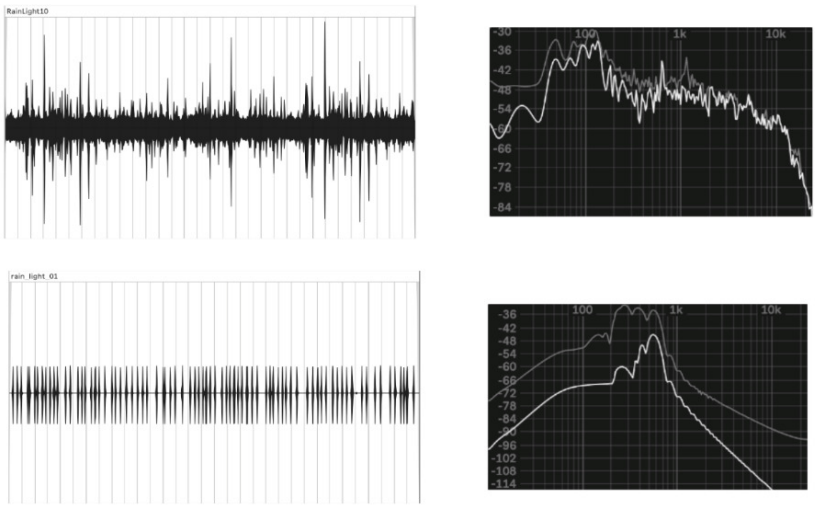


Abbildung 18: Wellenform und Frequenzbereich der beiden Regen-Audiosamples (»field recordings« oben, künstliches Sample unten)
Quelle: Eigene Darstellung

Die Hollywoodschaukel als statisches Objekt wird passgenau als virtuelles Objekt nachgebaut, um so beispielsweise die Sitzhöhe korrekt darzustellen. Hierbei muss lediglich der bewegbare Teil der Schaukel mit einem optischen Tracker erfasst werden. Die Rotation der Schwingung wird dadurch in VR dargestellt. Der feststehende Teil der Schaukel positioniert sich initial beim Starten der Anwendung am gleichen Tracker.

Bei der Platzierung der Interaktionspunkte wurde darauf geachtet, dass stets zwei Nutzende genug Raum haben, um die möglichen Handlungen auszuführen. Der Garten selbst ist visuell durch verschiedene Objekte wie eine Mauer, einen Zaun, ein Gartenhaus oder die Wasserkante begrenzt.

Klima-ACT! verzichtet auf das Konzept der VR-Teleportation (virtuelle Fortbewegung): Die Nutzenden bewegen sich physisch durch den Garten. Das dient auch dazu, die Immersion zu stärken. Das Multi-User-System »Fusion« der Firma Exit Games GmbH sorgt hierbei für die nötige Synchronisation zwischen den beiden Computersystemen der Nutzenden. Hierbei ist einer Person die Rolle des Hosts, allen anderen die Rolle des Clients zugeteilt. Objekte (wie etwa die Decke) werden somit nur auf einem System erfasst, die Positionsdaten dann aber vom Host an alle Clients verteilt. Auch die Sprache und der Zustand der VR-Welt werden über das System gesendet.

Personal vor Ort ist in Form eines Supervisor-Avatars ebenfalls in der virtuellen Welt vertreten. Da die Nutzenden nicht durch eine dritte Person in der

VR-Welt gestört werden sollen, wurde ein System programmiert, wodurch der Supervisor-Avatar erst dann sichtbar wird, wenn er sich den Nutzenden auf eine bestimmte Distanz genähert hat.

6. Showcase und User Study 2

Der finale Showcase ist eine raumgreifende, multisensorische Ausstellung inklusive der VR-Anwendung. Die Besuchenden finden Zugang über Ausstellungsposter sowie die textilen Objekte im Raum (siehe Abbildungen 19–21). Die Poster informieren über klimawissenschaftliche Modellberechnungen und Prognosen zu zukünftigen globalen klimatischen Veränderungen sowie Beispiele dafür, wie sich der Klimawandel zukünftig auf städtische Umgebungen und die Lebensbedingungen (am Beispiel Hamburg) auswirken wird. Die textilen Objekte sind offen zugänglich und können angefasst werden. Zusätzlich gibt es einen Tisch und Stühle mit speziellen Spielkarten samt Nutzungsvorschlägen sowie eine Feedbackwand, an der die Spielenden Eindrücke und Äußerungen hinterlassen können. Es gibt keine Vorgaben, in welcher Reihenfolge die Installationsbestandteile erkundet werden sollen.



Abbildung 19: Ausstellungsansicht mit Postern und Tisch
Quelle: Eigene Fotografie



Abbildung 20: Ausstellungsansicht mit Tisch, Schaukel, Wearable und Decke im Hintergrund
Quelle: Eigene Fotografie



Abbildung 21: Wearable und Schaukel im Hintergrund
Quelle: Eigene Fotografie



Abbildung 22: Spielende auf der Schaukel mit Decke
Quelle: Eigene Fotografie

6.1 Aufbau der VR-Installation

Die VR-Installation ist für zwei Personen konzipiert und nutzt eine Fläche von ca. 6 x 6 m. Dies ist der Bereich, in dem das Trackingsystem die Personen gut erkennen kann.

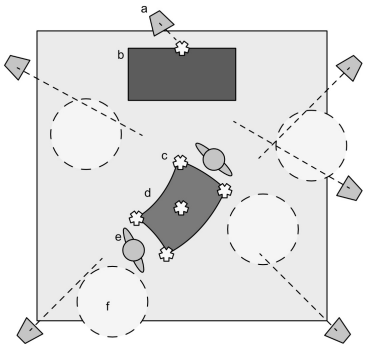


Abbildung 23: Grundriss des virtuellen Schrebergartens mit Interaktionspunkten (b: Hollywoodschaukel, f.1: Sandkasten, f.2: Tomatenpflanze, f.3: Schildkrötengehege, f.4: Gartenmauer)
Quelle: Eigene Darstellung

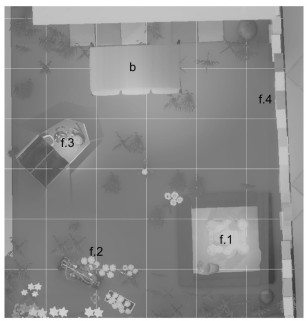


Abbildung 24: Schematischer Aufbau der VR-Area; 6 x 6 m (a: Basestation, b: Hollywoodschaukel, c: Tracker, d: Decke, e: Nutzender, f: Interaktionspunkt in VR)
Quelle: Eigene Darstellung

Angeordnet in der Installationsfläche finden sich eine Hollywoodschaukel, die getrackte Decke, zwei baugleiche Wearables und zwei VR-Headsets. Zwei Mitarbeitende betreuen die Installation und erklären den Nutzenden die Technik sowie den narrativen Kontext. Die Lage des virtuellen Schrebergartens in der HafenCity direkt am Elbufer wird von Häuserfassaden, einem asphaltierten Parkplatz,

einem Kraftwerk am gegenüberliegenden Ufer sowie benachbarten Schrebergärten flankiert.

6.2 Ablauf der VR-Installation

Entscheiden sich zwei Besuchende der Klima-ACT!-Ausstellung dazu, die VR-Installation zu nutzen, werden sie beim Anziehen der Wearables unterstützt und zur Hollywoodschaukel geführt, um dort nebeneinander Platz zu nehmen. Die getrackte Decke wird als Objekt der Interaktion vorgestellt und über die Beine der beiden Nutzenden gelegt. Sie erhalten Informationen über den Ablauf, die technischen Eigenschaften der Textilien und die Sicherheitsvorkehrungen sowie eine Einführung in die Story des Erlebnisses: Die Spielenden werden sich in wenigen Augenblicken in ihrem persönlichen Schrebergarten direkt an der Hamburger Elbwaterkant wiederfinden.

Sobald die Nutzenden ihre Bereitschaft kommunizieren, das virtuelle Erlebnis zu starten, werden die VR-Brillen aufgesetzt. Zunächst sehen die beiden lediglich ihre eigenen Avatare samt Wearable, die Decke über ihren Beinen sowie die Schaukel in einer abgedunkelten Umgebung.

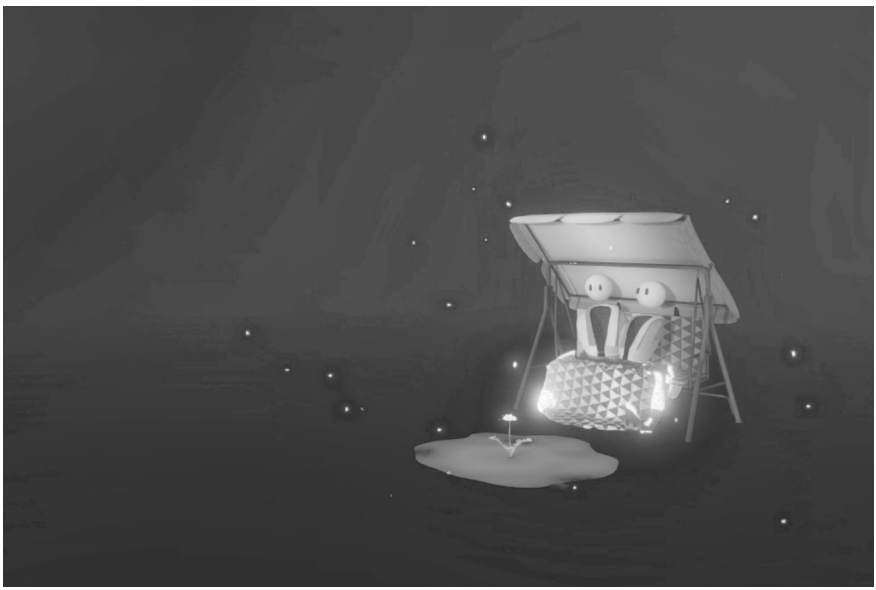


Abbildung 25: Screenshot der Anfangssequenz mit Schaukel, Nutzenden und sich langsam lichtendem Garten
Quelle: Eigene Darstellung

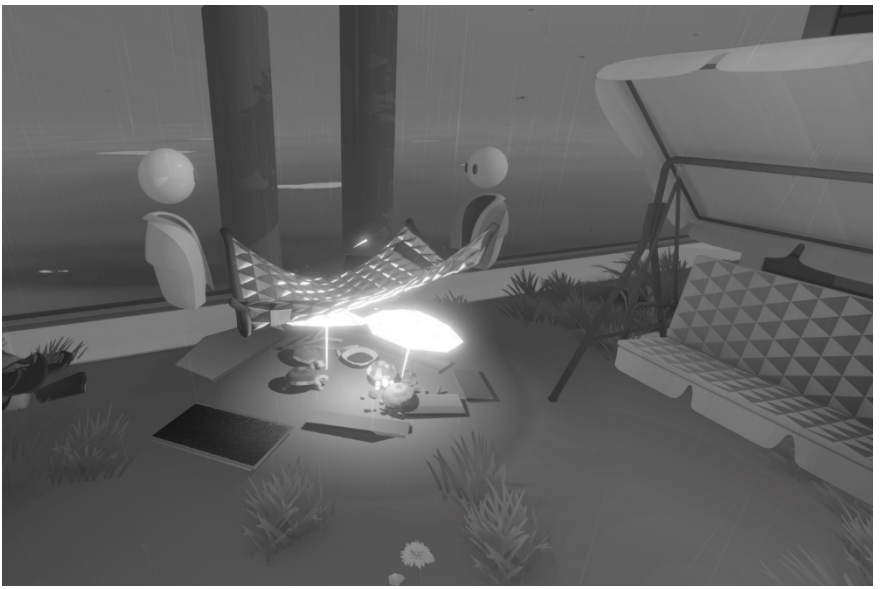


Abbildung 26: Screenshot von funkelnder Decke bei korrekt ausgeführter Handlung
Quelle: Eigene Darstellung

Langsam lichtet sich die dunkle Szenerie und der Schrebergarten erscheint. Die Atmosphäre ist friedlich. Nach einer Weile beginnt es, leicht zu regnen. Ein Ausrufezeichen weist auf die erste mögliche Interaktion hin: Es gilt, eine Sandburg zu schützen, die möglicherweise durch den Regen zerstört wird. Auf der Decke geben die Piktogramme Auskunft darüber, dass die Sandburg mit der darüber gehaltenen Decke geschützt werden kann. Wenn diese Interaktion ausgeführt wird, fängt die Decke an zu funkeln und die Sandburg baut sich magisch von selbst wieder auf.

Aber die Interaktion löst das Problem nur für den Moment. An mehreren anderen Interaktionspunkten verändern sich die virtuellen Piktogramme im Laufe der Zeit entsprechend und zeigen verschiedene Interaktionsmöglichkeiten an. Währenddessen ändert sich das Wetter von angenehm zu unangenehm bis stürmisch. Pflanzen werden vom Wind umgeworfen und das Dach des Schildkrötengeheges wird wegweht. Der starke Regen verursacht auf der versiegelten Fläche außerhalb des Gartens eine Überschwemmung. Die letzte Interaktionsmöglichkeit für die Spielenden besteht darin, ein Loch in der Mauer zu schließen, um das Eindringen von Wasser zu verhindern. Auch diese Interaktion ist unlösbar und die Nutzenden stehen vor dem Dilemma, was sie zuerst retten sollen: sich selbst, die Tiere oder die Pflanzen? Nach einer bestimmten Zeit stürzt die Mauer unweigerlich ein und der Garten wird überflutet. Die Gartenschaukel lädt die Nut-

zenden dazu ein, in den wohligen Unterstand zurückzukehren. Ein Supervisor-Avatar erscheint, um die beiden in die physische Welt zurückzubegleiten.

Nach dem Ablegen der Wearables und der VR-Headsets werden die beiden Spielenden eingeladen, an dem bereitgestellten Tisch in Austausch zu treten. Auf dem Tisch liegt ein Satz Karten, die Fragen dazu enthalten, wie man auf persönlicher, kommunaler und politischer Ebene klimapositiv handeln könnte, sowie mehrere positive und negative Handlungsoptionen, die beispielhaft für alltägliche und persönliche Entscheidungen stehen. Die Handlungskarten unterliegen zudem einer Bewertung, welche die Handlungen im Sinne eines positiven Klimahandelns einordnet. Die Karten sollen als Diskussionsanregung zwischen den Spielenden und möglichen weiteren Besuchenden der Ausstellung dienen. Auch hier bilden die Anwendungshinweise lediglich ein Angebot.



Abbildung 27: Verschiedene Montagen aus VR-Ansicht und Ausstellungsfoto
Quelle: Eigene Fotografie

Zudem können mittels Blankokarten eigene Handlungskarten erstellt werden, die wiederum in die Diskussion einfließen, aber auch nachfolgenden Ausstellungsbesuchenden zugänglich gemacht werden können.

6.3 User Study 2

Die zweite User Study teilte sich in Abschnitt 1, in dem die Motivationsbildung in Bezug auf Klimahandeln durch die Klima-ACT!-Installation mit der Motivationsbildung durch bloße Informationsbereitstellung verglichen wurde, und in Abschnitt 2, in dem explorativ untersucht wurde, welche Aspekte der Anwendung bestimmte Emotionen ausgelöst haben, auf.

In Abschnitt 1 wurden zwei Gruppen verglichen: Gruppe 1 wurde zur Handlungsmotivation mithilfe eines Fragebogens befragt, erlebte dann die Klima-



Abbildung 28: Ansicht des Kartenspiels
Quelle: Eigene Fotografie

ACT!-Installation (wie unter Kapitel 6.2 beschrieben) und wurde anschließend erneut zur Handlungsmotivation mithilfe eines Fragebogens befragt. Gruppe 2 (Vergleichsgruppe) beantwortete den Fragebogen zur Handlungsmotivation online, dann wurden ihr Informationen zum Klimawandel in Form von Text und Bild zur Verfügung gestellt (entsprechend den Ausstellungspostern des Prototyps). Anschließend wurde erneut die Handlungsmotivation abgefragt. Gruppe 1 wurde während der Nutzung der VR-Installation zudem vom Forschungsteam mit Blick auf Verständlichkeit der Anwendung, der Handhabung der Textilien etc. beobachtet. Unter Beobachtung des Forschungsteams folgte der VR-Anwendung eine ca. 15-minütige Diskussion am Tisch. Die Teilnehmenden erörterten dort untereinander Verhaltensveränderungen, die in ihrem persönlichen Leben möglich wären.

An Abschnitt 2 nahm nur Gruppe 1 teil. Mit jeweils zwei Teilnehmenden wurde qualitativ erarbeitet, welche Elemente des Prototyps besondere emotionale Reaktionen hervorgerufen haben und welche zur Motivationsbildung beigetragen haben könnten. Hierfür wurde den Teilnehmenden ein Zeitstrahl des VR-Ablaufs sowie »PLEX-Karten« (Lucero et al., 2013; Lucero et al., 2014) zur Verfügung gestellt, um Emotionen zu identifizieren, die innerhalb der Anwendung durch Erlebtes oder getätigte Interaktionen zu bestimmten Zeitabschnitten (bildlich verdeutlicht durch Wettersituationen) hervorgerufen wurden. Das Gespräch zwi-

schen den Teilnehmenden und dem Forschungsteam wurde per Audioaufnahme, der Zeitstrahl mit den ausgewählten PLEX-Karten per Foto dokumentiert.



Abbildung 29: Zeitstrahl mit Zeitabschnitten und gelegten PLEX-Karten
Quelle: Eigene Darstellung

6.4 Öffentliche Ausstellung und Studiendurchführung

Die Klima-ACT!-Installation wurde drei Wochen lang öffentlich und kostenlos zugänglich in zwei zentralen Designmuseen ausgestellt: vom 13.–26. Mai 2024 im Rahmen der Ausstellung »Water Pressure« im Museum für Kunst und Gewerbe Hamburg sowie vom 13.–23. Juni 2024 im Rahmen der »Textil Week Hamburg 2024« im Design Zentrum Hamburg. Die User Study 2 wurde vom 2.–4. Mai 2024 mit 22 Teilnehmenden durchgeführt.

7. Vorläufige Ergebnisse

An den Ergebnissen der PLEX-Karten-Befragung lässt sich ablesen, dass der gesamtgestalterische Ansatz von Klima-ACT! eine emotionale Betroffenheit auslöste: Positive (z. B. Entspannung, Sensation, Humor) und negative (z. B. Frustration, Grausamkeit, Leiden) Emotionen entlang der Zeitachse wurden deutlich. Daraus schließen wir, dass gesamtgestalterisch konzipierte Installationen mit interaktiven virtuellen und realen Räumen komplexe Themen (wie die Auswirkungen des Klimawandels) greifbar machen können und für Menschen niedrigschwellig zugänglich sind. Spielerische Mechaniken, die durch die kollaborative Nutzung von textilen Objekten geformt wurden, konnten ein Zusammengehörigkeitsgefühl schaffen, das sich auf die physische Welt übertrug.

Die Entzerrung in einen realen Ausstellungsteil und einen VR-Erlebnisteil sowie der Verzicht einer Win/Loose-Vorgabe führte dazu, dass sich die Teilnehmen-

den einerseits mehr auf das Erleben selbst, andererseits auf die thematische Auseinandersetzung mit dem Klimawandel fokussieren konnten.

In den meisten Diskussionsgruppen waren sich die Teilnehmenden einig, dass das eigene Handeln einen begrenzten Wirkungsgrad habe und Veränderungen vor allem aufseiten der Industrie, gegebenenfalls angeregt durch politisches Handeln, stattfinden müssen, um einen globalen Effekt zu erzeugen. Konträr zu den Angaben im Fragebogen sahen die Teilnehmenden nur eine geringe Handhabe, Einfluss auf Industrie oder Politik ausüben zu können.

Die Auswertung der Fragebögen bildete hingegen eine Tendenz zur Verhaltensänderung im realen Alltag (Vergleich vor und nach der Anwendung/Diskussion) bzw. einer geänderten Einschätzung bezüglich der Folgen des Klimawandels auf das persönliche Leben ab. Am deutlichsten wurde, dass die Teilnehmenden nach der Anwendung bereit sind, in Bereich Mobilität auf öffentliche Verkehrsmittel umzusteigen. Des Weiteren lassen die Angaben auf die Frage, ob die Teilnehmenden meinen, im gemeinsamen Kampf gegen den Klimawandel etwas erreichen zu können, darauf schließen, dass die immersive Kollaboration in VR einen positiven Effekt hinterlassen hat. Auch hier ist eine Verschiebung hin zu positiveren Angaben im Vorher-nachher-Vergleich erkennbar.

Laut Fragebogenauswertung haben sich fast alle Teilnehmenden bereits mit dem Thema Klimawandel und dem eigenen Klimahandeln auseinandergesetzt. In den Diskussionsrunden wurde deutlich, dass die meisten Teilnehmenden ihr eigenes Verhalten bereits geändert hatten, meist in den Bereichen Mobilität (Verzicht auf ein Auto) oder Ernährung (vegetarische oder vegane Ernährung). Eine Auswertung der Ergebnisse aus der Befragung der Gruppe 2 (Vergleichsgruppe) ist in dieser Publikation noch nicht enthalten, weshalb ein Vergleich mit Gruppe 1 noch aussteht.

8. Diskussion und Fazit

Klima-ACT! konzentrierte sich auf die designgeleiteten multisensorischen Möglichkeiten textiler Objekte und gamifizierter VR. Dies ermöglichte es, die Aufmerksamkeit und das emotionale Erleben der Nutzenden durch gestalterische Details gezielt zu lenken. Dabei wurde besonderes Augenmerk auf den Übergang von der physischen Welt in die virtuelle Welt gelegt, der immersions- und präsenzsteigernd wirkte, um dadurch einen niedrigschwelligen Einstieg in einen technisch aufwendigen Installationsaufbau zu gewährleisten.

Die Forschungsergebnisse von Klima-ACT! zeigen, dass gesamtgestalterische Konzepte die physische und die virtuelle Welt mit immersiven Technologien zusammenführen können, um emotionale Betroffenheit hervorrufen. Sie machen

es den Nutzenden leichter, zwischen diesen Welten zu navigieren und die virtuellen Erlebnisse auf ihren Alltag zu beziehen.

Zusätzlich öffnete das Design die Gruppe der Nutzenden für weniger technikaffine und VR-erfahrene Menschen. Die Größe und das Aussehen sowie die Familiarität der realen Objekte und somit der niedrigschwellige und spielerische Umgang damit machten das Klima-ACT!-Erlebnis für Nutzende verschiedener Altersgruppen, Bildungsstände, technischer Vorerfahrung und körperlicher Eigenschaften (beispielsweise Körpergröße) zugänglich.

Die Ausstellung und die User Study 2 von Klima-ACT! wurden im Rahmen der Ausstellung »Water Pressure« durchgeführt und beworben, die sich ihrerseits ebenfalls mit den Folgen des Klimawandels auseinandersetzt (MKG, 2024). Es liegt also nahe, dass sich Teilnehmende mit bereits bestehendem Grundinteresse an der Problematik des Klimawandels zu der Studie angemeldet haben. Ob Menschen mit weniger Kenntnissen oder Zweifeln gegenüber dem menschengemachten Klimawandel ähnlich klimapositiv beeinflusst werden, war in dieser Untersuchung nicht festzustellen. Eine diversere und größere Teilnehmendengruppe ist für weiterführende Studien notwendig. Dafür soll Klima-ACT! zukünftig auch in anderen Kontexten (z. B. Schulen, Einkaufszentren) gezeigt werden.

Es wurde auch deutlich, dass der Fokus auf individuelles Verhalten und dessen Änderung Motivationsgrenzen aufzeigt. So sahen die Teilnehmenden zwar mögliche positive Auswirkungen ihres eigenen Handelns in Bezug auf den Klimawandel, waren aber auch überzeugt, dass globale Veränderung nur eingeschränkt von ihrem persönlichen Handeln abhängt.

Für weitere Entwicklungen sollte sich Klima-ACT! auch darauf fokussieren, wie diese erlebte Betroffenheit in aktives politisches Handeln und nicht nur in eine persönliche Verhaltensänderung überführt werden kann. Ein virtuelles Workshopformat ist bereits in Planung, in dem die gestalterischen Erkenntnisse aus dem VR-Erlebnis genutzt werden, um diese von einer Mikro- (persönliches Handeln) auf eine Mezzo- und Makroebene zu lenken.

Förderhinweis

Das Projekt Klima-ACT! wurde vom Bundesministerium für Bildung und Forschung (BMBF)[1] gefördert.

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Literatur

Ahn, Soun Joo, Bailenson, Jeremy N. & Park, Dooyeon (2014). Short- and Long-term Effects of Embodied Experiences in Immersive Virtual Environments on Environmental Locus of Control and Behavior. In: *Computers in Human Behavior,* Vol. 39, Elsevier Science Publishers B.V., Amsterdam. https://doi.org/10.1016/j.chb.2014.07.025

Albers, Anni (2017). *On Weaving*, Princeton University Press, Princeton.

Arrasvuori, Jura et al. (2011). Applying the PLEX Framework in Designing for Playfulness. In: *Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces*, ACM, New York. https://doi.org/10.1145/2347504.2347531

Barad, Karen (2018). Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter. In: *A Feminist Companion to the Posthumanities*, Springer International Publishing, Basel. https://doi.org/10.1007/978-3-319-62140-1_19

Bekker, Tilde et al. (2015). Investigating Perspectives on Play: The Lenses of Play Tool. In: *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, ACM, New York. https://doi.org/10.1145/2793107.2810328

Buso, Alice et al. (2022). The Unfolding of Textileness in Animated Textiles: An Exploration of Woven Textile-forms. In: *DRS2022: Bilbao*, Bilbao. https://doi.org/10.21606/drs.2022.612

Döring, Tanja (2016). A Materials Perspective on Human-Computer Interaction, Universität Bremen, Bremen.

Eron, Kathryn et al. (2020). Weighted Blanket Use: A Systematic Review. In: *The American Journal of Occupational Therapy*, Vol. 74, No. 2, American Journal of Occupational Therapy, Bethesda. https://doi.org/10.5014/ajot.2020.037358

Frauenberger, Christopher (2020). Entanglement HCI The Next Wave? In: *ACM Transactions on Computer-Human Interaction*, Vol. 27, No. 1, ACM, New York. https://doi.org/10.1145/3364998

Frayling, Christopher (1993). Research in Art and Design. In*: Royal College of Art Research Papers,* Vol. 1 No. 1, Royal College of Arts, London.

Gowrishankar, Ramya, Bredies, Katharina & Ylirisku, Salu (2017). A Strategy for Material-Specific e-Textile Interaction Design. In: *Smart Textiles. Fundamentals, Design, Interaction*, Springer Nature, Berlin. https://doi.org/10.1007/978-3-319-50124-6_11

Hatwell, Yvette, Streri, Arlette & Gentaz, Edouard (2003). *Touch for Knowing*, John Benjamins Publishing Company, Amsterdam. https://doi.org/10.1075/aicr.53

Hobye, Mads & Löwgren, Jonas (2011). Touching a Stranger: Designing for Engaging Experience in Embodied Interaction. In: *International Journal of Design*, Vol. 5, No. 3, International Journal of Design, Taipei City. https://www.ijdesign.org/index.php/IJDesign/article/view/976/364

Homewood, Sarah et al. (2021). Tracing Conceptions of the Body in HCI: From User to More-than-human. In: *Proceedings of the Conference on Human Factors in Computing Systems*, ACM, New York. https://doi.org/10.1145/3411764.3445656

Honauer, Michaela, Yavuz, Secil Uğur & Kuusk, Kristi (2020). WORM-E: An Interactive Toy Enriching Children's Bodily and Social Play. In: *Companion Publication of the 2020 ACM Designing Interactive Systems Conference*, ACM, New York. https://doi.org/10.1145/3393914.3395814

Huizinga, Johan (1981). *Homo Ludens: Vom Ursprung der Kultur des Spiels*, aus dem ndl. v. Hans Nachod, Rowohlt Verlag, Hamburg.

Höök, Kristina (2018). *Designing with the Body: Somaesthetic Interaction Design*, The MIT Press, Cambridge. https://doi.org/10.7551/mitpress/11481.001.0001

Höök, Kristina et al. (2021). Unpacking Non-dualistic Design: The Soma Design Case. In: *ACM Transactions on Computer-Human Interaction*, Vol. 28, No. 6, ACM, New York. https://doi.org/10.1145/3462448

Jerald, Jason (2015). *The VR Book: Human-Centered Design for Virtual Reality*, ACM, New York. https://doi.org/10.1145/2792790

Karga, Valentina (2023). Well Beings. In: *Museum für Kunst und Gewerbe*, 24.3. – 3.9.2023, Hamburg. https://www.mkg-hamburg.de/ausstellungen/valentina-karga (1.9.2025)

Kauhanen, Otto et al. (2017). Assisting Immersive Virtual Reality Development with User Experience Design Approach. In: *Proceedings of the 21st International Academic Mindtrek Conference*, ACM, New York. https://doi.org/10.1145/3131085.3131126

Ley-Flores, Judith et al. (2019). Altering Body Perception and Emotion in Physically Inactive People Through Movement Sonification. In: *8th International Conference on Affective Computing and Intelligent Interaction*, ACII, Cambridge. https://api.semanticscholar.org/CorpusID:209320354

McGonigal, Jane (2010). Gaming can make a Better World. In: *TED2010*, Oxford. https://www.ted.com/talks/jane_mcgonigal_gaming_can_make_a_better_world [Zugriff 09.01.2025]

Meiklejohn, Elizabeth, Devendorf, Laura & Posch, Irene (2024). Design Bookkeeping: Making Practice Intelligible Through a Managerial Lens. In: *Proceedings of the 2024 ACM Designing Interactive Systems Conference*, AMC, New York. https://doi.org/10.1145/3643834.3660754

Meyer zur Capellen, Thomas (2001.) *Lexikon der Gewebe*, Deutscher Fachverlag, Frankfurt.

MK&G – Museum für Kunst und Gewerbe & Jane Withers Studio (2024). Water Pressure – Gestaltung für die Zukunft. In: *Museum für Kunst und Gewerbe*, 15.3. – 13.10.2024, Hamburg. https://www.mkg-hamburg.de/ausstellungen/water-pressure [Zugriff 09.01.2025]

Mueller, Florian et al. (2020). Towards Designing Bodily Integrated Play. In: *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, New York. https://doi.org/10.1145/3374920.3374931

Poulsen, Mathias (2020). Designing for Playful Citizenship. In: *Framing Play Design – A Hands-on Guide for Designers, Learners & Innovators*, BIS Publishing, Amsterdam. https://www.bispublishers.com/framing-play-design.html [Zugriff 09.01.2025]

Redström, Maria, Redström, Johan & Mazé, Ramia (2010). *IT + Textiles*, Edita Publishing Oy, Helsinki.

Robles, Erica & Wiberg, Mikael (2010). Texturing the »Material Turn« in Interaction Design. In: *Proceedings of the 4th International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, New York. https://doi.org/10.1145/1709886.1709911

Scholz, Barbro et al. (2022). Designing Wearable Interactions Through Playful On-body Explorations. In: *Short Paper Proceedings of the 5th European Tangible Interaction Studio*, CEUR Workshop Proceedings, Bielefeld. https://ceur-ws.org/Vol-3328/workshop5.pdf

Seifi, Hasti, Zhang, Kailun & MacLean, Karon E. (2015). VibViz: Organizing, Visualizing and Navigating Vibration Libraries. In: *2015 IEEE World Haptics Conference*, IEEE, New York. https://doi.org/10.1109/WHC.2015.7177722

Shusterman, Richard (1999). Somaesthetics: A Disciplinary Proposal. In: *The Journal of Aesthetics and Art Criticism*, Vol. 57, No. 3, JSTOR, New York. https://doi.org/10.2307/432196

Strohmeier, Paul et al. (2020). bARefoot: Generating Virtual Materials Using Motion Coupled Vibration in Shoes. In: *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, ACM, New York. https://doi.org/10.1145/3379337.3415828

Studio MANND (2018). *X-Ray Fashion VR*, Copenhagen. https://www.mannd.dk/projects/xray-fashion [Zugriff 09.01.2025]

Takatalo, Jari et al. (2010). Presence, Involvement, and Flow in Digital Games. In: *Evaluating User Experience in Games*, Springer-Verlag, London. https://doi.org/10.1007/978-1-84882-963-3_3

Yavuz, Seçil Uğur et al. (2021). Design for Playfulness with Interactive Soft Materials: Description Document. In: *Proceedings of the 15th International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, New York. https://doi.org/10.1145/3430524.3442702

Vallgårda, Anna (2014). Giving Form to Computational Things: Developing a Practice of Interaction Design. In: *Personal and Ubiquitous Computing*, Vol. 18, No. 3, Springer-Verlag,London. https://doi.org/10.1007/s00779-013-0685-8

Vallgårda, Anna & Sokoler, Tomas. (2009). A Material Focus – Exploring Properties of Computational Composites. In: *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, ACM, New York. https://doi.org/10.1145/1520340.1520631

Vrellis, Ioannis et al. (2010). Measuring Presence in a Collaborative Physics Learning Activity in Second Life. In: *Proceedings of the 7th Pan-Hellenic Conference with International Participation »ICT in Education«*, Vol. 1, University of Peloponnese, Korinthos. https://www.researchgate.net/publication/267769793_Measuring_presence_in_a_collaborative_physics_learni ng_activity_in_Second_Life

Weber, Elke.U. (2006). Experience-Based and Description-Based Perceptions of Long-Term Risk: Why Global Warming Does Not Scare Us (Yet). In: *Climatic Change*, Vol. 77, Springer Nature, Berlin. https://doi.org/10.1007/s10584-006-9060-3

Whittle, Clayton et al. (2022). *The Environmental Game Design Playbook*, International Game Developers Association, Mount Royal.

Wiberg, Mikael (2018). *The Materiality of Interaction. Notes on the Materials of Interaction Design*, The MIT Press, Cambridge. https://doi.org/10.7551/mitpress/10427.001.0001

Wilde, Danielle, Vallgårda, Anna & Tomico, Oscar (2017). Embodied Design Ideation Methods: Analysing the Power of Estrangement, Conference on Human Factors in Computing Systems. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM, New York. https://doi.org/10.1145/3025453.3025873

Kulturelle und soziale Teilhabe aus der Ferne – Gestaltung und Entwicklung eines immersiven und sozialen virtuellen Theatererlebnisses – SocialSTAGE-VR

Frank Dittrich, Franziska Legler, Selina Palige, Angelika C. Bullinger, Joachim Keinert, Tobias Jaschke, Andree Josef, Anne Zeiler, Martin Krabbe

Zusammenfassung

Immersive und räumliche Computertechnologien, wie Virtual Reality, bieten neue Möglichkeiten, reale Bühnenveranstaltungen von zu Hause aus zu erleben. Im Projekt »SocialSTAGE-VR« der Förderrichtlinie des Bundesministeriums für Bildung und Forschung (BMBF)[1] »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft (VAR2)« wurden Ansätze für die virtuelle Teilnahme an realen Bühnenveranstaltungen entwickelt und beforscht. Im Beitrag wird ein technisches Konzept für das Erleben von Theateraufführungen mit Virtual-Reality-Headsets vorgestellt. Hierbei wurden reale Bühneninhalten von Theateraufführungen in ein virtuelles Theatermodell übertragen. Dabei kamen zwei Ansätze zum Einsatz: eine stereoskopische Videoaufnahme und eine Lichtfeldaufnahme. Im virtuellen Theatersaal ermöglichte die Integration eines Multiplayer-Frameworks das gemeinsame Erleben von lokal entfernten Personen, die über Avatare auf natürliche Weise miteinander interagieren können.

1. Einleitung

Theater- und Konzertbesuche sorgen für besondere Erlebnisse, die nicht nur eine kulturelle Teilhabe ermöglichen, sondern auch mit vertrauten Personen geteilt werden können. Persönliche sowie situationsbedingte Einschränkungen, aber auch pandemische Krisen können dieser Form der Teilhabe im Wege stehen oder

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

verhindern sie ganz. Mithilfe von Technologien der virtuellen Realität (VR) wurde im Projekt SocialSTAGE-VR in der Förderrichtlinie »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft (VAR2)« ein orts- und zeitunabhängiges Erleben kultureller Veranstaltungen beforscht. Reale Bühneninhalte wurden dafür anhand von stereoskopischen, also 3D-fähigen, Aufnahmen in virtuelle Veranstaltungsräume übertragen, die mittels VR-Brillen konsumiert werden können. Indem man virtuell präsent ist, soll man diese Erlebnisse auch mit vertrauten Personen teilen können. Dabei stand die Realisierung einer möglichst natürlichen Interaktion auf Basis von Avataren im Mittelpunkt. Zusammen mit Künstlerinnen und Künstlern wurden dafür Lösungen entwickelt, die neben einem möglichst realen Erleben auch Potenziale erschließen, die über die Erfahrung realer Veranstaltungen hinausgehen. Ob für örtlich getrenntlebende Personen, für Menschen mit mobilen oder körperlichen Einschränkungen oder einfach aufgrund von zeitlichen Defiziten – mittels der entwickelten Technologien sollte das gemeinsame Erleben realer Bühnenveranstaltungen dennoch möglich werden. Dies eröffnet neue Wege, kulturelle Inhalte zu schaffen und an diesen teilzuhaben. Im Beitrag werden wesentliche technische Projektergebnisse und wissenschaftliche Erkenntnisse vorgestellt. Dies geschieht am Beispiel der Entwicklung einer Software-Anwendung für die VR. Hierfür werden im Folgenden die ermittelten Nutzungsanforderungen, das technische Konzept und Evaluationsergebnisse vorgestellt.

2. Motivation

Der Besuch einer kulturellen Bühnenveranstaltung, etwa eines Konzerts oder eines Theaterstücks, ist ein besonderes Erlebnis. Neben den kulturellen Inhalten zählen u. a. die »Authentizität« und die Atmosphäre als wichtige Erlebenskomponenten (Schneider, 2008). Neben dem Erleben des Ereignisses an sich, sind Kulturveranstaltungen durch eine soziale Komponente des gemeinschaftlichen Erlebens geprägt. Ob durch eine direkte Interaktion mit vertrauten Personen oder durch das Gefühl mit anderen, auch unbekannten Personen das Gleiche zur gleichen Zeit zu erleben, von Bedeutung ist jeweils, ein Erlebnis mit anderen zu teilen. Denn wer geht gern allein ins Theater oder würde ein Konzert ohne Publikum besuchen?

Teilhabe, ob kulturell oder sozial, ist ein essenzieller Teil der Selbstbestimmung des Menschen (Rudolf, 1995: 14). Aber kulturelle Veranstaltungen sind nicht für alle gleichermaßen leicht zugänglich und gemeinsam erlebbar. So verhindern Faktoren wie räumliche Distanz, soziale Isolation oder individuelle Einschränkungen die Möglichkeit des Zustandekommens gemeinsamer kultureller Erleb-

nisse. Die Teilnahme an kulturellen Veranstaltungen ist zudem oft mit einem hohen finanziellen und zeitlichen Aufwand verbunden (Dollas, 1997). So zeigt sich, dass trotz des hohen gesellschaftlichen Wertes dieser Erlebnisse nur etwa fünf Millionen Personen in Deutschland regelmäßig Konzerte oder ähnliche Veranstaltungen besuchen (Statista, 2024: 1). Alternativen zum Besuch realer Kulturveranstaltungen finden sich heue bei Musikkonzerten. Hier gibt es eine steigende Anzahl von Livestreaming-Konzerten. Die geringe Nutzungshäufigkeit dieser Angebote (YouGov, 2020: 1) lässt jedoch darauf schließen, dass sich die positiven Effekte kultureller und sozialer Teilhabe beim Livestreaming nicht einstellen.

Digitale Medienformate können zwar die Zugänglichkeit entsprechender Inhalte aus der Ferne vereinfachen, bieten aber bisher nicht die mit realen Besuchen verbundenen Erlebniskomponenten. Ein Tele-Konsum (TV oder 360°-VR-Videos) erzeugt nicht das Gefühl, Teil einer bestimmten Gemeinschaft zu sein, denn man schaut nur von »außen« zu. Bei realen Veranstaltungen kann sich dagegen ein sogenanntes Flow-Gefühl einstellen – ein mentaler Zustand positiver Emotionen und ungeteilter Konzentration und Vertiefung in eine Aktivität, in der man vollständig »aufgeht« (Maier et al., 2018: 1). Dieser Zustand ist während der Interaktion von Künstler:innen (Deutsch, 2009: 79) mit Zuschauer:innen, aber auch im Publikum selbst wahrscheinlich. Als »Außenstehende« einer Live-Aufnahme ist dieser Zustand schwerer zu erreichen und mindert damit auch das Erlebnis der kulturellen Teilhabe.

Neue Ansätze hierzu bieten immersive und räumliche Computertechnologien und insbesondere VR. Immersive Technologien, konsumiert mittels sogenannter Head-Mounted-Displays, ermöglichen eine begeh- und erlebbare Verkörperung digitaler Informationen (Wilkinson et al., 2021: 1099). Dadurch wird das Gefühl räumlicher und sozialer Präsenz erzeugt, d. h. das Gefühl, selbst entgegen besseren Wissens an einem anderen Ort zu sein (Skarbez et al., 2017: 16) oder mit entfernten Menschen räumlich zusammen zu sein (Oh et al., 2018: 2). Dadurch wird es möglich, die Erlebniskomponenten kultureller und sozialer Teilhabe bei mediengestützter Remote-Teilnahme zu erzeugen, die bisher nur bei realen Veranstaltungen möglich waren. Dadurch ergibt sich für den regionalen Kultur- und Unterhaltungssektor die Chance, eine globale Zielgruppe zu erreichen und Inhalte orts- und zeitunabhängig verfügbar zu machen. Dies könnte zu einer Demokratisierung von kulturellen Bühnenformaten führen.

Ziel der hier vorgestellten Forschungsarbeit war es, eine VR-Anwendung inklusive einer Aufnahme- und Verarbeitungspipeline für das immersive Erleben real stattfindender Bühnenveranstaltungen am Beispiel von Theateraufführungen zu konzipieren und umzusetzen. Dabei stand eine möglichst getreue Überführung des realweltlichen Erlebens inklusive sozialer Faktoren für eine kulturelle und soziale Teilhabe im Vordergrund der Bemühungen.

3. Nutzungsanforderungen und technische Limitationen

Damit virtuelle Bühnenveranstaltungen erfolgreich sein können, muss ein ausreichend großer Mehrwert gegenüber derzeitigen Formen des Remote-Konsums, wie 2D-Streaming, erzeugt werden. Die Vorteile des räumlichen und sozialen Präsenzerlebens sollten deshalb aus technischer Sicht bestmöglich unterstützt werden. Für einen breitenwirksamen Erfolg des Ansatzes aufseiten der Veranstaltenden, muss die Technologie möglichst niederschwellig sein. Zudem sollten die realen Veranstaltungen sowie das Erleben für das Live-Publikum vor Ort nicht durch den technischen Aufbau eingeschränkt werden.

Die Aufzeichnung und Übertragung von 360°-Videos stellt derzeit den am häufigsten genutzten technisch niederschwelligen Ansatz dar, um real stattfindende Inhalte immersiv verfügbar zu machen (Matzen et al., 2017: 1; Weissig et al., 2012: 672). In der Praxis bieten diese Lösungen aufgrund von übertragungs- und darstellungsbedingten Limitationen der Gesamtauflösung oft eine reduzierte Immersion. Auch wenn dies zukünftig gelöst werden kann, ein gravierender Nachteil bleibt bestehen. 360°-Videoinhalte haben nur drei Freiheitsgrade (Three Degress of Freedom [3-DoF]). Translationsbewegungen der Nutzenden im Raum (vor – zurück, hoch – runter, links – rechts) können nicht abgebildet werden. Dies wirkt sich negativ auf die Immersion und das Erlebnis der räumlichen Präsenz aus (Debarba et al., 2022: 5; Chandrasekera et al., 2019: 520). Bei eigener Bewegung erhalten Nutzende nur dann das Gefühl, »Teil der Umgebung zu sein«, wenn sich, wie in der realen Welt, auch seine Perspektive ändert und Objekte in unterschiedlichen Entfernungen ihre Position zueinander verschieben (Parallaxe). Da dies bei 360°-Videoinhalten nicht gegeben ist, entsteht ein Eindruck des passiven Konsums von Inhalten aus der Ferne. Vor allem bei Live-Übertragungen stellt sich zudem das Problem, dass 360°-Aufzeichnungen auch das Publikum erfassen. Dies wirft Datenschutzbeschränkungen und Akzeptanzprobleme auf.

Im Gegensatz zu Videoinhalten ermöglichen computergenerierte virtuelle Umgebungen auch translatorische Bewegungen (Six Degress of Freedom [6-DoF]). In diesen können sich die Nutzenden frei bewegen, wodurch ein räumliches Präsenzgefühl entsteht, so als wäre man selbst vor Ort. Diese Umgebungen eignen sich auch für die Einbindung virtueller menschlicher Abbilder (Avatare) als Voraussetzung für ein gemeinsames Erlebnis und die Erzeugung eines sozialen Präsenzempfindens. Reale Objekte und Umgebungen können jedoch nicht einfach in 3D-Software gerendert werden. Damit real stattfindende Veranstaltungen und insbesondere die Bühneninhalte selbst in 3D-Software gerendert werden können, werden spezielle Studio-Setups notwendig. So können beispielsweise reale Bewegungen der Künstler:innen auf zuvor erstellte virtuelle Bilder übertragen werden (Xia et al., 2017: 537; Winkler et al., 2022: 6). Andere

Möglichkeiten bieten volumetrische Videos, die Inhalte aus vielen Perspektiven aufnehmen und daraus 3D-Inhalte berechnen. Während der Einsatz von volumetrischen Videos derzeit mit zu hohen Hardwareanforderungen für die Darstellung, aber auch für die Echtzeitverarbeitung verbunden ist, erfordern beide Ansätze auch im Bühnenbereich komplexe Aufnahmetechniken. Dies steht einer niederschwelligen und flexiblen Nutzung im Wege und könnte auch das Vor-Ort-Erlebnis der Veranstaltungsformate einschränken.

4. Konzept und technische Umsetzung

Um die Anforderungen eines hohen Präsenzerlebnisses und einer gleichzeitig niedrigschwelligen technischen Umsetzung zu erfüllen, wird im Folgenden ein hybrides Format vorgeschlagen, das virtuelle Inhalte mit Videoaufnahmen von der Bühne fusioniert. Forschungen in anderen Kontexten zeigten bereits gute Wirkungen auf das Präsenzerlebnis bei der Verschmelzung von computergenerierten 3D-Inhalten mit Videoaufnahmen (Debarba et al., 2022: 18). Der hier verfolgte Ansatz überträgt und erweitert dieses Konzept auf das Erleben von realen Bühnenereignissen und implementiert es in ein soziales Gesamterlebnis (vgl. Abbildung 1). Im Folgenden wird der Ansatz am Beispiel der Umsetzung einer VR-Anwendung für virtuelle Theatererlebnisse des Opernhauses Chemnitz vorgestellt.

4.1 Virtualisierung des Theatersaals

Damit sich die Nutzenden für eine möglichst hohe Immersion anhand aller sechs Freiheitsgrade in der Anwendung bewegen können, wurde für den gesamten Theaterraum eine 3D-Nachbildung, ein sogenannter digitaler Zwilling, erstellt. Damit zukünftig auch kleinere Kulturbetriebe davon profitieren können, muss der Aufwand zur Erstellung virtueller Event-Locations möglichst gering sein. Untersucht wurden hierfür verschiedene Herangehensweisen. In der Regel werden virtuelle Umgebungen mit 3D-Programmen, wie beispielsweise Autodesk 3D Studio Max oder Blender, modelliert. Dieser Prozess ist je nach Größe, Beschaffenheit des Raumes und benötigtem Realitätsgrad sehr aufwendig. Dies trifft besonders auf komplexe Veranstaltungsräume wie Opernhäuser zu, die sehr groß sind und sich durch viele Details auszeichnen. Die korrekten räumlichen Geometrien sind beispielsweise auf Basis von technischen Plänen oder 2D-Bildmaterial schwer zu ermitteln und die Erstellung von Nachbildungen kann fehlerhaft sein. Gerade aber für das hier vorgeschlagene Setup, der Über-

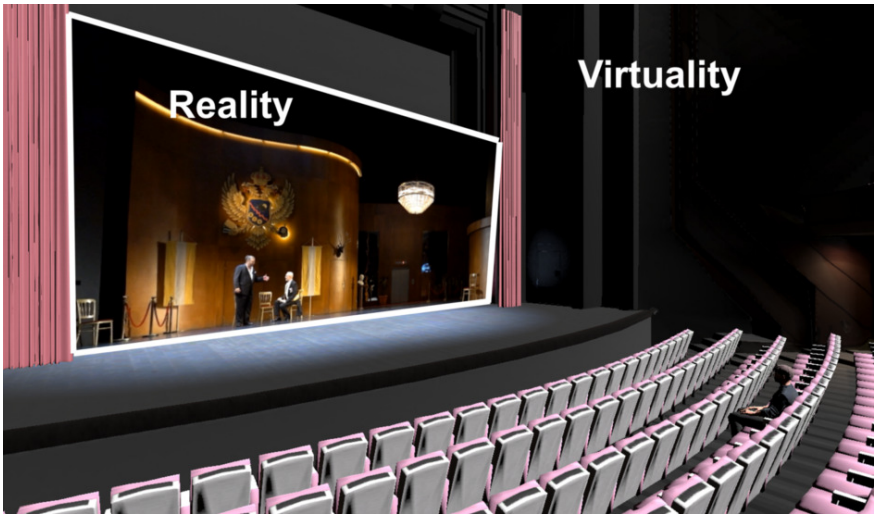


Abbildung 1: VR-Umgebung bestehend aus einem computergenerierten virtuellen Zuschauerraum und realen Videoaufnahmen des Bühneninhaltes
Quelle: SocialSTAGE-VR

tragung von Videoaufnahmen der Bühneninhalte in den virtuellen Raum, ist die Einhaltung exakter Geometrien der Bühne selbst und zur Blickachse der späteren Betrachtungsperspektive besonders wichtig. Alternativen bieten Verfahren, die die reale Umgebung automatisiert erzeugen, wie beispielsweise die Photogrammetrie. In einem ersten einfacheren Versuch wurden dafür 1194 Fotos aus unterschiedlichen Raumperspektiven aufgenommen. Programme, wie das hier verwendete »Reality Capture«, ermitteln aus den Parallaxenverschiebungen von Bildern aus verschiedenen Blickwinkeln eine räumliche Punktewolke und übertragen die Texturierung aus dem Bildmaterial, sodass ein 3D-Modell der Szenen entsteht. Das Ergebnis des Versuchs gab die grundsätzliche Geometrie des Opernhauses gut wieder und zeigte vereinzelt gute Texturierungen. Das Ergebnis enthielt aber sehr viele Fehler und Artefakte. Viele Stellen des Saals wurden trotz der hohen Bildanzahl aufgrund der Erreichbarkeit der Positionen nicht richtig erfasst (Abbildung 2).

In einem zweiten Versuch wurde eine angepasste Aufnahmestrategie ausgewählt. Dazu sollten in kreisförmigen Umläufen Aufnahmen aus der Totale in unterschiedlichen Höhen mit Detailaufnahmen der Wände, Böden und Decken kombiniert werden. Um die Erreichbarkeit der notwendigen Perspektiven zu erleichtern und zu beschleunigen, wurde dafür eine Drohne eingesetzt. Für die Detailaufnahmen flog die Drohne jeweils Geraden parallel zu den Wänden ab, um so möglichst gut Parallaxeverschiebungen zu dokumentieren (vgl. Abbildung

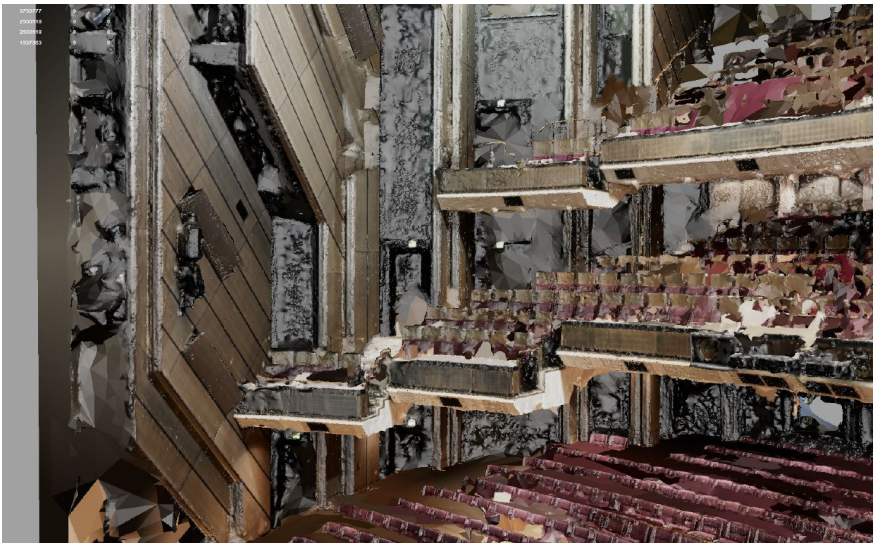


Abbildung 2: Das mittels Photogrammetrie erzeugte Modell zeigt deutliche Artefakte.
Quelle: SocialSTAGE-VR

3). Mit dem Einsatz einer Drohne können zudem sehr schnell akkurate Positionswechsel umgesetzt werden, was manuell nicht möglich wäre. In Summe entstanden so 1500 Fotos. Das daraus erzeugte Ergebnis war qualitativ deutlich hochwertiger und bot einen vergleichsweisen guten optischen Gesamteindruck. Die meisten relevanten Geometrien konnten korrekt erfasst werden.

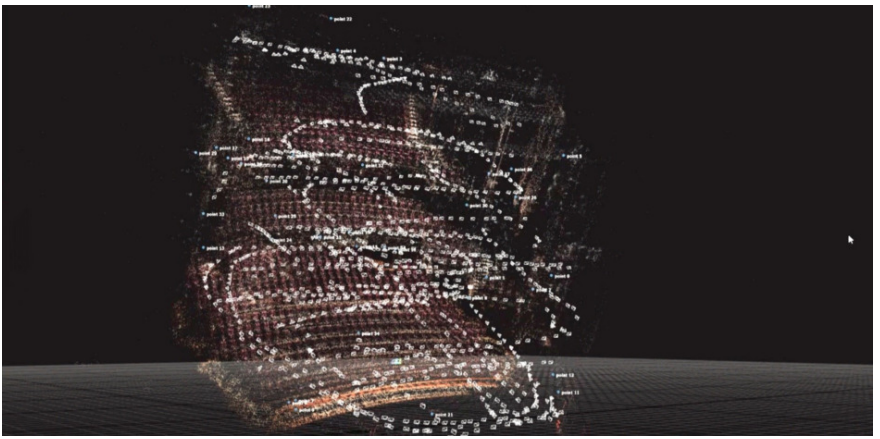


Abbildung 3: Aufnahmepositionen (helle Flecke bzw. Vierecke) und entstandene Punktewolke (schwache Punkte)
Quelle: SocialSTAGE-VR

Dennoch enthielt das Ergebnis gerade im Bereich von Verdeckungen, wie den Sitzreihen, noch viele Fehler. Besonders im Fall von glatten Oberflächen wurden aufgrund der ungenauen Tiefenschätzung unnötige Geometrien eingebaut, was optisch deutlich die Immersion reduzierte. Automatisierte Optimierungsverfahren sind derzeit noch nicht in der Lage, dieses Problem zufriedenstellend zu lösen. Die hohe Anzahl an Polygonen von einer Million Flächen ist außerdem für die Nutzung in VR-Anwendungen nicht performant genug (vgl. Abbildung 4).

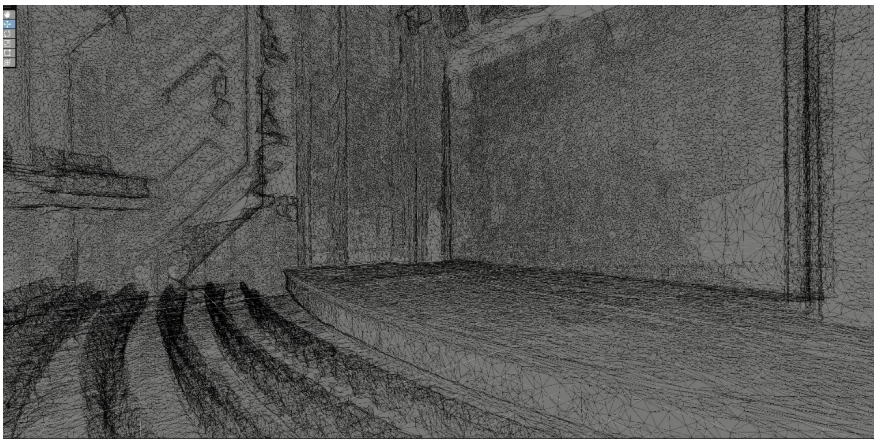


Abbildung 4: Netzwerk aus einer Million Polygonen
Quelle: SocialSTAGE-VR

Deshalb musste das Modell in einem weiteren Schritt manuell nachmodelliert werden. Durch die Überlagerung in einer 3D-Modellierungssoftware konnten auf Grundlage des vorhandenen Modells die Geometrien aufwandsarm nachgezeichnet und auf Basis des Bildmaterials nachtexturiert werden. Das optimierte Modell enthielt nach der Bearbeitung nur noch 4873 Polygone und wurde anschließend in ein Unity-Projekt integriert (vgl. Abbildung 5). Da die 714 Sitzplätze jeweils identisch waren, wurden diese über ein Skript automatisch als 3D-Modell im Raum dupliziert und positioniert. So befindet sich anstatt 714 Sitzplätzen nur ein Sitzplatz im Speicher der Grafikkarte, was die Performance deutlich erhöht. Durch eine automatische Platzierung besteht zudem die Möglichkeit, die Bestuhlung je nach Szenario individuell anpassen zu können. Ein Beleuchtungskonzept ermöglicht die kontextsensitive Anpassung der Lichtverhältnisse im Raum. Dabei wurden besondere Einstellmöglichkeiten auf den Bühnenraum gelegt, um die Integration der Videoaufnahme zu optimieren. So kann beispielsweise der in 3D modellierte Bühnenboden verschiedene Farbtöne und Helligkeiten annehmen, um sich dem realen Bühnenboden der Videoaufzeichnung anzugleichen. Mithilfe

eines sogenannten Shaders können zudem die Außenbereiche des Videos in eine zunehmende Transparenz überführt werden, was sichtbare Übergänge zwischen Video und modellierter Umgebung reduziert.



Abbildung 5: Endergebnis des virtualisierten Theatersaals
Quelle: SocialSTAGE-VR

4.2 Integration der Bühneninhalte per Videoaufzeichnung

Die Integration von Bühneninhalten realer Veranstaltungen in Form von Videoaufzeichnungen stellt eine einfache Möglichkeit dar, reale Theateraufführungen in virtuelle Räume zu erweitern und ortsunabhängig erlebbar zu machen. Dazu werden die Theateraufführungen mit Kamerasystemen vor Ort aufgezeichnet und in den virtuellen Bühnenraum integriert. Da einfache Videoaufnahmen in der Regel keine Tiefeninformationen enthalten, kann es bei der Einbindung entsprechender Aufnahmen in eine virtuelle 6-DoF-Umgebung zu Immersionsbrüchen kommen. Es besteht die Gefahr, dass die Inhalte aus Sicht der Betrachtenden wie eine Leinwand erscheinen und somit nicht als Teil des Raumes wahrgenommen werden. Um dennoch mittels Videoaufnahmen immersive Erlebnisse zu schaffen, wurden zwei Ansätze verfolgt.

Eine relevante Tiefeninformation ergibt sich aus dem stereoskopischen Sehen. Stereoskopie beschreibt die Darbietung von zwei Bildern für das linke und rechte Auge, welche in ihrer Perspektive leicht versetzt sind. Während die

computergenerierte Umgebung bereits für beide Augen entsprechende Bilder berechnet, setzt dies für die Videoaufzeichnung ein stereoskopisches Kamerasystem voraus (vgl. Abbildung 6 links). Die Vorteile dieses Ansatzes sind die gute Verfügbarkeit von Kameraaufbauten und der einfachere Übertragungsprozess, der auch gut in Echtzeit umgesetzt werden kann. Dieser Aufbau erfordert lediglich eine Kalibrierung vor Ort und eine sorgfältige Integration in die VR-Szene (vgl. Abbildung 6 rechts). So sorgen beispielsweise Shader für einen unauffälligen Übergang zwischen dem Videorand und dem Bühnenraum.



Abbildung 6: Stereoskopisches Kamerasystem (links) und stereoskopische Einbindung des Videos durch zwei versetzte Leinwände (rechts)
Quelle: SocialSTAGE-VR

Stereoskopische Aufnahmen bieten allerdings keinen Parallaxeneffekt. Das bedeutet, dass die Darstellung der Szene immer gleich bleibt, auch wenn sich die betrachtende Person räumlich bewegt. Allerdings können diese Effekte bei größeren Abständen zur Bühne und bei nur kleinen Bewegungen, wie dies beim Sitzen auf einem Sitzplatz der Fall ist, weniger stark ausgeprägt sein. Um dennoch auch die Parallaxe abbilden zu können, wurde als zweiter Ansatz eine angepasste Lichtfeldtechnik genutzt. Dazu wird die Szene aus mehreren leicht versetzten Perspektiven aufgenommen und anschließend maschinell verarbeitet. Im Idealfall wird eine große Anzahl von Kamerapositionen für die Aufnahme verwendet, damit alle von der Szene ausgehenden Lichtstrahlen erfasst werden können. So ließe sich für jede Betrachtungsposition ein korrektes Bild erzeugen. Dies ist jedoch sehr aufwendig und entspricht nicht den beschriebenen Anforderungen.

Da die Inhalte jedoch in der Regel im Sitzen konsumiert werden, ist ein korrektes Bildmodell aus allen Perspektiven nicht notwendig. Daher wird die Verwendung eines kleinen Kamera-Arrays vorgeschlagen, das leicht versetzte Perspektiven aus einer bestimmten Sitzposition aufnimmt (vgl. Abbildung 7 links). Die Unterschiede in den Bildern, die durch den Parallaxeneffekt entstehen, wer-

den zur Bestimmung der Tiefeninformation verwendet. Diese werden anschließend genutzt, um die Videoinhalte in ein 3D-Modell zu überführen (vgl. Abbildung 7 rechts). Durch die Anpassung der Größe des Kamera-Arrays an den gewünschten Bewegungsradius der Zuschauenden können auch mit kleinen Arrays gute Ergebnisse erzielt werden. Der Nachteil dieses Ansatzes ist der höhere Verarbeitungsaufwand und die deutlich höheren Datenmengen im Vergleich zu einer Stereoaufnahme. Dies stellt vor allem bei Echtzeitübertragungen bei Live-Veranstaltungen eine große Herausforderung dar. Zudem ist die Überführung der realen Aufnahmen in 3D-Modelle nicht frei von Fehlern, wodurch Artefakte entstehen, die negativ auf die Experience einwirken können.



Abbildung 7: Kamera-Array zur Aufnahme verschiedener leicht versetzter Perspektiven im Bewegungsradius des natürlichen Sitzens (links) und eine gekippte Szene des 3D-Videos mit Blick auf die Tiefenebenen und sichtbaren Verzerrungen (rechts)
Quelle: SocialSTAGE-VR

4.3 Multiplayer-Funktionen

Theatervorstellungen sind soziale Aktivitäten. Dabei ist das Gefühl eines zeitsynchronen gemeinsamen Erlebnisses mit anderen Menschen entscheidend für die Begeisterung während einer Live-Aufführung (Geige, 2018: 715). Um einen deutlichen Mehrwert für die breite Nutzung entsprechender virtueller Anwendungen zu schaffen, werden folgende Funktionen benötigt: (1) Bekannte Personen, wie Freund:innen und Familienangehörige, können das virtuelle Erlebnis ebenfalls besuchen, (2) es kann ein gegenseitiger Austausch stattfinden und (3) eine Zuschauerkulisse sollte einbezogen werden zur Vermeidung eines nahezu leeren Theatersaals. Andere soziale VR-Anwendungen ermöglichen es, alle Teilnehmenden als Avatare anzuzeigen, die miteinander interagieren können. Der Nachteil dabei ist, dass die Anzahl der Teilnehmenden zum einen aus Performance-Grün-

den schlecht skalierbar sind, was insbesondere bei großen Veranstaltungen zu einem schlechten Erlebnis führen würde. Zum anderen ist es im Falle eines Theaterbesuchs nicht kontextrelevant, mit allen Zuschauenden in einen direkten Austausch zu treten. Im Gegenteil, dadurch können störende Einflüsse auftreten, wie Gespräche von unbekannten Dritten, die das Erlebnis stark verschlechtern.

Aus diesen Gründen wurde in der vorliegenden Implementierung eine Trennung zwischen den vorgesehenen Interaktionspartner:innen, also den Begleitpersonen, und dem restlichen Publikum bestehend aus unbeteiligten Dritten vorgenommen und zwischen einem direkten und einem indirekten Kollaborationsszenario unterschieden. Beim direkten Kollaborationsszenario treffen sich vertraute Personen im virtuellen Raum als direkte kollaborative Gruppe, um die Veranstaltung gemeinsam zu erleben und sich auszutauschen. Das indirekte Szenario betrifft alle anderen Teilnehmenden in weiteren kollaborativen Gruppen, mit denen keine direkte Interaktion stattfinden soll.

Für das direkte Kollaborationsszenario wurden Multi-User-Funktionen zur direkten Übertragung von Bewegungsdaten und Audio auf Basis des Agora-Multiplayer-Frameworks integriert. Für die Darstellung der Personen sowie des eigenen Avatars wurde das Avatarsystem von ReadyPlayerMe verwendet (vgl. Abbildung 8). Es können beliebig viele kollaborative Gruppen teilnehmen. Eine gruppenübergreifende Interaktion findet aber nicht statt. Da die Bühneninhalte im vorliegenden Fall nur aus einer Perspektive aufgezeichnet wurden, entsteht ein Sweetspot, der das optimale immersive visuelle Erlebnis aus einem bestimmten Sitzbereich im Theatersaal bietet. Alle Gruppen sitzen demzufolge unbemerkt in verschiedenen Instanzen übereinander. Für die Visualisierung des restlichen Publikums wurden verschiedene animierte Avatare für die nähere Umgebung verwendet. Um die Performance zu optimieren, wurden für weiter entfernte Avatare perspektivisch korrekte 2D-Objekte verwendet. So konnte die verwendete Polygonenanzahl deutlich reduziert und die Performanz gesteigert werden.

Für das indirekte Kollaborationsszenario wurde eine Geräuscherkennung entwickelt und implementiert, um Rückmeldungen aus den einzelnen direkten kollaborativen Gruppen zuzulassen. Im vorliegenden Fall können Klatschgeräusche erkannt werden. Bei Erkennung entsprechender Ereignisse werden visuelle und akustische Informationen auf das simulierte Publikum in allen direkten kollaborativen Gruppen übertragen, was das gemeinschaftliche Erlebnis erhöhen soll.



Abbildung 8: Blick in den mit Avataren bevölkerten Theatersaal
Quelle: SocialSTAGE-VR

5. Menschzentrierte wissenschaftliche Begleitung und Evaluation

Mit dem Ziel eines breitenwirksamen Erfolgs der Anwendung erfolgte eine wissenschaftliche Begleitung der technischen Entwicklung. Dazu wurde der menschzentrierte Entwicklungsprozess angewendet. Um interaktive Systeme, wie VR-Anwendungen, menschzentriert zu gestalten, wurden potenzielle Nutzende der Systeme aktiv und iterativ in die Entwicklung einbezogen (Deutsches Institut für Normung, 2019: 12–13). Methodisch wurden dabei in frühen Entwicklungsphasen Stakeholder-Analysen, Workshops und Befragungen eingesetzt. Dabei konnte beispielhaft der Bedarf des Austauschs mit Begleitpersonen vor der eigentlichen Aufführung ermittelt werden sowie der Wunsch eines Einblicks in den Bühnenraum. Konkret folgte daraus die Implementierung einer virtuellen Lobby sowie eines virtuellen Erkundungsraums, der einen Blick von jeder beliebigen Position des Theatersaals, inklusive vom Bühnenraum aus, ermöglicht. In späteren Entwicklungsphasen erfolgten Nutzungsstudien zur formativen Evaluation der technischen Anwendung und Ableitung von konkreten Anpassungsbedarfen. So konnte beispielhaft in einer Studie mit 41 Teilnehmenden gezeigt werden, dass eine stereoskopische Videoaufzeichnung mit leicht erhöhtem Tiefeneindruck (Stereobasis) im Vergleich zum natürlichen stereoskopischen Sehen präferiert wurde und entsprechend in der weiteren Entwicklung der Anwendung bei Videoaufzeichnungen beibehalten wurde. Insgesamt wurden die Rückmeldungen und Daten von mehr als 500 Teilnehmenden zur Entwicklung der Anwendung einbezogen.

Um zu überprüfen, ob eine möglichst getreue Überführung des realweltlichen Erlebens inklusive sozialer Faktoren für eine kulturelle und soziale Teilhabe ermöglicht werden kann, erfolgte zudem ein Vergleich der Anwendung im Prototypstadium mit realen Theateraufführungen. Es konnte gezeigt werden, dass zwei Skalen zur Messung sozialer (*Mvirtuell* = 5.39, *Mreal* = 5.05, Beispiel-Item »Die Zeit des Zusammenseins mit meiner Begleitung empfand ich als bereichernd«) und kultureller Teilhabe (*Mvirtuell* = 5.49, *Mreal* = 5.15, Beispiel-Item »Durch die heutige Veranstaltung konnte ich etwas dazulernen«) ähnliche Mittelwerte auf einer Antwortskala von 1 bis 7 aufwiesen. Dagegen bleibt die virtuelle Anwendung im Prototypstadium hinsichtlich des Entertainments noch hinter realen Veranstaltungen zurück (*Mvirtuell* = 5.32, *Mreal* = 6.15, Beispielitem »Ich fühlte mich gut unterhalten«).

6. Fazit und Ausblick

Virtuelle Technologien bieten das Potenzial, die Vorteile digitaler Mediennutzung (aufwandsarme und flexible Zugänglichkeit von überall und zu jeder Zeit) und realer Veranstaltungen (hoher Erlebenswert und sozialer Austausch) zu verbinden. Dadurch wird es möglich, Erlebnisse ähnlich real-physischer Erfahrungen und sozialer Interaktion unter der Prämisse eines geringen Ressourceneinsatzes von zu Hause aus jederzeit zugänglich zu machen. So ist es beispielsweise möglich, mit Freunden, die in einer anderen Stadt leben, spontan ein Theaterstück am anderen Ende der Welt zu besuchen. Das virtuelle Erlebnis kopiert zwar nicht die Realität, aber die Immersion des virtuellen Raums kann dennoch das Gefühl vermitteln, vor Ort zu sein. Die Illusion der sozialen Präsenz anderer entfernter Personen, vermittelt durch interaktive Avatare, erzeugt das Gefühl eines echten Zusammenseins. Daraus ergeben sich vielschichtige Anwendungsmöglichkeiten sowohl für das hier erprobte Format der Theateraufführungen als auch für andere Bühnenveranstaltungen von Konzerten bis zu TV-Formaten. Das besondere Potenzial dieses Ansatzes besteht dabei in der Adressierung breiter Zielgruppensegmente. Denn nicht nur für mobil oder anderweitig eingeschränkte Personen bietet das Konzept viele Vorteile, sondern auch für einen breitenwirksamen Einsatz über alle Alters- und Nutzungsgruppen hinweg. Für die Betreiber von entsprechenden Kulturformaten bietet dieser Ansatz wiederum neue Möglichkeiten, ihre Inhalte ortsunabhängig zugänglich zu machen.

Auf Basis der vorgestellten Arbeiten konnte gezeigt werden, dass mit einem technologisch niederschwelligen Ansatz reale Bühnenveranstaltungen ortsunabhängig in virtuellen Räumen übertragen werden können und dabei eine kulturelle und soziale Teilhabe ähnlich realer Veranstaltungen erzeugt werden kann.

Ausschlaggebend dafür sind insbesondere immersiv gestaltete virtuelle Umgebungen und gute integrierte Bühneninhalte sowie soziale Interaktionsfunktionen, die räumliches und soziales Präsenzerleben erzeugen. Die 3D-Fähigkeit der Bühnenaufnahmen ist dabei nicht der entscheidende Faktor. So wurde beispielsweise deutlich, dass aufgrund der Entfernung zwischen Zuschauenden und der Bühne sowie des sitzenden Konsums über Parallaxe erzeugte Tiefeneffekte nur begrenzt wahrgenommen werden. Dennoch trägt ein stereoskopischer Effekt der Videoinhalte positiv zur Passfähigkeit in der virtuellen 3D-Umgebung bei.

Die Nutzung der mit VR verbundenen Potenziale für derartige virtuelle Veranstaltungsformate hängt zukünftig stark von der Verbreitung von VR-fähigen Endgeräten in der Bevölkerung ab. Hemmnisse hierbei sind insbesondere noch die Neuheit immersiver Anwendungen sowie schlechte ergonomische Eigenschaften entsprechender VR-Brillen. Außerdem ist offen, welche Betreiberkonzepte und Geschäftsmodelle derartige Angebote in die Breite bringen können.

Förderhinweis

Das Projekt SocialSTAGE-VR wurde vom Bundesministerium für Bildung und Forschung (BMBF)[2] gefördert.

Literatur

Chandrasekera, Tilanka, Fernando, Klinkini & Puig, Luis (2019). Effect of Degrees of Freedom on the Sense of Presence Generated by Virtual Reality (VR) Head-Mounted Display Systems: A Case Study on the Use of VR in Early Design Studios. In: *Journal of Educational Technology Systems*, Vol. 47(4), Sage Publishing, Los Angeles, S. 513–522. https://doi.org/10.1177/0047239518824862

Debarba, Henrique Galvan (2022). Content format and quality of experience in virtual reality. In: *Multimedia Tools and Applications*, Vol. 83, Springer Nature Ag & Co, KGaA, Berlin, S. 46481–46506. https://doi.org/10.1007/s11042-022-12176-9

Deutsch, Werner et al. (2009). Flow erleben und Flow erforschen. In: Bundesministerium für Bildung und Forschung (Hg.), *Pauken mit Trompeten: Lassen sich Lernstrategien, Lernmotivation und soziale Kompetenzen durch Musikunterricht fördern?*, Bildungsforschung Band 32, Bundesministerium für Bildung und Forschung, Bonn, Berlin, S. 71–87.

2 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Deutsches Institut für Normung (2019). DIN EN ISO 9241–210 Ergonomie der Mensch-System-Interaktion—Teil 210: Menschzentrierte Gestaltung Interaktiver Systeme, Beuth Verlag GmbH, Berlin.

Dollase, Rainer (1997). Musikpräferenzen und Musikgeschmack Jugendlicher. In: Baacke, Dieter (Hg.), *Handbuch Jugend und Musik*, VS Verlag für Sozialwissenschaften, Wiesbaden, S. 341–368.

Geigel, J. (2018). Creating a Theatrical Experience on a Virtual Stage. In: Cheok, Adrian David, Inami, Masahiko, Romão, Teresa (Hg.), *Advances in Computer Entertainment Technology,* ACE 2017, London, Lecture Notes in Computer Science, Vol. 10714, Springer, Cham. https://doi.org/10.1007/978-3-319-76270-8_49

Maier, Günter W. & Nissen, Regina (2018). Flow-Erleben. In: *Gabler Wirtschaftslexikon*. Springer, Wiesbaden. https://wirtschaftslexikon.gabler.de/definition/flow-erleben-35011/version-258501 [Zugriff 20.03.2025]

Matzen, Kevin et al. (2017). Low-Cost 360 Stereo Photography and Video Capture. In: *ACM Transactions on Graphics*, Vol. 36(4), Association for Computing Machinery, New York, S. 1–12. https://doi.org/10.1145/3072959.3073645.

Oh, Catherine S., Bailenson, Jeremy N. & Welch, Gregory F. (2018). A Systematic Review of Social Presence: Definition, Antecedents, and Implications. SYSTEMATIC REVIEW article. In: *Frontiers in Robot and AI,* Vol. 5, Frontiers Media S.A., Lausanne. https://doi.org/10.3389/frobt.2018.00114

Rudolf, Beate (1995). Teilhabe als Menschenrecht. In: Diehl, Elke (Hg.), *Teilhabe für alle?! Lebensrealitäten zwischen Diskriminierung und Partizipation,* bpb Bundeszentrale für politische Bildung, Bonn.

Schneider, Albrecht, Dammann, Lars & Kleist, Florian (2008). Live-Musik-Publikum in Hamburg. Empirische Studien zu einer urbanen Musik-Szene im digitalen Zeitalter, LIT Verlag, Münster.

Skarbez, Richard, Brooks, Frederick P., & Whitton, Mary C. (2017). A Survey of Presence and Related Concepts. In: *ACM Computing Surveys*, Vol. 50(6), Article 96, Association for Computing Machinery, S. 1–39. https://doi.org/10.1145/3134301

Statista Research Department (2024). *Anzahl der Personen in Deutschland, die Musikveranstaltungen oder Konzerte besuchen, nach Häufigkeit von 2019 bis 2024.* https://de.statista.com/statistik/daten/studie/171171/umfrage/besuchshaeufigkeit-musikveranstaltungen-konzerte/ [Zugriff 01.12.2024]

Weissig, Christian et al. (2012). The Ultimate Immersive Experience: Panoramic 3D Video Acquisition. In: *International Conference on Multimedia Modeling Advances in Multimedia Msodeling*, Klagenfurt, Lecture of Notes in Computer Schience, Vol. 7131, Springer Science+Business Media, Berlin/Heidelberg, S. 671–681. https://doi.org/10.1007/978-3-642-27355-1_72.

Wilkinson, Michael, Brantley, Sean & Feng, Jing (2021). A Mini Review of Presence and Immersion in Virtual Reality. In: *Proceedings of the 2021 HFES 65th International Annual Meeting*, Vol. 65(1), Human Factors and Ergonomics Society, Washington D.C., S. 1099–1103. https://doi.org/10.1177/1071181321651148

Winkler, Alexander, Won, Jungdam & Ye, Yuting (2022). QuestSim: Human Motion Tracking from Sparse Sensors with Simulated Avatars. In: *SIGGRAPH Asia 2022 Conference Papers*, Daegu, Association for Computing Machinery, New York, S. 1–8. https://doi.org/10.1145/3550469.3555411.

Xia, Shihong et al. (2017). A Survey on Human Performance Capture and Animation. In: *Journal of Computer Science and Technology*, Vol. 32(3), S. 536–554. https://doi.org/10.1007/s11390-017-1742-y.

XR-Technologie für die partizipative Stadtentwicklung – XR-Part: Potenziale und Herausforderungen am Beispiel der Städte Mannheim und Rostock

Svenja Rogoll, Heidi Sinning und Anja Wolter[1]

Zusammenfassung

Partizipative Verfahren finden in der kommunalen Stadtentwicklung immer häufiger Anwendung. Durch ihren Einsatz erhofft man sich mehr Akzeptanz und eine Legitimation von stadtplanerischen Maßnahmen sowie eine Qualifizierung der mit ihnen verbundenen Planungs- und Entscheidungsprozesse. Partizipative Verfahren sind jedoch mit Herausforderungen konfrontiert, welche in diesem Beitrag identifiziert werden. Anschließend wird aufgezeigt, inwieweit ergänzende, auf Extended Reality (XR) gestützte Partizipationsformate Lösungsansätze für diese Herausforderungen bieten, indem sie weitere Optionen demokratischer Teilhabe, eine Verbesserung der Vorstellungskraft, eine kreative, informierte Mitwirkung und eine konkretere Darstellung des Beteiligungsgegenstands ermöglichen. Dies wird durch eine Erörterung der Möglichkeiten zur Anwendung von XR-Technologien in partizipativen Stadtentwicklungsprozessen und der damit verbundenen Potenziale und Herausforderungen anhand der Forschungsergebnisse des Projektes »XR-Part« erreicht. Diese wurden durch einen iterativen Prozess von technischer Entwicklung von XR-Part-Partizipationsformaten, deren Erprobung in Maßnahmen der Stadtentwicklung in Rostock und Mannheim und einer anschließenden Evaluation generiert. Im Hinblick auf die Zukunft der Nutzung von XR-Technologie in partizipativen Stadtentwicklungsprozessen wird festgestellt, dass eine vermehrte Nutzung zu erwarten ist, wenn die identifizierten Barrieren abgebaut werden. Darüber hinaus wird die

1 Folgende Institutionen und Personen sind an dem Verbundforschungsprojekt XR-Part beteiligt: Institut für Stadtforschung, Planung und Kommunikation (ISP) der FH Erfurt: Prof. Dr.-Ing. Heidi Sinning, Svenja Rogoll, Anja Wolter; Fachrichtung Architektur der FH Erfurt: Prof. Yvonne Brandenburger, Sebastian Damek; Fachrichtung Angewandte Informatik der FH Erfurt: Prof. Rolf Kruse, Mohammadreza Maleki Moghadam, Lutz Nagy; TriCAT GmbH: Felix Gaisbauer, Fabian Kersting, Dr. Gregor Mehlmann, Patrick Reipschläger, Lena Schuler; Zebralog GmbH: Juliane Henn.

Notwendigkeit für weitere Forschung identifiziert, insbesondere bezogen auf die sozialen Auswirkungen des Einsatzes von XR-Technologien.

1. Einführung und Einordnung in den Stand der Wissenschaft und der Praxis

Der Beitrag thematisiert die Anwendungsmöglichkeiten von XR-Technologien in partizipativen Stadtentwicklungsprozessen anhand der wissenschaftlichen Erkenntnisse aus dem Verbundforschungsprojekt »XR-Part: XR-Partizipationsräume zur erweiterten sozialen Teilhabe in urbanen Transformationsprozessen«. Im Rahmen dieses Projektes arbeitete ein transdisziplinäres Forschungskonsortium aus Partner:innen der Wissenschaft (Fachhochschule Erfurt: Institut für Stadtforschung, Planung und Kommunikation [ISP], Fachrichtung Angewandte Informatik und Fachrichtung Architektur), der Wirtschaft (TriCAT GmbH und Zebralog GmbH) und der Kommunen (Modellstädte Mannheim und Rostock) zusammen. Darüber hinaus unterstützten die Städte Bonn und Köln als sogenannte »Tandempartner:innen« mit ihrer langjährigen Expertise den Forschungsverbund bei der Definition von Qualitätsstandards der XR-gestützten Partizipation. Gefördert wurde das Projekt durch das Bundesministerium für Bildung und Forschung (BMBF)[2] als Teil der Förderlinie »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft (VAR2)« (https://www.fh-erfurt.de/xr-part).

Das Ziel des Forschungsprojekts, dessen Innovationsansatz in einer integrativen Betrachtung der Mensch-Technik-Interaktion in virtuellen und realen Räumen besteht, war es, innovative XR-Partizipationsräume zu entwickeln, welche den Bewohner:innen eine zeit- und ortsunabhängige Teilnahme in stadtplanerischen Partizipationsprozessen ermöglichen, diese zu erproben und zu evaluieren. Die XR-basierten Beteiligungsformate sollten in crossmediale, partizipative Prozesse integriert werden, um die Qualität der Beteiligungsverfahren in Bezug auf Zugänglichkeit, Zielgruppenerreichbarkeit, Kooperationsfähigkeit und das emotionale Erleben städtischer Planungsprozesse zu verbessern und Kommunikationsbarrieren abzubauen. Dabei wurde der folgenden zentralen Fragestellung nachgegangen: Wie können räumliche und zeitliche Einschränkungen sowie Zugangsschwierigkeiten durch XR-gestützte Partizipationsformate überwunden werden, um Menschen zur digitalen Teilhabe an urbanen Transformationsprozessen zu befähigen? Das Projekt besteht aus den wesentlichen Bausteinen

2 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Technische (Weiter-)Entwicklung des XR-Systems, Erprobung und *Evaluation*. Im vorliegenden Beitrag werden vor allem zentrale Erkenntnisse aus den Erprobungen und den formativen Evaluationen dargestellt (s. a. Rogoll et al., 2024).

Zunächst erfolgt eine Einordnung in den aktuellen Stand der Wissenschaft und der Praxis zu XR-Technologien in der partizipativen Stadtentwicklung. Aufbauend auf den Zielen des Forschungsprojekts XR-Part und des Evaluationskonzepts werden anschließend die entwickelten XR-Part-Beteiligungsformate in ihrer Funktionalität und ihrem methodischen Design erörtert und deren zeitliche und konzeptionelle Einbettung in realen städtischen Entwicklungsprozessen der Modellstädte Mannheim und Rostock erläutert. Im Fokus des Beitrags steht die Diskussion der Potenziale und Herausforderungen der Beteiligungsinnovation für Bürger:innen und Kommunen, insbesondere hinsichtlich der Zielgruppenerreichbarkeit und Zugangsgerechtigkeit, ELS-Implikationen, der Förderung der Vorstellungskraft und der informierten Meinungsbildung, der Akzeptanz von XR-Technologien in partizipativen Stadtentwicklungsprozessen sowie des effizienten Einsatzes personeller und finanzieller Ressourcen. Abschließend werden die Anwendungsperspektiven von XR-Technologien in unterschiedlichen Arten von partizipativen Stadtentwicklungsprozessen, ihre Potenziale für die erweiterte soziale Teilhabe und die Kommunikation von Planungsszenarien dargelegt sowie die wesentlichen Erkenntnisse zusammengefasst.

1.1 XR-Technologien als Lösungsansatz für Herausforderungen der partizipativen Stadtentwicklung

Vor dem Hintergrund gegenwärtiger gesellschaftlicher Entwicklungen, zu denen der Klimawandel, Urbanisierungs- und Gentrifizierungsprozesse sowie die Digitalisierung und Verkehrswende zählen, und der damit einhergehenden komplexen Interessenlagen erlangen partizipative Verfahren in der Stadtentwicklung zunehmend an Bedeutung. Die Einbindung divergierender Perspektiven aus der Bürgerschaft, Politik, Verwaltung und Wirtschaft ermöglicht eine Qualifizierung von Planungs- und Entscheidungsprozessen, die Erhöhung der Akzeptanz sowie der Legitimation von Entscheidungen. Dabei gilt es, die Wünsche, Ideen und Anforderungen der Bürger:innen in Planungen und Konzepten zu berücksichtigen, um eine demokratischere, inklusivere und bedarfsgerechtere Stadtentwicklung zu fördern (vgl. Orthmann, 2017: 4 ff.; Le Blanc, 2020: 4).

Gleichzeitig sehen sich Kommunen bei der Durchführung von Beteiligungsprozessen mit verschiedenen Herausforderungen konfrontiert, die die Zielerreichung einer breiten Beteiligung und gesellschaftlichen Legitimation von Planungsvorhaben erschweren. Beispielsweise ist in Stadtplanungsprozes-

sen mit Öffentlichkeitsbeteiligung das sogenannte »Beteiligungsparadoxon« zu beobachten. Damit wird beschrieben, dass in frühen Planungsphasen, in denen noch ausreichend Entscheidungsspielraum besteht, eine geringe Beteiligung der Bürger:innen zu verzeichnen ist, während in späteren Phasen ein großes öffentliches Interesse und Mitbestimmungswille besteht, obwohl der Entscheidungsspiel- und -zeitraum bereits ausgeschöpft ist (Senatsverwaltung für Stadtentwicklung und Umwelt Berlin, 2012: 82; Hirscher, 2017: 323 f.).

Eine weitere Herausforderung besteht darin, dass bestimmte Zielgruppen nur schwer erreichbar sind, während andere Bevölkerungsgruppen in höherem Maße in der Lage sind, sich zu beteiligen (vgl. Orthmann, 2017: 12, 15). Beispielsweise beteiligen sich Menschen mittleren Alters, Menschen mit einem höheren Einkommen sowie Menschen, die in Teilzeitmodellen arbeiten, häufiger als die jüngeren und älteren Zielgruppen, Menschen mit niedrigerem Einkommen oder Menschen, die eine höhere Wochenarbeitszeit haben (vgl. Orthmann, 2017: 42; Hübler, 2024: 103). Dadurch wird es erschwert, die zielgruppenspezifischen Anforderungen und Interessen unterrepräsentierter Bevölkerungsgruppen in Planungsvorhaben zu berücksichtigen.

Schließlich erfordert die Beteiligung von Bürger:innen zunächst ein gemeinsames Verständnis der Planungsgegenstände. Da jedoch unterschiedliche Wissensstände und Expertisen der Beteiligten zu erwarten sind, entstehen oft verschiedene Interpretationen der vorgestellten Ideen (vgl. Rockmann et al., 2015). Ohne klare Kommunikation lassen sich Missverständnisse schwer vermeiden, was den Erfolg der Partizipation beeinträchtigen kann.

Seit einigen Jahren wird in Wissenschaft und Praxis diskutiert, inwiefern XR-gestützte Beteiligungsformate Lösungsansätze für die dargestellten Herausforderungen der Partizipation in der Stadtplanung bieten können. In Bezug auf die genannten Herausforderungen lassen sich aus den bisherigen Untersuchungen Potenziale, die mit XR-Technologien verbunden werden, identifizieren. Um dem Beteiligungsparadoxon und seinen Konsequenzen entgegenzuwirken, bieten laut Wolf et al. (2020: 125 f.) XR-gestützte Formate der E-Partizipation das Potenzial, die Verständlichkeit, Nachvollziehbarkeit, Kollaboration und Interaktion der Beteiligten in Planungsprozessen zu fördern und zu erhöhen.

Zudem können XR-Technologien neue Zugänge für bestimmte »stille« oder unterrepräsentierte Gruppen zu Beteiligungsangeboten zu schaffen, beispielsweise indem orts- und zeitunabhängige Beteiligungsformate geschaffen werden oder indem eine aufsuchende Beteiligung mit Augmented Reality (AR, »erweiterte Realität«) in den Quartieren stattfindet. Durch die Nutzung von digitalen Räumen für Beteiligung können außerdem jüngere Altersgruppen angesprochen werden, die einen höheren Digitalisierungsgrad und stärkere digitale Kompetenzen aufweisen (vgl. Initiative D21, o. J.: 18). Allerdings gibt es dabei auch Limi-

tierungen – »stille« Gruppen wie Ältere, Menschen mit geringerem Bildungsabschluss oder niedrigerem Einkommen weisen häufig einen geringeren Digitalisierungsgrad auf (vgl. Initiative D21, o. J.: 18 f.), sodass für sie auch bei der digitalgestützten Beteiligung Zutrittsbarrieren bestehen (vgl. Schöffmann, 2020: 43 f.).

Visualisierungen und insbesondere dreidimensionale Darstellungen sind ein wichtiges kommunikatives Instrument in der partizipativen Planung, um Planungsszenarien und deren Raumwirkung für Bürger:innen leichter verständlich und begreifbar zu machen (vgl. Brettschneider et al., 2017: 14; Schauppenlehner et al., 2018: 15). Studien belegen zudem, dass Technologien wie AR und Virtual Reality (VR, »virtuelle Realität«) partizipative Prozesse unterstützen können, indem sie den Interpretationsspielraum der Beteiligten reduzieren (vgl. Goudarznia et al., 2017). XR-Technologien bieten über die reinen Informationen hinaus große Potenziale für tatsächliche Beteiligungsverfahren, in denen Bürger:innen angehört werden, aktiv mitgestalten oder auch mitentscheiden können; dabei können XR-gestützte Formate an den Beteiligungsgrad angepasst und geeignete Beteiligungsmethoden eingesetzt werden (vgl. Brettschneider et al., 2017: 14; Hajrasouliha 2024: 560 f.). Simonofski et al. (2024: 4) stellten in ihrem Übersichtsartikel zum Einsatz von XR-Technologien in der Stadtplanung heraus, dass XR-Formate inzwischen deutlich häufiger zur Bürgerkonsultation als lediglich zur Information eingesetzt werden.

1.2 Anwendungen von XR-Technologien in der partizipativen Praxis

Aufgrund ihrer Potenziale für eine erweiterte gesellschaftliche Teilhabe und visuelle Kommunikation werden XR-Technologien zunehmend in partizipativen Prozessen genutzt, wenngleich ihre Verbreitung noch begrenzt ist. Vor allem im Forschungskontext konnten in der jüngsten Vergangenheit Pilotprojekte der Bürgerbeteiligung mit XR-Technologie realisiert werden (vgl. Simonofski, 2024: 4). Im deutschsprachigen Raum sind diesbezüglich Forschungsprojekte, wie beispielsweise Virtus, INSPIRER, GLARA, VR-Planning we're planning, PaKomm, Take Part und XR-Part zu nennen (vgl. Bundesministerium für Bildung und Forschung, o. J.; Virtual Dimension Center Fellbach, o. J.; GLARA Forschungskonsortium, 2021; Austrian Institute of Technology GmbH, o. J.; HafenCity University Hamburg, o. J.; Fegert et. al., 2021; Sinning et al., 2023).

Praktische Erprobungen zur Bürgerbeteiligung in der Stadtentwicklung sind vor allem mit AR-Technologien in verschiedenen Handlungsfeldern der Stadtentwicklung zu finden. Bei Formaten der Beteiligung mit AR können die Nutzer:innen vor Ort meist durch ein Tablet oder Smartphone als Endgerät oder durch spezielle AR-Brillen 3D-Objekte betrachten, die sich in die reale Umgebung einfügen.

Auf diese Weise konnten beispielsweise Anwohner:innen in der Bernhardgasse in Wien virtuelle Bäume im Straßenraum platzieren und über eine überlagernde Heatmap die veränderte Oberflächentemperatur sehen (siehe ÖGUT, 2020). Neben dem Anwendungsbereich der Mobilitätsplanung, sprich der Visualisierung von Straßenumgestaltungen und Parkraumkonzepten, kommt AR-Technologie auch bei der Umgestaltung von Quartiersplätzen, wie beispielsweise in der Stadt Essen (Neugestaltung Weberplatz), zum Einsatz.

Bei Beteiligungen im virtuellen Raum (Metaverse) oder mit Unterstützung von VR-Technologie tauchen die Nutzer:innen hingegen mittels einer VR-Brille vollständig in eine computergenerierte, dreidimensionale Umgebung ein (vgl. Dörner et al., 2019). Bürger:innen können in VR-Umgebungen beispielsweise alternative Szenarien erkunden, diese bewerten und kommentieren. Exemplarisch seien hier die zum Teil bereits abgeschlossenen Forschungsprojekte Virtus, INSPIRER oder VR-Planning genannt, welche in transdisziplinären Teams u. a. immersive virtuelle Umgebungen erschaffen, in denen Bürger:innen Entwicklungsvorhaben erleben und selbst gestalterisch aktiv werden können. Beteiligungsformate im virtuellen Raum bzw. Metaverse bieten darüber hinaus den Vorteil, dass verschiedene Akteur:innen kollaborativ miteinander in den Austausch treten können. So konnten zuletzt beispielsweise Vertreter:innen der Stadt Aalen und eines Planungsbüros bei einem virtuellen Bürgerdialog im Metaverse Bürger:innen im virtuellen Raum über die Entwicklung des örtlichen Kombibads informieren und gemeinsam darüber beraten (vgl. Presse- und Informationsamt Stadt Aalen, o. J.). Nach wie vor stellen sich in der Wissenschaft und Praxis im Zusammenhang mit den bisher vorliegenden Erfahrungen und Erkenntnissen zum Einsatz von XR-Technologien in der partizipativen Stadtentwicklung die Fragen, wie XR-Technologien zur Förderung der Bürgerbeteiligung konkret in realen Planungsprozessen effektiv eingesetzt werden können, welche Potenziale vor allem zum Tragen kommen und genutzt werden können und welche Herausforderungen sich aktuell (noch) stellen. Der vorliegende Beitrag geht vor dem Hintergrund der bisherigen Erkenntnisse diesen Fragen nach und reflektiert die Erkenntnisse aus dem Forschungsprojekt XR-Part.

2. Evaluationskonzept der XR-Part-Beteiligungsformate

Für die innovativen XR-basierten Partizipationsformate wurde im XR-Part-Projekt ein Evaluationskonzept entwickelt, welches zum einen die praktische Umsetzbarkeit dieser Formate in kommunalen Bürgerbeteiligungsprozessen in der Stadtplanung gewährleisten sollte und zum anderen ein Untersuchungsdesign zu den Auswirkungen der Innovationen auf die Chancen zur sozialen

Teilhabe darstellte. Die Evaluation bezieht sich auf das entwickelte XR-System, welches aus einem AR-Beteiligungsformat (XR-Part-Beteiligungstour) und einem virtuellen 3D-Begegnungs- und Erlebnisraum (XR-Part-Beteiligungsraum) besteht, sowie auf die Prozessinnovationen hinsichtlich der Integration der XR-Part-Beteiligungsformate in den kommunikativen Gesamtprozess. Die konzeptionell entwickelten XR-gestützten Beteiligungsformate wurden mehrfach in realen Partizipationsprozessen der Modellstädte Mannheim und Rostock mit Bürger:innen und Verwaltungsmitarbeiter:innen getestet. Die prozessuale sowie technische Entwicklung erfolgte nach einem iterativen Ansatz, bei dem die Phasen der Konzeption, Entwicklung, Erprobung und Evaluation wiederholt durchlaufen wurden. Die im Mixed-Method-Design verwendeten Methoden zur Evaluation umfassten Online-Befragungen der Nutzer:innen, teilnehmende Beobachtungen sowie leitfadengestützte Interviews mit den verschiedenen beteiligten Akteursgruppen.

Mithilfe der teilnehmenden Beobachtung (vgl. Thierbach & Petschick, 2022: 1564 f.) wurde das Verhalten der Teilnehmer:innen im Umgang mit der AR-Anwendung als auch ihr Verhalten im virtuellen XR-Part-Beteiligungsraum dokumentiert. Zudem wurden auftretende Schwierigkeiten, Gefahrensituationen und Lerneffekte der Nutzer:innen festgehalten.

Die XR-Part-Beteiligungsformate wurden des Weiteren durch eine standardisierte Online-Befragung evaluiert (vgl. Wagner & Hering, 2014: 661 ff.). Die dazu entwickelten Fragebögen füllten die Teilnehmer:innen im Anschluss an die XR-Part-Beteiligungstour bzw. die moderierte Veranstaltung im XR-Part-Beteiligungsraum aus. Durch offene und geschlossene Fragen konnten in der Nutzerbefragung umfangreichere Hintergründe und Kontexte zur Perspektive der Teilnehmer:innen ermittelt werden. Die insgesamt sechs thematischen Abschnitte des Fragebogens umfasste die Erhebung soziodemografischer Daten, Angaben zur technischen Affinität und Beteiligungserfahrung sowie Fragen zur Wahrnehmung der immersiven 3D-Visualisierungen und Akzeptanz der XR-Formate.

Ergänzende, leitfadengestützte Interviews ermöglichten spezifische, thematisch vertiefte Einblicke in Chancen, Herausforderungen und Verbesserungspotenziale aus den verschiedenen Perspektiven von Bürger:innen, Moderator:innen, städtischen Vertreter:innen und des personellen technischen Supports.

3. Erprobung der XR-Part-Beteiligungsformate

Im Folgenden werden zunächst die im Projekt entwickelten XR-Part-Beteiligungsformate in ihrer Funktionalität und ihrem methodischen Design erläutert

und deren zeitliche und konzeptionelle Einbettung in reale städtische Entwicklungsprozesse erörtert, bevor im Anschluss Potenziale und Herausforderungen der prozessualen Umsetzung von XR-Beteiligungsformaten anhand der Evaluationsergebnisse dargestellt werden.

3.1 Die XR-Part-Beteiligungsformate

3.1.1 XR-Part-Beteiligungstour

Die XR-Part-Beteiligungstour stellt eine AR-Anwendung dar, bei der an festgelegten Stationen mithilfe eines Tablets dreidimensionale Visualisierungen von Nutzungs- und Entwicklungsmöglichkeiten im realen Raum angezeigt werden, Bürger:innen diese kommentieren sowie selbst Gestaltungsobjekte aus einem Katalog platzieren können (siehe auch der nachfolgende Beitrag dieses Bandes »Nutzerzentrierte Gestaltung von XR-Technologien für die Transformation urbaner Partizipationsprozesse« mit ausführlicher Beschreibung von Umsetzung, Onboarding und Funktionalitäten der AR-Anwendung).



Abbildung 1: Anwendung der AR-Technologie bei der XR-Part-Beteiligungstour in Mannheim und Rostock
Quelle: ISP, TriCAT, 2022 und 2024

Bei den Erprobungen des Formats der XR-Part-Beteiligungstour wurde im Untersuchungsgebiet der betreffenden Modellstadt ein Informationsstand aufgebaut, an dem Vertreter:innen der Modellstädte sowie des XR-Part-Teams das Gespräch mit interessierten Bürger:innen suchten und sie zur Beteiligung motivierten. Um eine niedrigschwellige Zugänglichkeit zu gewährleisten, wurden den Teilnehmer:innen Tablets als Endgeräte ausgehändigt, um die 15- bis 20-minütige Tour zu absolvieren (siehe Abbildung 1). Personen des XR-Part-Teams standen als Support für Fragen zur Seite und unterstützten im Falle auftretender technischer Schwierigkeiten. Vertreter:innen der Modellstädte, darunter die Leiter:innen betreffender Fachämter, standen zudem für Fragen zum Beteiligungsgegenstand bereit.

3.1.2 XR-Part-Beteiligungsraum

Der XR-Part-Beteiligungsraum stellt einen virtuellen Begegnungsort (Metaverse) dar, welcher für die Umsetzung diverser digitaler Beteiligungsformate, wie beispielsweise Werkstattgespräche oder Dialogformate, genutzt werden kann. In diesem virtuellen Raum treffen sich die Bürger:innen, die Vertreter:innen der Stadtverwaltung sowie die Moderator:innen in Form von Avataren, also virtuellen Abbildern, um in gemeinschaftlichem Austausch an Themen des Planungsvorhabens zu arbeiten.

Neben verschieden großen und gestalteten Teilräumen für Gruppenarbeiten, Präsentationen und Diskursrunden liegt die Innovation vor allem in den integrierten virtuellen 3D-Modellen des betreffenden Planungsgebiets, die zum einen im Maßstab eines Tischmodells (siehe Abbildung 2) und zum anderen als begehbares Modell im Maßstab 1:1 (siehe Abbildung 3) zur kollaborativen Arbeit genutzt werden können.

Mithilfe eines 3D-Objektkatalogs und interaktiver Medienwände können in der virtuellen Umgebung und der 3D-Nachbildung des Planungsgebiets in spielerischer Weise Methoden wie Karten- und Punktabfragen, Mapping- und Baukasten-Methoden oder auch klassische Beteiligungsformate wie ein World-Café oder eine Fishbowl umgesetzt werden.

Die Teilnahme war für alle Interessent:innen unabhängig von ihrem Aufenthaltsort unter der Voraussetzung möglich, dass sie an dem betreffenden Ort über die nötige Ausstattung mit PC, Headset und Internetverbindung verfügten.

Der virtuelle XR-Part-Beteiligungsraum kann zum einen als Desktop-Anwendung heruntergeladen werden oder der Zugriff erfolgt über einen Internetbrowser. PC-Anwendungen sind derzeit für Nutzer:innen gut zugänglich, während die Nutzung einer VR-Brille die Verfügbarkeit dieser Technologie erfordert, was aktuell noch eine größere Hürde darstellt.



Abbildung 2: 3D-Tischmodell im XR-Part-Beteiligungsraum
Quelle: ISP, TriCAT, 2024



Abbildung 3: Begehbares Modell im XR-Part-Beteiligungsraum
Quelle: ISP, TriCAT, 2024

Des Weiteren besteht die Option, die Räumlichkeiten bei Bedarf dauerhaft zugänglich zu machen, sodass Bürger:innen beispielsweise 3D-Visualisierungen in einem interaktiven Ausstellungsformat ohne zeitliche Begrenzung erleben können.

3.2 XR-Beteiligungsformate in partizipativen Prozessen der Modellkommunen Mannheim und Rostock

3.2.1 Zeitliche Einbettung der Erprobungen

Die Erprobungen der XR-Part-Beteiligungsformate konnten in den zweijährigen Beteiligungsprozessen der Modellstädte zum einen in sehr frühen Prozessphasen stattfinden, in denen zunächst Bestandsanalysen durchgeführt und Anforderungsbedarfe ermittelt wurden sowie eine gemeinsame Auseinandersetzung mit Zukunftsleitbildern erfolgte. Zum anderen konnte eine zweite Erprobungsphase in einer fortgeschrittenen Phase durchgeführt werden, in der entwickelte Entwurfskonzepte und Gestaltungsideen bewertet, diskutiert und abgewogen wurden. Die Einbettung der XR-Part-Beteiligungsformate in diese zeitlich distanzierten Projektphasen brachte bedeutende Erkenntnisse, sowohl für die prozessuale Entwicklung als auch für die Ableitung von Empfehlungen für die Praxis (s. a. Rogoll et al., 2024). In Mannheim wie auch in Rostock wurden die XR-Part-Formate stets zeitlich an analoge Beteiligungsveranstaltungen im Prozess angegliedert, um diese zu ergänzen.

3.2.2 Partizipativer Planungsprozess in Mannheim

Gegenstand der Bürgerbeteiligung in der Modellstadt Mannheim war die Umgestaltung des Platzes vor der Uhlandschule im innenstadtnahen Quartier Neckarstadt-Ost. Dieser wurde zum Zeitpunkt des Beteiligungsprozesses als Parkplatz genutzt und bot eine geringe Aufenthaltsqualität. Das Ziel der Planung bestand in der weiteren Aufwertung des Platzes als zentraler Treffpunkt im Quartier. Die übergeordnete Fragestellung des Beteiligungsprozesses lautete: »Welche Funktionen soll der Vorplatz der Uhlandschule für das Quartier erfüllen?« (Stadt Mannheim, 2022a: 2).

Als Grundlage für die Diskussion zur Nutzung des Platzes wurden zunächst verschiedene Flächenanteile für gemeinschaftliche Nutzungen und Aneignungen visualisiert. Die unterschiedlich großen, den Raum einnehmenden Flächen fungierten dabei als visuelle Platzhalter für verschiedene Nutzungsarten, welche der Nutzung als Stellfläche für Pkws entgegenstehen (siehe Abbildung 4). Während

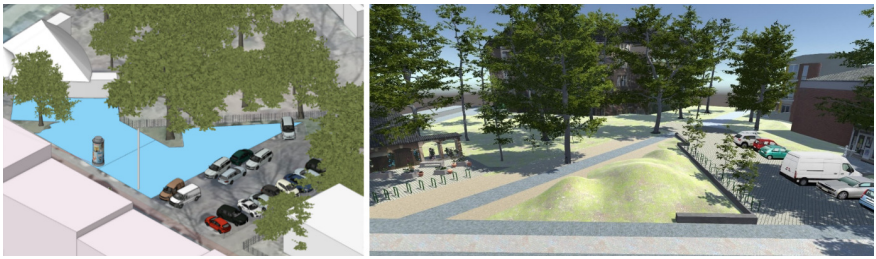


Abbildung 4: Gegenüberstellung der 3D-Visualisierungen in der ersten (links) und zweiten Erprobungsphase (rechts)
Quelle: ISP, TriCAT, 2024

des Workshops im virtuellen Raum bestand zudem die Möglichkeit, parkende Autos schrittweise auszublenden. Neben Fragen zur Wahrnehmung und Nutzung des Platzes wurden auch die klimafreundliche Gestaltung sowie die Wegeführung für Fußgänger:innen und Radfahrer:innen thematisiert.

Nach einer von der Stadt Mannheim anschließend durchgeführten Machbarkeitsstudie, wurden die daraus entstandenen Gestaltungsentwürfe in 3D-Visualisierungen übersetzt und mittels der XR-Formate präsentiert und diskutiert. In der zweiten Phase, die auf der allgemeinen Bedarfsabfrage für bestimmte Nutzungen und entsprechende Flächenbedarfe basierte, stand die spezifische Platzierung von Funktionen, z. B. von Spiel- und Sportbereichen, Treffpunkten und freien Bewegungsflächen sowie von Ausstattungselementen wie Bepflanzungselementen, Stadtmöbeln oder Ladesäulen, im Vordergrund (siehe Abbildung 4).

Die Zielgruppen, welche im Fokus der Beteiligung zur Neugestaltung des Platzes vor der Uhlandschule stehen, sind Anwohner:innen, Schüler:innen, Lehrer:innen und Eltern als Stakeholder:innen der anliegenden sozialen und wirtschaftlich orientierten Einrichtungen als auch die kommunalpolitische Ebene. 51,2 % der Anwohner:innen in der Neckarstadt-Ost sind zwischen 35 und 55 Jahre alt. 12,8 % der Anwohner:innen sind unter 18 Jahre alt und 17,8 % gehören zu der Gruppe der Älteren (65+) (vgl. Stadt Mannheim, 2022b).

3.2.3 Partizipativer Planungsprozess in Rostock

In der Modellstadt Rostock handelte es sich beim Gegenstand der Beteiligung um die Entwicklung eines ganzen Stadtgebiets, der Südstadt, für die ein Rahmenplan aufgestellt wurde. Rahmenpläne sind informelle Stadtentwicklungsinstrumente, die die Entwicklungsziele einer Stadt oder eines Stadtteils definieren und anschließend in der verbindlichen Bauleitplanung umgesetzt werden (vgl. Pahl-

Weber, 2018: 2410). Die Ergebnisse des Beteiligungsprozesses flossen in die Erarbeitung dieses Rahmenplans ein.

Der Entwicklungsbedarf ergibt sich aus vielfältigen Herausforderungen, wie dem Klimawandel, dem demografischen Wandel, sich veränderten Bewohnerzahlen sowie dem Wandel der Bewohnerbedürfnisse. Im Rahmen der Erprobung der XR-Beteiligungsformate wurde ein Schwerpunktgebiet in der Südstadt festgelegt, in dem alle wesentlichen Entwicklungsthemen des Rahmenplans behandelt wurden. Die Intention der Initiator:innen war es, verschiedene Entwicklungsmöglichkeiten für die Südstadt aufzuzeigen und diese mit den betroffenen Zielgruppen zu diskutieren. Unter der Leitfrage »Unter welchen Rahmenbedingungen kann in der Südstadt eine Weiter- und Innenentwicklung des Stadt- und Lebensraums stattfinden?« war es Ziel, mit der Bürgerschaft konkrete Kriterien für die Weiterentwicklung der Südstadt in allen relevanten Bereichen zu entwickeln.

In der ersten Erprobungsphase wurden ähnlich wie in der Stadt Mannheim zunächst verschiedene Bauvolumina als farbige dreidimensionale Geometrien visualisiert. Die stark abstrahierten Modelle zeigten u. a. drei Möglichkeiten zur Erweiterung der bestehenden Wohnbebauung auf. Mittels entwickelter 3D-Piktogramme konnten die Teilnehmer:innen zudem Aneignungsbedarfe auf den wohnraumnahen Flächen verorten, darunter beispielsweise Bereiche zum Spielen, Treffen oder Erholen. In der zweiten Erprobungsphase wurden die abstrakten 3D-Piktogramme durch detaillierte und realistische Gestaltungsobjekte (Bäume, Bänke, Spiel- und Sportgeräte) ersetzt, welche zur freien Gestaltung der Freiräume platziert und kommentiert werden konnten (siehe Abbildung 5). Auch die Varianten der Wohnraumerweiterung und der Unterbringung des ruhenden Verkehrs wurden mit realistischen Texturen visualisiert.

Zu den adressierten Interessengruppen des Beteiligungsprozesses in der Rostocker Südstadt zählen neben der vielfältigen Gruppe der Bürger:innen zudem ansässige Vereine und Initiativen – u. a. ein Kleingartenverein, das Stadtteilbegegnungszentrum, Vertreter:innen der Wohnungs- und Bauwirtschaft, Schulen sowie größere Bildungsstätten. Die Altersstruktur des Stadtteils ist durch einen signifikanten Anteil an Menschen im Alter von 15 bis 35 Jahren (30,6 %) gekennzeichnet, was in erster Linie auf die nahegelegene Universität zurückzuführen ist. Des Weiteren ist der Anteil der Personen ab 75 Jahren im Vergleich ebenfalls relativ hoch. Das Durchschnittsalter der Südstädter von 47,7 Jahren verdeutlicht zudem, dass es sich um einen Ortsteil im demografischen Wandel handelt, der zunehmend von einer Überalterung geprägt sein wird (vgl. Hanse- und Universitätsstadt Rostock, 2023: 251, 256).



Abbildung 5: Visualisierung zur Unterbringung des ruhenden Verkehrs in der Rostocker Südstadt
Quelle: ISP, TriCAT, 2024

4. Potenziale und Herausforderungen der prozessualen Umsetzung von XR-Beteiligungsformaten

Partizipative Stadtplanung und Stadtentwicklung stehen angesichts der zunehmenden Etablierung von XR-Technologien vor der Herausforderung, diese in die bestehende Beteiligungspraxis zu integrieren (vgl. Sinning et al., 2023). Im Folgenden werden Herausforderungen und Lösungsansätze beschrieben, die sich im Verlauf der Erprobungen im Projekt XR-Part ergaben. Aus den Evaluationsergebnissen werden zudem Potenziale abgeleitet, welche durch die XR-Beteiligungsformate für die Praxis ausgeschöpft werden können.

4.1 Zielgruppenerreichbarkeit und Zugangsgerechtigkeit

Wie eingangs beschrieben, zielte die Entwicklung der XR-Part-Beteiligungsformate u. a. auf die Schaffung einer neuen Qualität der Zugänglichkeit städtischer Planungsprozesse, indem die Beteiligungsfähigkeit der Bürger:innen verstärkt wird. Die Evaluation der insgesamt vier Erprobungen zeigte, dass zwar *neue Zielgruppen, die bisher noch nicht an Beteiligungsprozessen der Städte mitgewirkt haben*, erreicht wurden. Jedoch stellt sich die Zugänglichkeit bzw. Zugangsgerechtigkeit der beiden Formate unterschiedlich dar, was bereits die Teilnahmestatistik aufzeigt. Während bei den Erprobungen der XR-Part-Beteiligungstour (AR-Anwendung) vor Ort insgesamt 211 Personen teilnahmen, erschienen zu den mo-

derierten Beteiligungsveranstaltungen im virtuellen XR-Part-Beteiligungsraum 32 Bürger:innen.

Vor allem das Format der XR-Part-Beteiligungstour motivierte durch den aufsuchenden Charakter Bürger:innen ohne Beteiligungserfahrung zur Teilnahme. Im Rahmen der Befragung gaben 55,2 % der 143 der Teilnehmenden an, zuvor noch nie ein Beteiligungsangebot der Modellstädte wahrgenommen zu haben. Ebenfalls zeigen die Befragungsergebnisse ein breites Altersspektrum der Teilnehmer:innen auf, da alle Alterskategorien von 18 bis über 65 Jahren[3] vertreten waren. Während bei der Erprobung der AR-Anwendung in Mannheim mehr Personen mittleren Alters zwischen 35 bis 55 Jahren vertreten waren, nahmen in Rostock entsprechend der Einwohnerstatistik des Stadtteils auch vermehrt Senior:innen ab 65 Jahren teil. Der Vergleich der erhobenen demografischen Angaben mit städtischen Bewohnerdaten zeigt in beiden Modellstädten, dass sich diese stark annähern. Für eine umfassendere Darstellung der hier beschriebenen und darüber hinausgehend erhobenen Daten sei an dieser Stelle auf den Evaluationsbericht des Projekts verwiesen (vgl. Rogoll et al., 2024).

Entsprechend der Demografie der Teilnehmer:innen zeigten sich bei den Erprobungen beider Formate zudem starke *Unterschiede in der technischen Affinität*. Den Beobachtungsdaten als auch den Selbsteinschätzungen der Befragten zufolge verfügte rund ein Drittel zunächst über geringe Kompetenzen im Umgang mit Technologien. Eine Herausforderung in der Konzeption der XR-Part-Beteiligungsformate besteht darin, diesem differenten Vorwissen mit einem angemessenen Onboarding zu begegnen. Das *Onboarding* erfolgte zunächst als *Tutorial* innerhalb der Anwendung, mit der Zielsetzung, die Steuerung, den Ablauf sowie die zentralen Funktionen der Anwendung zu erläutern. Bei der Entwicklung eines intuitiven Onboarding-Verfahrens wurde angestrebt, die Informationen leicht verständlich und mit wenigen Textinhalten zu vermitteln. Dabei sollte das Tutorial zudem adaptiv gestaltet sein, sodass Personen, die bereits im Umgang mit der Anwendung geübt sind, dieses bei Bedarf verkürzen oder überspringen konnten. In den Erprobungen der beiden entwickelten XR-Formate zeigte sich, dass ein ergänzender *personeller Support* in der praktischen Umsetzung von hoher Bedeutung ist, um den Unterstützungsbedürfnissen der Nutzer:innen Rechnung zu tragen. Sowohl im virtuellen Raum als auch bei der AR-Anwendung wurde durch das XR-Part-Team eine personelle Einführung und Unterstützung angeboten, welche die Teilnehmer:innen in der anschließenden Befragung überwiegend positiv bewerteten. In den teilnehmenden Beobachtungen beider XR-Part-Beteiligungsformate zeigten sich auch bei den zunächst ungeübten Personen schnelle

3 Aus datenschutzrechtlichen Gründen konnten Kinder und Jugendliche unter 18 Jahren, welche die XR-Beteiligungsangebote wahrnahmen, nicht befragt werden.

Lerneffekte im Umgang mit der Anwendung bzw. in der Steuerung des Avatars im virtuellen Raum.

Zum Onboarding im Prozessschritt der Vorbereitung bzw. Prozessplanung gehörte auch eine der Zielgruppe entsprechenden *Öffentlichkeitsarbeit*, um verständlich die Funktionsweise der XR-Technologien zu erklären und eine gegebenenfalls vorhandene Distanziertheit zum unbekannten digitalen Format abzubauen. Im Projektverlauf wurden aus diesem Grund kurze *Erklärvideos* produziert, welche neben den herkömmlichen Medien der Öffentlichkeitsarbeit auf der städtischen Website oder auf Social-Media-Plattformen anschaulich die Funktionsweise und Teilnahmebedingungen vermittelten.

Eine weitere *Zugangsbarriere*, der es während der Erprobungen zu begegnen galt, war neben dem unterschiedlichen technischen Know-how auch die *Hardwareverfügbarkeit* unter den interessierten Bürger:innen. Um auch Personen mit einem schlechteren Zugang zu digitalen Technologien die Teilnahme zu ermöglichen, wurden für das Angebot der XR-Part-Beteiligungstour *Tablets als Leihgeräte* zur Verfügung gestellt. Die Bereitstellung von Endgeräten bewährte sich aus organisatorischen, Datenschutz- und Performancegründen. Durch die Bereitstellung des virtuellen XR-Part-Beteiligungsraums als Desktopanwendung konnte zudem eine niedrigschwelligere Zugänglichkeit gegenüber einer VR-Anwendung mit entsprechend nötiger Hardware (Head-Mounted Displays [HMDs]) gewährleistet werden. Dennoch konnten sowohl mit der XR-PartBeteiligungstour (44,0 %) als auch mit der moderierten Beteiligungsveranstaltung (66,7 %) insbesondere Personen mit einem Hochschulabschluss, also einem hohen Bildungsgrad, erreicht werden. Wissenschaftliche Untersuchungen legen einen Zusammenhang zwischen den sozio-ökonomischen Hintergründen, dem Bildungs- und Digitalisierungsgrad sowie den Teilhabechancen an demokratischen Prozessen nahe, denen zufolge die Verfügbarkeit oder auch Nichtverfügbarkeit dieser Ressourcen häufig zusammen auftritt (vgl. Statistisches Bundesamt, 2023; Kramer, 2023: 228; Schelisch & Spellerberg, 2021: 54 f.; Initiative D21, o. J.: 14, 16 f.). Insbesondere die aufsuchenden und begleiteten XR-Beteiligungsformate bieten hier jedoch eine Chance, diese Barriere in der Zugangsgerechtigkeit gezielt abzubauen.

Die Angaben und Ausführungen zum Teilnahmegrund der Befragten legten dar, dass das Interesse an den neuartigen digitalen Formaten und an den 3D-Visualisierungen eine Motivation darstellte, das Beteiligungsangebot wahrzunehmen. Die städtischen Vertreter:innen führten im Rahmen der Interviews aus, dass sie die XR-Beteiligungsformate aufgrund ihres *spielerischen Charakters* als geeignet für die Adaption in der Kinder- und Jugendbeteiligung erachteten. Mit einer altersgerechten inhaltlichen Ausgestaltung der XR-Beteiligungsformate

und gezielt eingesetzten Gamification-Elementen können XR-Technologien speziell bei der Beteiligung jüngerer Zielgruppen empfehlenswert sein.

Grundsätzlich bieten die konzipierten XR-Beteiligungsformate das Potenzial einer *zeitlich flexiblen Teilnahme*, indem entweder die XR-Part-Beteiligungstour nach dem Download auf ein privates Endgerät jederzeit selbstständig durchgeführt werden kann oder der virtuelle XR-Part-Beteiligungsraum als persistentes Angebot längerfristig zur freien Erkundung zur Verfügung gestellt wird. Die volle oder teilweise Zustimmung von zusammengerechnet 71,1 % in Rostock und 78,2 % in Mannheim verdeutlicht, dass eine zeitlich flexible Teilnahmeoption für die Befragten von großer Bedeutung ist. Als Grund, die flexible und individuell wählbare Teilnahme gegenüber einer festen Veranstaltungszeit zu bevorzugen, geben die betreffenden Personen in der Befragung ihre Berufstätigkeit, persönlichen Verpflichtungen als auch Freizeitaktivitäten an. Auch diese Daten können im Evaluationsbericht des Projekts in detaillierterer und ausführlicherer Form abgerufen werden (vgl. Rogoll et al., 2024). Die erprobten moderierten und damit zeitlich gebundenen, aber *ortsunabhängigen* Veranstaltungen im virtuellen XR-Part-Beteiligungsraum bieten den Interessent:innen den Vorteil, von jedem Ort aus mit den entsprechenden technischen Voraussetzungen teilnehmen zu können. Dies ermöglicht u. a. Eltern, welche ihre Kinder betreuen, oder auch Personen, die in ihrer Mobilität eingeschränkt sind, die Teilhabe am Prozess. Die interviewten Bürger:innen führten diesbezüglich aus, dass die Teilnahme von zu Hause aus für sie leichter in den Alltag integrierbar war, da u. a. keine zusätzliche Zeit für die An- und Rückfahrt notwendig war.

4.2 Ethische, rechtliche und soziale Implikationen (ELSI)

Obwohl XR-Beteiligungsformate durch die Überwindung raumzeitlicher Barrieren die Teilhabe potenziell erleichtern können, ist es ein relevanter Schritt in der Planungsphase, ethische, rechtliche und soziale Implikationen der gewählten Beteiligungsangebote zu bedenken.

Die Definition der Zielgruppen für den Beteiligungsprozess erfordert eine differenzierte Betrachtung. Einerseits ist zu beschreiben, welche Gruppen gezielt durch XR-Beteiligungsformate angesprochen werden sollen. Andererseits sind diejenigen zu identifizieren, die entweder von der Nutzung der XR-Technologien ausgeschlossen sind oder deren Zugang zu diesen Technologien stark eingeschränkt ist (vgl. Huning, 2014: 39 f.; Orthmann, 2017: 12 f.). In Bezug auf die Zugangsgerechtigkeit ist es erforderlich, z. B. für Personen mit erschwerter Verfügbarkeit nötiger Hardware oder Menschen mit motorischen Beeinträchtigungen alternative Beteiligungsangebote bereitzustellen. XR-Be-

teiligungsformate sollten deshalb stets in *Kombination mit weiteren analogen und digitalen Angeboten in einem crossmedialen Prozess* angeboten werden.

In Bezug auf ELS-Implikationen ist zudem die *Vermeidung von Schäden* zu berücksichtigen. In diesem Sinne ist es erforderlich, XR-Anwendungen so zu konzipieren, dass sowohl *physische* als auch *psychische Belastungen und Gefahrensituationen* auf ein Minimum reduziert werden. Vor der Nutzung der XR-Technologie sollten mündige und selbstbestimmte Nutzer:innen über mögliche Risiken und Konsequenzen informiert werden. Bei der inhaltlichen Ausgestaltung der XR-Part-Beteiligungstour wurde zu Beginn ein Warnhinweis gegeben, um die Nutzer:innen dazu anzuhalten, die reale Umgebung während der Nutzung der AR-Anwendung im Blick zu behalten. Die Beobachtungen ergaben, dass in etwa 20 % der Fälle Gefahrensituationen auftraten, hauptsächlich durch Kollisionsgefahren mit anderen Verkehrsteilnehmer:innen oder durch übersehene Bordsteine aufgrund des Blickfokus auf den Tablet-Bildschirm. Schäden konnten jedoch in allen Fällen verhindert werden.

Zur Prävention von Risiken ist es ferner erforderlich, *negative Emotionen und Reaktionen wie Angst, Stress und mögliche Auslöser von Motion Sickness zu identifizieren und bei der (technischen) Gestaltung* der XR-Beteiligungsformate zu *berücksichtigen*. Diesbezüglich wurde bei der Erstellung der immersiven 3D-Modelle im virtuellen XR-Part-Beteiligungsraum darauf geachtet, Ängsten, wie Akrophobie (Höhenangst) oder Klaustrophobie (Platzangst), durch sichtbare und unsichtbare Barrieren entgegenzuwirken, die das Herunterfallen des Avatars von hohen Strukturen verhindern.

Durch die Umsetzung eines *barrierefreien Designs*, die Einbindung von Untertiteln oder die Bereitstellung alternativer Eingabemethoden kann die Zugänglichkeit von XR-Beteiligungsformaten für eine möglichst große Nutzergruppe erzielt werden (vgl. Dirks & Bühler, 2018: 8 f.). Im Entwicklungsprozess der XR-PartBeteiligungstour konnte durch die Integration einer *Text-to-Speech-Funktion* als auch durch die Nutzung der *Spracheingabe* die *Bedienungsfreundlichkeit* gezielt *verbessert* werden. Die Umfrageteilnehmer:innen beschrieben, dass ihnen die Spracheingabe die Abgabe von Kommentaren deutlich erleichterte. Auch die teilnehmenden Beobachtungen bestätigen, dass die Nutzer:innen es als schwierig empfanden, Beiträge über die Tablet-Tastatur einzugeben. Die erläuterten Textinhalte, welche im Verlauf der Tour eingeblendet wurden, ermöglichten es zudem zwei gehörlosen Bürger:innen, die Tour selbstständig zu durchlaufen.

In der Verknüpfung virtueller und realer Räume liegt das Potenzial, bestehende Kommunikationsbarrieren zu reduzieren und den Austausch zwischen verschiedenen Stakeholder:innen und der kommunalen Verwaltung zu fördern. Bei den moderierten Diskursen zwischen verschiedenen Interessengruppen im virtuellen XR-Part-Beteiligungsraum war es möglich, anhand konkreter und

immersiver Visualisierungen die Perspektiven anderer zu verstehen und in der eigenen Meinungsbildung zu berücksichtigen. Eine *transparente Kommunikation mittels der bildlichen Sprache eindrücklicher 3D-Varianten* und ergänzender Hintergrundinformationen kann auf diese Weise das *Verständnis* und die *Akzeptanz* für städtebauliche Veränderungsprozesse *steigern*.

4.3 Förderung der Vorstellungskraft und Meinungsbildung

Die adäquate visuelle Darstellung von Planungsinhalten erfordert eine sorgfältige Abwägung einer *angemessenen Detailtreue und Abstraktionsstufe der 3D-Objekte*. Die Darstellung des aktuellen Zustands des Beteiligungsgegenstandes in 3D-Stadtmodellen sollte in einer Weise erfolgen, die eine Wiedererkennung ermöglicht. Dabei sind korrekte Größenverhältnisse, Proportionen sowie Form- und Farbgestaltung von entscheidender Bedeutung, um eine präzise Darstellung zu gewährleisten (vgl. Brettschneider et al., 2017: 39). Der Grad der Abstraktion und die Detailgenauigkeit der 3D-Modelle sollten sich nach dem Fortschritt des Planungsprozesses und den Zielen der Beteiligung richten. In frühen Planungsphasen können grobe Modelle und einfache 3D-Piktogramme zur Visualisierung verwendet werden. In späteren Phasen sind detailliertere Modelle notwendig, um eine umfassende Beurteilung der Entwicklungsvarianten zu ermöglichen (siehe Abbildung 4). Diesbezüglich ist festzuhalten, dass ein gewisser Abstraktionsgrad erforderlich ist, um die Rahmenbedingungen und Grenzen des Projekts adäquat darzustellen und eine Ablenkung durch unwichtige Details zu vermeiden (vgl. u. a. Brettschneider et al., 2017: 39). Im Sinne einer informierten Meinungsbildung ist es erforderlich, *Hintergrundinformationen begleitend zu vermitteln, um eine Missinterpretation der Visualisierungen zu vermeiden*. Gleichermaßen ist jedoch ebenfalls auf einen ausgewogenen Anteil von Textinhalten innerhalb der Anwendung zu achten.

In den Erprobungen in den Städten Mannheim und Rostock gaben die Bürger:innen an, dass sie durch die anfänglich abstrakten Darstellungen von Bauvolumen und Flächenanteilen (wie beispielsweise Wohnraumerweiterungen) ein *besseres Verständnis für die räumlichen Auswirkungen und Sichtbeziehungen* aus der Egoperspektive gewinnen konnten. In der zweiten Erprobung wurden Diskussionen über konkrete Gestaltungsmöglichkeiten und die Nutzung des Raums insbesondere durch detaillierte Entwicklungsvarianten und die kreative Verwendung realistischer Stadtmöbel aus den Objektkatalogen angeregt. Des Weiteren bot die Auswahl an vegetativen Elementen (beispielsweise Bäume, Büsche, Blumenwiesen und Hochbeete) die Möglichkeit, sich intensiver mit den

Themen Klimaanpassung, z. B. die verschattende und kühlende Wirkung, und Biodiversität auseinanderzusetzen.

In den geführten Interviews wurde seitens der Bürger:innen sowie der kommunalen Vertreter:innen das *Potenzial des virtuellen Modells* betont. Die Simulation des Schattenwurfs zu unterschiedlichen Tageszeiten ermöglicht es, z. B. die Wirkung von Gebäudeaufstockungen oder Baumpflanzungen zu bewerten. Des Weiteren wurden die Möglichkeiten erörtert, klimatische Effekte, Beleuchtungskonzepte, Geräuschkulissen sowie indirekte Auswirkungen, wie beispielsweise der erhöhte Stellplatzbedarf aufgrund eines Bevölkerungszuwachses im Stadtteil, visuell zu vermitteln.

Grundsätzlich bieten XR-Beteiligungsformate den Vorteil, dass 3D-Objekte, Annotation und Kommentare georeferenziert verortet werden können, was eine raumbezogene Auswertung, z. B. in Form einer Heatmap ermöglicht, um Entwicklungsschwerpunkte zu identifizieren.

4.4 Akzeptanz von XR-Technologien

Die Analyse der begleitenden Befragungen zeigt eine hohe Akzeptanz der Bürger:innen gegenüber den erprobten XR-Part-Beteiligungsformaten. Dies zeigt sich an der mehrheitlichen Bereitschaft, auch zukünftig Beteiligungsangebote wahrnehmen zu wollen, bei denen XR-Technologie zum Einsatz kommt. Lediglich wenige Befragungsteilnehmer:innen gaben an, Beteiligungsangebote mit XR-Technologie nicht noch einmal wahrnehmen zu wollen. Dies waren zum einen Personen, welche kein Vertrauen in die Legalität des Prozesses und die Realisierung der Visualisierungen hatten, oder Senior:innen im höheren Alter, welche sich aufgrund ihrer Distanziertheit gegenüber neuartigen Technologien nicht vorstellen konnten, die XR-Formate selbstständig in Anspruch zu nehmen.

Eine überwiegende Mehrheit der Befragten führte hingegen aus, dass sie die *immersiven Visualisierungen als anregend und bereichernd* empfunden haben und die Möglichkeit einer *zeit- bzw. oder ortsunabhängigen Teilhabe* positiv bewerten. Aus der Nutzerperspektive wurde zudem der *spielerische Charakter der Formate* sowie die Kurzweiligkeit der XR-Beteiligungstour als positiv und motivierend herausgestellt. Eine detailliertere Darstellung der Daten kann ebenfalls dem Evaluationsbericht des Projektes entnommen werden (vgl. Rogoll et al., 2024).

4.5 Effizienter Einsatz von Ressourcen

Eine Herausforderung für die Umsetzung von XR-gestützten Beteiligungsformaten stellt der finanzielle Rahmen der Verwaltungen dar, um personellen Bedarf zu decken und Aufträge an externe Dienstleister zu vergeben. Eine effiziente Planung und Durchführung der Beteiligungsprozesse erfordert darüber hinaus den Austausch zwischen verschiedenen involvierten Akteursgruppen auf der Seite der Initiatoren, darunter das Amt bzw. die Verwaltungsabteilung für Öffentlichkeitsbeteiligung, das Amt für Geoinformation und Vermessung, die Datenschutzabteilung sowie extern beauftragte Planungsbüros, Beteiligungsdienstleister und nicht zuletzt ein entsprechender Technologieanbieter. Eine weitere Herausforderung für XR-Technologien stellt die *Anschlussfähigkeit an städtische Strukturen und die Beteiligungspraxis* dar. Diese Thematik betrifft u. a. die *Datenverfügbarkeit, verwaltungstechnische Prozesse und die technischen Schnittstellen bei den Kommunen*. Der technische Aufwand, z. B. für die Modellierung, kann je nach Größe des Beteiligungsvorhabens beträchtlich sein, was einen entsprechenden Finanzierungsrahmen voraussetzt (vgl. Schauppenlehner et al., 2018: 22). Vertiefende Ausführungen zur Aufgabenkoordination und Kalkulation von Kosten sind im Leitfaden zur Anwendung von XR-Beteiligungsformaten in städtischen Partizipationsprozessen zu finden, welcher ein zentrales Forschungsprodukt des Projekts darstellt (vgl. Rogoll et al.,2025).

Weitere Evaluationsergebnisse und Empfehlungen für die Umsetzung von XR-Beteiligungsformaten in der kommunalen Praxis wurden im Evaluationsbericht zum Projekt dokumentiert (vgl. Rogoll et al., 2024). Dieser erörtert u. a. die Kommunikationsqualität und -atmosphäre, die Gewährleistung einer niedrigschwelligen Bedienbarkeit als auch die methodische Ausgestaltung der XR-Beteiligungsformate.

5. Fazit und Ausblick

In diesem Beitrag wurden die XR-gestützten Beteiligungsformate im Projekt XR-Part, ihre Einbettung in die realen Beteiligungsverfahren zur Stadtentwicklung in den Modellstädten Mannheim und Rostock sowie die Erkenntnisse aus der begleitenden Evaluation der XR-Formate dargestellt. Die Ergebnisse zeigen, dass XR-Beteiligungsformate eine sinnvolle Ergänzung in Beteiligungsprozessen darstellen können, um eine erweiterte soziale Teilhabe zu ermöglichen. Insbesondere durch das aufsuchende Beteiligungsangebot der XR-PartBeteiligungstour konnten Bürger:innen zur Mitwirkung motiviert werden, die bisher noch keine Beteiligungsangebote wahrgenommen haben. Die

Potenziale von XR-Formaten für eine zeit- und ortsunabhängige Beteiligung fanden bei den Bürger:innen großen Anklang. Weiterhin zeigte sich, dass mit den XR-Beteiligungsformaten ein breiter Altersquerschnitt sowie Menschen mit unterschiedlich großer technischer Vorkenntnis bzw. Affinität erreicht werden konnten. Als hilfreich hat sich hierbei eine adaptive Onboarding-Strategie erwiesen, mit der auf die unterschiedlichen Bedürfnisse und Vorkenntnisse der Teilnehmer:innen eingegangen werden kann. Der Einsatz von Gamification-Elementen kann zudem die Nutzer:innen zusätzlich motivieren.

Technologien wie AR, VR bzw. virtuelle Kollaborationsräume eröffnen Bürger:innen und anderen Stakeholdern erweiterte Möglichkeiten der demokratischen Teilhabe sowie der verbesserten Vorstellungskraft und fördern eine kreative und informierte Mitwirkung an Entwicklungsfragen. Dies zeigen sowohl die Erkenntnisse aus dem XR-Part-Forschungsprojekt als auch weitere abgeschlossene Untersuchungen (vgl. u. a. Wolf et al., 2020: 125 f.). Der Einsatz von XRbasierten Beteiligungsformaten birgt also das Potenzial, öffentliche Mitwirkungsmöglichkeiten in städtischen Entwicklungsprozessen zu erweitern.

Die Nutzung von XR-Beteiligungsformaten erlaubt eine Teilnahme unabhängig von Ort und Zeit, wodurch sich neue, effiziente Beteiligungsoptionen für verschiedene Bevölkerungsgruppen eröffnen. Durch die Verwendung immersiver Visualisierungen wird es Bürger:innen mit unterschiedlichem Vorwissen erleichtert, komplexe Planungsthemen zu verstehen, was eine fundierte Meinungsbildung fördert. Insbesondere in frühen Entwicklungsphasen von Planungsvorhaben – zu einem Zeitpunkt, zu dem noch Einflussnahme möglich ist – haben XR-Technologien das Potenzial, durch die erweiterten Möglichkeiten der visuellen Kommunikation (v. a. 3D-Darstellungen) dabei zu unterstützen, den Beteiligungsgegenstand konkreter darzustellen. Damit belegen die aus den praktischen Erprobungen gewonnenen Erkenntnisse des Projektes die bereits ausgeführten Forschungsergebnisse von Brettschneider et al. (2017: 14). Beteiligte, denen die Vorstellung von abstrakteren Entwicklungszielen und Ideen schwerfällt, können sich so leichter in das Beteiligungsverfahren einbringen. Dabei kann XR-Technologie bei verschiedenartigen Vorhaben, von der Freiraumgestaltung über die Mobilitätsplanung bis hin zu Hochbauvorhaben, eingesetzt werden. Bei der Wahl der XR-Technologie hat sich gezeigt, dass AR besser geeignet ist, um Objekte im realen Raum visuell zu ergänzen (z. B. Gebäudeerweiterung, Gestaltungselemente im Freiraum). Der virtuelle Raum ermöglicht es währenddessen, den Ist-Zustand umfassender zu verändern und Elemente zu entfernen bzw. zu ersetzen, statt diese nur zu überlagern (z. B. Flächenverteilung, Eingriff in die Morphologie).

In der Zukunft könnte es zu einer breiteren Anwendung dieser Technologien kommen, insbesondere wenn die Barrieren in Bezug auf Kosten, Zugang, Daten-

verfügbarkeit und technische Komplexität überwunden werden. Weiterhin ist zu erwarten, dass Fortschritte in der Benutzerfreundlichkeit und die zunehmende Verbreitung von XR-kompatiblen Geräten die Teilnahme an solchen Prozessen zunehmend erleichtern. Dennoch ist der Einsatz von XR-Technologien in Beteiligungsverfahren immer nur als Ergänzung zu analogen Formaten zu betrachten, da nicht alle Bürger:innen mithilfe von XR-gestützten Beteiligungsformaten mitwirken können oder wollen.

Gleichzeitig ist die Forschung gefordert, weiterhin die sozialen Auswirkungen dieser Technologien zu untersuchen, insbesondere im Hinblick auf Chancengleichheit und die Sicherstellung, dass alle Bürger:innen gleichermaßen von diesen Innovationen profitieren können, insbesondere Menschen mit niedrigerem Bildungs- oder Digitalisierungsgrad bzw. sozial benachteiligte Menschen. Der aufsuchende Charakter der XR-Part-Beteiligungstour bietet hier wichtige Potenziale, um zielgruppengerecht die Zugänglichkeit von Beteiligungsangeboten für benachteiligte Zielgruppen zu erhöhen. Datenschutz, ethische Fragen und die Sicherung der Integrität von Beteiligungsprozessen werden ebenfalls von zentraler Bedeutung in der künftigen Forschung zu XR-Technologien sein. Zukünftige Weiterentwicklungspotenziale bestehen auch in der stärkeren Verknüpfung mit anderen innovativen Technologien, beispielsweise Künstlicher Intelligenz oder digitalen Zwillingen.

Insgesamt zeigt sich, dass XR-Technologien Potenziale haben, die partizipative Stadtentwicklung zu transformieren, indem sie eine stärkere Vorstellbarkeit von räumlichen Veränderungen bzw. Szenarien und eine erweiterte Teilhabe der Bevölkerung in Planungen ermöglichen. Die erfolgreiche Implementierung hängt jedoch stark von der Überwindung technischer und sozialer Hemmnisse ab, wozu Wissenschaft und Praxis gemeinsam gefordert sind, Lösungen zu finden. Hierzu ist auch die Erweiterung des Know-hows sowie die Erprobung und Reflexion von XR-Beteiligungsformaten in den Kommunen und bei Beteiligungsdienstleistern notwendig. Gleichzeitig sollten auch die Beteiligungsstandards, Leitlinien bzw. Regelwerke für die kommunale Bürgerbeteiligung um XR-gestützte Beteiligung weiterqualifiziert werden (vgl. Rogoll et al.,2025).

Förderhinweis

Das Projekt XR-Part wurde vom Bundesministerium für Bildung und Forschung (BMBF)[4] gefördert.

4 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Literatur

Austrian Institute of Technology GmbH (Hg.) (o.J.). Projektsteckbrief. VR-Planning. Virtual Reality für eine partizipative Planung und Evaluierung bedarfsgerechter und aktiver Mobilitätsumgebungen, Wien. https://www.ait.ac.at/fileadmin/mc/technology_experience/Infomaterialien/Projektsteckbrief_VR-Planning_barrierefrei.pdf [Zugriff 01.10.2024]

Brettschneider, Frank, Spieker, Arne & Wenzel, Günter (2017). *Bauprojekte visualisieren. Leitfaden für die Bürgerbeteiligung*, Schriftenreihe der Baden-Württemberg Stiftung, Nr. 86, Baden-Württemberg Stiftung (Hg.), Stuttgart. https://www.bwstiftung.de/fileadmin/bwstiftung/Publikationen/Forschung/Forschung_Bauprojekte_Visualisieren_Nr._86.pdf [Zugriff 26.08.2024]

Brysch, Armin (2023). Extended Reality (XR) im Tourismus – Erlebnisse durch Augmented Reality und Virtual Reality. In: *Digital Leadership im Tourismus – Digitalisierung und Künstliche Intelligenz als Wettbewerbsfaktoren der Zukunft*, Springer Gabler, Wiesbaden, S. 461–475.

Bundesministerium für Bildung und Forschung (Hg.) (o.J.): *VIRTUS. Mit Virtual Reality alle an einen Planungstisch bringen*, Mainz. https://www.interaktive-technologien.de/projekte/virtus [Zugriff 01.10.2024]

Dirks, Susanne & Bühler, Christian (2018). Barrierefreiheit, Usability und Universelles Design in der Softwareentwicklung. In: Drachselt, R.; Weber, G. (Hg.), *Mensch und Computer 2018 – Workshopband*, 2.-5. September 2018, Gesellschaft für Informatik e.V., Bonn, S. 7–17. https://doi.org/10.18420/muc2018-ws01-0250 [Zugriff 28.06.2024]

Dörner, Ralf et al. (2019). Virtual und Augmented Reality (VR/AR). Grundlagen und Methoden der Virtuellen und Augmentierten Realität. Springer Viewig, Berlin.

Fegert, Jonas et al. (2021). Ich sehe was, was du auch siehst. Über die Möglichkeiten von Augmented und Virtual Reality für die digitale Beteiligung von Bürger:innen in der Bau- und Stadtplanung. In: *HMD Praxis der Wirtschaftsinformatik*, Vol. 58, Springer Vieweg, Heidelberg, S. 1180–1195. https://doi.org/10.1365/s40702-021-00772-6 [Zugriff 01.10.2024]

GLARA Forschungskonsortium (Hg.) (2021). GLARA. Green Living Augmented+virtual ReAlity. Projekterkenntnisse & Empfehlungen für Beteiligungsprozesse unter Anwendung digitaler Vermittlungstools im öffentlichen Raum, Wien. https://jimdo-storage.global.ssl.fastly.net/file/d6f5cef2-c908-4bfb-8a3e-0c484a72a6c0/GLARA%20White%20Paper.pdf [Zugriff 01.10.2024]

Goudarznia, Toomaj, Pietsch, Matthias & Krug, René (2017). Testing the Effectiveness of Augmented Reality in the Public Participation Process: A Case Study in the City of Bernburg. In: *Journal of Digital Landscape Architecture*, Wichmann Verlag, Berlin, S. 244–251. https://doi.org/10.14627/537629025 [Zugriff 26.08.2025]

HafenCity University Hamburg (Hg.) (o.J.). *PaKomm*, Hamburg. https://pakomm.de/ [Zugriff 01.10.2024]

Hajrasouliha, Amir H. (2024). Applications, Approaches, and Ethics of the Extended Reality in Urban Design and Planning. In: *Journal of the American Planning Association*, Vol. 90, No. 3, Taylor & Francis, American Planning Association, London/Chicago, S. 551–567. DOI: 10.1080/01944363.2023.2275123 [Zugriff 08.10.2024]

Hanse- und Universitätsstadt Rostock (Hg.) (2023). *Stadtbereichskatalog. Stadtbereich L: Südstadt*, Rostock, https://rathaus.rostock.de/Statistik/Stadtbereichskatalog/2023%20Stadtbereich%20L.pdf [Zugriff 11.10.2024]

Hirscher, Ruthard (2017). Beteiligungsparadoxon in Planungs- und Entscheidungsverfahren. In: *vhw FWS*, Vol. 6, vhw-Dienstleistung GmbH, Bonn, S. 323–326. https://www.vhw.de/fileadmin/user_upload/08_publikationen/verbandszeitschrift/FWS/2017/6_2017/FWS_6_17_Beteiligungsparadoxon_in_Planungs_und_Entscheidungsverfahren_R._Hirschner.pdf [Zugriff 08.10.2024]

Hübler, Olaf (2024). Bürgerbeteiligung in Deutschland – Wer beteiligt sich wofür mit welchen Auswirkungen? In: *AStA – Wirtschafts- und Sozialstatistisches Archiv*, Vol. 18(1), Springer Verlag, Berlin/ Augsburg. S. 99–116. https://doi.org/10.1007/s11943-024-00336-2

Huning, Sandra (2014). Wer plant für wen? Partizipation im Kontext gesellschaftlicher Differenzierung. In: Küpper, Patricket al. (Hg.), *Raumentwicklung 3.0 – Gemeinsam die Zukunft der räumlichen Planung gestalten*, Arbeitsberichte der ARL, Vol. 8, Verlag der ARL, Hannover, S. 33–43. https://d-nb.info/1047124173/34 [Zugriff 26.08.2024]

Initiative D21 e.V. (Hg.) (o.J.): *D21-Digital-Index 2023/24. Jährliches Lagebild zur digitalen Gesellschaft*, Berlin. https://initiatived21.de/uploads/03_Studien-Publikationen/D21-Digital-Index/2023-24/d21digitalindex_2023-2024.pdf [Zugriff 24.09.2024]

Kramer, Christine (2023). Teilhabe in der Kultur der Digitalität. Kinderarmut und die Chancen digitaler Bildung. In: *MedienPädagogik – Zeitschrift für Theorie und Praxis der Medienbildung*, Vol. 52, o.O., S. 217–236. https://doi.org/10.21240/mpaed/52/2023.02.11.X

Le Blanc, David (2020). *E-Participation: A Quick Overview of Recent Qualitative Trends,* United Nations Department of Economic and Social Affairs (DESA) Working Paper, Vol. 163, New York. https://doi.org/10.18356/0f898163-en

Orthmann, Thomas (2017). *Wegweiser breite Bürgerbeteiligung. Argumente, Methoden, Praxisbeispiele,* Bertelsmann Stiftung, Allianz Vielfältige Demokratie (Hg.), Gütersloh. https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/Vielfaeltige_Demokratie_gestalten/Wegweiser_breite_Beteiligung_FINAL.pdf [Zugriff 13.08.2024]

ÖGUT – Österreichische Gesellschaft für Umwelt und Technik (Hg.) (2020). *GLARA: Partizipative Stadtgestaltung mit Augmented + Virtual Reality*, Wien. https://partizipation.at/praxisbeispiele/glara-partizipative-stadtgestaltung-mit-augmented-virtual-reality/ [Zugriff 08.10.2024]

Pahl-Weber, Elke (2018). Städtebaulicher Rahmenplan. In: ARL – Akademie für Raumforschung und Landesplanung (Hg.): *Handwörterbuch der Stadt- und Raumentwicklung*, Hannover, S. 2409–2413. [Zugriff 08.10.2024]

Presse- und Informationsamt Stadt Aalen (o.J.). *Virtueller Bürgerdialog*, Aalen. https://www.aalen.de/virtueller-buergerdialog.208940.25.htm [Zugriff 01.10.2024]

Rockmann, Lisa, Raabe, Susanne & Adler, Simon (2015). Augmented-Reality als Erweiterungs-Tool des partizipativen Austausches in Planungsprozessen zum Ziel einer integrativen städtebaulichen Entwicklung. REAL CORP 2015. PLAN TOGETHER – RIGHT NOW – OVERALL. From Vision to Reality for Vibrant Cities and Regions. In: *Proceedings of 20th International Conference on Urban Planning*, Regional Development and Information Society, Kompetenzzentrum für Stadtplanung und Regionalentwicklung, Wien, S. 83–92.

Rogoll, Svenja, Sinning, Heidi & Wolter, Anja (2024). XR-Part Beteiligungsformate in kommunikativen Planungsprozessen der Städte Rostock und Mannheim. Evaluationsergebnisse im Rahmen des Forschungsprojekts XR-Partizipationsräume zur erweiterten sozialen Teilhabe in urbanen Transformationsprozessen, ISP-Schriftenreihe, Bd. 20, ISP – Institut für Stadtforschung und Kommunikation der Fachhochschule Erfurt, Erfurt.

Rogoll, Svenja, Sinning, Heidi & Wolter, Anja (2025). XR-Beteiligungsformate in planerische Partizipationsprozesse integrieren. Handreichung für Kommunen, ISP-Schriftenreihe, Bd. 22, ISP – Institut für Stadtforschung, Planung und Kommunikation der Fachhochschule Erfurt, Erfurt.Schauppenlehner, Thomas et al. (2018). Anwendungserfahrungen von Virtual Reality als Kommunikationswerkzeug in partizipativen Planungsprozessen. In: *Journal für Angewandte Geoinformatik*, Vol. 4, Wichmann Verlag, Berlin/ Offenbach, S. 15–24. https://gispoint.de/fileadmin/user_upload/paper_gis_open/AGIT_2018/537647003.pdf [Zugriff 08.10.2024]

Schelisch, Lynn & Spellerberg, Annette (2021). Digital Divide. Soziale Aspekte der Digitalisierung. In: Spellerberg, Annette (Hg.), *Digitalisierung in ländlichen und verdichteten Räumen*, Arbeitsberichte der ARL 31. ARL – Akademie für Raumentwicklung in der Leibniz-Gemeinschaft, Hannover, S. 53–62.

Schöffmann, Dieter (2020). Bürgerbeteiligung – Die Lauten, die Stillen und die Unsichtbaren. In: *Alternative Kommunal Politik*, Vol. 3, Bielefeld, S. 43–45.

Senatsverwaltung für Stadtentwicklung und Umwelt Berlin (Hg.) (2012): *Handbuch zur Partizipation*, 2. Auflage, Berlin.

Sinning, Heidi et al. (2023). Partizipative Stadtentwicklung mit XR-Technologien. XR-Partizipationsräume als Beitrag zur erweiterten Teilhabe in urbanen Transformationsprozessen. In: *Forum Wohnen und Stadtentwicklung*, Vol. 2, vhw-Dienstleistung GmbH, Bonn, S. 88–92. https://isp.fh-erfurt.de/fileadmin/Dokumente/ISP/Publikationen/2023_Sinning_et_al_Partizipative_Stadtentwicklung.pdf [Zugriff 08.10.2024]

Simonofski, Anthony, Rohde Johannessen, Marius & Stendal, Karin (2024). Extended reality for citizen participation: A conceptual framework, systematic review and research agenda. In: *Smart Cities and Society*, Vol. 113, Art. 105692, IOS Press, Amsterdam. https://doi.org/10.1016/j.scs.2024.105692 [Zugriff 08.10.2024]

Stadt Mannheim (2022a). *Uhlandschule Vorplatz Beteiligungskonzept*, Mannheim. https://mannheim-gemeinsam-gestalten.de/sites/default/files/files/beteiligungskonzept-vorplatz-uhlandschule.pdf [Zugriff 08.10.2024]

Stadt Mannheim (2022b). *Statistische Daten 2022 Neckarstadt-Ost*, Mannheim. https://web2.mannheim.de/statistikatlas/pdf/03_neckarstadt-ost.pdf [Zugriff 08.10.2024]

Statistisches Bundesamt (Hg.) (2023). *Gehaltsvergleich 2022: Neben dem Beruf ist der Bildungsabschluss entscheidend*, Pressemitteilung Nr. 200 vom 23. Mai 2023, Wiesbaden. https://www.destatis.de/DE/Presse/Pressemitteilungen/2023/05/PD23_200_62.html [Zugriff 04.03.2024]

Thierbach, Cornelia & Petschick Grit (2022). Beobachtung. In: Baur, Nina, Blasius, Jörg (Hg.), *Handbuch Methoden der empirischen Sozialforschung,* Springer Fachmedien, Wiesbaden, S. 1563–1579.

Virtual Dimension Center Fellbach (Hg.) (o.J.). *Projekt. Partizipation in Stadtplanungsprozessen in virtuellen und realen Räumen (INSPIRER)*. https://projekt-inspirer.de/ [Zugriff 01.10.2024]

Wagner, Pia & Hering, Linda (2014): Online-Befragung. In: Baur, Nina & Blasius, Jörg (Hg.), *Handbuch Methoden der empirischen Sozialforschung*, Springer Fachmedien, Wiesbaden, S. 661–673. Online verfügbar unter: https://doi.org/10.1007/978-3-531-18939-0 (Zugriff 2024–10-07).

Wolf, Mario, Söbke, Heinrich & Wehking, Florian (2020). Mixed Reality Media-Enabled Public Participation in Urban Planning. In: Jung, Timothy, tom Dieck, M. Claudia & Rauschnabel,

Philipp A. (Hg.), *Augmented Reality and Virtual Reality. Changing Realities in a Dynamic World*, Progress in IS, Springer, Cham, S. 125–138. https://doi.org/10.1007/978-3-030-37869-1_11

Nutzerzentrierte Gestaltung von XR-Technologien für die Transformation urbaner Partizipationsprozesse – XR-Part: Herausforderungen und Lösungen für Geolokalisierung, Interaktion und Visualisierung zur Förderung der Bürgerbeteiligung

Yvonne Brandenburger, Sebastian Damek, Rolf Kruse, Mohammadreza Maleki Moghadam und Patrick Reipschläger

Zusammenfassung

Das Forschungsprojekt »XR-Part« untersucht den Einsatz von Extended-Reality-Technologien (XR-Technologien), insbesondere Augmented Reality (AR) und Virtual Reality (VR), zur Förderung von Bürgerbeteiligung in städtischen Planungsprozessen. Ziel ist es, soziale Teilhabe zu verbessern, bestehende Barrieren wie soziale Exklusivität zu überwinden und neue interaktive sowie immersive Beteiligungsformate zu schaffen. Hierfür wurden zwei Ansätze entwickelt und getestet: ein virtueller Beteiligungsraum und eine AR-gestützte Beteiligungstour. Während der Beteiligungsraum eine ortsunabhängige Interaktion in einer virtuellen 3D-Umgebung ermöglicht, erlaubt die Beteiligungstour vor Ort die Visualisierung und Gestaltung von Planungsinhalten direkt in der realen Umgebung.

Ein besonderer Fokus des Projekts liegt auf der Benutzerfreundlichkeit und Barrierefreiheit der Anwendungen, einschließlich der Unterstützung durch intuitive räumliche Bedienkonzepte, Sprachausgabe und adaptive Visualisierungstechniken. Technisch-gestalterische Herausforderungen wie präzises Tracking, Lokalisierung und die Einflüsse von Witterungsbedingungen auf die Bedienbarkeit wurden ebenso bearbeitet wie organisatorische Herausforderungen bei der Erstellung immersiver 3D-Modelle sowie der Integration neuer Technologien in bestehende Beteiligungsformate.

Die praktischen Erprobungen in zwei Modellstädten zeigen, dass XR-basierte Beteiligungsformate die Reichweite und Inklusion deutlich steigern können, insbesondere durch zeit- und ortsunabhängige Teilnahmemöglichkeiten sowie die verständliche Visualisierung komplexer Planungsinhalte. Trotz verbleibender

Herausforderungen, wie der Integration in bestehende Prozesse und der Unterstützung technikferner Nutzergruppen, bestätigt das Projekt das hohe Potenzial von XR-Technologien, Bürgerbeteiligung nachhaltiger und inklusiver zu gestalten.

1. Einleitung

Beteiligungsverfahren an städtischen Planungsprozessen leiden aktuell unter sozialer Exklusivität und einer eingeschränkten Repräsentativität, wodurch bestimmten Bevölkerungsgruppen u. a. aufgrund zeitlicher, räumlicher oder kognitiver Hürden die Teilnahme erschwert wird bzw. sie gänzlich von der Mitbestimmung ausgeschlossen werden. Das Forschungsprojekt XR-Part verfolgt das Ziel, die soziale Teilhabe mithilfe eines XR-Systems zu verbessern, welches in bestehende Prozesse integriert werden kann. Zu diesem Zweck wurden immersive und interaktive Beteiligungsräume geschaffen, die durch die Verbindung von VR- und AR-Technologien bestehende Barrieren überwinden sollen, indem in verschiedenen Formaten neue visuelle, räumliche und interaktive Formen der Kommunikation und Diskussion ermöglicht werden. Durch die zugänglichere, breitere und effizientere Beteiligung sollen gesellschaftliche Gestaltungsprozesse nachhaltig transformiert werden. Um diese Ziele mit den Anforderungen der Praxis abzugleichen und praktisch zu erproben, wurden innerhalb des Projekts zwei reale Beteiligungsverfahren in angeschlossenen Modellstädten begleitet. Ziel war die Erhebung der Anforderungen von Kommunen und Bürger:innen sowie die Evaluierung der Effizienz und Attraktivität der neuen Formate unter realen Bedingungen.

Bürgerbeteiligungen in städtischen Planungsprozessen bestehen regelmäßig aus verschiedenen Formaten, die sich an unterschiedliche Zielgruppen richten oder unterschiedliche Inhalte thematisieren. XR-Part ergänzt diese Formate durch XR-Anwendungen in Form von Bausteinen, die anpassbar und mit klassischen Formaten sowie untereinander kombinierbar sind (siehe Abbildung 1). Nach eingehender Analyse der Anforderungen und Rahmenbedingungen für die Erprobung unter heutigen Praxisbedingungen wurden zwei Formate priorisiert, konzipiert und entwickelt. Zum einen wurde ein von jedem Ort mit Internetzugang zugänglicher Beteiligungsraum als dreidimensionale virtuelle Umgebung entworfen, der einen Begegnungsraum für Information und Diskurs über die Planung darstellt, aber auch das Eintauchen in eine Simulation des Planungsgebietes mit interaktiven Gestaltungsmöglichkeiten ermöglicht. Das zweite Format, die Beteiligungstour, ermöglicht die Visualisierung und Interaktion vor Ort auf Basis von AR-Technologien. Innerhalb dieses Formates

ist es Bürger:innen möglich, sich verschiedene Varianten des Beteiligungsgegenstandes anzuschauen, diese zu kommentieren und auch selbst durch die Platzierung von 3D-Objekten im realen Raum gestalterisch tätig zu werden. In diesem Beitrag werden Aspekte der technischen Realisierung, der Bedienung und der visuellen Darstellung des Beteiligungsrundgangs mit AR-Funktionen erläutert. Weiterhin werden die Erfahrungen beschrieben, die zu diesen Themen innerhalb des Projekts gesammelt wurden, sowie die Lösungsansätze, welche das Projekt letztendlich verfolgt hat, um die Herausforderungen zu bewältigen. Für detaillierte Darstellungen zu den Entwicklungen und Erprobungen im Rahmen realer Beteiligungsprozesse sei auf den Beitrag der Kolleginnen Svenja Rogoll, Heidi Sinning und Anja Wolter in diesem Buch verwiesen.

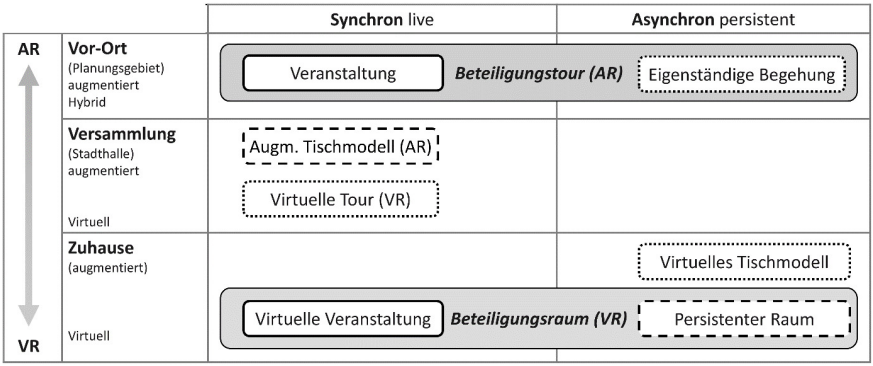


Abbildung 1: Bausteine des XR-Part-Systems. Die Beteiligungstour und der Beteiligungsraum wurden im Rahmen von realen Partizipationsveranstaltungen mehrfach praktisch erprobt.
Quelle: Projekteigene Darstellung FH Erfurt, Angewandte Informatik (FHE AI), 2023

2. Entwicklungsprozess und technische Grundlagen

Das Design des XR-Systems erfolgte unter Berücksichtigung der Bedürfnisse und Anforderungen der beteiligten Bürger:innen sowie der kommunalen Verwaltungen. Durch dieses nutzerzentrierte, iterative Verfahren sollte gewährleistet werden, dass die entwickelten Techniken existierende Beteiligungsformate aufgreifen und sinnvoll ergänzen können sowie innerhalb tatsächlicher Nutzungsszenarien erprobt und evaluiert werden können. Die Nutzerzentrierung wurde durch kontinuierliche Rückmeldungen und Interaktionen mit den Zielgruppen umgesetzt. Dies umfasste Workshops und mehrfache Erprobungen in Modellstädten. Dabei wurden die XR-Technologien (AR, VR, XR) in realen Planungsprozessen

getestet, um die Nutzbarkeit und Akzeptanz direkt zu evaluieren und die Anwendungen entsprechend anzupassen. Ein solches Vorgehen erhöht die Zugänglichkeit und Akzeptanz der Technologien, indem es sicherstellt, dass die XR-Anwendungen für unterschiedliche Bevölkerungsgruppen nutzbar sind, unabhängig von technischen Vorkenntnissen oder sozialen Barrieren.

In diesem Sinne spielt auch die Auswahl der Plattform für die Anwendung eine wesentliche Rolle (siehe Abbildung 2). Um den möglichst niederschwelligen Zugang für die teilnehmenden Bürger:innen zu ermöglichen, wurde für den Beteiligungsraum auf eine Desktop-Anwendung gesetzt, die mit handelsüblichem Desktop-PC, Laptop oder Mac ausgeführt werden kann und klassisch über Maus und Tastatur gesteuert wird. Hierdurch sollte gewährleistet werden, dass einem möglichst hohen Anteil der Bevölkerung der Zugang zu den Beteiligungsformaten ermöglicht wird. Für die AR-Anwendung wurde auf Mobilgeräte gesetzt. Erste stichprobenartige Erhebungen bei Kommunen sowie mithilfe verfügbarer Statistiken ergaben eine sehr hohe Verfügbarkeit von Smartphones in der Bevölkerung (2024: 82 %, vgl. Statista, 2024) und mit 25 % eine mittlere von Tablets (vgl. IfD Allensbach, 2023). Bei den Betriebssystemen dieser mobilen Geräte liegt aktuell Android von Google mit 66 % vor iOS von Apple mit 33 % (vgl. Kantar, 2024). Die Entwicklung der Anwendung erfolgte mit Unity 3D, was eine Ausbringung sowohl auf Android und iOS sowie auch zahlreiche andere Plattformen ermöglicht. Die Nutzung von AR- und VR-Brillen wurde im Projekt umfassend diskutiert. Eine Verwendung wäre hinsichtlich eines optimalen Nutzererlebnisses für die angestrebten Ziele wie der besseren Veranschaulichung und natürlichen Interaktion sehr attraktiv. Im Hinblick auf das Vorhaben, die Erprobung der XR-Formate nicht nur im Labor, sondern vor allem im Rahmen realer Beteiligungsverfahren und mit realen Bürger:innen durchzuführen, wurde die Verfügbarkeit von XR-Brillen innerhalb dieser Zielgruppe als sehr gering eingeschätzt und der Einsatz von XR-Brillen daher letztlich nur explorativ im Labor untersucht.

Die AR-Anwendung für die Beteiligungstour selbst besteht aus einer Kombination von konventionellen Fragebögen mit Single-, Multiple-Choice- und Freitext-Fragen sowie immersiven Inhalten. Nutzer:innen können verschiedene Stationen in der realen Umgebung ansteuern und sich beispielsweise verschiedene Varianten eines geplanten Beteiligungsgegenstands anschauen, kommentieren und bewerten. Außerdem war es Nutzer:innen möglich, an ausgewählten Stationen eigene Inhalte aus einem Katalog von 3D-Modellen in der realen Umgebung zu platzieren, um existierende Varianten zu modifizieren oder eigene zu gestalten. Antworten von Nutzer:innen sowie Art und Position platzierter Gegenstände wurden in anonymisierter Form gespeichert, um die spätere Auswertung der Einschätzungen, Wünsche und Vorschläge, aber auch des Interaktionsverhaltens zu ermöglichen. Für die erleichterte Bedienung unterstützt die Anwendung Sprach-



Abbildung 2: Einsatzszenarien mit verschiedenen XR-Geräteklassen
Laptops, Tablets und Smartphones sind besser verfügbar und deren Bedienung vertraut, haben aber eine geringere Immersion als VR-/AR-Brillen
Quelle: Projekteigene Darstellung FHE AI 2021 mit Material von TriCAT, ksg Architekten Köln/rendertaxi, Niantic, Vodafone GmbH, Meta, Magic Leap

ausgabe für alle Texte sowie die Möglichkeit der Texteingabe über Spracherkennung.

3. Verortung innerhalb der realen Umgebung

Da XR-Part ein Projekt zur Bürgerbeteiligung im Bereich der Stadtplanung ist und die Verortung virtueller Inhalte in der realen Umgebung eine der zentralen Merkmale von AR darstellt, ist es wichtig, dass diese Inhalte ihre Position relativ zur physischen Umgebung beibehalten, auch wenn Nutzer:innen ihre Position oder ihren Blickwinkel in der realen Umgebung ändern.

Hierfür sind zwei unterschiedliche technische Voraussetzungen wichtig: Zum einen das kontinuierliche Verfolgen der Position und Orientierung von Nutzer:innen in Relation zur realen Umgebung, allgemein als Tracking bezeichnet, sowie die initiale Verortung der Nutzer:innen innerhalb der realen Umgebung zur Bestimmung der Transformation zwischen virtuellem und realem Koordinatensystem, welche als Lokalisierung bezeichnet wird. Innerhalb dieser beiden Bereiche haben sich innerhalb von XR-Part einige Herausforderungen ergeben, wie bestehende Technologien auf die konkreten Anforderungen des Projekts angewendet werden können, um eine funktionale und robuste Lösung zu schaffen.

3.1 Kontinuierliche Verortung innerhalb der realen Umgebung

Tracking bezeichnet die kontinuierliche Verortung eines Gerätes in Relation zur Umgebung. XR-Part baut hierbei auf den existierenden Technologien der gängigen Plattformen auf, *ARCore* auf *Android* und *ARKit* auf *Apple* Geräten, die auch auf Geräten der Mittelklasse mittlerweile gut verfügbar sind (vgl. Google, 2024; Apple, 2024). Dabei handelt es sich um einen SLAM-Ansatz, bei welchem simultan eine Karte der initial unbekannten Umgebung erstellt (Mapping) sowie zeitgleich das Gerät innerhalb dieser Karte verortet wird (Lokalisierung). Als Folge des zur Laufzeit erfolgenden Mappings muss die globale Position des Geräts mithilfe anderer Verfahren ermittelt werden, welche im nächsten Kapitel vorgestellt werden. Die von der jeweiligen Plattform zur Verfügung gestellten Tracking-Verfahren zeichnen sich durch ihre einfache Verwendung und universelle Verfügbarkeit aus. Allerdings ermöglichen sie auch keinen Einblick in die darunterliegende technologische Basis und bieten entsprechend auch wenig Möglichkeiten für Anpassungen. Das Tracking ist im gewöhnlichen Anwendungskontext robust und Probleme mit Drift, d. h. dem unerwünschten Auseinanderdriften des realen und virtuellen Koordinatensystems, sind zwar vorhanden, aber vergleichsweise gering.

Sowohl ARCore wie auch ARKit verwenden einen bildbasierten Ansatz, bei welchem, vereinfacht gesagt, ein Algorithmus interessante Punkte in einem Bild ermittelt und versucht, diese in folgenden Bildern wiederzufinden. Dies ermöglicht Rückschlüsse auf die relative Änderung von Position und Orientierung zwischen zwei Bildern und damit die kontinuierliche Verortung des Geräts. Bildbasierte Ansätze liefern gute Ergebnisse bei gleichbleibenden Bedingungen der Umgebung, was vor allem Innenräume unter künstlichem Licht begünstigt. Urbane Planungsprozesse, für die das XR-Part-System vorgesehen ist, erfordern als Anwendungsgebiet allerdings fast ausschließlich Außenbereiche innerhalb großer Areale. Hier zeigen sich auch die Grenzen dieser Technologie: Außenbereiche sind stark wechselnden Lichtverhältnissen abhängig von Tageszeit und Witterung unterworfen, was die Wiedererkennung der Bild-Features extrem erschwert. Die Witterung beeinflusst aber nicht nur die Lichtverhältnisse, sondern auch Oberflächeneigenschaften. So erzeugen nasse Oberflächen bei Regen zusätzliche Reflexionen, während sich bei Schneefall im Prinzip die komplette Umgebung verändert und die Wiedererkennung nahezu unmöglich wird. Durch die großen Areale, welche in XR-Part begangen werden, potenzieren sich auch vergleichsweise geringe Abweichungen als Folge von Drift, also der Akkumulation von Ungenauigkeiten der Positionserfassung über die Zeit, schnell zu merklichen Abweichungen bei der Positionierung von virtuellen Inhalten. Als Folge finden Nutzer:innen beispielsweise platzierte Objekte nicht mehr an der gleichen Position wieder, wenn sie in ein zuvor bearbeitetes Areal zurückkehren. Dies ist

problematisch, wenn Nutzer:innen gefordert sind, das Areal eigenständig durch die Platzierung von Objekten zu gestalten.

3.2 Absolute Lokalisierung innerhalb der realen Umgebung

Während Tracking standardmäßig nur relative Positionsänderungen liefert, ist für AR eine initiale absolute Positionierung erforderlich, damit Inhalte an ihrer exakten Position in der realen Umgebung dargestellt werden können. Im Folgenden wird dieser Prozess als Lokalisierung bezeichnet.

Die benötigte Genauigkeit der Lokalisierung ist abhängig von den Zielen der AR-Anwendung. Je größer die angezeigten virtuellen Inhalte sind und je weiter der Nutzer oder die Nutzerin von ihnen entfernt ist, umso mehr Ungenauigkeiten bei der Lokalisierung können toleriert werden. Im Allgemeinen ist für eine überzeugende und immersive AR-Erfahrung jedoch eine Lokalisierung im Zentimeterbereich erforderlich. Obwohl es zahlreiche Technologien für die absolute Positionserkennung gibt, existiert aktuell kein einheitlicher Standard in diesem Bereich für XR-Systeme. Verfügbare Technologien sind entweder auf bestimmte Plattformen beschränkt oder erreichen nicht die nötige Präzision und Zuverlässigkeit. Im Folgenden werden einige der infrage kommenden Technologien erläutert sowie ihre entsprechenden Vor- und Nachteile diskutiert.

Eine weit verbreitete Methode zur Positionsbestimmung ist GPS (Global Positioning System). GPS bietet den Vorteil einer nahezu universellen und plattformunabhängigen Verfügbarkeit, da das System in nahezu jedes moderne Mobilgerät integriert ist. Darüber hinaus benötigt GPS keine Verbindung zu Mobilfunknetzen oder WiFi, was eine weiträumige Verwendung im Außenbereich zulässt. Gleichsam ist die Verwendung von GPS in Innenräumen schwierig bis unmöglich, insbesondere in mehrstöckigen Gebäuden, bei denen eine genaue Bestimmung der Etage relevant ist. Obwohl mit GPS in der Praxis unter guten Bedingungen die Positionierung mit einer Präzision im Bereich von 30 cm (Beran et al., 2007: 53–63) möglich ist, reicht dies für detaillierte AR-Visualisierungen oft nicht aus. Viel wichtiger ist allerdings, dass GPS zwar die Ermittlung der Position zulässt, aber keine Bestimmung der Orientierung. Dennoch kann der Einsatz von GPS nützlich sein, um eine grobe Positionserkennung durchzuführen und den initialen Suchradius für die Verwendung präziserer Verfahren zu minimieren.

Eine weitere verbreitete Methode ist ein featurebasierter Ansatz, bei dem interessante Punkte in einem Kamerabild erkannt werden. Parallel wird ein sogenannter Anker in der Szene platziert, welcher ein lokales Koordinatensystem definiert. In folgenden Sitzungen können diese interessanten Punkte wiedergefunden werden, woraufhin der entsprechende Anker platziert wird. Das lokale Koor-

dinatensystem befindet sich somit wieder am gleichen Ort wie zuvor, genauso wie die in ihm platzierten virtuellen Objekte. Somit kann eine persistente, sitzungsübergreifende Verortung von Inhalten erreicht werden. Wie auch das featurebasierte Tracking ist dieses Verfahren jedoch ausgesprochen anfällig für wechselnde Umweltverhältnisse, die eine verlässliche Erkennung erschweren oder sogar unmöglich machen. Obwohl für den Innenbereich durchaus geeignet, stellt dieser Ansatz für den Außenbereich keine verlässliche Lösung dar.

Zuletzt sei noch der Ansatz des markerbasierten Trackings besprochen. Hierbei wird ein Marker mit bekannten physischen Eigenschaften hinsichtlich Größe und Aussehen in der realen Umgebung platziert. Diese Marker können beispielsweise die Form eines QR-Codes oder auch eines featurereichen Bildes annehmen. Ein Algorithmus versucht die Marker innerhalb des Kamerabildes des AR-Gerätes zu erkennen und deren Position und Orientierung im Verhältnis zur Kamera zu bestimmen. Ist die globale Position des Markers bekannt, kann somit auch die Position des Geräts bestimmt werden. Ein offensichtlicher Nachteil dieser Methode ist die Notwendigkeit der Ausbringung von Markern in die zu augmentierende Umgebung. Im Gegenzug zeichnet sie sich allerdings durch ein hohes Maß an Robustheit und vergleichsweise geringe Abhängigkeit von Umwelteinflüssen aus.

3.3 Lösungsansatz für die AR-Lokalisierung in XR-Part

Im Verlauf des Projektes wurden verschiedene Verfahren wie GPS und imagebasiertes Tracking im Zuge der an das Projekt angegliederten Beteiligungsverfahren praktisch erprobt. Aufgrund der zuvor beschriebenen Probleme der beiden ersten Verfahren fiel die Entscheidung letztendlich zugunsten eines markerbasierten Ansatz aus. Dies diente vor allem der Erreichung eines Maßes an Zuverlässigkeit, das einen möglichst reibungslosen Ablauf der Vor-Ort-Beteiligung gewährleisten kann. Insbesondere die Unempfindlichkeit gegenüber wechselnden Witterungsverhältnissen war hierbei ausschlaggebend. Zu diesem Zweck wurden Marker in entsprechenden Aufstellern platziert (siehe Abbildung 3a), bei denen sich Nutzer:innen initial verorten konnten. Die vorgesehene Position der Aufsteller innerhalb der realen Umgebung wurde zuvor bei der Planung der Veranstaltung bestimmt. Vor Ort wurde die genaue Lokalisierung pro Marker dadurch realisiert, dass in einem Kalibrierungsschritt der Offset ermittelte wurde, in dem die virtuelle und echte Umgebung voneinander abweichen. Dieser Schritt erfolgte manuell anhand eines optischen Abgleichs. Die ermittelten Offsets wurden persistent gespeichert, sodass beim erneuten Einlesen eines Markers die virtuelle Umgebung mit korrekter Position und Orientierung platziert werden konnte.

Um Problemen mit Drift und Ungenauigkeiten bei der Kalibrierung entgegenzuwirken, welche durch die Größe der involvierten Areale bei ersten Tests zu merklichen Abweichungen geführt haben, wurde ein Stationskonzept gewählt (siehe Abbildung 3b). Für jeden *Point of Interest* (POI) des Beteiligungsverfahrens wurde ein entsprechender Aufsteller mit AR-Marker platziert und als eigenständige Station ausgezeichnet. Abhängig von den gewählten Fragestellungen im jeweiligen Beteiligungsverfahren wurden Anwender:innen entweder in vordefinierter Reihenfolge zu den einzelnen Stationen geleitet oder konnten selbstständig entscheiden, welche Station sie als nächstes erkunden möchten. Um die Navigation zwischen Stationen zu erleichtern, wurde zum einen eine *Mini-Map* innerhalb der Anwendung angezeigt, welche die eigene Position und die von umliegenden POIs visualisiert. Zum anderen wurde nach der Auswahl einer Station ein AR-3D-Pfad angezeigt, dem Nutzer:innen zur Station folgen konnten. An der jeweiligen Station angekommen, wurde wieder der entsprechende Marker eingescannt, sodass die korrekte Verortung der Inhalte gewährleistet war.



Abbildung 3a: Aufsteller mit AR-Image-Marker in Rostock
Quelle: FHE AI 2024

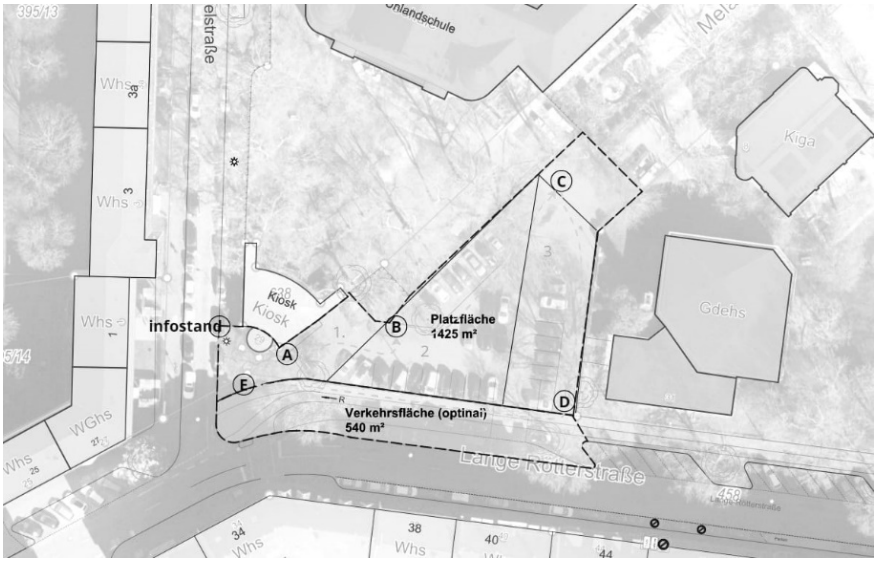


Abbildung 3b: Karte zum Stationskonzept
Fünf Stationen (A–E) der Beteiligungstour am Platz an der Uhlandschule in Mannheim mit verschiedenen inhaltlichen Schwerpunkten und unterschiedlichen Visualisierungs-, Interaktions- und Befragungsformen
Quelle: Projekteigene Darstellung FHE AI mit Material der Stadt Mannheim 2022

4. Bedienung und Usability der AR-Anwendung

Die AR-App wurde unter dem Gesichtspunkt der einfachen und niederschwelligen Verwendbarkeit in Bürgerbeteiligungsprozessen mit dem Ziel entwickelt, ein hohes Maß an Usability zu bieten und die kognitive Belastung der Teilnehmer:innen zu minimieren. Durch regelmäßiges Feedback der Zielgruppe und den in den praktischen Erprobungen gesammelten Erfahrungen wurde die Nutzeroberfläche (User Interface [UI]) und das Look-and-Feel der Anwendung, also ihre grafische Gestaltung und ihr Verhalten, kontinuierlich verbessert. Eine Herausforderung bestand hierbei darin, die Anwendung so zu gestalten, dass sie ohne begleitende Betreuung von Teilnehmer:innen mit unterschiedlichsten technischen und sozialen Hintergründen verwendet werden kann.

4.1 Technische Einführung und Hilfestellungen in der Bedienung

Eine weitere besondere Herausforderung bei der Entwicklung der AR-Anwendung ist die Inklusion eines möglichst großen Teils der Bürgerschaft und die damit verbundene Heterogenität der Nutzergruppe. Dies betrifft Aspekte wie die technischen Vorkenntnisse, aber auch physische Aspekte wie eingeschränkte Sicht oder limitierte Mobilität. Aufgrund dieser Heterogenität wurde auch die Notwendigkeit einer Einführung zu AR und der Bedienung der Anwendung identifiziert, da davon auszugehen war, dass Bürger:innen nur vereinzelt Vorerfahrungen in diesem Bereich haben. Zu diesem Zwecke wurden beim Start der Anwendung einige Einführungsseiten gezeigt, welche die wesentlichen Funktionen erläutern. Anfänglich handelte es sich dabei um textbasierte Erläuterungen (siehe Abbildung 4). Innerhalb der ersten Erprobungen wurde durch strukturierte Befragungen deutlich, dass diese für viele Teilnehmer:innen zu abstrakt sind und übersprungen wurden (Details zur Studie im Beitrag von Rogoll, Sinning, Wolter in diesem Band). Daher wurden die Erläuterungen in folgenden Erprobungen durch Animationen ergänzt (siehe Abbildung 5a), was das Verständnis bei Teilnehmer:innen deutlich verbessert hat. Trotz dieser Einführung benötigten viele Teilnehmer:innen im Verlauf der Beteiligungstour weiterhin Unterstützung, insbesondere bei der Nutzung spezifischer Funktionen wie der Platzierung von Objekten oder der Kommentarfunktion. So gaben nur 44,8 % der Befragten an, dass ihnen die Einführung geholfen habe, die Tour selbstständig zu bewältigen. Vielen Teilnehmer:innen viel es schwer, mit einer konkreten Aufgabe konfrontiert sich an die entsprechenden Informationen aus der Einführung zu erinnern. Um dieses Problem zu adressieren, wurde eine interaktive Hilfefunktion in die Anwendung integriert. Mittels einer entsprechenden Bedienfläche konnten Teilnehmer:innen Hilfestellung zu ihrer aktuellen Aufgabe erhalten, welche wieder mit Animationen unterstützt wurden.

Ein weiteres Problem, das sich bei den ersten Erprobungen zeigte, war die Schwierigkeit vieler Teilnehmer:innen, Text mithilfe der Bildschirmtastatur einzugeben. Einige Teilnehmer:innen brauchten für diesen Prozess entweder sehr lange oder delegierten ihn gänzlich an umstehende Begleiter:innen. Um dieses Problem zu adressieren, wurde eine *Speech-to-Text*-Funktion in die Anwendung integriert, die es Anwender:innen ermöglichte, ihre Anmerkungen über das Mikrofon zu diktieren. Um auch Personen mit eingeschränkter Sehfähigkeit nicht auszuschließen, besaß die Anwendung von Anfang an eine Sprachausgabe, welche alle relevanten Texte vorgelesen hat. Es war zu beobachten, dass diese Funktion vor allem von älteren Anwender:innen geschätzt und genutzt wurde, indem sie beispielsweise dem Gerät von vornherein nur zuhörten, anstatt die Texte auf dem Bildschirm zu lesen. Jüngere Teilnehmer:innen wiederum haben vor allem



Abbildung 4: Textbasierte Einführungsseiten
a: Die erste Erprobung in Rostock. b: Die zweite Erprobung in Mannheim
Quelle: Projekteigene Darstellung 2024

die Texte gelesen. Eine interessante Erkenntnis in diesem Zusammenhang ist, dass viele Teilnehmer:innen, obwohl die vorgelesenen Texte jederzeit abgebrochen werden konnten, diese dennoch bis zum Ende der Ausgabe abgewartet haben, auch wenn sie die entsprechenden Texte bereits gelesen hatten. In der Beobachtung und Befragung der Teilnehmer:innen wurde deutlich, dass hierbei vor allem soziale Protokolle ursächlich waren, da es als unhöflich empfunden wurde, die Ausgabe zu unterbrechen.

Für die Durchführung der praktischen Erprobungen wurde in der Anwendung mit einem Stationskonzept gearbeitet, bei dem sich Teilnehmer:innen zwischen mehreren POIs bewegen konnten. Hier wurde auch die Notwendigkeit der Unterstützung bei der Navigation deutlich, um Teilnehmer:innen erfolgreich zwischen den Stationen zu leiten. Der Ansatz, der hierbei verfolgt wurde, umfasst drei Aspekte: Zum einem wurde eine *Mini-Map* (siehe Abbildung 5b) innerhalb der Anwendung gezeigt, welche die aktuelle Position der Teilnehmenden sowie die umgebenden POIs zeigt. Dies ermöglichte es Teilnehmer:innen, sich in der Umgebung zu orientieren. Des Weiteren wurden die Stationen durch farbige Säulen repräsentiert (siehe Abbildung 5d), die in AR an der festgelegten Position in der realen Welt platziert waren und diese somit aus größerer Entfernung sichtbar gemacht haben. Sollte eine bestimmte Station angesteuert werden, so wurde ein 3D-Pfad (siehe Abbildung 5c) in der Anwendung visualisiert, der Teilnehmer:innen zur entsprechenden Station leitet, und zwar über bestehende Wege, um beispielsweise potenziell gefährliche Verkehrssituationen bei der Überquerung von Straßen oder Fahrradwegen zu vermeiden.

Die Summe dieser Maßnahmen führte zu einer allgemeinen Verbesserung der Nutzbarkeit für die Teilnehmer:innen. Dies zeigte sich auch in der Bewertung der Anwendung, die in beiden Städten überwiegend positiv ausfällt: 79,3 % der Teil-

nehmenden in Mannheim und 86 % in Rostock empfanden die Nutzung in den letzten Erprobungen als einfach.

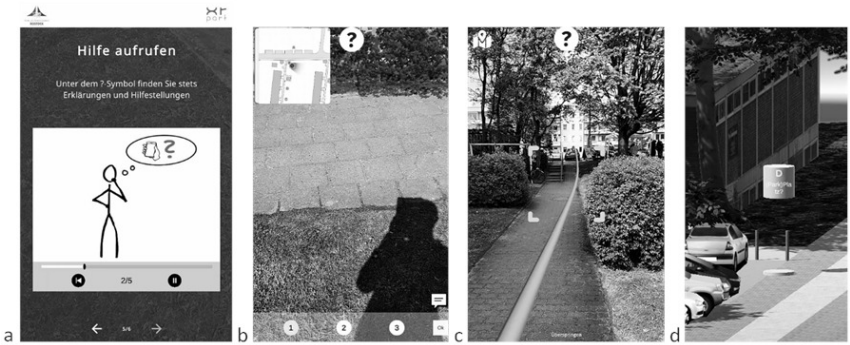


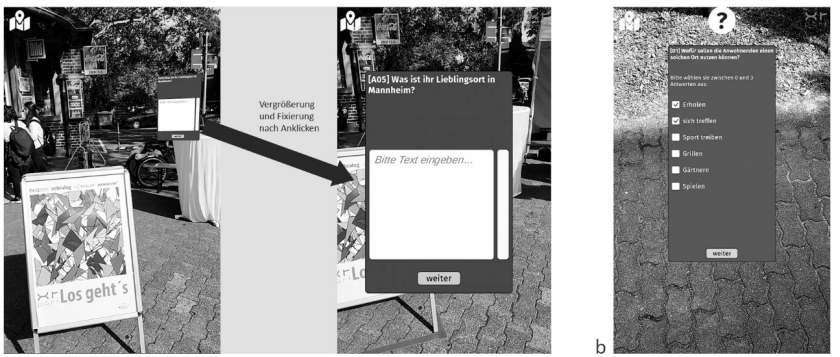
Abbildung 5: Elemente der grafischen Bedienoberfläche
a: Einführungsanimationen zu den Hauptfunktionen und Sicherheitshinweisen. b: Mini-Map-Funktion, die den aktuellen Standort und die Bewegungsrichtung anzeigt. c: 3D-Pfad für die Navigation zwischen den Stationen. d: Stationsmarkierung mit 3D-Pylonen.
Quelle: Projekteigene Darstellung FHE AI 2024

4.2 Platzierung von 2D-Inhalten in der Umgebung

Ein Aspekt, um die Anschlussfähigkeit der immersiven Beteiligungstour zu bestehenden Beteiligungsformaten zu verbessern, war die Integration von Informationstafeln und Fragebögen. Da es sich vor allem bei den Fragebögen um nicht inhärent dreidimensionale Inhalte handelt, stellte sich die Frage, wie diese in die AR-Erfahrung der Anwendung integriert werden können. Um dennoch ein gewisses Maß an Immersion zu bieten, war der zunächst verfolgte Ansatz, sowohl Fragebögen als auch Informationsmaterialien als feste Tafeln in der realen Umgebung zu verorten, ähnlich wie bei einer physischen Infotafel. Bei der praktischen Erprobung dieser Variante zeigte sich jedoch, dass Nutzer:innen große Probleme hatten, die Tafeln zu identifizieren, die darauf abgebildeten Inhalte zu lesen sowie bei den Fragen die gewünschten Antworten auszuwählen. Ursächlich dafür war bei einigen Nutzenden das Problem, das Tablet auf einen bestimmten Inhalt gerichtet zu halten und gleichzeitig mit dem Bildschirm zu interagieren. Außerdem zeigte sich bei einigen Nutzer:innen eine gewisse Zurückhaltung, sich frei in der Umgebung mit dem Tablet zu bewegen, um beispielsweise einen besseren Blick auf die Infotafeln zu erlangen. Stattdessen versuchten einige Anwender:in-

nen mit der *Pinch-to-Zoom*-Geste, die sie aus konventionellen Anwendungen kannten, die Tafeln zu vergrößern.

Basierend auf diesen initialen Erfahrungen wurde für die nächste Erprobung ein hybrider Ansatz entwickelt, bei welchem die Tafeln weiterhin fest in der Umgebung verortet waren, die Tafel aber beim ersten Tap an den Bildschirm geheftet wurde – ein Ansatz, der *SnapPanel* getauft wurde (siehe Abbildung 6a). Hierdurch konnte mit den Tafeln wie mit konventionellen UIs interagiert werden. Dies verbesserte deutlich deren Lesbarkeit und Bedienung. Viele Nutzer:innen hatten in der Praxis aber weiterhin große Probleme, die Tafeln innerhalb der realen Umgebung zu lokalisieren. Aus diesem Grund wurde in den letzten zwei Erprobungen eine rein bildschirmzentrierte Version (siehe Abbildung 6b) entwickelt, bei der die Inhalte ausschließlich wie konventionelle mobile UI dargestellt wurden und es keine virtuelle Verankerung mehr in der realen Umgebung gab. Hierdurch wurde es notwendig, Inhalte, die vorher im Raum exploriert werden konnten, nun sequenziell anzuzeigen, sodass Nutzer:innen diese nicht mehr wahlfrei konsumieren konnten. Auch wurde hierdurch die Immersion der restlichen AR-Inhalte negativ beeinträchtigt, da somit ein Bruch zwischen im Raum verorteten und bildschirmzentrierten Inhalten entstanden ist. Dennoch zeigte die praktische Erprobung, dass die Bürger:innen mit dieser Variante deutlich besser und fehlerfreier Eingaben durchführen konnten, was die zuvor erwähnten Nachteile aus Sicht des Projektes aufwiegt.



a b

Abbildung 6: Screenshots

a: Screenshots der SnapPanel-Funktion: Umfrage-Tafeln werden für den Zeitraum der Eingabe aus ihrer räumlichen Verankerung (Georeferenzierung) zentriert an den Bildschirm geheftet. b: Screenshot einer bildschirmzentrierten Version für eine Multiple-Choice-Fragetafel aus einer Erprobung in Rostock.

Quelle: Projekteigene Darstellung FHE AI 2024

4.3 Einfluss des Wetters auf die Bedienbarkeit

Alle Erprobungen von XR-Part wurden unter freiem Himmel und zu unterschiedlichen Zeiten des Jahres durchgeführt. Dies hatte zur Folge, dass die Anwendung unter sehr unterschiedlichen Witterungsbedingungen erprobt und deren Einfluss auf die Bedienbarkeit der Anwendung erfasst wurde. Praktische Erfahrungen haben gezeigt, dass die ideale Witterungsbedingung für AR-Anwendungen ein bewölkter Himmel ohne direkte Sonneneinstrahlung ist. Im Folgenden werden die Probleme erläutert, die bei starkem Sonnenschein, Regen, starkem Wind oder Schneefall bestehen, und welche möglichen Lösungen hierfür infrage kommen.

Direktes Sonnenlicht beeinflusst stark, wie gut Inhalte auf dem Display der Mobilgeräte wahrgenommen werden können. Selbst bei maximaler Helligkeitseinstellung können die Displays so dunkel wirken, dass Inhalte nur schwer oder gar nicht mehr wahrgenommen werden können. Ein weiteres Problem besteht in Spiegelungen, die durch Sonneneinstrahlung auf dem Display erzeugt werden und die Wahrnehmung zusätzlich erschweren. Um den Spiegelungen entgegenzuwirken, wurde im Projekt mit matter Folie auf den Displays experimentiert. Diese brachte aber nur eine geringfügige Verbesserung und hat die Displays noch dunkler gemacht. Eine unmittelbare Abhilfe bei den praktischen Erprobungen wurde durch die Verwendung von Regenschirmen als Schattenspender erreicht, welcher durch eine zweite, begleitende Person gehalten wurde. Hierdurch war es zwar möglich, die Beteiligungstouren weiter durchzuführen, aber für die selbstständige Verwendung der Anwendung stellte dies keine praktikable Lösung dar, da beide Hände für die Verwendung des Mobilgeräts erforderlich sind.

Ein weiteres Problem stellt Regen dar. Regen auf dem Display führt nicht nur zu Reflexionen, sondern beeinflusst auch die Touch-Eingabe der Geräte. Hierdurch werden Touches nur schlecht oder gar nicht erkannt, was die Bedienung der Anwendung deutlich erschwert. Ähnlich wie zuvor ist auch hier die Verwendung von typischen Hilfsmitteln wie Regenschirmen problematisch, da für die Bedienung der Anwendung beide Hände benötigt werden. In der praktischen Durchführung mussten einige Touren durch starken Regen abgebrochen bzw. bis zu einer Verbesserung der Wetterlage verschoben werden.

Die Durchführung von Beteiligungstouren unter winterlichen Bedingungen bietet ebenfalls einige Herausforderungen. Die Nutzung der Touch-Eingabe auf den Mobilgeräten ist typischerweise nicht mit Handschuhen möglich. Entsprechend klagten einige Bürger:innen über kalte Hände und Finger, was nicht nur unangenehm war, sondern auch die Bedienung der Anwendung erschwerte. Im Projekt wurde mit Handschuhen experimentiert, deren Fingerkuppen leitfähig sind und somit die Verwendung kapazitiver Touch-Screens ermöglichen. Auch

wenn die Touch-Erkennung damit schlechter ist als ohne Handschuhe, so stellt dies dennoch ein geeignetes Mittel dar, die Verwendung auch in kaltem Wetter zu ermöglichen. Allerdings kann nicht vorausgesetzt werden, dass Bürger:innen selbst solche Handschuhe besitzen. Schneefall bietet eine weitere Herausforderung, da hiermit eine komplette optische Metamorphose der physischen Umgebung einhergeht. Dies kann es erschweren, die Beteiligungsinhalte zu vermitteln, da die virtuellen Objekte nicht mehr als Teil der Umgebung wahrgenommen werden, sondern stattdessen kontextlos auf einer weißen Decke erscheinen. Abhilfe verschaffte hier der Schritt hin zu mehr Virtualität, indem die nun nicht mehr sichtbaren Teile der Umgebung ebenfalls als virtuelle 3D-Objekte abgebildet wurden. Innerhalb von XR-Part war dieser Schritt leicht zu vollziehen, da es ohnehin eine VR-Version der Umgebung für das Format des Beteiligungsraumes gab.

5. Visualisierung von 3D-Inhalten in digitalen Beteiligungsverfahren

Die Bedeutung von Visualisierungen im Kontext städtebaulicher Beteiligungsverfahren lässt sich bis in die Anfänge dieser Verfahren zurückverfolgen. Ursprünglich erfolgte die Erstellung von Visualisierungen durch Handzeichnung. Mit der kontinuierlichen Weiterentwicklung der Computergrafik und der Verfügbarkeit von entsprechenden digitalen Werkzeugen für Planer:innen erweiterte sich das Spektrum an Darstellungsmöglichkeiten. In den vergangenen Jahren wurden zunehmend aufwendigere Technologien wie VR und AR in Bürgerbeteiligungsprozessen genutzt, z. B. in Form virtueller Begehungen des Planungsgebiets im Rahmen von Beteiligungsveranstaltungen oder durch die Darstellung des Vorhabens als visuelle Überlagerung (Augmentierung) vor Ort mit einem Tablet. Durch die interaktive und räumliche Visualisierung ermöglicht XR eine bessere Verständlichkeit komplexer Planungsinhalte (vgl. Wolf et al., 2020: 10). Die genannten Technologien eröffnen die Möglichkeit einer direkteren und interaktiven Einbindung der Bürgerinnen und Bürger in Planungsprozesse (vgl. Gordon & Manosevitch, 2010: 75–95). Gleichwohl ist ihr Einsatz nach wie vor vergleichsweise selten.

Um die immersiven Möglichkeiten des im Rahmen des Projekts entwickelten XR-Systems optimal auszunutzen, wurden verschiedene 3D-Modelle erstellt sowie untersucht, inwiefern immersive Umgebungen bei Bürger:innen das Verständnis für räumliche Planungen und städtische Transformationsprozesse fördern. Im Folgenden wird dieser Prozess im Detail dargelegt.

5.1 Rahmenbedingungen der visuellen Modellerstellung

Die Visualisierung von 3D-Modellen im Kontext der Bürgerbeteiligung unterliegt drei wesentlichen Faktoren: dem Beteiligungsziel, der Datengrundlage und den technischen bzw. finanziellen Rahmenbedingungen, die zur Verfügung stehen. Technische Rahmenbedingungen beziehen sich auf die Verfügbarkeit und Wahl der Endgeräte (Mobiltelefon, Tablet, Laptop etc.), während die finanziellen Rahmenbedingungen unmittelbar mit dem Personeneinsatz für Entwicklung und Betreuung verbunden sind. Um diese Bedingungen berücksichtigen zu können, wurde der Ansatz des Reallabors gewählt, sodass zu bestimmten Terminen und Situationen entsprechende Prototypen entwickelt wurden. Durch den Ansatz, zwei zentrale XR-Technologien als XR-Part-Beteiligungstour (AR-App) bzw. XR-Part-Beteiligungsraum (virtuelle 3D-Welt) zu nutzen, ergaben sich innerhalb der Projektlaufzeit acht verschiedene Modellprototypen, die den oben genannten Faktoren (Beteiligungsziel, Datengrundlage sowie technische und finanzielle Rahmenbedingungen) unterworfen waren. Die Entwicklung der Prototypen beinhaltet verschiedene Modellierungs- und Darstellungsansätze der räumlichen Umgebung sowie einzelner Objekte.

Beim Beteiligungsziel stellt sich die grundlegende Frage: Was möchte man von den Bürgerinnen und Bürgern erfahren, welche Informationen möchte man ihnen vermitteln oder wie kann man sie zu einem aktiven Diskurs anregen? Aus der Richtline VDI-MT 7001 geht hervor, dass eine klare Faktenlage über das Mögliche gegeben sein muss (Verein Deutscher Ingenieure, 2021: 15) und, dass die Darstellungen das zeigen sollten, was in dem Verfahren gefordert wird und dem Planungsstand entspricht (Spieker et al., 2017: 61). Darüber hinaus lässt sich aus der in 2022 durchgeführten Studie von Boos et al. ableiten, dass ein höherer Detaillierungsgrad (Level of Detail [LoD]) zur Verständlichkeit der Modelle beiträgt (Boos et al., 2023: 71–74).

5.2 Der visuelle Kontext am Beispiel des ersten Modellprototyps

Diese beiden Erkenntnisse, nur das darzustellen, was dem Planungsstand entspricht, und mehr Details für ein besseres Verständnis bereitzustellen, bildeten den Grundsatz für die Visualisierung des ersten Prototyps (siehe Abbildung 7) im Beteiligungsverfahren »Neugestaltung Vorplatz Uhlandschule« in Mannheim ab dem Jahr 2022 (vgl. Mannheim, 2024). Ziel der Bürgerbeteiligung war es, ein erstes Stimmungsbild zu erfragen, welche Bedürfnisse im konkreten Stadtraum bestehen und welche Gestaltungswünsche bei den betroffenen Bürgerinnen und Bürgern vorhanden sind. Der Einsatz von XR-Technologie innerhalb dieser

frühen Phasen des Beteiligungsverfahrens wird durch das sogenannte Beteiligungsparadox erschwert, da es zu diesem Zeitpunkt des Verfahrens häufig nur wenige konkrete Inhalte gibt, die visualisiert werden können. Für ein interaktives und immersives Erleben des Beteiligungsgegenstandes, welches eine wesentliche Stärke von XR-Technologien ist, werden allerdings detaillierte 3D-Modelle benötigt (vgl. Wolf et al., 2020: 3 f.). Daher bestand der Ansatz darin, den Gegenstand der Beteiligung in eine möglichst nachvollziehbare virtuelle Umgebung einzubinden. Ziel war es, mithilfe verschiedener Techniken wie z. B. Modelage, UV-Mapping, Fotocollage eine Realitätsnähe zu erreichen, die den Ort des Beteiligungsverfahrens erkennbar macht und zugleich den technischen und ressourcenbedingten Vorgaben entspricht.



Abbildung 7: Ausschnitt aus dem XR-Part-Beteiligungsraum Mannheim 2022
Das helle Feld in der Bildmitte stellt die Fragestellung dar.
Quelle: Projekteigene Darstellung FHE AI 2024

5.3 Aufgabenbezogene Gestaltungsansätze

Die Erprobungen der jeweiligen Prototypen im Forschungsprojekt XR-Part wurden realpraktisch im Rahmen von Beteiligungsformaten als Teil eines Rahmenplanverfahrens zur Entwicklung der Südstadt Rostock durchgeführt. Die zeitliche Achse des Forschungsprojekts erforderte zwei parallele Entwicklungsstränge für die Ausarbeitung der Objektdarstellungen, um am gleichen Tag erprobt und

diskutiert werden zu können. Auf der einen Seite stand das beauftragte Planerkollektiv, auf der anderen Seite das Forschungstem von XR-Part, das nicht immer mit ausreichendem Vorlauf über mögliche Planungsergebnisse informiert werden konnte. Definiert durch die Rahmenbedingungen der Stadt und ein gemeinsames Auftaktgespräch wurden die einzelnen partizipativen Fragestellungen unterschiedlich visuell entwickelt.

Für das Verfahren in Rostock wurden die Erkenntnisse der visuellen Darstellungstechniken aus Mannheim fortgeführt. Auf Grundlage der von der Stadt definierten Referenzbilder wurden 3D-Objekte erstellt und in der Game Engine in eine gemeinsame Szene überführt. Hierbei wurde ein weiterer Entwicklungsschritt vorgenommen: Während bislang eine Gesamtszene für die jeweilige Erprobung erstellt wurde, entstanden in Rostock aufgrund der Größe und der Vielzahl darzustellender Objekte zwei verschiedene Szenen in unterschiedlichen Detaillierungsgraden mit sich überschneidenden Objekten – einmal für die AR-Tour und einmal für den virtuellen Beteiligungsraum. Die visuellen Unterschiede dieser beiden Entwicklungsschritte in Bezug auf die Qualität der Beteiligung werden in einem gesonderten Verfahren weiter untersucht.

Dabei ist zu betonen, dass unterschiedliche Dateiformate zusätzlich über die grundsätzliche Unterscheidung zwischen Polygonmodellen (Darstellung besteht aus ebenen Begrenzungsflächen) und NURBS-Modellen (hier werden Freiformen verwendet) hinaus für Export- und Transformationsherausforderungen sorgen können. Betrachtet man die Daten, die im Gesamtsystem der gebauten Umwelt entstehen, gibt es sowohl Überschneidungen als auch Herausforderungen beim Export von Informationen. Der Fokus von Daten aus Geografischen Informationssystemen (GIS), der großräumliche Bezug, unterscheidet sich grundlegend von dem der – im Vergleich tendenziell kleinräumlichen – Planungsdaten während des Bauprozesses, obwohl beide Elemente zu unterschiedlichen Zeitpunkten Teil des Gesamtsystems der gebauten Umwelt waren.

Darüber hinaus zeichnet sich ein Trend zur Entwicklung interaktiver virtueller Beteiligungswelten ab, bei denen weniger die planerische Genauigkeit als vielmehr die visuelle Attraktivität im Zusammenspiel mit ethischen, rechtlichen und sozialen Implikationen (ELSI) im Vordergrund steht. Dies ermöglicht eine vielfältige und realistische Darstellung von Menschen, wie sie derzeit nur mit 3D-Engines möglich ist, und gleichzeitig den direkten Import von Planungsdaten. Dieser Ansatz wurde in Ansätzen mit dem Import von IFC-Dateien aus CAD-Software erprobt und bietet vermutlich noch großes Potenzial für eine strukturierte Weiterentwicklung.

6. Fazit und Ausblick

Das XR-Part-Projekt hat durch viele praktische Erprobungen mit realen Bürger:innen wesentliche Erkenntnisse darüber gewonnen, wie transformative Technologien wie AR und VR Bürgerbeteiligungsprozesse unterstützen und bereichern können. Dieser Beitrag hebt die wichtigsten Herausforderungen und Lösungsansätze hervor, die sich bei der Anwendung dieser Technologien ergeben haben. Dazu zählen insbesondere die Verortung der Teilnehmenden in der realen Umgebung, die Auswirkungen von Witterungsbedingungen auf Tracking und Bedienbarkeit sowie die besonderen Anforderungen einer heterogenen Nutzergruppe. Zudem wurde der komplexe Prozess der Erstellung immersiver 3D-Inhalte beleuchtet, die trotz ihrer hohen Detailgenauigkeit den dynamischen Anforderungen eines Beteiligungsverfahrens gerecht werden müssen.

Die Ergebnisse zeigen, dass benutzerfreundliche Gestaltung und eine zugängliche technische Einführung Schlüsselfaktoren für den Erfolg von XR-gestützten Bürgerbeteiligungsprozessen sind. Während viele Teilnehmer:innen die Nutzung der Anwendungen als intuitiv empfanden, wird deutlich, dass technikferne und ältere Nutzer:innen zusätzliche Unterstützung benötigen. Trotz bedeutender Fortschritte in der Barrierefreiheit und Benutzerfreundlichkeit besteht in diesem Bereich weiterhin erhebliches Verbesserungspotenzial (vgl. Sassmannshausen et al., 2022: 250–265).

XR-Part demonstriert, dass XR-Technologien das Potenzial haben, Bürgerbeteiligung auf transformative Weise zu erweitern. Die orts- und zeitunabhängige Nutzung dieser Formate ermöglicht es Bürger:innen, ihre Teilnahme besser in ihren Alltag zu integrieren, während die immersive und interaktive Natur der 3D-Visualisierungen komplexe Planungsinhalte zugänglicher macht. Insbesondere die Veranschaulichung und interaktive Gestaltung von Planungsprojekten hebt XR-Anwendungen deutlich von konventionellen Beteiligungsformaten ab.

Abschließend lässt sich festhalten, dass XR-Technologien nicht nur die Inklusion und Reichweite von Bürgerbeteiligung erhöhen, sondern auch eine neue Qualität der Interaktion und Visualisierung bieten. Dennoch erfordern diese innovativen Ansätze weitere Forschung und Entwicklung, insbesondere zur Überwindung technischer und sozialer Barrieren. Die innerhalb von XR-Part gesammelten Erfahrungen und erzielten Fortschritte bilden eine solide Grundlage für zukünftige Projekte und belegen das transformative Potenzial digitaler Technologien, Bürgerbeteiligung nachhaltiger und inklusiver zu gestalten.

Förderhinweis

Das Projekt XR-Part wurde vom Bundesministerium für Bildung und Forschung (BMBF)[1] gefördert.

Literatur

Apple. (2024). *Verifying Device Support and User Permission*. Abgerufen am 25. September 2024, von https://developer.apple.com/documentation/arkit/verifying_device_support_and_user_permission

Beran, T., Langley, R., Bisnath, S., & Serrano, L. (2007). *High-Accuracy Point Positioning with Low-Cost GPS Receivers*. Navigation, 54: 53–63. Institute of Navigation, Manassas. https://doi.org/10.1002/j.2161-4296.2007.tb00395.x

Boos, U.C., Reichenbacher, T., Kiefer, P., & Sailer, C. (2023). *An augmented reality study for public participation in urban planning*. Journal of Location Based Services, 17(1), 48–77. Taylor & Francis, London. https://doi.org/10.1080/17489725.2022.2086309

Google (2024) *Mit ARCore unterstützte Geräte*. Abgerufen am 25. September 2024, von https://developers.google.com/ar/devices?hl=de

Gordon, E., & Manosevitch, E. (2010). *Augmented deliberation: Merging physical and virtual interaction to engage communities in urban planning*. Journal of Information Technology & Politics, 7(2-3), 75–95. Taylor & Francis, London.

IfD Allensbach. (2023). *Anzahl der Personen in Deutschland, die einen Tablet-PC (z. B. iPad) im Haushalt nutzen, von 2019 bis 2023 (in Millionen)*. In Statista. Abgerufen am 25. September 2024, von https://de.statista.com/statistik/daten/studie/268052/umfrage/tablet-pc-besitz-in-deutschland/

Kantar. (2024). *Marktanteile der mobilen Betriebssysteme am Absatz von Smartphones in Deutschland im 2. Quartal 2024*. In Statista. Abgerufen am 25. September 2024, von https://de.statista.com/statistik/daten/studie/198435/umfrage/marktanteile-der-smartphone-betriebssysteme-am-absatz-in-deutschland/

Mannheim. (2024). *Neugestaltung Vorplatz Uhlandschule*. Abgerufen am 28. November 2024, von https://mannheim-gemeinsam-gestalten.de/vorplatz-uhlandschule

Sassmannshausen, S.M., Radtke, J., Bohn, N., Hussein, H., Randall, D., & Pipek, V. (2021). *Citizen-centered design in urban planning: How augmented reality can be used in citizen participation processes.* Proceedings of the 2021 ACM Designing Interactive Systems Conference, 250–265. ACM, New York. https://doi.org/10.1145/3461778.3462075

Spieker, A., Wenzel, G., & Brettschneider, F. (2017). *Bauprojekte visualisieren: Leitfaden für die Bürgerbeteiligung* (Schriftenreihe der Baden-Württemberg Stiftung, Nr. 86). Baden-Württemberg Stiftung gGmbH, Stuttgart. https://www.bwstiftung.de/fileadmin/bw-stiftung/Publikationen/Forschung/Forschung_Bauprojekte_Visualisieren_Nr._86

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Statista. (2024). *Anteil der Smartphone-Nutzer* in Deutschland in den Jahren 2012 bis 2023 und Prognose bis 2030*. In Statista. Abgerufen am 25. September 2024, von https://de.statista.com/statistik/daten/studie/585883/umfrage/anteil-der-smartphone-nutzer-in-deutschland/

Verein Deutscher Ingenieure. (2021). *Kommunikation und Öffentlichkeitsbeteiligung bei Bau- und Infrastrukturprojekten: Standards für die Leistungsphasen der Ingenieure (VDI-MT 7001)*. Verein Deutscher Ingenieure, Düsseldorf.

Wolf, Mario & Söbke, Heinrich & Wehking, Florian. (2020). *Mixed Reality Media-Enabled Public Participation in Urban Planning*. In: Jung, T., tom Dieck, M.C., Rauschnabel, P.A. (eds) Augmented Reality and Virtual Reality. Progress in IS. Springer, Cham. https://doi.org/10.1007/978-3-030-37869-1_11

IV.
Transformative Technologien im Gesundheitsbereich

Human-Centered Design of Mixed Reality Applications in Medical Education – GreifbAR

Robert Luzsa, Moritz Queisner, Christopher Remde, Igor Sauer, Nadia Robertini and Susanne Mayr

Project and Chapter Description

The project »Tangible Reality – Interaction of Hands with Physical Tools in Mixed Reality« investigates how the user's hands can be integrated into virtual reality applications to learn and train work-related tasks with physical objects. As part of the »Interactive systems in virtual and real spaces« program funded by the German Federal Ministry of Education and Research (BMBF)[1], the project developed mixed reality (MR) prototypes to improve and facilitate the acquisition of practical skills. Based on a cooperation between surgery (Charité – Universitätsmedizin Berlin), psychology (Universität Passau), artificial intelligence (German Research Center for Artificial Intelligence) and interaction design (NMY Mixed-Reality Communication GmbH) we found that spatial and tangible interaction in mixed reality systems helped trainees to acquire practical knowledge in medical training when learning tasks such as tying knots or using instruments. This chapter first introduces the project. Then, it gives an overview of methods that were used to assess the perspectives of trainees and experts during the project. These are integrated into a generalized model for the human-centered design of MR-based learning tools that may inspire similar projects. Finally, three studies carried out as part of the project are presented and their results and implications discussed.

1 At the time of printing: Federal Ministry of Research, Technology and Space (BMFTR)

Abstract

The capabilities of mixed reality (MR) applications to combine visualizations with a user's real environment offer benefits for the training of complex surgical skills like the tying of surgical knots. By wearing a virtual reality headset combined with a passthrough function, users can see additional information that adapts to their body or physical objects, such as visual instructions to perform a work task or they can receive individualized feedback based on their movement. To ensure usability, user acceptance and effective learning outcomes of MR applications, it is essential to assess the knowledge of domain experts as well as to derive requirements from state-of-the-art training concepts. This chapter provides an overview of tools and challenges associated with integrating users into the development of medical MR applications. Drawing on findings and insights gained during the development of an MR-based knot-tying trainer in the Tangible Reality project, the chapter offers recommendations for effectively incorporating users into the design of similar MR applications.

1. Human-Centered Design for Surgical Skills Training

In the development of innovative digital applications, various approaches have emerged that highlight the importance of involving users and other stakeholders in the design and development process. The terms user-centered design (UCD) and human-centered design (HCD) were coined to describe a way of designing systems that are not primarily defined by the capabilities of a technology but instead optimized to fit the needs of its users (Norman, 1986: 32 ff.). A key consequence of this conceptualization of design is that users and stakeholders should not only evaluate a system once it is fully developed. Instead, they should be involved in all phases of the design process – from identifying needs, generating ideas, and developing concepts through to iterative testing of increasingly high-fidelity prototypes (International Organization for Standardization, 2019: 6; Stanton et al., 2013: 2 ff.).

The integration of user perspectives in all phases of development is particularly important in scenarios where specialized expert user groups are involved and where comprehensibility, precision and reliability of the system are critical. A prime example for this is the field of surgical skills training, where systems are developed to teach surgeons skills that are necessary, for example in laparoscopic surgery and knot tying (Hagelsteen et al., 2019: 309 ff.). To design effective surgical training applications, it is essential to consider not only explicit information about the surgical skills to be taught but also implicit standards in surgical prac-

tice which are not explicitly outlined in textbooks. These implicit standards encompass, for example, the knowledge and experience that experts utilize and rely on in their daily work. Additionally, it is crucial to understand the specific needs of the user group and the capabilities of the training technology used. Yet, what is the best way to obtain all this information, and what are the best methods for user involvement?

This chapter provides answers to this question by describing the process and methods of user involvement in the development of a mixed reality-based (MR) training application for surgical knot tying.

2. The Tangible Reality Project

The Tangible Reality project aims to bridge the gap between the real and virtual worlds. It focuses on the creation of mixed reality environments that are tangible and graspable, incorporating both virtual and augmented reality. Our major goal in the project is to enable users to interact with real objects, virtually recreated real objects and purely virtual objects using their bare hands seamlessly and naturally, without the need for additional controllers.

This chapter describes a case study that focuses on the development of an MR training application that teaches surgical knot-tying techniques to medical students. The skilled handling of sutures is one of the basics of any surgical procedure. Usually, surgical knotting techniques are learnt through personal instruction by experienced surgeons, diagrams, and video-based tutorials. Learning surgical knot tying techniques through participatory observation and direct instruction by experienced surgeons is costly and difficult to scale. This type of training is increasingly reaching its limits in day-to-day hospital operations, due to the changing economic, social, and regulatory conditions of surgical practice. Students and trainees, as well as professionals in training, are therefore faced with the problem of applying and practicing the theoretical knowledge they have acquired. Sometimes, universities offer workshops in which experienced surgeons demonstrate their knotting techniques. Although such workshops are didactically valuable due to the intensive support in small groups and individual feedback, they are time-consuming and resource-intensive, making them impractical on a larger scale.

Students may rely on textbooks for theoretical knowledge and videos for observational learning. Text- and image-based media allow a scalable acquisition of theoretical knowledge independent of time and place. Yet, this way of learning lacks interactivity, personalization, feedback, and expert guidance, making it insufficient for effective skill acquisition. Knot tying gestures and work steps can

only be passively observed and then imitated. Furthermore, learning success cannot be measured and verified. In the project, we concluded that MR-based training could combine the benefits of self-guided, autonomous, and instructor-led learning: Users could wear an MR headset while tying knots on their own, and the headset would display instructions and provide individualized feedback based on hand and finger tracking. This would allow users to train independently, without the need for an instructor, while still receiving individual guidance.

Yet, how should such an MR-based learning system be designed so that learners benefit from it, accept it, and enjoy working with it? And what information should be conveyed, that is, what knotting techniques surgeons use and teach in workshops, and with which means (visual, acoustic etc.) should the system present information and provide feedback?

These questions can only be answered by involving users and experts. For this reason, in the project, existing models and procedures of user involvement were first examined and integrated into a generalized model for the user-centered design of the MR-based learning systems. This model and its scientific foundations are presented in sub-chapter »Stages of Human-Centered MR Development«. Based on this, two studies to assess user and expert perspectives were conducted, which inspired the development of different design variants of the learning system. Finally, students rated and evaluated these variants in an online experiment. The results and the technical details of the developed systems are presented in the final sub-chapter »Case Study: Development and User Testing of a Mixed Reality Knot Tying Trainer«.

3. Stages of Human-Centered MR Development

3.1 Existing Models and Approaches

There are various procedural models for the human-centered development of products and services that can be applied in the context of MR-based surgical learning. However, they all focus only on certain parts of the development process and reflect the research and design tradition from which they originate. To effectively guide the development of surgical training applications, these perspectives must be integrated and combined. To achieve this, the following section introduces three models with different focuses, which are subsequently merged into a comprehensive model. This model encompasses all stages of design and development, from analyzing the current state to evaluating the individual and organizational impacts of application deployment.

3.1.1 ISO 9241–210

The ISO 9241–210 suggests four stages for the human-centered design of digital systems. These should be followed to achieve efficient and effective systems, usability, and well-being of users while minimizing adverse effects on their health, safety, and performance (International Organization for Standardization, 2019: 10 f.). The stages are:

1. *Context analysis:* The usage context of the systems needs to be defined. This includes, for example, characteristics of future users, activities to be performed, workplace conditions and environments, such as available space, lighting, and noise levels. Literature research, e. g., on similar projects and products, should also be conducted at this stage.
2. *Definition of requirements:* Concrete requirements are formulated, from the perspective of future system users, the organization in which the system will be deployed, and other stakeholders (e. g., patients in the domain of medicine).
3. *Design:* Based on contextual knowledge and requirements, increasingly high-fidelity prototypes are created, from pen-and-paper mock-ups and wireframes up to the final working product.
4. *Evaluation:* The evaluation is conducted iteratively, alternating with step three. Users and stakeholders test prototypes and provide feedback, which is integrated into design.

3.1.2 The Double Diamond Model

A similar approach is the double diamond model. It originates from design thinking and aims to involve users and stakeholders to solve problems and develop ideas that meet user needs and market demands (Lindberg et al., 2011: 4 ff.). In the model, divergent and convergent thinking alternate: In the first stage, ideas are generated (divergent *discover* stage), which are then selected (convergent *define* stage), after which solutions are developed (divergent *develop* stage) and finally implemented (convergent *deliver* stage). Hence the term double diamond, as the divergent stages can be visualized as arrows pointing outwards and the convergent stages as arrows pointing inwards, forming two diamonds when put together. The stages are iterative, meaning that stages can be repeated with additional insights until a satisfactory final product or system is reached.

Both models emphasize the importance of stakeholder integration and iterative development. Following these steps is especially important during the development of surgical training tools: Designers and developers are usually unfamiliar with surgical procedures and terms, and learners' surgical knowledge and experience need to be considered. Previous studies have shown that a human-cen-

tered design process can improve the usability of training applications in surgery, for example in ventriculostomy (Reyes et al., 2022: 52 ff.) and minimally invasive surgery (Freudenthal et al., 2011: 198 ff.).

3.1.3 The Kirkpatrick Evaluation Model

The models discussed so far provide helpful approaches for application development yet overlook one key aspect: What happens after a system is developed and deployed? How does it impact users and the organization in real-world work or learning contexts? This is especially crucial for medical and surgical learning applications: These should not only offer benefits for users and patients but must also be successfully integrated into the hospital or academic teaching organization. For this reason, a further model was used in the project that focuses on the successful roll-out of developments in real-world environments, namely Kirkpatrick's four-level evaluation model (Kirkpatrick, 1996: 55 ff.; Smidt et al., 2009: 267 ff.). It assesses four levels of effects:

1. *Reaction*: Users' immediate reactions to the training application.
2. *Learning*: The effectiveness of the system in teaching relevant skills.
3. *Behavior*: How users apply learned skills in their work.
4. *Results*: The organizational impact, such as increased productivity or reduced errors.

Combining these levels of effects with the development stages of the models described above, it becomes clear that reactions can primarily be measured in the context of early iterative user tests with mock-ups and prototypes. Learning effectiveness can be measured with summative evaluations with the final prototype or product. However, long-term changes in behavior and organizational outcomes can only be observed after the product has been rolled out in the organization and integrated into learning routines and contexts. For medical and surgical training scenarios, this means assessing the on-the-job performance of nurses or surgeons on tasks they have learned using a new training application. In addition, evaluations from different user groups can be gathered (e. g. patient satisfaction, employee satisfaction, supervisor ratings) and performance measures such as number of errors may be used (Praslova, 2010: 219 ff.). Based on this, recommendations can be made for the use of the system and any necessary steps such as further qualifications of users.

3.2 A Generalized Model of Human-Centered Design and Testing

The three models described above can be combined into a generalized, three-stage, iterative process model that encompasses all phases of the design, development, testing, and implementation of a system (see figure 1).

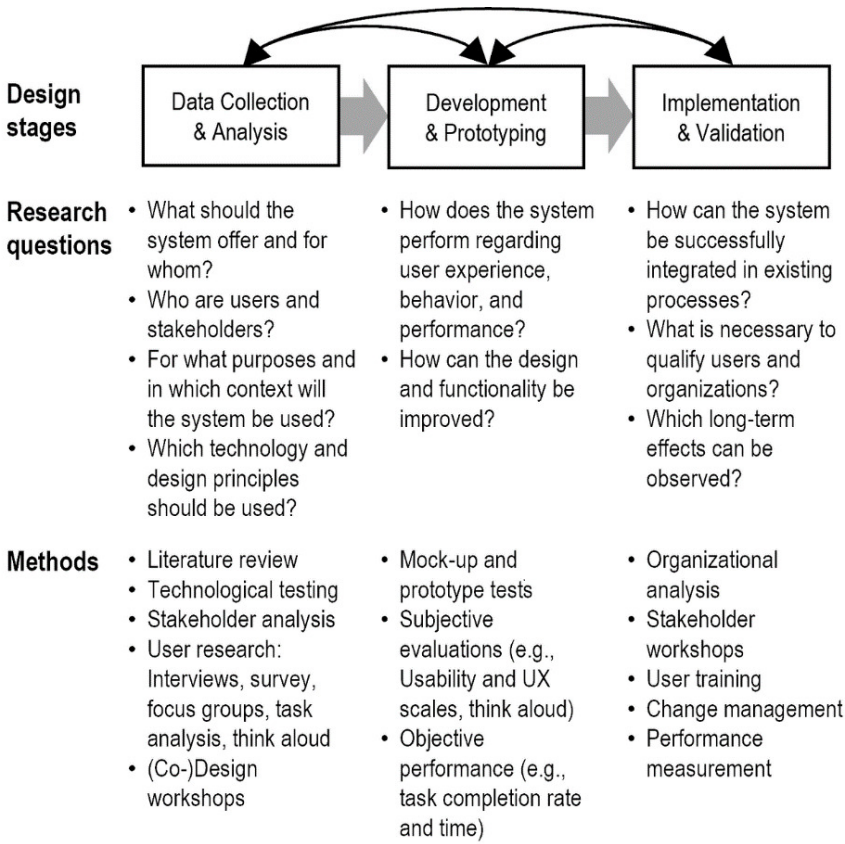


Figure 1: A generalized model of human-centered design and testing
Source: Own figure

The first stage of the model – data collection and analysis – encompasses all activities that focus on learning about users, their needs, perspectives, workflows, and usage scenarios, as well as exploring the technological possibilities for the development project. It summarizes the first two stages of the double diamond model and the ISO 9421–210. The second phase – development and prototyping – focuses on the actual development, ideally of different variants, which are then

subjected to user testing and compared concerning different acceptance and success criteria. It summarizes the third and fourth stages of the above models. In the final stage – implementation and validation – the system is introduced into working and learning environments, to measure its real-world impact on and to derive recommendations for further action. There may be feedback loops between the stages, for example lessons learned during development and testing which may require additional data collection and field research, or the identification of additional needs during implementation inspiring further developments. These feedback loops are shown as arrows in figure 1.

4. Methods for Human-Centered MR Development

Once the relevant stages of development have been identified, suitable methods need to be selected to address the research questions in each stage, for example observational techniques, user testing or ideation methods. In addition, measures of success must be selected that provide valid, objective, and reliable information about the performance and user acceptance of developed mock-ups and prototypes, usually in the form of standardized questionnaire scales (DeVellis & Thorpe, 2021: 6 ff.; Stanton et al., 2013: 11 ff.).

There are several overviews of methods for user-centered product development that researchers and designers of MR-based learning systems can refer to. Stanton et al. (2013: 39 ff.) describe methods from human factors research and engineering psychology that are suitable for analyzing user tasks, interaction patterns with technical systems, and testing prototypes and user interfaces. These methods include, for example, hierarchical task analysis, which observes and breaks down work tasks into smaller steps to optimize technological solutions, the human error template, which identifies possible sources of error in activities, and heuristic analysis, where usability experts evaluate a system using checklists and templates.

Stickdorn et al. (2018: 36 ff.) and Tidwell et al. (2020: 1 ff.) describe methods from design thinking, which primarily incorporate user perspectives into design decisions. Prominent methods here include customer journey maps, which record and visualize user interactions with a system, the development of personas, detailed characterizations of user groups regarding their characteristics (e. g., prior knowledge, skills, socio-demographics, requirements), and co-creation workshops, in which users design drafts for user interfaces.

The following four methods are described in detail as they are general and applicable to the development of almost any product or technological system. They also appear particularly suitable for the design of MR-based surgical training ap-

plications because their combined use can provide both extensive qualitative and quantitative insights during all phases of development and because they can be easily adapted towards different target groups such as expert surgeons or students.

4.1 Stakeholder Analysis

The first stage of design is to identify the stakeholders, that is, any person, group, or organization that is directly or indirectly involved or affected by the development project (Bryson, 2004: 22). This is especially important for a surgical training application, as stakeholders include not only trainee surgeons using the system but also experienced surgeons and teachers who may provide learning content, and patients, whose needs should also be reflected in training materials. A common method for identifying stakeholders is stakeholder mapping, where all possible stakeholders are brainstormed and clustered according to different criteria. The involvement of each stakeholder group in the development process is discussed, for example whether they should provide information or evaluate prototypes.

4.2 Think Aloud

Think aloud can be used in the first design stage to understand complex tasks, such as surgical skills, and in the second stage to evaluate users' interactions with prototypes. Participants perform a task or interact with a system while continually verbalizing their thoughts and impressions. The verbalizations are recorded and analyzed (e. g., via qualitative content analysis, see Kuckartz, 2022: 15 ff.), providing insights into detailed task descriptions for integration into a training application or lists of usability problems to be addressed.

Compared to post-hoc interviews, think aloud provides more comprehensive and detailed insights as participants verbalize their thoughts while performing tasks. It is also very well suited to understanding procedural knowledge and actions that experts perform automatically and that they would not normally verbalize because they take them for granted. This is the case in many medical fields, including surgical knot tying, where think aloud can help identify the knot tying steps and understand requirements for correct knotting.

Detailed instructions and probing questions are essential for think aloud, as participants are often not used to verbalizing their thoughts continuously. Researchers might demonstrate the concept by thinking aloud themselves (e. g.,

counting windows in a building) and prepare probing questions for when participants stop verbalizing (Charters, 2003: 71 ff.).

4.3 The Technology Acceptance Model

The Technology Acceptance Model (TAM) (Davis, 1989: 319 ff.) provides measures of success to determine users' acceptance of prototypes and their perceived strengths and weaknesses. Moreover, it can assess relevant characteristics of the user group.

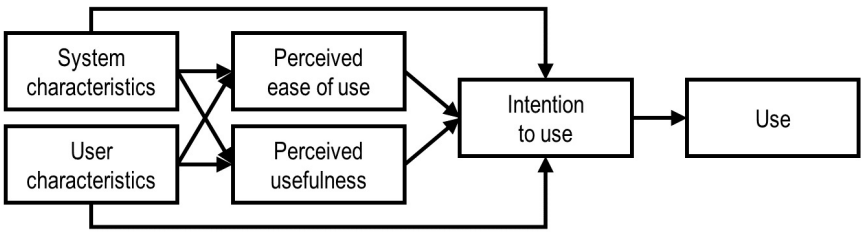


Figure 2: The technology acceptance model including influencing factors
Source: Own figure, adapted from Venkatesh & Bala, 2008: 276

TAM conceptualizes users' intention to use a system as determined by perceived ease of use (how easy it is to use and interact with the system) and perceived usefulness (how efficiently and effectively the system helps users achieve their goals, see figure 2). Using standardized scales, these facets of technology acceptance can be measured during prototype evaluation to identify weaknesses and areas for improvement, such as whether perceived ease of use is rated high or low.

Additionally, the facets of technology acceptance are influenced by factors that originate from both the users (e. g., how competent they rate themselves when using technology) and the system (e. g., how well it provides instructions and feedback). Measuring these influences can help to identify relevant user needs and user groups. For example, if users' acceptance depends on their technological competence, the system could be optimized to provide different interaction modes and functions tailored to the needs of more or less technologically competent users. A general overview of user- and system-related antecedents of technology acceptance and scales to measure them is given by Venkatesh and Bala (2008: 275 ff.) and Mlekus et al. (2020: 275 ff.). Moreover, for MR-based training applications, Graser and Böhm (2022: 3 ff.) provide a review of 33 studies

that identify factors influencing system acceptance, ranging from the system's interface style to learners' perceptions of self-efficacy and teachers' attitudes towards integrating technology into training. Granić (2022: 9733 ff.) also provides a general overview of research on factors influencing acceptance of educational technology.

4.4 The User Experience Questionnaire

Finally, the User Experience Questionnaire (UEQ) (Laugwitz et al., 2008: 63 ff.) is another standardized questionnaire scale, which can be used during prototype testing to get a more detailed picture of the usability and user experience of the system. Usability refers to the degree to which a product can support the user in a specific context to achieve goals in an effective, efficient, and satisfactory manner (International Organization for Standardization, 2018: 6). The concept of user experience is broader and encompasses user perceptions and reactions resulting from the use of a product, system or service (International Organization for Standardization, 2019: 7). It adds emotional and hedonic aspects such as enjoyment to the concept of usability. The standard version of the UEQ uses a semantic differential to measure six facets of usability and user experience: Attractiveness (e. g. annoying vs. enjoyable), perspicuity (e. g. not understandable vs. understandable), dependability (e. g. unpredictable vs. predictable), efficiency (e. g. obstructive vs. supportive), novelty (e. g. dull vs. creative), and stimulation (e. g. boring vs. exciting)

After averaging the items of each scale, a profile can be created and the rating of each aspect can be compared to identify strengths and weaknesses. The UEQ also facilitates the comparison of several variants, e. g. to find the best scoring variant of a user interface. Additionally, there are benchmark values for each scale that are based on 246 product evaluations (Schrepp et al., 2017a: 42 ff.). The results of a user test can be compared with the benchmark values and conclusions can be drawn as to whether the usability and user experience dimensions are significantly below average, and whether the system needs to be revised accordingly.

If only a general assessment of the user experience is required, and the 26 questions of the full UEQ are too comprehensive, the short version of the UEQ-S with only eight questions can be used as provided by Schrepp et al. (2017b: 105). This measures pragmatic and hedonic quality and an overall assessment of the system.

Some scales of UEQ and TAM measure closely related concepts (Mlekus et al., 2020: 276 f.), therefore it is recommendable to decide for each development

project which TAM and UEQ scales are most relevant, and to combine them into an appropriate questionnaire.

5. Case Study: Development and User Testing of a Mixed Reality Knot Tying Trainer

This chapter focuses on how the design and evaluation methods were applied in the Tangible Reality project for the development of an MR-based knot tying trainer. It gives an overview of three studies that took place in stage one (data collection and analysis) and stage two (development and prototyping) of development.[2] It also describes the MR-based learning system and how it builds on findings from user research and available technology.

5.1 Status Quo: Students' and Experts' perspectives on surgical knot tying

5.1.1 User Study: Assessing students' perspectives

The first study investigated how medical students currently learn to tie surgical knots, which knot-tying tasks and techniques they find most important and difficult, and what they expect from an MR-based knot-tying training application.

Method: We conducted an exploratory online questionnaire study with 80 medical students from Charité – Universitätsmedizin Berlin in October and November 2022. The participants were mostly female (56 female, 22 male, 2 diverse), with a mean age of 24.56 years ($SD = 4.5$, *range* = 18–41). Most were in the practice-oriented phase of their studies (44 in the fourth to tenth semesters), with the rest in the theory-oriented phase (18 in the first to fourth semesters) or the final practical year (16).

Participants were informed about the aims and procedures of the study and gave their consent for their responses to be used for scientific purposes. They first answered closed-ended questions on socio-demographics, study phase, knot-tying experience, learning and training methods, preferred types of training, and openness to and perceived advantages and disadvantages of MR-based training. Those with previous knot tying experience also answered open-ended questions about the most challenging and important steps in knot tying.

2 We cannot report on results in implementation and validation as development is still going on at the time of writing.

The open responses were evaluated according to Kuckartz's content-structuring approach (2022: 129 ff.) by first categorizing the statements and then counting the frequencies of the categories.

Results: Most students initially learned knot tying by attending a workshop where instructors give guidance and feedback on knotting (see figure 3A): 59 out of 80 students had already attended such a workshop, while the rest mostly learned from online videos ($n = 7$ mentions) or had no prior knotting experience ($n = 9$). However, many students who had attended a workshop did not continue practicing knot tying afterwards ($n = 24$, see figure 3B). Of those who did continue practicing, most did so on strings attached to objects ($n = 21$) or from re-tying knots shown in tutorial videos ($n = 17$), and only rarely on professional practice benches ($n = 14$) or on patients ($n = 8$).

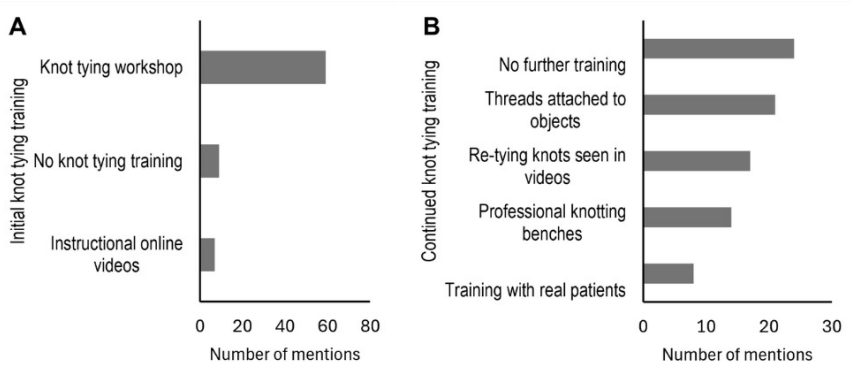


Figure 3: Reported frequency of strategies used in initial knot training (3A) and continued knot training (3B)
Source: Own figure

Regarding prior experience with simulators in general and MR, 33 reported having interacted with non-MR medical simulators, whereas only 15 had been using MR or VR at all. Despite this lack of first-hand experience, students were quite open to learning and training medical and surgical tasks with simulators ($M = 4.59$, $SD = 0.57$, on a scale of 1 to 5, with 5 indicating high openness). They were also rather interested in learning knot tying with MR ($M = 4.18$, $SD = 1.00$), citing the possibility of realistic yet risk-free training ($n = 16$), the ability to train repeatedly and at any time ($n = 7$) and a general fascination with this new technology ($n = 18$) as main advantages. Still, they considered instructor support essential for successful learning ($M = 4.06$, $SD = 0.89$) and mostly did not agree that simulators could be used to train without human instructors ($M = 3.10$, $SD = 1.12$).

Other responses also underlined the impression that students value the personal guidance and feedback from an instructor: Learning with instructor guidance was rated as the most helpful learning method ($M = 4.84$, $SD = 0.37$), while independent practice on training benches without further aids and independent tying of knots were rated as the least helpful ($M = 2.84$, $SD = 1.07$ and $M = 3.11$, $SD = 0.88$). Students also expressed concerns that an MR-based training application could not fully substitute human instructors and their direct support ($n = 5$ mentions).

The most important and most difficult steps in knot tying, according to the students, were tight and correct thread guidance throughout the entire tying process ($n = 11$), correct hand coordination ($n = 10$), correct tightening of the partial and complete knots ($n = 9$), and correct tying of counter knots, i.e., knots tied in opposite directions to stabilize the overall knot ($n = 8$).

Conclusions: The first survey showed that an MR-based system could support knot-tying training, which is hardly done systematically, especially outside of workshops. Students are open and interested but want the system to provide comprehensive support, guidance, and individualized feedback similar to a human instructor. It should teach, among other things, correct thread handling, hand coordination, knot tightening, and the correct sequence of counter knots. Additionally, the system must be user-friendly, especially for users without MR experience.

5.1.2 Expert Study: Analyzing Surgeons' Knot Tying Techniques

The second study focused on the perspectives of practicing surgeons as knot tying experts. It aimed to identify the knot tying technique used in surgical practice, the individual steps of the knot tying process, and experts' recommendations and requirements for an MR-based knot tying trainer.

Method: Four expert interviews including think aloud were conducted with physicians from Charité – Universitätsmedizin Berlin (three male, one female; two assistant physicians, two surgeons; self-reported number of knots tied between 100 and 1,000,000) in November and December 2022.

Following Wassermann's (2015: 55 ff.) recommendations for expert interviews, a semi-structured format was used and recorded via audio. The interviews covered three topics: the experts' knot tying techniques, quality criteria for surgical knots, and their teaching experiences and recommendations. The interviewees also demonstrated knot tying using a knot bank and braided threads, thinking aloud and explaining their steps while being recorded on video. For the second knot, they wore the Varjo XR-3 XR glasses used in the project to allow them to provide feedback, for example on its weight and pass-through display quality.

Data analysis involved transcribing the interviews and applying qualitative content analysis using a mixed deductive-inductive approach, with three top-level categories (knot technique, quality criteria, teaching recommendations) and additional sub-categories. Two independent coders developed a category system from one interview, which was used to code the remaining interviews. Video recordings were analyzed to identify knot tying steps.

Results: All experts reported using Granny Knots, which is a kind of slip knot formed by first crossing the two threads, and then tying two knots. Both knots should be tied in the same direction, but the initial crossing of the threads results in the first knot being inverted. Opposing knots are necessary to achieve a neat and flatly packed knot, minimizing irritations of the tissue. It can be tightened further after tying, which prevents it from loosening easily and makes it resistant to untight air knots (see figure 4). Finally, it is secured by one or more counter knots each tied in the opposite direction to the previous one. The experts report to use up to four counter knots when using monofilament threads but usually only one when using polyfilament threads.

Three of them preferred to tie the first two knots of the Granny Knot with the middle finger and the opposite knot with the index finger, and one did it the other way round. Two experts emphasized the importance of crossing the threads at the beginning, to avoid twisting the first knot and reduce the force necessary for tightening the knots. Two experts tie knots with their dominant, two with their non-dominant hand, mostly because there is less risk of injury when using a needle.

Regarding quality criteria, the interviewees identified two main categories: Criteria regarding the knot tying process, and criteria regarding the final knot.

During the knot tying process, maintaining thread tension is crucial for keeping the tissue in position ($n = 4$). Keeping the correct course of the thread is important, as a twisted run increases the force needed to tighten the knot ($n = 2$). Movements and tensile force should be minimized to avoid tissue damage ($n = 3$). Additional aspects include high tying speed ($n = 1$), injury-free needle handling ($n = 2$), and awareness of the specifics of the used thread material ($n = 3$).

For the final knot, the strength of the knot is paramount. It should not come undone ($n = 4$) and should fit tightly ($n = 2$). The knot should also consist of as few individual knots as possible to reduce foreign material in the body and potential inflammation ($n = 1$).

Conclusions: Based on these expert insights, it was decided that the MR-based knot tying trainer should teach a knot consisting of two single knots in the same direction, tied with the middle finger and secured by a counter knot with the index finger, using a polyfilament thread. A visual indicator of tensile force was developed to indicate if too much force is being used.

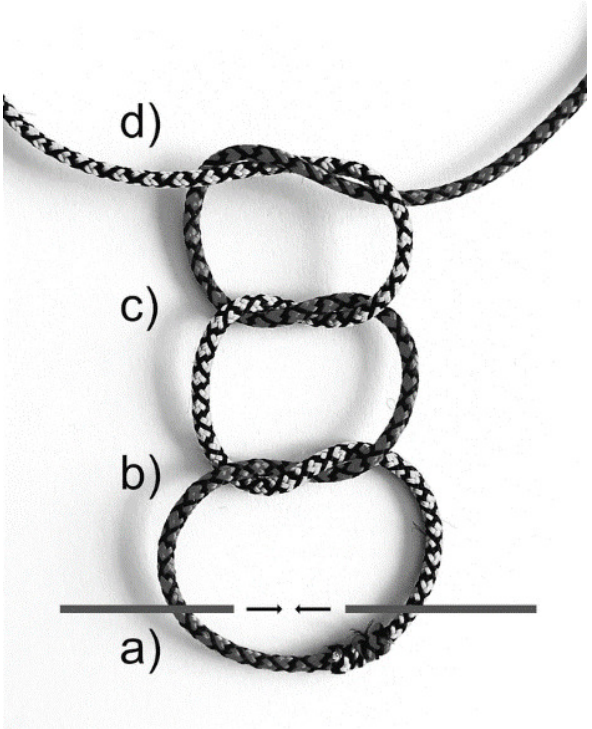


Figure 4: Granny Knot used by experts and taught by the MR training application.
a) Subcutaneous thread section. Tissue layer symbolized by red lines.
b) First partial knot, tied forwards with middle finger technique. c) Second partial knot, tied backwards using the middle finger technique.
d) Third partial knot, tied forwards using the index finger technique.
Source: Own figure

5.2 Implementation: Developing the MR-based Surgical Knot Tying Trainer

Based on the requirements analysis with students and experts, a concept for the digital knot tying trainer was developed and made available to the developers in the form of a design document. The studies' findings suggested the development of a prototype that combines real-world physical elements with virtual instructions and automated feedback to meet students' needs for a realistic and practice-based yet risk-free and interactive system. The derived learning concept is based on three main principles: Firstly, the learning content should be displayed three-dimensionally, as knot tying requires a particularly distinctive presentation of spatial relations. Secondly, the learning instructions should be applied to the

learner's own body and movements, which is achieved by spatially mapping the instructions on the learner's hands. And thirdly, the learning content should be automatically adapted to the learner's progress and learning speed. A prototype app, based on this concept was implemented by the developers, as seen in figure 5.



Figure 5: A student uses the knot tying trainer prototype implementation.
The volumetric video-based learning instructions are visible in the center. Elements for controlling the application are seen on the right side.
Source: Own figure

The prototype of the knot tying trainer consists of three components:

a) Knotting bench: Learners use a physical knotting bench with two strings that must be knotted together with a provided thread. The bench can digitally measure and display the thread tension and tensile strength of the knot during the exercise, allowing learners to avoid excessive pull that could damage patients' tissue.

b) VR goggles with hand tracking: We use Varjo XR3 VR goggles, a state-of-the-art VR/MR headset that offers a high resolution and good visual fidelity for the stereoscopic visualization and three-dimensional capture of the training scene.

Using the goggles' pass-through function, the learner's real environment can be overlaid with virtual elements. The hand tracking technology UltraLeap accurately registers the hand and finger movements of users in real-time.

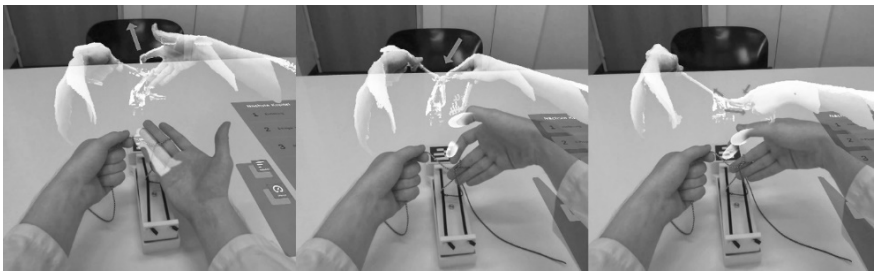


Figure 6: An exemplary sequence of the volumetric video-based learning instructions, in the order from left to right.
The arrows highlight the movements which need to be performed.
Source: Own figure

c) Learning software: The digital learning application provides step-by-step instructions in the form of virtual instructor hands that demonstrate the Granny Knot favored by experts. For this purpose, volumetric videos of experts' knot tying techniques were recorded and integrated into the system, superimposed on the learner's movements (figure 6). The steps can then be imitated by the learner using visual instructions mapped on the hands via superimposed arrows and a voice generated by text-to-speech (figure 7). The system uses hand and finger tracking to recognize individual learning progress based on key gestures and adapts the learning content accordingly. During a knot tying task, the application can overlay visual feedback over learners' fingers, for example by color-highlighting wrongly positioned fingers. The system also provides comprehensive onboarding and step-by-step guidance to assist users without prior MR experience.

During development, questions arose that could only be answered by user testing: It was necessary to know how much visual feedback is useful – do instructor hands, arrows, and colored highlights on the learner hands all support learning, or should individual elements be removed to avoid distraction and visual clutter? Several design variants were developed for this purpose, which included all or only individual elements and were to be evaluated by users. Additionally, we wanted to receive general feedback on the system and improvement potential from the user's perspective.

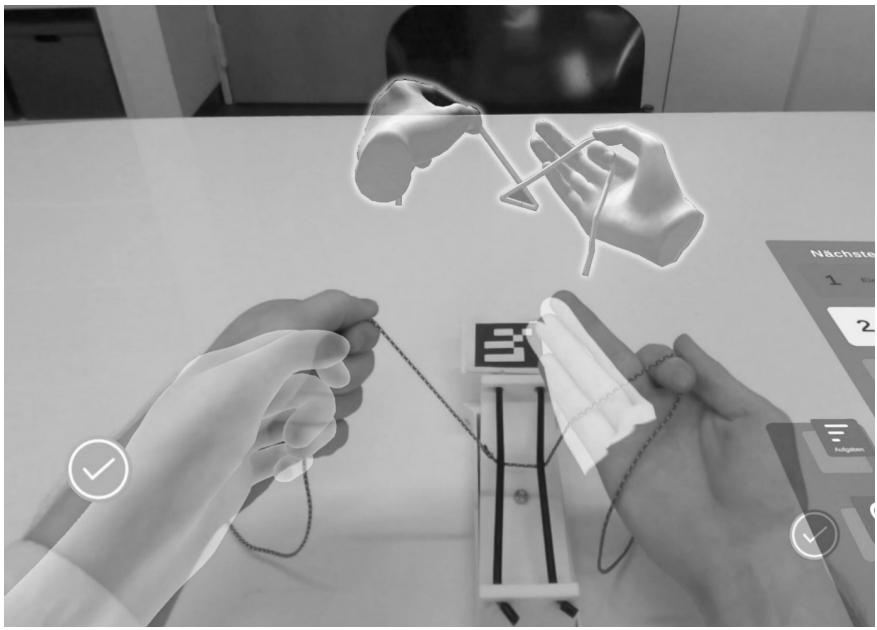


Figure 7: Superimposed instructions and information, which is projected onto the hands of the user. The green augmentation on the left indicates that the user has successfully executed a step. The blue overlay highlights fingers which need to be moved to perform the next step.
Source: Own figure

5.3 Testing: Online Evaluation of the Proof-of-Concept Prototype

Therefore, in a third, experimental study, three design variants of the MR-based learning system were evaluated by medical students at Charité – Universitätsmedizin Berlin. The goal was to identify the best design for further development and to get feedback on the system's usability, user experience, and learnability.

Method: In November 23, fifty-five medical students participated in the online experiment, of whom 46 remained after exclusions (30 women, 16 men, M_{age} = 23.5 years, *SD* = 4.6, *range* = 19–42). Participants watched videos of the app from the viewpoint of someone wearing mixed reality glasses. A between-subject experimental design with three conditions was used. Participants were randomly assigned to one of three app variants:

– *Complete:* The app included virtual instructor hands, verbal instructions, visual feedback on tensile force, arrows indicating direction of movement, and hand highlighting (*n* = 11).

– *Without arrows:* The app included all elements except arrows ($n = 16$).
– *Basic:* The app included only basic elements, no arrows, and no highlighting ($n = 19$).

Participants first answered questions on socio-demographics and knot tying experience. Then, they watched the assigned video, split into three parts: An interactive tutorial (2 min), knot bank preparation (2 min), and the first steps of knot tying (1 min). After each part, they evaluated the functions and elements of the application using rating scales and open questions.

Dependent variables included user experience (short User Experience Questionnaire [UEQ-S]) and willingness to use (taken from the TAM), using 5- and 7-point Likert scales. Open questions addressed strengths, weaknesses, and improvements for the system.

For analysis, mean values of dependent variables were calculated and compared descriptively. Due to the exploratory nature of the analysis and the small sample size, only descriptive statistics are reported. Open-ended responses were categorized inductively.

Results: Regarding user experience, the complete design variant was overall rated best ($M_{UEQS} = 2.20$, $SD_{UEQS} = 0.27$, with values ranging between -3 and 3) and better than the variant without arrows ($M_{UEQS} = 1.90$, $SD_{UEQS} = 0.73$) and the basic variant ($M_{UEQS} = 1.58$, $SD_{UEQS} = 0.89$). A similar pattern of differences emerged for pragmatic and hedonic quality (see figure 8). When compared to the UEQ-S benchmark values, all variants are ranked at least above average, with the complete variant and the one without arrows ranked as excellent.

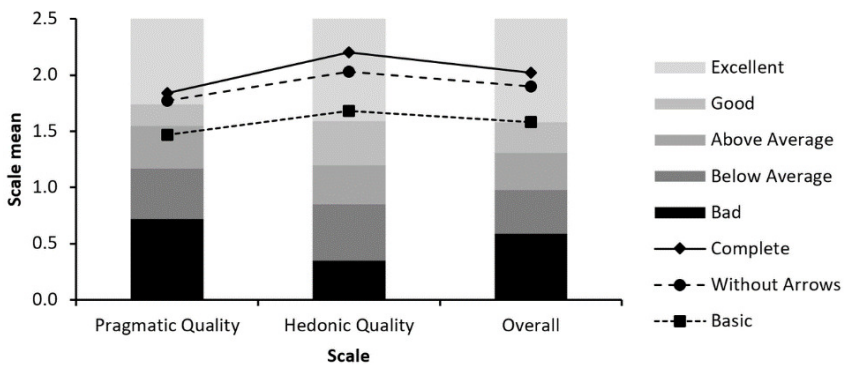


Figure 8: Benchmark of UEQ-S mean values of the three design variants
The colored bars indicate the 25th, 50th, 75th and 90th percentile of values in the benchmark sample of studies to which the values of this study are compared.
Source: Own figure

The most helpful system element, according to participants' ratings, was the visual feedback on tensile force, that is, whether sufficient or too much tension is applied to the thread ($n = 19$). Other positive features mentioned were the virtual instructor hands that demonstrated the knot ($n = 7$) and the colored highlighting of the users' hands when mistakes are made ($n = 6$).

The main point of criticism was the poor visual representation of the virtual instructor hands and the virtual thread. Respondents wished for higher contrast, more defined, sharper, less shaky visualizations, which also take color blindness into account ($n = 22$). Participants criticized the text-to-speech-based computer voice as too monotonous, slow and detrimental for concentration and wished for a more human sounding voice ($n = 14$). Concerning the user interface, participants wished for more self-explanatory or better labelled buttons ($n = 7$), a menu to select the knot to be trained ($n = 3$) and overall, for a more visually appealing and modern design ($n = 3$).

Participants reported the greatest intention to use for the variant without arrows ($Mintention = 4.62$, $SDintention = 0.62$, with values ranging between 1 and 5) and the complete variant ($Mintention = 4.36$, $SDintention = 0.50$), and a much lower for the basic variant ($Mintention = 3.53$, $SDintention = 1.39$).

Conclusions: The system variant with all elements – instructor hands, arrows and highlighting of user hands – was rated best and was therefore developed further in the project. To improve the display quality of the hands, new volumetric videos with additional cameras were recorded. Additionally, high-resolution still images of 3D modelled hands and threads were added to the instruction videos, illustrating the final state of a knot step in greater detail than still images from volumetric video.

6. Discussion and Conclusion

This chapter explored the human-centered design of a mixed reality (MR) application for surgical knot tying, integrating different design models and methods, and reporting on three studies conducted in the Tangible Reality project.

A generalized iterative process model for the design, development, and evaluation of medical and surgical training applications was proposed, which integrates the ISO 9241–210, the double diamond model, and the Kirkpatrick evaluation model. This model emphasizes understanding user contexts, defining requirements, iterative prototyping, and evaluating both short- and long-term effects of the system.

Based on this model, an overview of methods for integrating users and stakeholders in the design and development process was given, highlighting four

methods that appear to be particularly suitable for development in medical training: Stakeholder analysis, think aloud, and questionnaires measuring technology acceptance and user experience.

Three studies illustrated the use of these methods in the context of the development of an MR-based surgical knot tying trainer: Inspired by an initial stakeholder analysis, study one measured medical students' requirements and perspectives on MR-based knot tying training, while study two assessed experienced physicians' knot tying strategies and common quality criteria. These were integrated in different design variants of the training application. In study three, first-person video mock-ups of three design variants were rated by medical students in terms of user experience and acceptance to identify the best variant and additional requirement for further design.

In conclusion, the findings of the studies and their interplay illustrate that a human-centered, iterative design approach is essential for developing effective MR-based medical training tools. By taking into account stakeholder perspectives and integrating user feedback and expert insights throughout the design process, the development of training solutions that enhance the skills and confidence of medical professionals can be ensured.

Funding information

The Project GreifbAR is funded by the German Federal Ministry of Education and Research (BMBF)[3] *within the funding line »Interactive systems in virtual and real spaces – Innovative technologies for the digital society«.*

Literature

Bryson, John M. (2004). What to do when Stakeholders matter: Stakeholder Identification and Analysis Techniques. In: *Public Management Review,* vol. 6(1), p. 21–53. https://doi.org/10.1080/1471903041000167572

Charters, Elizabeth (2003). The Use of Think-aloud Methods in Qualitative Research. An Introduction to Think-aloud Methods. In: *Brock Education Journal,* vol. 12(2), p. 68–82. https://doi.org/10.26522/brocked.v12i2.38

Davis, Fred D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. In: *MIS Quarterly,* vol. 13(3), p. 319–340. https://doi.org/10.2307/249008

3 At the time of printing: Federal Ministry of Research, Technology and Space (BMFTR)

DeVellis, Robert F. & Thorpe, Carolyn T. (2021). *Scale Development: Theory and Applications*, SAGE Publications, Thousand Oaks.

Freudenthal, Adinda et al. (2011). Collaborative co-design of emerging multi-technologies for surgery. In: *Journal of Biomedical Informatics*, vol. 44(2), p. 198–215. https://doi.org/10.1016/j.jbi.2010.11.006

Granić, A. (2022). Educational Technology Adoption: A systematic review. In: *Education and Information Technologies*, vol. 27(7), p. 9725–9744. https://doi.org/10.1007/s10639-022-10951-7

Graser, Stefan & Böhm, Stephan (2022). A Systematic Literature Review on Technology Acceptance Research on Augmented Reality in the Field of Training and Education. In: *CENTRIC 2022: The Fifteenth International Conference on Advances in Human-Oriented and Personalized Mechanisms, Technologies, and Services, 2022*, vol. 12, IARIA, p. 20–28. https://doi.org/10.48550/arXiv.2411.13946

Hagelsteen, Kristine et al. (2019). Performance and perception of haptic feedback in a laparoscopic 3D virtual reality simulator. In: *Minimally Invasive Therapy & Allied Technologies*, vol. 28(5), p. 309–316. https://doi.org/10.1080/13645706.2018.1539012

International Organization for Standardization (2018). *Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts. (ISO Standard No. 9241–11:2018)*. https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en [Zugriff 20.03.2025]

International Organization for Standardization (2019). *Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems (ISO Standard No. 9241–210:2019)*. https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en [Zugriff 20.03.2025]

Kirkpatrick, Donald (1996). Great ideas revisited. Techniques for Evaluationg Training Prgrams. Revisiting Kirkpatrick's Four-Level-Model. In: *Training & Development*, vol. 50(1), p. 54–57.

Kuckartz, Udo (2022). Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung, Beltz Juventa, Weinheim, Basel.

Laugwitz, Bettina, Held, Theo & Schrepp, Martin (2008). Construction and Evaluation of a User Experience Questionnaire. In: *HCI and Usability for Education and Work*, Springer, Berlin, Heidelberg, p. 63–76. https://doi.org/10.1007/978-3-540-89350-9_6

Lindberg, Tillmann, Meinel, Christoph & Wagner, Ralf (2011). Design Thinking: A Fruitful Concept for IT Development? In: *Design Thinking*, Springer, Berlin, Heidelberg, p. 3–18. https://doi.org/10.1007/978-3-642-13757-0_1

Mlekus, Lisa et al. (2020). How to raise technology acceptance: User experience characteristics as technology-inherent determinants. In: *Gruppe. Interaktion. Organisation. Zeitschrift Für Angewandte Organisationspsychologie (GIO)*, vol. 51(3), p. 273–283. https://doi.org/10.1007/s11612-020-00529-7

Norman, Donald Arthur (1986). Cognitive Engineering. In: *User Centered System Design*, CRC Press, Boca Raton, p. 31–62.

Praslova, Ludmila (2010). Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in Higher Education. In: *Educational Assessment, Evaluation and Accountability*, vol. 22(3), p. 215–225. https://doi.org/10.1007/s11092-010-9098-7

Reyes, Jonatan et al. (2022). User-Centered Design for Surgical Innovations: A Ventriculostomy Case Study. In: Ethical and Philosophical Issues in Medical Imaging, Multimodal Learning and Fusion Across Scales for Clinical Decision Support, and Topological Data Analy-

sis for Biomedical Imaging, Springer, Cham, p. 51–62. https://doi.org/10.1007/978-3-031-23223-7_5

Schrepp, Martin, Hinderks, Andreas & Thomaschewski, Jörg (2017a). Construction of a Benchmark for the User Experience Questionnaire (UEQ). In: *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4(4), p. 40–44. https://doi.org/10.9781/ijimai.2017.445

Schrepp, Martin, Hinderks, Andreas & Thomaschewski, Jörg (2017b). Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). In: *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4(4), p. 103–108. https://doi.org/10.9781/ijimai.2017.09.001

Smidt, Andy et al. (2009). The Kirkpatrick model: A useful tool for evaluating training outcomes. In: *Journal of Intellectual & Developmental Disability*, vol. 34(3), p. 266–274. https://doi.org/10.1080/13668250903093125

Stanton, Neville A. et al. (2013). *Human Factors Methods: A Practical Guide for Engineering and Design*. Ashgate Publishing Ltd., Farnham.

Stickdorn, Marc et al. (2018). This Is Service Design Doing: Applying Service Design Thinking in the Real World, O'Reilly Media Inc., Sebastopol.

Tidwell, Jennifer, Brewer, Charles & Valencia-Brooks, Anne (2020). *Designing Interfaces: Patterns for Effective Interaction Design (3rd Edition)*, O'Reilly Media, Sebastopol.

Venkatesh, Viswanath & Bala, Hillol (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. In: *Decision Sciences*, vol. 39(2), p. 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Mixed Reality zum Anfassen: Innovation für virtuelle Haptik durch Gewichts- und Kraftsimulation – interaktionsKRAFT

Apostolos Vrontos, Sophie Lambert, Martin Klasen, Saša Sopka, Lina Vogt, Alexander Mertens und Christopher Brandl

Zusammenfassung

Die Immersion bei Anwendungen der virtuellen (VR), erweiterten (AR) und gemischten Realität (MR) ist häufig dadurch eingeschränkt, dass Einflüsse und Effekte auf die Nutzenden nicht auch haptisch wahrnehmbar sind. Könnten etwa verschiedene Gewichte oder Berührungen simuliert werden, könnte das Eintauchen in die virtuellen Welten für Menschen noch realer gestaltet und ein besseres Situationsbewusstsein ermöglicht werden. Die hierfür benötigten sogenannten propriozeptiven Empfindungen können mittels Elektromyostimulation (EMS) der Muskeln generiert werden. Hierbei bedarf es einer komplexen Steuerung elektrischer Signale, damit propriozeptive Empfindungen (z. B. ein Gewicht von 3 kg einer »virtuellen Hantel«) zielgerichtet erzeugt werden können. Anhand von zwei konkreten Nutzungsszenarien wird die Umsetzung einer solchen innovativen Haptikkomponente dargestellt. Zum einen wird die Simulation virtueller Gewichte mittels EMS für eine VR-Anwendung dargestellt. Hierzu sind die Antagonisten der bei der Haltearbeit beteiligten Muskeln zu stimulieren. Zum anderen wird die Korrektur der Bewegungsausführung bei Kniebeugen in einer virtuellen Physiotherapie beschrieben, bei der die EMS um Bewegungsinformationen durch Motion Capture erweitert wird. Hierfür wird anhand eines Lehr-Lern-Konzepts für das Extended-Reality(XR)-basierte Bewegungslernen aufgezeigt, wie immersive VR-, AR- oder MR-Systeme angewendet werden können, um motorische Fähigkeiten gezielt und speziell im Anwendungsbereich der Physiotherapie zu üben. Die Simulation von Gewicht und Kraft durch elektrische Muskelstimulation erweitert haptische Empfindungen und verbessert damit die Immersion in MR-Anwendungen. Unabhängig von einer reinen Nutzung für VR-, AR- und MR-Anwendungen kann diese Technologie auch für weitere Anwendungen eingesetzt werden, bspw. als haptische Anzeige in Mensch-Technik-Schnittstellen, für das Krafttraining ohne reale Gewichte und zur Stabilisierung von Körperhaltungen bei dynamischen Krafteinwirkungen.

1. Einleitung

Die physischen Interaktionsmöglichkeiten in der Mixed Reality (MR), einschließlich VR und AR, sind bislang noch stark eingeschränkt. Vor allem die haptische Wahrnehmung, die eine wesentliche Rolle für die Immersion in virtuellen Umgebungen spielt, stellt eine große Herausforderung dar. Haptik ist für die Nutzenden ein entscheidendes Element, das die virtuelle Welt greifbar und erlebbar macht. In physischen Interaktionen ist das haptische Feedback grundlegend, um das Gewicht, die Textur und die Widerstände von Objekten wahrzunehmen und somit die eigene Bewegung zu steuern. In der virtuellen Realität fehlt diese Rückmeldung jedoch größtenteils oder ist nur rudimentär vorhanden.

Derzeitige VR- und AR-Lösungen bieten kaum Optionen, um die propriozeptive Wahrnehmung – also das Bewusstsein für die eigene Körperposition, Bewegung und die Einwirkung externer Kräfte auf den Körper – realistisch zu simulieren. Stattdessen liegt der Fokus auf visuellen und auditiven Rückmeldungen, wodurch die haptische Dimension vernachlässigt wird. Die Fähigkeit, die eigene Position im Raum und die auf den Körper wirkenden Kräfte wahrzunehmen, ist jedoch entscheidend, um die virtuelle Welt als real und überzeugend zu erleben. Dies führt dazu, dass viele immersive Systeme nicht das volle Potenzial der virtuellen Interaktion ausschöpfen.

Ein Beispiel für innovative Ansätze, die diese Lücke zu schließen versuchen, ist der Einsatz von Echtzeit-Feedbacksystemen, die auf markerlosem Motion Capturing basieren. Eine Studie von Brandl et al. (2022) demonstriert die Nutzung eines Echtzeit-Feedbacks zur Verbesserung der Hebetechnik in manuellen Materialhandhabungsaufgaben, wobei die L5/S1-Kompressionskraft gemessen wurde. Solche Systeme können auch auf VR/AR-Anwendungen übertragen werden, um physikalische Kräfte und Belastungen für die Nutzenden nachvollziehbar zu machen und so ein realistisches Feedback zu erzeugen. Gleichzeitig zeigt eine andere Untersuchung die Eignung solcher markerlosen Systeme für ergonomische Risikoanalysen, indem die Genauigkeit der Bewegungsinformationen überprüft wurde (Brunner et al., 2022). Dies belegt die Anwendbarkeit solcher Technologien in der Praxis.

Gerade die Haptik ist ein entscheidender Bestandteil eines immersiven VR-Systems, da sie die Interaktion mit virtuellen Objekten intuitiver und effizienter gestaltet. Studien haben gezeigt, dass haptisches Feedback, sei es aktiv durch mechanische Stimulationen oder passiv durch sensorische Eingaben, das Gefühl der »Präsenz« in der virtuellen Welt signifikant verstärkt. Das Empfinden, tatsächlich in der virtuellen Welt zu sein, wird oft als Schlüsselfaktor für erfolgreiche VR- und AR-Anwendungen angesehen. Ohne haptisches Feedback bleibt diese Erfahrung jedoch unvollständig.

Zudem ist das haptische Feedback nicht nur für die individuelle Präsenz entscheidend, sondern auch für soziale Interaktionen in virtuellen Umgebungen. Die Wahrnehmung von Berührungen und physikalischen Kräften zwischen Avataren trägt zur sozialen Präsenz bei, was die immersive Erfahrung weiter verstärkt. In Mehrbenutzerumgebungen ist diese Art von Feedback besonders wichtig, da es die Wahrnehmung der Interaktion mit anderen Benutzer:innen realer macht und so eine engere Bindung zur virtuellen Welt ermöglicht.

Es ist bekannt, dass Adhärenz von Therapien durch mobile Applikationen und digitale Unterstützung gesteigert werden kann. Eine Studie von Mertens et al. (2016) zeigte beispielsweise, dass eine mobile Anwendung die Therapie-Adhärenz bei älteren Patient:innen in der Rehabilitation signifikant verbessern kann, indem die Dokumentation über ein iPad statt über papierbasierte Verfahren erfolgt. Solche Ansätze könnten auch auf VR/AR-basierte Rehabilitationssysteme angewendet werden, um deren Wirksamkeit weiter zu steigern.

Trotz dieser Erkenntnisse bleibt die Entwicklung effizienter und kostengünstiger Technologien zur Bereitstellung realistischer haptischer Rückmeldungen in VR/AR eine große technische Herausforderung. In den letzten Jahren wurden verschiedene Ansätze, wie z. B. mechanische Exoskelette, Vibrationsmotoren und neuere Technologien wie EMS, erforscht, um diese Lücke zu schließen. EMS bietet eine interessante Alternative, da es die Möglichkeit schafft, direkte Muskelstimulationen zu nutzen, um das Gefühl von Gewicht und Widerstand zu simulieren. Damit könnte EMS das haptische und propriozeptive Feedback in VR/AR erheblich verbessern und somit die Immersion und Interaktion in virtuellen Umgebungen auf ein neues Niveau heben.

Ein weiterer Aspekt, der in der Entwicklung solcher Technologien berücksichtigt werden sollte, betrifft die ethischen, rechtlichen und sozialen Implikationen (ELSI). Eine methodische Grundlage dafür bietet das AMICAI-Framework, das auf Risikobewertungen basiert und Responsible Research and Innovation (RRI) in die Arbeit von Forschungs- und Innovationspraktiker:innen integriert (Brandl et al., 2020). Dies gewährleistet eine verantwortungsvolle Entwicklung von Technologien, die die Nutzer:innen in den Mittelpunkt stellen.

2. Hintergrund

Die propriozeptive Wahrnehmung, auch als Propriozeption bekannt, beschreibt die Fähigkeit des Menschen, die Position und Bewegung seiner Körperteile durch interne Sensoren zu erfassen. Diese Propriozeptoren, wie Muskelspindeln und Golgi-Sehnenorgane, überwachen Muskelspannung und Bewegungen und übermitteln diese Informationen an das Gehirn (Kaya et al., 2018). Im Gegensatz zur

haptischen Sensitivität, die sich auf externe Reizwahrnehmungen durch Hautrezeptoren konzentriert, erfasst die Propriozeption das Bewusstsein für die eigenen Bewegungen und die auf den Körper wirkenden Kräfte.

In VR-Umgebungen wird die Propriozeption bislang weitgehend vernachlässigt, während haptische Technologien wie Force-Feedback-Handschuhe oder kraftrückgekoppelte Steuerelemente primär auf die Simulation taktiler Reize abzielen. Dies schafft ein Defizit, da effiziente Lösungen fehlen, um die Wahrnehmung der eigenen Körperbewegungen und Muskelspannungen in MR-Systemen realistisch darzustellen (siehe Abbildung 1 für eine Übersicht der verwendeten Stimulationssysteme).

Forschungsergebnisse zeigen, dass haptisches Feedback für das Gefühl der Präsenz mindestens genauso wichtig ist wie visuelle Rückmeldungen. Dieses Präsenzgefühl steigt proportional zur Qualität des haptischen Feedbacks, da es nicht nur das individuelle Gefühl der »Präsenz« verstärkt, sondern auch die soziale Präsenz in virtuellen Umgebungen fördert (Gibbs et al., 2021).

Ein innovativer Ansatz, der propriozeptives Feedback ermöglicht, ist die EMS. EMS aktiviert durch elektrische Impulse die Muskeln, was zu Muskelkontraktionen und einer gezielten Stimulation der Propriozeptoren führt. Dies bietet klare Vorteile gegenüber mechanischen Systemen, da EMS weniger Platz benötigt, energieeffizienter ist und die Muskulatur auf natürliche Weise nutzt, um propriozeptive Rückmeldungen zu erzeugen (siehe Abbildung 2 zur Anbringung der Elektroden am Arm). EMS hat sich bereits in der Rehabilitation etabliert, insbesondere als funktionelle elektrische Stimulation (FES) bei neurologischen Störungen. Beispielsweise beschreiben Schauer und Vrontos (2019) die gezielte Kombination von künstlich und freiwillig induzierten Muskelkontraktionen mittels FES, um Bewegungen wie die Schulterabduktion präzise zu steuern. Solche Ansätze zeigen, wie EMS sowohl funktional als auch therapeutisch eingesetzt werden kann.

In VR- und AR-Anwendungen erweitert EMS diese Technologie um die Fähigkeit, haptisches Feedback in Echtzeit zu liefern, was die Immersion deutlich steigert. Die Anpassungsfähigkeit der EMS-Parameter, wie Pulsform, Intensität und Frequenz, ermöglicht eine maßgeschneiderte haptische Erfahrung und schafft die Grundlage für neue Interaktionsparadigmen in virtuellen Umgebungen.

Die Kombination von EMS mit Motion-Capture-Technologien bietet weitere Möglichkeiten, insbesondere im Bewegungslernen und der Rehabilitation. Hier können Nutzer:innen nicht nur visuell, sondern auch propriozeptorisch in die virtuelle Umgebung eingebunden werden. Dies ist besonders relevant in der medizinischen Anwendung, wie z. B. beim Erlernen neuer Bewegungsmuster oder in der Sportrehabilitation.

Trotz dieser Fortschritte fokussieren sich viele MR-Systeme nach wie vor auf visuelle und akustische Reize, während die Simulation externer Kräfte und Gewichte oft vernachlässigt wird. EMS könnte hier eine Schlüsselrolle spielen, um die fehlenden physischen Interaktionen zu simulieren und die immersive Erfahrung in MR-Systemen erheblich zu verbessern.

Insgesamt bleibt die Propriozeption ein zentraler, aber oft unterschätzter Aspekt in der Gestaltung immersiver VR- und MR-Systeme. Die Integration von EMS und weiteren Technologien verspricht, realistischere und effektivere Interaktionen in verschiedenen Anwendungsfeldern zu ermöglichen, sei es im Bewegungslernen, der Rehabilitation oder in interaktiven MR-Erfahrungen.



Abbildung 1: Das Stimulationssystem P24 Science von HASOMED zur EMS-Generierung
Quelle: Projekteigene Darstellung

3. Stand der Forschung

Der Stand der Forschung zur EMS in AR/VR-Umgebungen wurde umfassend von Vrontos et al. (2024) untersucht und ist Open Access verfügbar. Die Arbeit bietet wertvolle Einblicke in die aktuellen Entwicklungen und Anwendungen von EMS zur Verbesserung des propriozeptiven Feedbacks und deckt dabei verschiedene Einsatzmöglichkeiten sowie die Herausforderungen der Technologie auf. Die zentralen Erkenntnisse lassen sich in vier Hauptbereiche unterteilen:



Abbildung 2: Anbringung der EMS-Elektroden am Arm des Nutzers zur gezielten Muskelstimulation in VR-Anwendungen
Quelle: Projekteigene Darstellung

3.1 EMS als Alternative für propriozeptives Feedback in AR/VR

Die Forschung zeigt zunehmend, dass EMS eine vielversprechende Alternative zu herkömmlichen mechanischen haptischen Geräten darstellt, insbesondere in AR und VR. Durch die Fähigkeit von EMS, erhebliche Kräfte zu erzeugen, wird eine realistischere Simulation physischer Interaktionen ermöglicht. Diese Technologie wurde in verschiedenen Studien erprobt, die sich mit der Simulation von Reibungskräften und dem Erleben von Widerständen in mobilen Spielen beschäftigen. Dabei berichten Nutzer:innen, dass EMS-Feedback im Vergleich zu herkömmlichen vibrotaktilen Rückmeldungen immersiver und realistischer ist. Eine Studie zeigte sogar, dass EMS das Gefühl von Präsenz, Beteiligung und Immersion in virtuellen Umgebungen deutlich steigern kann (Sallnäs, 2010). Forschungen von Lopes et al. (2018) veranschaulichen darüber hinaus das Potenzial von EMS

zur Schaffung haptischer Empfindungen wie Gewicht, Aufprall und Reibung, was die Vielseitigkeit und den Wert dieser Technologie im Bereich AR/VR weiter hervorhebt.

In Szenarien, in denen das Gefühl einer physischen Verbindung mit der virtuellen Welt besonders wichtig ist, erweist sich EMS als geeignetes Mittel, um die immersiven Erlebnisse weiter zu verstärken. Durch die gezielte Stimulation von Muskeln, wie etwa dem Trizeps oder Bizeps, können realistische Kraftempfindungen erzeugt werden, die es den Nutzer:innen ermöglichen, virtuelle Objekte ähnlich wie physische Objekte wahrzunehmen. Besonders interessant ist die Entwicklung genauerer EMS-Modelle, die in der Lage sind, Kraft- und Druckparameter präziser zu steuern, was zur weiteren Optimierung solcher Anwendungen beiträgt.

3.2 Praktische Anwendungen von EMS

EMS hat sich in der Forschung als äußerst vielseitig erwiesen und wurde erfolgreich in einer Vielzahl von Aufgaben eingesetzt, die den Umgang mit Objekten in virtuellen Umgebungen betreffen. Ein markantes Beispiel ist die Simulation von Interaktionen mit virtuellen Wänden. Hier konnten Lopes et al. (2018) zeigen, dass durch die gezielte Stimulation der Muskeln eine realistische haptische Rückmeldung erzeugt wird, die verhindert, dass der Arm des/der Nutzer:in durch die virtuelle Wand hindurchdringt. Dieses Gefühl, eine physische Barriere zu berühren, obwohl es sich nur um ein virtuelles Objekt handelt, eröffnet neue Möglichkeiten für AR/VR-Erlebnisse. In einer anderen Studie, die sich auf den Umgang mit virtuellen Objekten konzentrierte, konnten 76 % der Teilnehmer:innen die Nutzung von durch EMS kommunizierten Objekten korrekt erraten. Dies zeigt, dass EMS den Einsatz von AR/VR weit über das bloße visuelle und auditive Erlebnis hinaus erweitern kann.

Darüber hinaus wird EMS in Lernprozessen als effektiv betrachtet, etwa beim Erlernen von Musikrhythmen oder bei der Unterstützung von Trainingsaufgaben wie der Speicherung von Arbeitsabläufen. Eine Studie ergab, dass EMS den Teilnehmer:innen dabei half, sich besser an Abfolgen von Tastenanschlägen zu erinnern. Dies unterstreicht das Potenzial von EMS, sowohl in der Ausbildung als auch in der industriellen Schulung als Werkzeug zur Verbesserung der Gedächtnisleistung und der motorischen Fertigkeiten eingesetzt zu werden.

3.3 Vergleich von EMS und vibrotaktiler Rückmeldung

Der Vergleich von EMS und vibrotaktiler Rückmeldung hat gezeigt, dass EMS in bestimmten Anwendungen erhebliche Vorteile bietet. Besonders in Situationen, die schnelle und präzise motorische Reaktionen erfordern, erwies sich EMS als effektiver. In mehreren Experimenten konnte nachgewiesen werden, dass EMS die Reaktionszeiten verkürzt und gleichzeitig die Fehlerquote reduziert. Dies macht die Technologie in Bereichen wie medizinischen Simulationen, der Luftfahrt und virtuellen Laboren besonders wertvoll. In einer Studie, die das schnelle Stoppen von Handbewegungen untersuchte, stellte sich heraus, dass EMS sowohl visuelle als auch vibrotaktile Hinweise übertraf, da es präzisere Steuerungen ermöglichte und weniger Fehler verursachte (Korres et al., 2022).

Dieser Vorteil von EMS ist besonders in Bereichen nützlich, in denen Präzision und Geschwindigkeit entscheidend sind, wie etwa in der Robotik, der chirurgischen Ausbildung oder sicherheitskritischen Simulationen. Durch die direkte Stimulation der Muskeln kann EMS eine sofortige und klare Rückmeldung erzeugen, was die Effizienz und Genauigkeit der ausgeführten Aktionen verbessert. Damit bietet sich EMS als überlegenes Feedbacksystem in anspruchsvollen Umgebungen an.

3.4 Herausforderungen und zukünftige Forschung

Trotz der beeindruckenden Fortschritte, die EMS in der Forschung zu AR/VR gemacht hat, bestehen nach wie vor wesentliche Herausforderungen. Eine der größten Schwierigkeiten besteht in der uneinheitlichen Berichterstattung der verwendeten EMS-Parameter. In verschiedenen Studien variierten die Amplituden, Impulsbreiten und Frequenzen erheblich, was die Vergleichbarkeit der Ergebnisse erschwert. So wurden Amplitudenbereiche von 10 mA bis 150 mA verwendet, während die Impulsbreiten von 25 µs bis 800 µs reichten. Diese Variabilität macht deutlich, dass ein standardisierter Ansatz bei der Anwendung von EMS in der Forschung notwendig ist, um die Reproduzierbarkeit und Vergleichbarkeit der Ergebnisse sicherzustellen.

Ein weiteres Hindernis stellt die demografische Verzerrung der Studienteilnehmenden dar. Die Mehrheit der Teilnehmenden besteht aus jungen Männern zwischen 20 und 30 Jahren, was die Anwendbarkeit der Ergebnisse auf andere Bevölkerungsgruppen einschränken könnte. Insbesondere im Hinblick auf unterschiedliche physiologische und psychologische Reaktionen auf EMS besteht die Notwendigkeit, eine größere Vielfalt in den Studienpopulationen zu berücksichtigen. Zukünftige Forschung sollte daher auf eine breitere Inklusion von Teilneh-

menden mit unterschiedlichen Hintergründen abzielen, um sicherzustellen, dass die Ergebnisse auf eine größere Population übertragbar sind. Verbesserte Standards in der Berichterstattung und eine diversifizierte Auswahl der Teilnehmenden könnten dazu beitragen, das volle Potenzial von EMS in AR/VR vollständig auszuschöpfen.

4. VR/AR-Anwendungen in der Unreal Engine 4

Im Rahmen des Forschungsprojekts »interaktionsKRAFT« wurden zwei Anwendungsszenarien entwickelt und untersucht:

4.1 Simulation von VR-Gewichten mittels EMS an den Antagonisten

Beim Halten eines physischen Gewichts entstehen Drehmomente, insbesondere im Handgelenk und im Ellbogengelenk, die das Gewicht spürbar machen. Diese Drehmomente fehlen jedoch, wenn ein virtuelles Objekt in der VR gehalten wird. Um dieses fehlende physische Feedback zu kompensieren, wird EMS an den Antagonistenmuskeln angewendet, um ein realistisches Gefühl von Gewicht zu erzeugen. In einer durchgeführten Studie wurde untersucht, wie verschiedene EMS-Einstellungen das propriozeptive Feedback und die Wahrnehmung des Gewichts beeinflussen.

Dabei spielt die Stimulation der Antagonistenmuskeln eine zentrale Rolle: Während die Agonistenmuskeln das Objekt »halten«, wird durch die Stimulation der Antagonisten ein Drehmoment induziert, das die Teilnehmenden ausgleichen müssen. Dies führt dazu, dass das virtuelle Objekt sich für den/die Benutzer:in schwerer anfühlt, ähnlich wie bei einem realen Gewicht. Die Studie untersuchte, wie die Stimulation bestimmter Muskelgruppen das wahrgenommene Gewicht eines Objekts beeinflusst und wie realistisch diese Gewichtswahrnehmung durch EMS erzeugt werden kann.

Ein entscheidender Faktor für das realistische Feedback sind die EMS-Stimulationsparameter, die sorgfältig ausgewählt und angepasst werden müssen. Dazu gehören die Stimulationsamplitude, die bestimmt, wie stark die elektrischen Impulse sind; die Stimulationsfrequenz, die angibt, wie oft die Impulse pro Sekunde gesendet werden; und die Pulsbreite, die die Dauer jedes einzelnen Impulses festlegt. Um die optimale Stimulation für jeden Teilnehmenden zu gewährleisten, war ein Kalibrierungsprozess notwendig, da die individuellen Empfindlichkeiten gegenüber EMS variieren. Die Kalibrierung begann mit einer niedrigen Stromstärke von 1 mA, die schrittweise um 1 mA pro Sekunde erhöht wurde. Der soge-

nannte Motor Threshold wurde ermittelt, indem sichtbare Muskelkontraktionen oder Zuckungen beobachtet wurden. Die Schmerzgrenze wurde durch verbales Feedback der Teilnehmenden festgestellt, bei dem sie angaben, wann die Stimulation die Grenze ihres Komforts erreichte.

In der Studie wurden zwei Versuchsbedingungen getestet: Bei der ersten wurde nur der *Trizeps brachii* des dominanten Arms stimuliert (Ein-Muskel-Stimulation), während bei der zweiten Bedingung sowohl der *Trizeps brachii* als auch der *Extensor carpi ulnaris* des dominanten Arms simultan stimuliert wurden (Zwei-Muskel-Stimulation). Die Kombination dieser Parameter beeinflusste maßgeblich die Wahrnehmung der Gewichtszunahme und das realistische Gefühl von physischen Kräften in der virtuellen Umgebung.

Die vorläufigen Ergebnisse der Studie zeigen, dass durch die Feinjustierung der EMS-Stimulationsparameter und die individuelle Kalibrierung ein realistisches Gewichtsempfinden in VR-Umgebungen simuliert werden kann. Dies eröffnet neue Möglichkeiten für die Entwicklung immersiver VR-Erfahrungen, insbesondere in Bereichen wie Training, Simulationen und Rehabilitation, bei denen das Gefühl von physischen Objekten eine entscheidende Rolle spielt.

Als Teil des Demonstrators wurden fünf greifbare virtuelle Objekte mit unterschiedlichen simulierten Gewichten entwickelt, die es den Nutzenden ermöglichen, durch die Variation der EMS-Intensitäten ein realistisches Gefühl von Gewicht zu erleben (siehe Abbildung 3). Diese Objekte unterscheiden sich nicht nur visuell, sondern auch in der Stärke der EMS-Stimulation, wodurch das Gefühl von unterschiedlichen Gewichten in der virtuellen Realität erzeugt wird. Dieser Demonstrator dient zur Veranschaulichung der Möglichkeiten von EMS für haptisches Feedback in VR, insbesondere zur Simulation von physikalischen Eigenschaften wie Gewicht. Die Anwendung stellt jedoch keine formale Studie dar, sondern soll das Potenzial und die Funktionalität von EMS als Feedbackmechanismus verdeutlichen.

4.2 Korrektur der Bewegungsausführung einer Kniebeuge in einer VR-Physiotherapie mittels EMS

In diesem Anwendungsfall wird EMS genutzt, um die korrekte Ausführung einer Kniebeuge in einer VR-Physiotherapie sicherzustellen. Die Körperhaltung des/der Benutzer:in wird mittels markerloser Motion-Capture-Technologie erfasst, sodass die Bewegungen des Körpers präzise und ohne den Einsatz von physischen Markern überwacht werden können. Ein solches System muss echtzeitfähig sein, wie es beispielsweise in der Studie von Brandl et al. (2022) demonstriert wurde. Dort wurde ein Echtzeit-Feedbacksystem entwickelt, das

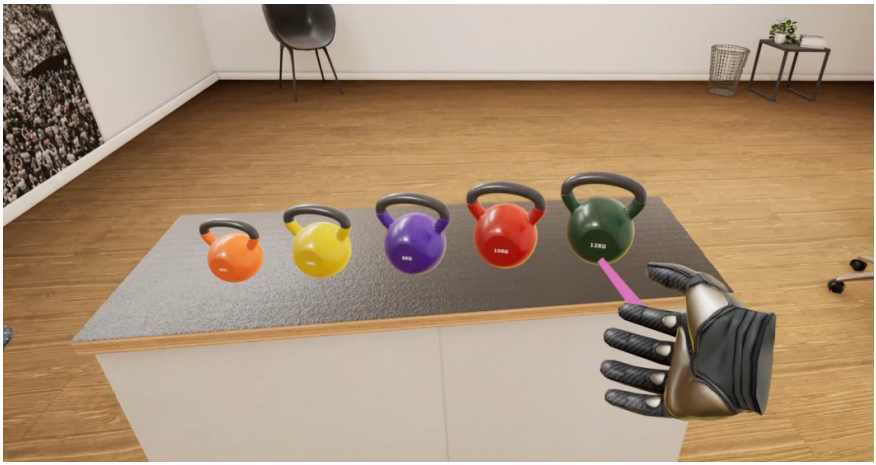


Abbildung 3: Fünf greifbare VR-Objekte mit unterschiedlichen virtuellen Gewichten, die durch unterschiedlich starke EMS-Stimulationsintensitäten simuliert werden.
Quelle: Projekteigene Darstellung

die L5/S1-Kompressionskraft mithilfe markerloser optischer Motion-Capture-Technologie überwacht, um die Hebetechnik bei manueller Materialhandhabung zu verbessern. Die Übertragung solcher Ansätze auf VR-gestützte Physiotherapie zeigt das Potenzial, Bewegungsanalysen in Echtzeit zu ermöglichen und direktes Feedback zu geben.

Eine ausreichende Genauigkeit der erfassten Bewegungsinformationen ist dabei Voraussetzung für eine funktionierende und sichere Steuerung. Wissenschaftliche Untersuchungen bestätigen die Genauigkeit solcher markerlosen Systeme auch für komplexe ergonomische Risikoanalysen in der Arbeitswelt (Brunner et al., 2022). Diese Erkenntnisse unterstreichen die Zuverlässigkeit der Technologie, Bewegungen der Nutzenden präzise zu erfassen und darauf basierend angemessenes Feedback zu generieren.

Bei Abweichungen von der idealen Bewegung, beispielsweise wenn die Kniebeuge zu tief oder zu flach im Vergleich zu den Vorgaben eines virtuellen Trainers ist, greift das System in Echtzeit ein. Die Intensität der EMS-Stimulation wird proportional zur Abweichung zwischen der tatsächlichen Bewegung (Ist-Wert) und der idealen Bewegung (Soll-Wert) angepasst: je größer die Differenz, desto stärker die Muskelstimulation. Dadurch erhält der/die Benutzer:in in Echtzeit ein direktes Feedback zur korrekten Ausführung der Übung, was nicht nur die Effizienz, sondern auch die Sicherheit der Bewegungsausführung erhöht. In Abbildung 4 ist Ausschnitt aus der visuellen Darbietung für die Nutzenden dargestellt.

Dieses System ersetzt das unmittelbare Feedback eines menschlichen Trainers und ermöglicht es Benutzer:innen, autonom von zu Hause aus Kniebeugen

durchzuführen. Insbesondere für Personen, die ihre Übungen eigenständig und ohne permanente Betreuung durch eine Physiotherapeutin oder einen Physiotherapeuten ausführen möchten, bietet diese Lösung eine effektive Möglichkeit, ihre Bewegungsausführung kontinuierlich zu verbessern und Fehler in Echtzeit zu korrigieren. Die Anwendung läuft in der Unreal Engine 4 und stellt eine fortschrittliche Methode dar, physiotherapeutische Übungen in virtuellen Umgebungen effizienter zu gestalten.



Abbildung 4: Vergleich der Kniebeugen-Ausführung in einer VR-Physiotherapie Persönliches Avatar (links) und VR-Trainer (rechts) zur Korrektur der Bewegungsform.
Quelle: Projekteigene Darstellung

5. Entwicklung eines Lehr-Lern-Konzepts für XR-basiertes Bewegungslernen

XR-basiertes Bewegungslernen bietet eine innovative Möglichkeit, motorische Fertigkeiten in einer sicheren, immersiven und realitätsnahen Umgebung zu entwickeln und zu verbessern. Im Rahmen des vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Projekts interaktionsKRAFT wurde diese Technologie durch die Einbeziehung einer Haptikkomponente, die auf EMS basiert, weiter optimiert. Diese Haptikkomponente korrigiert mittels elektrischer Stimulation der betroffenen Muskelgruppen sanft Bewegungen und verhindert so das Erlernen falscher Bewegungsmuster. Somit können therapeutische Übun-

gen im häuslichen Umfeld präzise und korrekt ausgeführt werden, ohne dass es der Anwesenheit einer Physiotherapeutin oder eines Physiotherapeuten bedarf.

Das Konzept des XR-basierten Bewegungslernens eignet sich besonders gut für den Bereich der Physiotherapie, wo präzises und korrektes Bewegungslernen für den Therapieerfolg entscheidend ist. Im Folgenden wird ein Leitfaden zur Entwicklung eines Lehr-Lern-Konzepts für XR-Umgebungen am Beispiel des interaktionsKRAFT-Projekts dargestellt. Dieses Konzept orientiert sich am Zyklus der Curriculumsentwicklung nach Kern (2009) und integriert Prinzipien des Bewegungslernens nach Peyton (Peyton & Walker, 1998).

5.1 Anwendungsbeispiel interaktionsKRAFT

Am Beispiel des interaktionsKRAFT-Projekts wird ein Leitfaden für die Gestaltung eines Lehr-Lern-Konzepts zum Bewegungslernen für XR-Umgebungen gezeigt. Diese Inhalte lassen sich jedoch auch auf andere Lehr-Lern-Szenarien außerhalb des Bewegungslernens übertragen. Konkrete Anwendungsbeispiele aus dem interaktionsKRAFT-Projekt sind in Boxen dargestellt.

5.2 Problemidentifikation

Bevor man mit der Entwicklung eines Lehr-Lern-Konzepts beginnen kann, bedarf es einer umfangreichen Auseinandersetzung mit der ausgewählten Thematik und der jeweiligen Zielgruppe, für die die Lernumgebung zugeschnitten sein soll. Für die Vorbereitung der Konzeption eignen sich verschiedene Methoden, wie beispielhaft unten dargestellt.

Recherche

Für eine anfängliche Recherche zum gewünschten Thema des Lehr-Lern-Konzepts eignen sich verschiedene Datenbanken. Hieraus lassen sich Informationen ableiten, ob bereits ein ähnliches Lehr-Lern-Konzept existiert oder sogar bereits erprobt wurde.

Beispiele von Datenbanken:

- Google Scholar
- PubMed
- CINAHL
- Cochrane

Beispiel für die Suchstrategie: ("motor learning" OR "physical therapy" OR "sports" OR "motor skills") AND ("virtual reality" OR "extended reality" OR "mixed reality")

Austausch mit Expert:innen

Sollten relevante Informationen fehlen, die Sie zur Entwicklung eines Lehr-Lern-Konzepts benötigen, eignet sich zur Vorbereitung auch ein Gespräch mit Endnutzer:innen oder Expert:innen zu diesem Thema. Hierbei gilt es insbesondere herauszufinden, was den potenziellen Endnutzer:innen an Lernmöglichkeiten aktuell fehlt und welche Fertigkeiten gezielt mit dem XR-basierten Training verbessert werden können, die sich unter Umständen im Alltag nicht trainieren lassen.

Relevante Ansprechpersonen am Beispiel des interaktionsKRAFT-Lehr-Lern-Konzepts:

- Physiotherapeut:innen
- Patient:innen
- Sportler:innen
- Entwickler:innen

5.3 Zielgerichtete Bedarfsanalyse: Wissenschaftliche Erhebungen

1 Qualitative Interviews

Bei Bedarf können Sie umfangreiche Interviews mit mehreren Personen durchführen. Diese können sowohl in Form von Einzelinterviews als auch in einer (Fokus-)Gruppe erfolgen und haben zum Ziel, die Meinung mehrerer Personen in die Vorbereitung und Entwicklung des Lehr-Lern-Konzepts einfließen zu lassen und so die Grundlage für eine bedarfsgerechte Technologie zu schaffen.

2 Quantitative Fragebogenerhebung

Für anonyme Rückmeldungen einer breiteren Stichprobe wird häufig ein Fragebogendesign verwendet. Hierbei können Trainingsbedarfe und -wünsche an die Technologie offengelegt und anhand einer breiteren Stichprobe abgefragt werden.

Im Rahmen des interaktionsKRAFT Projekts wurde nach einem ersten formlosen Austausch mit den Konsortialpartner:innen aus der Physiotherapie ein Leitfaden für ein semistrukturiertes Interview erstellt (siehe Abbildung 5) und Interviews mit Patient:innen bis zur inhaltlichen Sättigung durchgeführt.

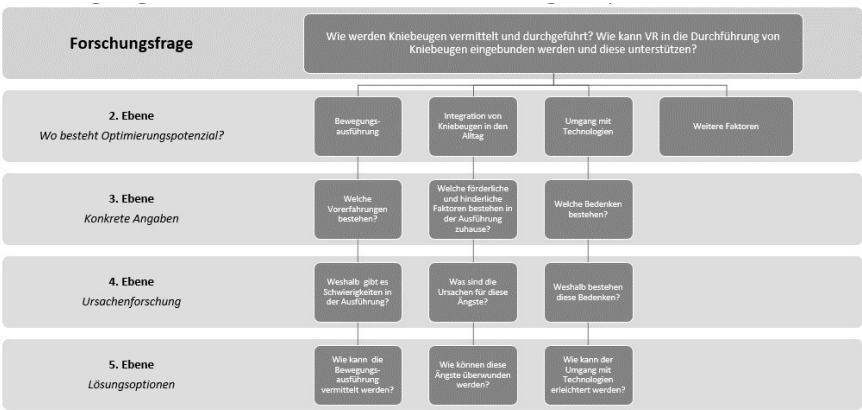


Abbildung 5: Leitfaden für ein semistrukturiertes Interview
Quelle: Projekteigene Darstellung

Auf Grundlage der Interviewergebnisse wurden sowohl hinderliche als auch förderliche Faktoren identifiziert.

3 Hinderliche Faktoren

Ängste vor Beschwerdeverschlimmerung: Patient:innen befürchteten, dass die fehlende Korrektur durch eine Physiotherapeutin oder einen Physiotherapeuten zu einer Verschlechterung ihrer Beschwerden führen könnte. Dieses Problem wurde in der Technologieumsetzung durch die Integration eines Bewegungsmonitorings und durch visuelle sowie haptische Rückmeldungen adressiert. Diese Funktionen helfen dabei, den Patient:innen präzise Anweisungen zu geben und sicherzustellen, dass Übungen korrekt ausgeführt werden.

4 Förderliche Faktoren

Beobachtung von Leistungssteigerung: Patient:innen berichteten von einem erwarteten Anstieg ihrer Motivation, der durch die Möglichkeit zur Verfolgung von

Wiederholungen gegebenenfalls verstärkt würde. Das Tracking der ausgeführten Wiederholungen ermöglicht es den Nutzenden, Fortschritte zu sehen, und motiviert sie zusätzlich zur aktiven Teilnahme an den Übungen.

Diese Erkenntnisse trugen zu der Entwicklung der interaktionsKRAFT-Technologie bei und gewährleisten, dass sowohl die Bedenken der Patient:innen als auch ihre positiven Erfahrungen berücksichtigt wurden.

5.4 Lernziele

Bei der Entwicklung des Lehr-Lern-Konzepts ist es wichtig, auf Basis der zuvor identifizierten Bedarfe und Probleme konkrete Lernziele zu definieren. Die Lernziele definieren, was die Teilnehmer:innen am Ende der Lernerfahrung wissen, verstehen bzw. anwenden können.

Klar definierte Lernziele schaffen eine Struktur für das Training und tragen zu einer einheitlichen Trainingsqualität bei. Eine gute Hilfe bei der Definition von Lernzielen stellen die SMART-Kriterien dar (Doran, 1981). Demnach sind gute (Lern-)Ziele **S**pezifisch, **M**essbar, **A**usführbar, **R**ealistisch und **T**erminiert.

Bei der Definition von Lernzielen können Sie sich ferner z. B. an der Taxonomie nach Bloom et al. orientieren, die verschiedene Ebenen des Lernens beschreiben: 1. Wissen, 2. Verstehen, 3. Anwenden, 4. Analysieren, 5. Bewerten und 6. Schaffen (Bloom et al., 1956).

Auf Basis der in Schritt 2 identifizierten Themen wurde folgendes Hauptlernziel entwickelt (Bloom's Taxonomie *Anwenden*):

Die Lernenden führen eine Kniebeuge durch, indem sie die Bewegungsabläufe mit der korrekten Technik ausführen. Dies umfasst die richtige Beckenpositionierung, Fußstellung sowie Knie- und Oberkörperausrichtung.

5.5 Edukative Methoden

Die Peyton-Methode (Peyton & Walker, 1998) bietet eine strukturierte Vorgehensweise zur Vermittlung von *Bewegungslernen*. Die Methode umfasst vier wesentliche Schritte:

1 Demonstration

In der XR-Umgebung zeigt der Avatar die Kniebeuge mit der korrekten Technik. Der Fokus liegt darauf, die gesamte Bewegung in normaler Geschwindigkeit von Anfang bis Ende visuell darzustellen. Der Avatar führt die Übung langsam und präzise durch, um die korrekte Körperhaltung und Bewegungsausführung zu verdeutlichen.

2 Dekonstruktion

Der Avatar erklärt detailliert die einzelnen Schritte der Kniebeuge bei einer verlangsamten Bewegungsdurchführung. Diese umfassen neben der Beckenpositionierung auch die Fußstellung und die Haltung des Oberkörpers. Diese Erklärungen erfolgen in der XR-Umgebung, um die Bewegungen zusätzlich zu den visuellen und haptischen Komponenten zu verdeutlichen.

3 Verständnis

Die Lernenden geben Anweisungen an den Avatar, um die korrekte Ausführung der Kniebeuge zu simulieren. Dies geschieht durch Auswahl von Optionen in Multiple-Choice-Fragen, die verschiedene Aspekte der Kniebeuge abdecken.

4 Durchführung

Die Lernenden führen die Kniebeuge eigenständig in der XR-Umgebung durch. Während der Ausführung erhalten sie kontinuierliches haptisches und visuelles Feedback von der XR-Technologie.

5.6 Evaluation und Weiterentwicklung

Ob ein Lehr-Lern-Konzept erfolgreich ist und wie es weiterentwickelt werden kann, lässt sich mittels einer Evaluation hinsichtlich der Nutzung herausfinden. Hierzu empfiehlt es sich, Teilnehmende beispielsweise einen Fragebogen im Nachgang ausfüllen zu lassen und ihr Feedback und ihre Wahrnehmungen sowie Erfahrungen aus dem Training einzuholen. Dieser kann sowohl die Möglichkeit für positive Anmerkungen als auch für Kritik und Anregungen zur Verbesserung des Lehr-Lern-Konzepts bieten. Mit der Integration von Wissensfragen kann

zudem eine Lernwirksamkeitskontrolle durchgeführt werden. Des Weiteren bietet ein solcher Fragebogen die Möglichkeit, auch Aspekte wie die Nutzbarkeit, Immersion oder Akzeptanz des Systems quantitativ zu erheben. Hierzu bieten sich standardisierte Skalen wie beispielsweise die System Usability Scale (SUS) (Brooke, 1996) an. Die Ergebnisse können auch als wertvolle Grundlage für die schrittweise Weiterentwicklung der Technologie dienen.

Darüber hinaus bietet eine Evaluation die große Chance der Selbstreflexion. Außerdem bietet sie die Möglichkeit zur Überarbeitung und Weiterentwicklung des Lehr-Lern-Konzepts, indem das Feedback iterativ eingearbeitet wird. Neue Trainingsdurchläufe ermöglichen somit eine stetige Anpassung und Weiterentwicklung des Trainings.

6. Fazit

Die vorliegende Arbeit zeigt deutlich das Potenzial von EMS als zukunftsweisende Technologie zur Verbesserung des propriozeptiven Feedbacks in AR und VR. Der Überblick über den aktuellen Stand der Forschung verdeutlicht, dass EMS herkömmliche haptische Rückmeldesysteme durch die direkte Stimulation von Muskeln ersetzen oder ergänzen kann, um realistischere und immersivere Interaktionen in virtuellen Umgebungen zu ermöglichen.

Die im Projekt interaktionsKRAFT entwickelten Technologien leisten einen Beitrag zur BMBF-Förderlinie »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft«. Durch EMS-basiertes propriozeptives Feedback eröffnen sie neue Möglichkeiten für realistische Interaktionen in AR/VR-Umgebungen, die herkömmliche haptische Rückmeldesysteme erweitern oder ersetzen. Diese Technologie fördert nicht nur die Entwicklung immersiver Anwendungen im Bildungs- und Gesundheitsbereich, sondern unterstützt auch selbstgesteuertes Training und Lernen in digitalen Räumen.

Die Einbindung von EMS in AR/VR-Anwendungen reflektiert die Kernziele der Förderlinie: innovative Lösungen zu schaffen, die den Menschen in vielfältigen Lebensbereichen unterstützen. EMS-basierte Systeme, die gezielt körperliches Feedback bieten, können wichtige gesellschaftliche Aufgaben adressieren. In der Rehabilitation etwa ermöglicht EMS es Menschen, ihre Therapie flexibel und individuell zu gestalten, unabhängig von einem oder einer direkten Trainer:in oder Therapeut:in. Solche Technologien tragen zur Autonomie und Lebensqualität bei, indem sie verlässliche und anpassbare technische Assistenz bieten, die sich an die Bedürfnisse und Fähigkeiten der Nutzenden anpassen lässt. Es ist seit langem bekannt, dass die Adhärenz von Therapien durch mobile Applikationen

und digitale Unterstützung gesteigert werden kann. Eine Studie zeigt, dass eine mobile Anwendung die Therapie-Adhärenz bei älteren Patient:innen in der Rehabilitation signifikant verbessern kann, indem die Dokumentation über ein iPad statt über papierbasierte Verfahren erfolgt (Mertens et al., 2016).

Für die digitale Gesellschaft bedeutet dies einen innovativen Schritt hin zu interaktiven XR-Erfahrungen, die neben der Unterhaltung insbesondere in Rehabilitation und Ausbildung von langfristigem Nutzen sind. Durch die Möglichkeit, präzise Bewegungsabläufe zu Hause unter Echtzeit-Feedback zu üben, trägt EMS zur Teilhabe bei und ermöglicht eine effektivere Nutzung digitaler Gesundheits- und Bildungstechnologien. Ein weiterer wichtiger Beitrag des Projekts interaktionsKRAFT liegt in der Förderung eines inklusiven Technologieverständnisses. Durch die gezielte Anpassung der EMS-basierten Interaktionen an individuelle Bedürfnisse eröffnet sich das Potenzial für personalisierte Anwendungen in der digitalen Gesellschaft, die Menschen unabhängig von Alter oder körperlichen Einschränkungen nutzen können. Gerade im Hinblick auf den demografischen Wandel bietet EMS die Möglichkeit, barrierefreie und unterstützende Lösungen zu entwickeln, die Menschen aktiv in digitale Umgebungen einbinden und ihnen ein hohes Maß an Selbstständigkeit und Teilhabe ermöglichen. Die Forschung zeigt zudem, dass auch immer mehr ältere Menschen Informations- und Kommunikationstechnologie in ihr Leben integrieren (Mertens et al., 2022). Damit ist eine wichtige Grundlage für zukünftige MR-Anwendungen, wie etwa der virtuellen Physiotherapie, vorhanden.

Darüber hinaus ist die Berücksichtigung ethischer, rechtlicher und sozialer Implikationen (ELSI) von zentraler Bedeutung, um die verantwortungsvolle Entwicklung und Nutzung solcher Technologien sicherzustellen. Das AMICAI-Framework bietet eine methodische Grundlage, um Risiken zu analysieren und Responsible Research and Innovation (RRI) in den Entwicklungsprozess zu integrieren (Brandl et al., 2020). Dies gewährleistet, dass Innovationen nicht nur technisch, sondern auch gesellschaftlich nachhaltig gestaltet werden.

Die Arbeit unterstreicht somit, dass EMS nicht nur technische Innovationen für immersive AR/VR-Erlebnisse liefert, sondern auch die gesellschaftliche Relevanz interaktiver Systeme für eine diverse Nutzergruppe in den Mittelpunkt rückt. Solche Ansätze können langfristig dabei helfen, die Akzeptanz und Nutzbarkeit digitaler Technologien in allen Bevölkerungsschichten zu stärken und den Zugang zu innovativen Bildungs- und Gesundheitslösungen zu erweitern.

Die im Rahmen des Projekts interaktionsKRAFT entwickelten Nutzungsszenarien bestätigen die Vielseitigkeit von EMS in verschiedenen Anwendungsbereichen. Die Simulation von VR-Gewichten mittels EMS zeigt, wie präzise Muskelstimulation genutzt werden kann, um physische Empfindungen wie das Halten von Gewichten in VR zu simulieren. Ebenso bietet die Kombination von EMS

und Motion Capture in der VR-Physiotherapie eine innovative Methode zur Bewegungsanalyse und -korrektur, die das autonome Üben von Nutzenden zu Hause effektiv unterstützt.

Das vorgestellte Lehr-Lern-Konzept demonstriert darüber hinaus, wie EMS in XR-basierten Bewegungslernprozessen erfolgreich eingesetzt werden kann, um motorische Fähigkeiten gezielt zu trainieren und falsche Bewegungsmuster zu korrigieren. Besonders im Bereich der Physiotherapie und Rehabilitation bietet dieses Konzept vielversprechende Ansätze für die zukünftige Integration von EMS in praxisnahe Trainingsumgebungen.

Insgesamt zeigt die Arbeit, dass EMS das Potenzial hat, immersive Erlebnisse in AR/VR zu erweitern und sowohl in der Unterhaltungsbranche als auch in der Medizin vielfältige Anwendungsmöglichkeiten zu bieten. Zukünftige Forschung sollte sich auf die Optimierung der Stimulationsparameter und die Erhöhung der Diversität der Nutzenden fokussieren, um die Technologie weiterzuentwickeln und für eine breitere Zielgruppe zugänglich zu machen.

Dank

Wir danken unseren Projektpartnern LavaLabs, MedAix und Velamed für die wertvolle Zusammenarbeit und Unterstützung im Rahmen dieses Projektes. Ihre Expertise und ihr Engagement haben maßgeblich zur erfolgreichen Umsetzung der Forschungsziele beigetragen.

Förderhinweis

Das Projekt interaktionsKRAFT wurde vom Bundesministerium für Bildung und Forschung (BMBF)[1] gefördert.

Literatur

Bloom, Benjamin S. (1956). Taxonomy of educational objectives: the classification of educational goals, Handbook I: Cognitive domain. In: *Taxonomy of educational objectives: the classification of educational goal*, McKay, David, David McKay Company Inc., Longmans, Green & Co Ltd., London, S. 1–111.

Brandl, Christopher et al. (2020). AMICAI: A Method Based on Risk Analysis to Integrate Responsible Research and Innovation into the Work of Research and Innovation Practitioners.

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

In: *Science and engineering ethics,* Vol. 26 (2), Dordrecht, Netherlands: Springer. S. 667–689. https://doi.org/10.1007/s11948-019-00114-2

Brandl, Christopher et al. (2022). Using real-time feedback of L5/S1 compression force based on markerless optical motion capture to improve the lifting technique in manual materials handling. In: *International Journal of Industrial Ergonomics,* Vol. 91, Elsevier Science, Amsterdam, S. 103350. https://doi.org/10.1016/j.ergon.2022.103350

Brooke, John (1996). SUS: A quick and dirty usability scale. In: *Usability Evaluation in Industry*. S. 4–7. CRC Press, London.

Brunner, Oliver et al. (2022). Accuracy of a markerless motion capture system for postural ergonomic risk assessment in occupational practice. In: *International journal of occupational safety and ergonomics*, JOSE 28 (3), Taylor & Francis, London. S. 1865–1873. https://doi.org/10.1080/10803548.2021.1954791

Doran, George T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. In: *Management Review*, Vol. 70(22), American Management Association, New York. S. 35–36.

Gibbs, Janet, Gillies, Marco & Pan, Xueni (2021). A Comparison of the Effects of Haptic and Visual Feedback on Presence in Virtual Reality. In: *International Journal of Human-Computer Studies*, Vol. 157, 102717, Elsevier, Oxford, S. 1–12. https://doi.org/10.1016/j.ijhcs.2021.102717

Kaya, Defne, Yosmaoglu, Baran & Doral, Mahmut Nedim (2018). *Proprioception in Orthopaedics, Sports Medicine and Rehabilitation*, Springer, Cham, S. 3–4.

Kern, David E., Thomas, Patricia A. & Hughes, Mark T. (Hg.) (2009). A six-step approach to curriculum development. In: *Curriculum development for medical Education a six-step approach*, The Johns Hopkins University Press, Baltimore, S. 5–9.

Korres, Georgios, Park, Wanjoo & Eid, Mohamad (2022). A Comparison of Vibrotactile Feedback and Electrical Muscle Stimulation (EMS) for Motor Response During Active Hand Movement. In: *IEEE Trans. Haptics 2022*, 15, IEEE, New York, S. 74–78.

Lopes, Pedro et al. (2018). Adding Force Feedback to Mixed Reality Experiences and Games using Electrical Muscle Stimulation. In: *Proceedings of the 2018 Chi Conference on Human Factors in Computing Systems (Chi 2018),* Montreal, Vol. 446, Association for Computing Machinery, New York, S. 13. https://doi.org/10.1145/3173574.3174020

Mertens, Alexander et al. (2016). A mobile application improves therapy-adherence rates in elderly patients undergoing rehabilitation: A crossover design study comparing documentation via iPad with paper-based control. In: *Medicine,* Vol. 95 (36), e4446, Wolters Kluwer Health, Philadelphia, S. 1–8. http://dx.doi.org/10.1097/MD.0000000000004446

Mertens, Alexander et al. (2022). Health Technology Use in Germany Among Older Adults (Part I): Short Time Changes in Information and Communication Technology. In: *International Conference on Human-Computer Interaction*, Springer International Publishing, Cham, S. 112–128.

Peyton, R., & Walker, M. (1998). Teaching in the theatre. In: Peyton J. W. R. (Hg.), *Teaching and Learning in Medical Practice*, Manticore Publishers Europe Ltd., Oxford, S. 171–180.

Sallnäs, Eva-Lotta (2010). Haptic Feedback Increases Perceived Social Presence. In: *Proceedings of the Haptics: Generating and Perceiving Tangible Sensations*, Amdsterdam, Springer-Verlag, Berlin, Heidelberg, S. 178–185.

Schauer, Thomas & Vrontos, Apostolos (2019). Modeling of mixed artificially and voluntary induced muscle contractions for controlled functional electrical stimulation of shoulder ab-

duction. In: *IFAC-PapersOnLine*, Vol. 51(34), Elsevier Ltd., Amsterdam, S. 284–289. https://doi.org/10.1016/j.ifacol.2019.01.038

Vrontos, Apostolos, Nitsch Verena & Brandl, Christopher (2024). Electrical Muscle Stimulation for Kinesthetic Feedback in AR/VR: A Systematic Literature Review. In: *Multimodal Technologies and Interaction*, Vol. 8(2):7, MDPI, Basel, S. 1–21. https://doi.org/10.3390/mti8020007

Zusammen Erleben, Immersiv Teilhaben: Ein Praxisbeispiel zur Entwicklung immersiver VR-Technologien zur Steigerung der sozialen Teilhabe älterer Menschen – ZEIT

Felix Krooß, Danchen Zhang, Simon Kimmel, Wilko Heuten, Eric Landwehr, Yuliya Sobol, Fabian Hemmert, May-Britt Rohden; Anja Härtel, Clemens Winter und Arya Dhar

Zusammenfassung

Im Projekt »ZEIT« wurde eine immersive Virtual-Reality(VR)-Technologie zur Förderung der sozialen Teilhabe älterer Menschen entwickelt. Die Kombination aus einer interaktiven VR-Umgebung, Emotionserkennung und taktilem Feedback schuf ein System, das ältere Menschen dabei unterstützt, soziale Interaktionen auf neue Weise zu erleben. Durch die enge Einbindung der Zielgruppe in den Entwicklungsprozess wurde sichergestellt, dass die Technologie auf ihre Bedürfnisse abgestimmt ist. Erste Studien zeigen eine hohe Akzeptanz des Systems sowie das Potenzial, soziale Bindungen zu stärken und Einsamkeit zu reduzieren. Die Ergebnisse unterstreichen das Potenzial der virtuellen Realität als innovative Technologie für die digitale Gesellschaft.

1. Einleitung

Studien zeigen, dass ältere Menschen, die in ein unterstützendes soziales Netzwerk aus Familie, Freunden, Nachbarn und Gemeinschaftsgruppen eingebunden sind, seltener an Depressionen leiden und ein höheres Maß an Wohlbefinden berichten (Haslam et al., 2015: 2). Einsamkeit hingegen macht krank. In Zeiten der Corona-Pandemie wurde es einmal mehr deutlich, wie wichtig die soziale Verbundenheit für uns ist. Von geliebten Menschen getrennt und nicht in sozialer Kommunikation zu sein, ist schmerzhaft für uns. Das Risiko für das Auftreten von klinisch relevanten depressiven Symptomen im Nachgang einer Quarantänemaßnahme liegt bei ca. einem Drittel (Röhr et al., 2020: 185). Gerade mit zunehmendem Alter nehmen soziale Netzwerke und die soziale Teilhabe nachweislich

ab. Nach Angaben des Statistischen Bundesamtes leben 16,9 Millionen Menschen alleine in ihrer Wohnung, 5,8 Millionen sind dabei 65 Jahre alt oder älter (Destatis, 2021: 58). 2040 wird voraussichtlich jeder vierte Mensch in Deutschland alleine leben. Die Folge: Einsamkeit.

Aktuelle Kommunikationstechnologien überwinden für viele diese Einsamkeit, werden von älteren Menschen jedoch oft wenig akzeptiert. Hemmnisse sind die nichtzielgruppengerechte Ansprache und fehlende Adaption der Technologie auf die Einschränkungen der Zielgruppe. Wenig technikaffine Menschen, insbesondere ältere Bürgerinnen und Bürger, sind es nicht gewohnt, wie z. B. Jugendliche, über einen kleinen Bildschirm zu kommunizieren und lehnen dies ab. Hinzu kommt oft eine erschwerte Ablesbarkeit durch Sehschwächen. Eine Förderung von sozialer Teilhabe bedeutet damit auch eine Stärkung der Verbindung von älteren Bürgerinnen und Bürgern in ihren sozialen Netzwerken von Familien und Freund:innen.

Zur Vermittlung von Nähe über Distanz hat das im Projekt ZEIT forschende Team neue Möglichkeiten untersucht. Das daraus entstandene System basiert auf einer VR-Umgebung, in der sich mehrere Nutzende treffen können, um Spiele zu spielen und gemeinsam neue Erlebnisse zu schaffen. Seniorinnen und Senioren navigieren in einem Pflegesessel, der an die Zielgruppe angepasste Steuerelemente enthält, durch die VR-Umgebung. Im Pflegesessel ist zudem einerseits Sensorik verbaut, die es mittels Künstlicher Intelligenz (KI) ermöglicht, Emotionen zu erfassen und zu visualisieren. Andererseits verfügt der Sessel über smarte Textilien, die haptisches Feedback an die Nutzenden übertragen. So soll die virtuelle Nähe auch in der Realität spürbar werden. Im weiteren Verlauf werden die Entwicklungen ausführlich beschrieben.

1.1 Einsamkeit als gesellschaftliche Herausforderung

In der heutigen Gesellschaft ist die soziale Verbundenheit sowohl im privaten als auch im beruflichen Bereich gefährdet. Flexible Arbeitsregelungen und die Nutzung des Homeoffice erfahren eine stetig steigende Beliebtheit (vgl. Barrero, 2021: 11). Gleichzeitig führt die fortschreitende Globalisierung, welche den Menschen die Möglichkeit bietet, ihren Lebensmittelpunkt in andere Regionen oder Länder zu verlegen, zu einer Mobilität in nie gekanntem Ausmaß (vgl. Barrero, 2021: 25–27; United Nations International Migration Report, 2013, S1-18; McAuliffe, 2021: 21–58). Diese Entwicklung führt zu einer signifikanten Zunahme des Anteils der in Einpersonenhaushalten lebenden Personen, der im Jahr 2021 in Deutschland einen beachtlichen Wert von 41 % erreicht (vgl. Bundeszentrale für politische Bildung, 2021; Sozialbericht, 2021: 52). Des Weiteren hat

die Corona-Pandemie zu einer verstärkten räumlichen Distanzierung geführt, welche durch soziale Distanzierungsmaßnahmen und Quarantäneprotokolle zusätzlich intensiviert wurde (vgl. Ozaslan, 2020: 101). Die sozialen Folgen dieser Entwicklung sind bereits heute sichtbar: Rund 11 % der Bundesbürger sind von Einsamkeit betroffen, wobei es vor der Pandemie noch knapp 7 % waren (vgl. Bundesministerium für Familie, Senioren, Frauen und Jugend, 2024: 17–25). Der Anstieg der sozialen Isolation kann sowohl das psychische als auch das physische Wohlbefinden erheblich beeinträchtigen und möglicherweise Symptome wie Depressionen, Stress und Angstzustände verstärken (vgl. Jacob, 2019: 12; Smith, 2020: 1–3). Dies unterstreicht die dringende Notwendigkeit, sich mit den potenziellen Folgen dieses gesellschaftlichen Wandels auseinanderzusetzen und diesem u. a. durch technologische Innovation entgegenzuwirken.

2. Lösungsansatz

In diesem Projekt soll die soziale Verbundenheit von getrennt bzw. entfernt voneinander lebenden Menschen durch VR und soziale Signale gestärkt werden. Im Mittelpunkt dabei stehen die emotionalen Aspekte der zwischenmenschlichen Kommunikation, wie z. B. die Übertragung von Freude und Glück, Zufriedenheit, Unwohlsein, Stress und Sorgen, aber auch die sonst in der Kommunikation per Video und Audio fehlenden typischen physischen Interaktionen, wie die Berührung, Umarmung, der Handschlag zur Begrüßung oder das Schulterklopfen zum Lob und zur Bestätigung. Der verfolgte Lösungsansatz besteht aus fünf Teilaspekten und soll als Praxisbeispiel zur Entwicklung immersiver VR-Technologien zur Steigerung der sozialen Teilhabe älterer Menschen dienen. Die jeweiligen Teilaspekte sind in Abbildung dargestellt und werden im weiteren Verlauf dieses Kapitels detailliert beschrieben.

2.1 Emotionserkennung

Ein Teilaspekt der entwickelten Technologie ist ein System zur Erkennung von Emotionen, welches auf Basis von Vitalparametern den emotionalen Zustand der Nutzenden in Echtzeit erkennt.

Biomarker werden von Unternehmen und Wissenschaft zur Meldung von akuten Krankheitsbildern (z. B. Herzinfarkt) und Krankheitstendenzen erforscht. Herausforderungen sind die Langzeitaufnahme und Erhebung von Daten von spezifischen Zielgruppen. Daten älterer Menschen werden insbesondere in Krankphasen aufgenommen, d. h., gesunde Daten oder Informationen

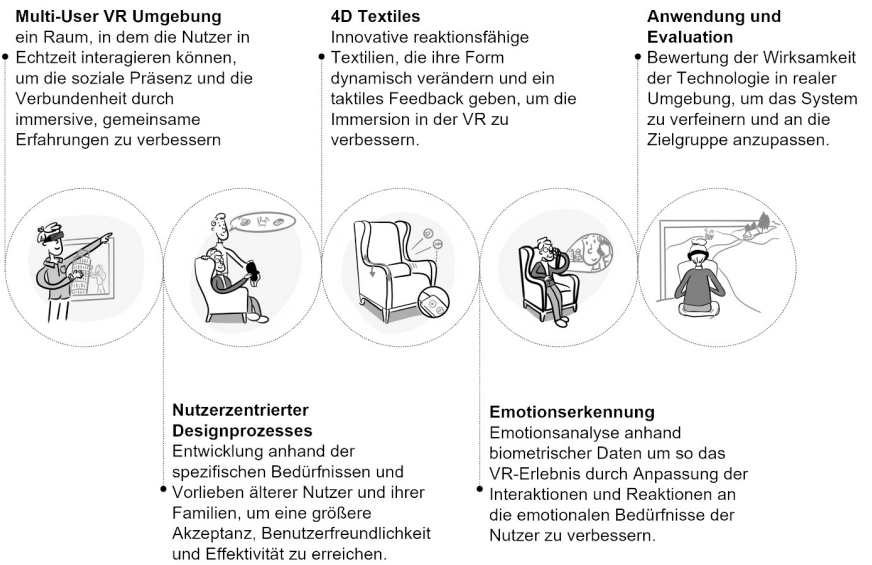


Abbildung 1: Teilaspekte zur Entwicklung einer immersiven VR-Technologie zur Steigerung der sozialen Teilhabe älterer Menschen
Quelle: Eigene Darstellung

über den emotionalen Zustand werden selten in Gesundphasen aufgenommen. In diversen Forschungsprojekten wurde der emotionale Zustand über Vitalparameter ermittelt, um situativ visuelle Hilfestellungen zu geben. Die Forschung des Affective Computing kennt in der Klassifizierung die Emotionen Freude, Verachtung, Zufriedenheit, Peinlichkeit, Aufregung, Schuld, Stolz, Erleichterung, Zufriedenheit, Sinnesfreude und Scham (Ekman, 1999: 45). Das Erkennen von emotionalen Biomarkern über Vitalparameter setzt klare emotionale Reaktionen voraus.

Als Grundlage zur Datenerfassung werden im Projekt empfindliche Ballistokardiografie-Sensoren (BCG) verwendet. Diese Sensoren werden in ein Sitzkissen integriert, um Biomarker für die Herzfrequenz und die Herzfrequenzvariabilität abzuleiten. Beide Parameter können zur Bestimmung von Emotionen verwendet werden, indem das Erregungsniveau einer Person aus diesen Daten abgeleitet wird. Niedrige oder hohe Erregungsniveaus beschreiben dabei eine potenziell ruhige oder angespannte Person. Die gemessenen Parameter werden in einem Machine-Learning-Modell ausgewertet und bilden das Fundament der Emotionsanalyse. Das Modell leitet Emotionen und Stress auf die zwei Dimensionen Arousal (Erregung von ruhig bis hoher Stress) und Valence (negativ bis positiv) ab. Die gemessenen Daten werden in Echtzeit durch einen Einplatinencomputer

verarbeitet und an die VR-Umgebung im Gesamtsystem zur Verfügung gestellt. Der Ablauf der Datenverarbeitung ist in Abbildung 2 dargestellt.

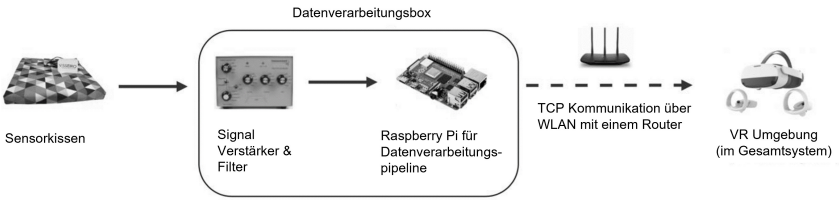


Abbildung 2: Datenverarbeitung der Emotionserkennung
Quelle: Eigene Darstellung

In diversen Studien wurden mithilfe verschiedener VR-Szenarien Emotionen und Stresssituationen ausgelöst. Parallel dazu erfolgte eine Datenerfassung der Herzaktivität mit dem zu entwickelnden Sensorkissen als auch mit einem EKG-Brustgurt. Im Anschluss der VR-Szenarien erfolgte eine Selbsteinschätzung der Emotionen und des Stresslevels mit zwei in der Literatur anerkannten sprachfreien Fragebögen, dem Self-Assessment Mannikin (SAM) und dem Positive and Negative Affect Schedule (PANAS).

In der anschließenden Auswertung wurden aus den beiden physiologischen Datenaufnahmen vom Sensorkissen und dem Referenz-EKG-Brustgurt die Herzrate und die Herzratenvariabilität (HRV) abgeleitet und mit der Selbsteinschätzung auf Korrelation über die vier Szenarien analysiert. Die Ergebnisse zeigen wie subjektiv und individuell Emotionen und Stressempfinden sind. Die physiologischen Daten deuten auf leichte Veränderungen der HRV während des Stressszenarios hin. Jedoch zeigen die Ergebnisse aus der Selbsteinschätzung, dass sich die Teilnehmenden zum Teil nicht unbedingt gestresst fühlten. Zwar zeigt sich bei einigen Teilnehmenden sehr wohl eine positive Korrelation zwischen emotionaler Selbsteinschätzung und der gemessenen physiologischen Herzratenvariabilität, jedoch eben nicht bei allen. Dieses hohe Maß an Subjektivität ist eine der Hauptherausforderungen bei der KI-gestützten Emotionserkennung.

Eine explorative Datenanalyse weiterer frei verfügbarer Datensätze im Forschungsbereich der Emotions- und Stresserkennung wurde durchgeführt. Hierbei bestätigten die Analysen den statistischen Zusammenhang zwischen situativen Emotionen und Stressempfinden und den physiologisch gemessenen Herzraten und Herzratenvariabilität. Ferner zeigten sie aber auch, dass Biomarker, welche im Projekt nicht aufgenommen werden können, einschließlich Pupillendilatation, EEG (Elektroenzephalografie) und GSR (Galvanische Hautleitfähigkeit),

gegebenenfalls signifikantere Aussagen erlauben könnten. Die Aussagekraft von HRV wie im Projekt ist oft limitiert.

Die verschiedenen extrahierten HRV-Parameter wurden vorverarbeitet, um sie für das Training mit einer tiefen neuronalen Netzwerkmethode, dem LSTM-Modell (Long Short-Term Memory), vorzubereiten. Das Modell wurde auf einem offenen Datensatz trainiert und auf einem anderen getestet. Nach der Abstimmung der Hyperparameter für das Modell wurde das Modell mit der besten Leistung gespeichert und auf den Raspberry Pi geladen, um am Ende der Pipeline Vorhersagen zu machen. Die Vorhersagen wurden dann mittels TCP-Kommunikation (Transmission Control Protocol) über WLAN mit einem Router an die VR-Umgebung übertragen.

2.2 Textile Aktuatoren

Zur Steigerung der Immersion in der VR Umgebung werden Textilien entwickelt, welche Berührungen an die Nutzenden übertragen können. Diese Textilien können ihre Form und ihre Temperatur verändern und werden auch als 4D-Textilien bezeichnet. Im Folgenden wird das grundlegende Prinzip der textilen Aktuatoren detailliert erläutert.

Die Aktuatoren sind darauf ausgelegt, ein komfortables taktiles Feedback zu erzeugen. Diese 4D-Textilien entstehen durch die Kombination von additiver Fertigung und flexiblen Textilien. Der 3D-Druck auf Textilien ermöglicht dabei einen Übergang von einer zweidimensionalen Konfiguration zu einer dreidimensionalen Morphologie.

Vor dem Druckprozess werden die Textilien gezielt vorgespannt, was zur Speicherung potenzieller Energie führt. Nach der Freisetzung dieser Spannung tritt eine rückstellende Kraft innerhalb des Textils auf, die der intrinsischen elastischen Steifigkeit in den gedruckten Bereichen entgegenwirkt. Dieser Rückkehrprozess zu einem energieoptimalen Gleichgewicht bewirkt eine Metamorphose der zweidimensionalen Druckstruktur hin zu einem räumlichen Design.

Die Steuerung dieser 4D-Textilien erfolgt durch den Einsatz von Garnen aus Formgedächtnispolymeren (engl. Shape Memory Polymeren [SMP]), die einen signifikanten Fortschritt in der Entwicklung elastischer Garne und Textilien darstellen und für dieses Projekt entwickelt wurden. Im Gegensatz zu herkömmlichen elastischen Materialien, die nach einer Deformation sofort ihre ursprüngliche Form zurückerlangen, behalten Formgedächtnismaterialien ihre veränderte Struktur bei, bis sie durch spezifische Reize aktiviert werden. Diese Reize können sowohl thermisch sein als auch alternative Stimuli wie elektromagnetische

Felder, Feuchtigkeit oder Licht bilden. Diese Anpassungsfähigkeit erlaubt eine gezielte Modifikation für verschiedenste Anwendungen.

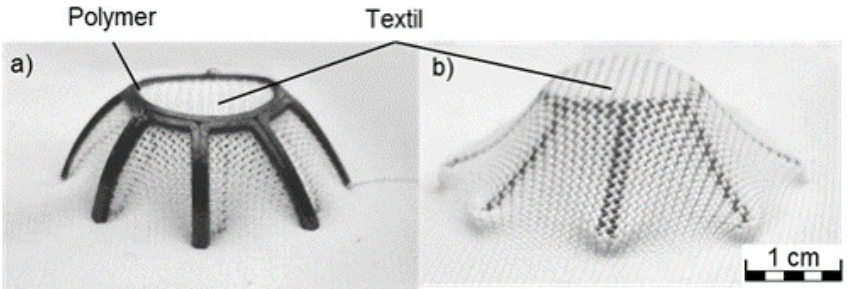


Abbildung 3: Sonne-Struktur mit Bistabilität
a) stabiler Zustand 1, b) stabiler Zustand 2
Quelle: Eigene Darstellung

In diesem Projekt wurde eine 4D-Struktur, genannt »Sonne«, mit zwei stabilen Zuständen entwickelt. Diese Struktur zeichnet sich durch eine zuverlässige Bistabilität aus und bietet einen ausreichend flexiblen Schaltweg für taktiles Feedback. Dabei wird sichergestellt, dass der Druck auf den Körper nicht zu stark ist, während gleichzeitig genügend Kraft erzeugt wird, um das Feedback spürbar zu machen. Abbildung 3 zeigt die Sonnenstruktur sowie die beiden stabilen Zustände der 4D-Textilien.

Das entwickelte SMP-Garn reagiert auf thermische Reize und wird bei einer Temperatur von etwa 40°C aktiviert. In Abbildung 44 ist zu sehen, wie das SMP-Garn ein Objekt anhebt, wenn es durch elektrisch erzeugte Wärme aktiviert wird.

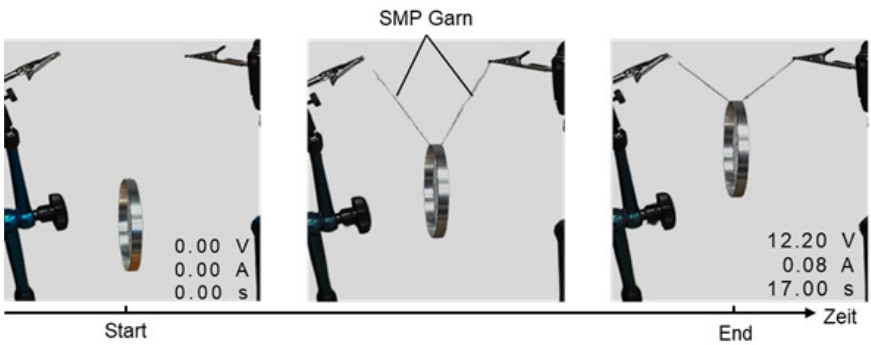


Abbildung 4: Bei der Aktivierung zieht sich das SMP-Garn zusammen.
Quelle: Eigene Darstellung

Um die Aktivierung der 4D-Textilien zu optimieren, wurde die Stickart des Garns auf der Struktur untersucht. Die in Abbildung dargestellte Stickart erweist sich als effektiv: Das Garn umschlingt jedes Bein der Struktur und kreuzt sich innerhalb derselben. Bei Aktivierung zieht sich das Garn zusammen und erzeugt dadurch eine Kraft, die die Beine der Struktur zusammenführt und somit die Umwandlung des Zustands realisiert.

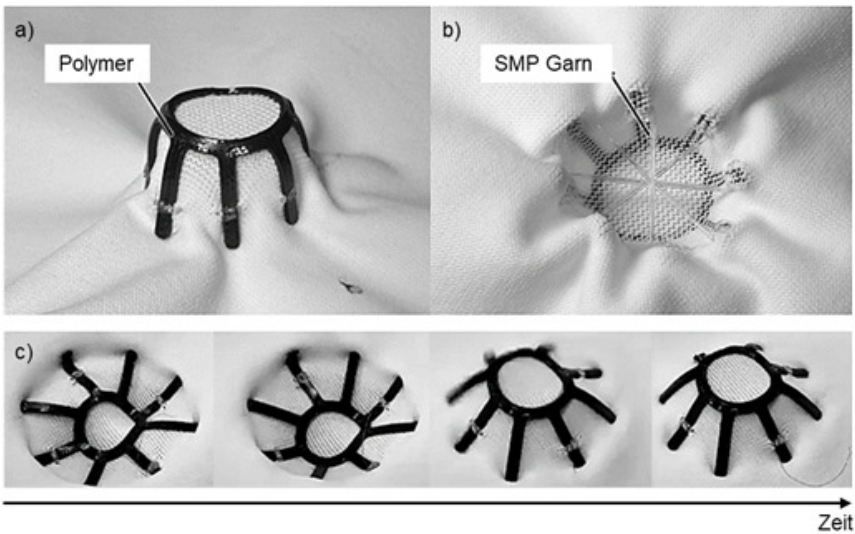


Abbildung 5: Optimierung der Aktivierung der 4D-Textilien
a) 4D-Textilien: Sonne-Struktur, b) Stickart des SMP-Garns, c) Aktivierung der 4D-Struktur
Quelle: Eigene Darstellung

Um das Garn thermisch zu aktivieren, wurde eine Heizfolie unter der textilen Struktur platziert, die durch ein Arduino gesteuert wird. Sechs textile Aktuatoren sind in das Nackenkissen integriert (siehe Abbildung 6).

2.3 Multi-User-VR-Umgebung

Um in unserem Projekt eine synchrone Kommunikation in Echtzeit über Distanzen hinweg zu ermöglichen, haben wir eine Multi-User-VR-Umgebung implementiert. Diese stellt ein technologisches Kommunikationsmedium dar, das die gemeinsame Interaktion in virtuellen Räumen und mit virtuellen Objekten ermöglicht. Im Gegensatz zur einfachen Videotelefonie, die seit der Pandemie an Popularität gewonnen hat, bietet die VR-Umgebung eine umfassendere Erfah-

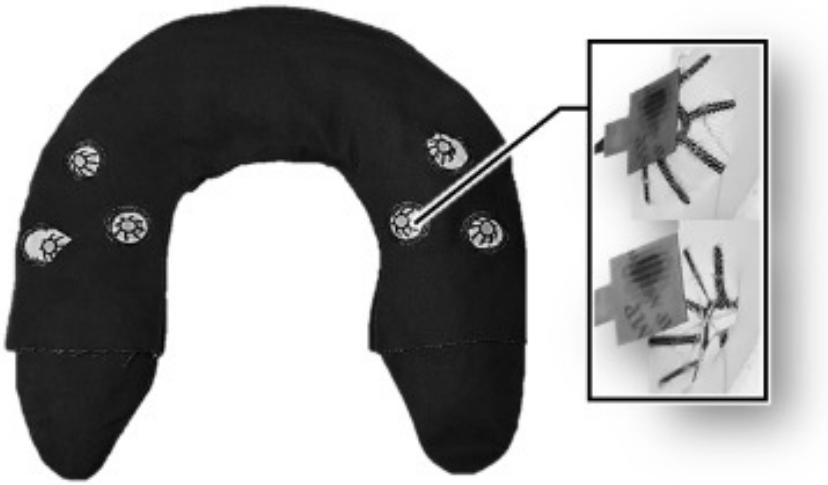


Abbildung 6: Das Nackenkissen mit sechs textilen Aktuatoren
Quelle: Eigene Darstellung

rung, die es ermöglicht, vergangene Erlebnisse zu teilen und gemeinsam neue Erfahrungen zu machen.

Die Schaffung der VR-Umgebung für mehrere Benutzer war mit verschiedenen technologischen Herausforderungen verbunden, die im Folgenden beschrieben werden.

2.3.1 Aufbau der Multi-User-Infrastruktur unter Berücksichtigung eines hohen Datenschutzniveaus

Um die Multi-User-VR-Umgebung zu erstellen, mussten zunächst die Anforderungen an das VR-Headset definiert werden. Ein Headset, das mobile Einsatzmöglichkeiten bietet, mit der Unity Game Engine entwickelt werden kann, Handtracking unterstützt und der Datenschutzgrundverordnung entspricht, wurde gesucht. Nach einer sorgfältigen Auswahl haben wir uns für das Pico Neo 3 Pro als Hardware entschieden.

Ein grundlegendes Framework für eine Multi-User-VR-Umgebung wurde entwickelt, das es zwei VR-Benutzer:innen ermöglicht, sich über das Internet zu verbinden und miteinander zu kommunizieren. Als Netzwerkinfrastruktur wurde das System »Netcode« gewählt, welches eine kostenlose und sichere Möglichkeit bietet, mehrere Benutzer:innen über einen selbst gehosteten Server miteinander zu verbinden. Um die Sicherheit zu erhöhen und einen vollständigen Datenschutz im Rahmen der Experimente zu gewährleisten, ist der Server

nur über einen speziell gesicherten VPN-Zugang (Virtual Private Network) erreichbar.

2.3.2 Integration einer Nutzendenrepräsentation

Um ein hohes Maß an sozialer Präsenz unter den Nutzenden zu erzeugen, wurde eine Nutzendenrepräsentation in die VR-Umgebung integriert. Es wurden verschiedene technologische Ansätze zur Erstellung virtueller Avatare untersucht und schließlich die Entscheidung für die Integration des Ready-Player-Me-Systems für weniger fotorealistische Avatare sowie für die Integration von Avaturn, das die Erstellung möglichst fotorealistischer Nutzendenrepräsentation ermöglicht, getroffen.

Für die Integration von Avaturn wurde eine eigene Anwendung, die auf Basis der Avaturn-API aus drei Bildern einer Person einen Avatar erstellt und diesen automatisch in die VR-Anwendung laden kann, entwickelt. Neben dem realistischen Aussehen wurde auch besonderer Wert auf ein besonders realistisches Verhalten der Avatare in der VR-Umgebung gelegt. Die eigenen Arbeiten im Projekt haben immer wieder gezeigt, dass dieser erhöhte Verhaltensrealismus besonders wichtig ist, um bei den Nutzenden das Gefühl der sozialen Präsenz zu erzeugen, also tatsächlich mit einer anderen Person an einem Ort zu sein (Kimmel, 2022: 2; Kimmel, 2023: 12; Kimmel, 2024: 20).

Basierend auf unseren Ergebnissen wurden Handtracking, Inverse Kinematics und Lip-sync-Technologie implementiert, um Mund, Hand und andere Körperbewegungen so realistisch wie möglich zu simulieren und damit den Verhaltensrealismus so weit wie möglich zu erhöhen.

2.3.3 Erstellung virtueller Interaktionsumgebungen

Die Erstellung virtueller Interaktionsumgebungen stellte einen wesentlichen Schritt bei der Entwicklung der Multi-User-VR-Umgebung dar. Die Ergebnisse der Userevaluationen wurden genutzt, um Designkonzepte für die zu erstellenden VR-Umgebungen zu entwickeln. Diese Konzepte werden im Abschnitt 2.4 »Nutzerzentrierter Designprozess« präsentiert. In Abbildung 7 sind die verschiedenen im Experiment verwendeten virtuellen Szenarien dargestellt.

Flur: Der Flur dient als Ausgangspunkt für die Anwendung. In diesem Raum sind die Nutzenden noch nicht mit anderen Nutzenden verbunden. Hier können sie sich mit der Steuerung vertraut machen und einen Avatar als Vertretung der Nutzenden auswählen. Der Flur ist ein wichtiger Teil der Anwendung, da er es den

Nutzenden ermöglicht, sich mit der Umgebung vertraut zu machen, bevor sie in die gemeinsame Erfahrung einsteigen.

Wohnzimmer: Das Wohnzimmer dient als Ausgangspunkt für die gemeinsame VR-Erfahrung. Von hier aus können die gemeinsamen VR-Erlebnisse ausgewählt werden. Zu diesem Zweck kann während der Nutzung jederzeit in das Wohnzimmer zurückgekehrt werden. Das Wohnzimmer ist ein wichtiger Teil der Anwendung, der es den Nutzenden ermöglicht, ihre Erlebnisse zu planen und diese gemeinsam zu beginnen.

360-Grad-Videoraum: Der 360-Grad-Videoraum ermöglicht es den Nutzenden, gemeinsam 360-Grad-Videos anzusehen, die vom Konsortium aufgenommen wurden. Diese Videos beinhalten Aufnahmen aus dem Wuppertaler Zoo, einem Wald, einem griechischen Strand und Paris. Der 360-Grad-Videoraum ist ein wichtiger Teil der Anwendung, da er es den Nutzenden ermöglicht, Erfahrungen auszutauschen und gemeinsam zu lernen.

Erklärspiel: Das Erklärspiel ermöglicht es den Nutzenden, ein Spiel zu spielen, in dem sie sich gegenseitig Begriffe mit Worten oder Grafiken erklären. Indem das Spiel ein hohes Maß an Kollaboration erfordert, Emotionen auslöst sowie das Erinnern vergangener Erlebnisse fördert, trägt das Spiel dazu bei, soziale Verbindungen zwischen den Nutzenden zu fördern.

Reiseraum: Der Reiseraum ermöglicht es, an bekannte Orte zu »reisen«. Derzeit unterstützt der Reiseraum eine Reise nach Paris. Dabei sind beide Nutzenden auf einem Balkon mit Blick auf den virtuellen Eifelturm positioniert. Der Reiseraum soll es den Nutzenden ermöglichen, gemeinsame Erlebnisse zu schaffen und sich an vergangene Reisen zu erinnern.

Fotogalerie: Die Fotogalerie ermöglicht die Navigation durch eine VR-Fotogalerie mit persönlichen Bildern. Zu Demonstrationszwecken und aus Datenschutzgründen sind hier nur lizenzfreie Stockbilder eingebettet, die das Potenzial einer solchen Umgebung zur Darstellung persönlicher Bilder widerspiegeln sollen.

Bauklötze-Spiel: In einem weiteren Spiel bieten wir den Nutzenden die Möglichkeit, gemeinsam Figuren aus Bauklötzen zusammenzusetzen. Dabei sind die Nutzenden auf die Mithilfe ihres/ihrer Interaktionspartner:in angewiesen, da jeweils ein/eine Spieler:in aktiv Bauklötze bewegen kann, während das Gegenüber die Bauanleitung einsehen kann. Das Spiel fördert die räumliche Bewegung

in der VR-Umgebung und soll durch seinen kollaborativen Charakter die soziale Bindung zwischen den Nutzenden erhöhen.



Abbildung 7: Übersicht über alle virtuellen Interaktionsumgebungen der ZEIT-Anwendung
Quelle: Eigene Darstellung

Die diversen im Rahmen des Projektes entwickelten virtuellen Interaktionsumgebungen stellen einen essenziellen Kern für die ZEIT-Anwendung dar. Sie ermöglichen es den Nutzenden, gemeinsam zu interagieren, Erfahrungen zu teilen und Erinnerungen wiederzuerleben. Zur Verbesserung der Umgebungen wurden diese im Rahmen des Projektes iterativen Tests unterzogen und fortlaufend verbessert.

2.3.4 Implementierung zielgruppengerechter Interaktion und Navigation in den virtuellen Räumen

Um den Nutzenden einen einfachen und intuitiven Wechsel zwischen den virtuellen Umgebungen zu ermöglichen, wurde im Rahmen des Projektes ein zielgruppengerechtes Interaktionskonzept entwickelt und umgesetzt.

Im virtuellen »Wohnzimmer« wurden verschiedene Interaktionsobjekte wie Bilder an der Wand oder andere markante Objekte platziert, die jeweils eine andere virtuelle Umgebung repräsentieren (siehe Abbildung 8). Durch längeres Zeigen auf die verschiedenen Objekte kann ein Wechsel in die entsprechende Umgebung initiiert werden. Ebenso kann aus allen anderen Umgebungen durch längeres Zeigen auf ein kleines Häuschen, das sich in jedem Raum befindet, in das

Wohnzimmer gewechselt werden. Aufgrund der Schwierigkeiten bei der Interaktion mit Controllern, die in unserer Zielgruppe vorherrschen, wurde die Interaktion mit der Umgebung für das native Handtracking des Pico Neo 3 Pro konzipiert und so umgesetzt. Dieses erfasst die Handbewegungen sowie die Zeigerichtung des Nutzenden präzise.



Abbildung 8: Implementierte Interaktion zur Auswahl einer virtuellen Umgebung
Sicht von außen (links) und aus Kameraperspektive (rechts)
Quelle: Eigene Darstellung

Neben der Interaktion und der Auswahl weiterer Räume wurde auch ein zielgruppengerechtes Navigationskonzept zur Bewegung in der VR-Umgebung benötigt. Das in unserem Projekt entwickelte Konzept sieht vor, dass ein umgedrehter Controller des VR-Systems, der mithilfe von 3D-gedruckten Komponenten am Sessel befestigt wird, als Joystick zur Bewegung in der VR-Umgebung genutzt werden kann. Dieses Design hat u. a. den Vorteil, dass der (umgedrehte) Controller bereits automatisch vom System getrackt wird und die Positionierung des Controllers ohne die Integration weiterer Tracking-Technologien für die Nutzenden genutzt werden kann. Das Design baut zusätzlich auf die Metapher eines motorisierten Rollstuhls, der in der Zielgruppe sehr bekannt ist. So ermöglicht unser Design den Nutzenden eine intuitive und natürliche Navigation in der VR-Umgebung.

2.3.5 Zusammenführen aller Teilkomponenten in ein Gesamtsystem

Über die Darstellung der virtuellen Umgebung und die Vernetzung der Benutzer:innen hinaus sollte das Multi-User-VR-System auch dazu genutzt werden, die zuvor dargestellten Teilkomponenten zu einem Gesamtsystem zu vereinen. Dazu wurden im Rahmen des Projektes Kommunikationsschnittstellen zwischen dem VR-System und dem Aktuator-Textil bzw. der Emotionserkennung identifiziert und eine Systeminfrastruktur entwickelt, in der die Einzelkomponenten drahtlos zu einem Gesamtsystem zusammengeführt werden können.

Die entwickelte Systeminfrastruktur basiert auf einem 5G-Router mit WiFi-6-Unterstützung. Der Router ermöglicht eine schnelle drahtlose Vernetzung der Einzelkomponenten.

Für die Implementierung einer drahtlosen Kommunikationsschnittstelle zwischen VR-System und Aktuatortextil wurde eine TCP-Verbindung entwickelt und evaluiert, die bei Auftreten einer »virtuellen« Berührung im Schulterbereich innerhalb der Multi-User-VR-Umgebung ein Signal an den Mikrocontroller des Aktuatortextils sendet, welcher daraufhin die aktorische Ausgabe an das Gegenüber initiiert.

Der entsprechende Bereich, der von den Nutzenden berührt werden kann, wird durch leuchtende Funken über der Schulter symbolisiert, die bei Berührung die Farbe wechseln (siehe Abbildung 9, rechte Darstellung). Zur Integration der drahtlosen Kommunikation zwischen VR-System und Emotionserkennung wurde eine drahtlose TCP-Verbindung zwischen dem VR-System und einem Raspberri Pi entwickelt und evaluiert. Über diese Schnittstelle werden die erkannten emotionalen Zustände an die VR-Brille übertragen, wo sie im nächsten Schritt den Nutzenden präsentiert werden. Die Visualisierung zeigt eine Farbwolke über dem Avatar der Nutzenden, deren Farbe sich nach dem Arousal richtet. Dabei steht die Farbe Weiß für ein niedriges Arousal und eine bunte Farbe für ein hohes Arousal. Die Darstellung der emotionalen Zustände erfolgt alle 30 Sekunden oder wenn eine Veränderung erkannt wird (siehe Abbildung 9, linke Darstellung).



Abbildung 9: Visuelle Repräsentation des Emotionszustandes (links) und der Berührung des Gegenübers (rechts)
Quelle: Eigene Darstellung

Sowohl im Rahmen des Integrationstests als auch bei der abschließenden Evaluation wurde das Gesamtsystem implementiert und seine Schnittstellen evaluiert. Die Ergebnisse der Evaluation zeigen, dass die Kommunikationsschnittstellen funktionsfähig sind und das Gesamtsystem erfolgreich implementiert wurde.

2.4 Nutzerzentrierter Designprozess

Im Rahmen des Projekts wurde eine immersive VR-Technologie entwickelt, bei der Seniorinnen und Senioren aktiv in den Entwicklungsprozess einbezogen wurden, um ihre Bedürfnisse und Wünsche zu berücksichtigen. Der verfolgte Entwicklungsansatz wird als nutzerzentrierter Designprozess bezeichnet und zielt darauf ab, eine hohe Benutzerfreundlichkeit und Effektivität der zu entwickelnden Technologie zu gewährleisten. Der gesamte Entwicklungsprozess wurde zusätzlich durch Workshops zu ethischen, legalen und sozialen Aspekten (ELSA) unterstützt. Ziel der Workshops war es, eine detaillierte Analyse der Vulnerabilität der Nutzenden zu ermöglichen.

Zu Beginn des Projektes wurden semi-standardisierte Interviews mit Senioren und Seniorinnen und deren Angehörigen durchgeführt, um ein besseres Verständnis für ihre Bedürfnisse und Anforderungen zu erhalten. In den Interviews mit zehn Senioren und Seniorinnen im Alter von 74 bis 96 Jahren zum Thema der Generationenkommunikation wurden Einblicke zu wichtigen Erinnerungen, Kommunikationsformen und früheren Aktivitäten gesammelt. Dabei spiegelten die Erzählungen und die Erinnerungen der Teilnehmenden sowohl positive als auch schmerzhafte Erfahrungen wider. Sensible Themen wie Erlebnisse während des Zweiten Weltkriegs wurden meist vermieden, während schöne Erinnerungen an Familie, Freund:innen und Reisen gern geteilt wurden. Viele Senioren und Seniorinnen äußerten, dass sie die mit diesen Erinnerungen verbundenen Erlebnisse, wie Familientreffen, Hochzeiten oder Theater- und Konzertbesuche, sehr vermissen. Besonders das Reisen spielt in ihren Erinnerungen eine zentrale Rolle.

Auf Basis dieser Interviews wurde eine erste Studie durchgeführt, bei der Senioren und Seniorinnen über VR-Brillen 360-Grad-Videos mit Natur- und Urlaubsszenen gezeigt wurden, wobei die Inhalte auf den Ergebnissen der vorherigen Interviews basierten. Die VR-Erfahrung wurde von den Teilnehmenden als »Mini-Urlaub« beschrieben. Die Senioren und Seniorinnen berichteten, dass sie sich in eine andere Welt versetzt gefühlt hätten. Szenen wie ein Wald, Pinguine im Zoo und ein Strand wurden besonders positiv wahrgenommen. Die Ergebnisse dieser Studie verdeutlichen sowohl das Potenzial als auch die Akzeptanz neuer Technologien bei älteren Menschen.

In einer zweiten Studie mit 15 Teilnehmenden (11 Frauen und 4 Männer) im Alter von 20 bis 29 Jahren wurde die Benutzerfreundlichkeit des Systems untersucht. Ziel war es, den Demonstrator zunächst mit einer jüngeren Zielgruppe (Enkelkinder) zu testen, um sicherzustellen, dass das System intuitiv bedienbar und leicht verständlich ist. Eindrücke der Studie sind in Abbildung 10 zu sehen. Es wurde festgestellt, dass die Teilnehmenden sich in der Anwendung sicher fühlten

und die Möglichkeit der sozialen Interaktion positiv bewerteten. Die Bewegungen der Lippen, die Gestik, die Stimme sowie die Möglichkeit, die andere Person an der Schulter zu berühren, vermittelten das Gefühl, dass die andere Person wirklich da ist. Allerdings gab es auch Kritikpunkte und Verbesserungsvorschläge. Beispielsweise wurde der Mangel an akustischen Signalen, die im realen Leben allgegenwärtig sind, bemängelt. Auch die Navigation mit Objekten sowie das gesamte Navigationskonzept wurden infrage gestellt und als schwer bedienbar bewertet, vor allem für die primär anvisierte Zielgruppe.

Die Anmerkungen und Vorschläge der Teilnehmenden wurden ausgewertet und in die nächste Version des Demonstrators integriert, um die Nutzerfreundlichkeit des Demonstrators weiter zu optimieren.



Abbildung 10: Teilnehmende während Usability-Demonstrator-Studie
Quelle: Eigene Darstellung

2.5 Anwendung und Evaluation

Das entwickelte Gesamtsystem besteht aus einem Pflegesessel, der mit verschiedenen Komponenten ausgestattet ist, darunter eine VR-Brille bzw. -Applikation, ein Sensorkissen und ein Nackenkissen. Das System ist in Abbildung 11 dargestellt.

Das auf Basis der zuvor beschriebenen Schritte entwickelte System wurde in zwei Phasen getestet. Zunächst wurde das System ausschließlich mit Senioren und Seniorinnen getestet, wobei die Versuchsleiterinnen die Rolle der Angehörigen übernahmen. In der zweiten Phase nahmen die Senioren und Seniorinnen gemeinsam mit ihren Angehörigen teil.

Am ersten Teil der Studie beteiligten sich zehn Senioren und Seniorinnen im Alter von 53 bis 86 Jahren (Abbildung 12). Die Tests wurden in den Pflegeeinrichtungen der AWO durchgeführt. Alle Teilnehmenden verbrachten zwischen 20 und 39 Minuten in der VR-Welt. Sechs der Teilnehmenden hatten keine vorherige VR-Erfahrung, während vier die VR-Brille bereits in einer früheren Studie verwendet hatten.

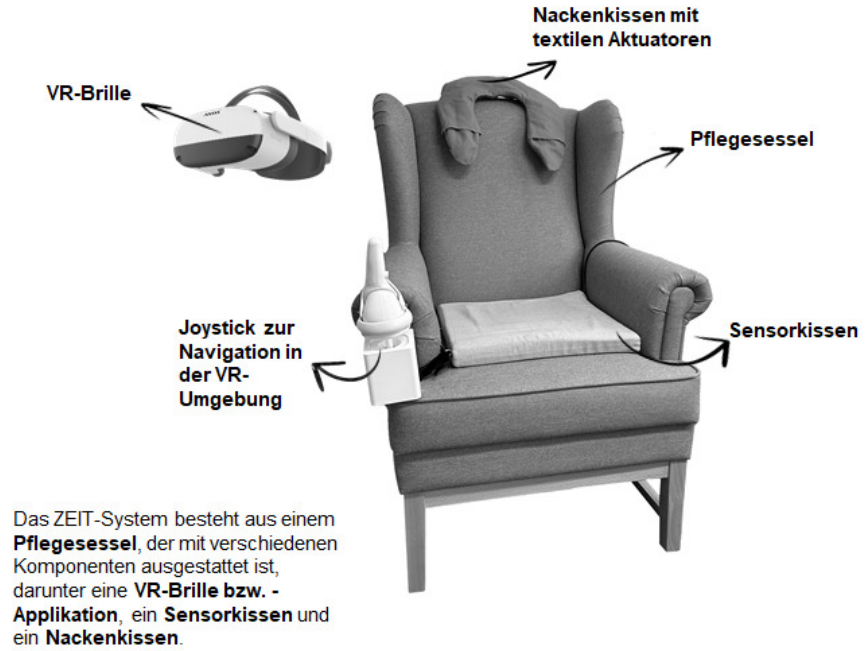


Abbildung 11: Gesamtsystem zur Steigerung der sozialen Teilhabe älterer Menschen
Quelle: Eigene Darstellung



Abbildung 12: Senioren und Seniorinnen während der Demonstrator-Studie
Quelle: Eigene Darstellung

Die Erfahrung wurde von allen Teilnehmenden als sehr positiv bewertet und sie waren mit dem System zufrieden. Eine Seniorin beschrieb die VR-Erfahrung als *»wie im Traum, wie in einer anderen Welt« (TN2, w, 86 Jahre alt)*, während eine andere Teilnehmerin ergänzte: *»Eine tolle Idee, weil ich glaube, dass das wirklich eine gute Möglichkeit ist. Gerade, wenn die Familie weit weg ist, kann man so etwas gemeinsam erleben.« (TN3, w, 53 Jahre alt)*

Die 360-Grad-Videos wurden besonders gut angenommen, da sie den Senioren und Seniorinnen das Gefühl gaben, an einem anderen Ort zu sein und neue Erlebnisse zu genießen. Auch die interaktiven Spiele sorgten für viel Freude. Alle Teilnehmenden spielten mindestens zwei Runden des Spiels »Begriffe mündlich erklären«. Ein Teilnehmer sagte dazu: »*Das Spiel fand ich besonders toll. Das kann ich mir echt gut vorstellen. Es passt perfekt in diese Umgebung, weil hier oft niemand zum Spielen da ist. Hier leben zwar 90 oder 95 andere Leute, aber keiner setzt sich mit dir an den Tisch, um zu spielen.*« *(TN4, m, 66 Jahre alt)*

Acht von zehn Teilnehmenden berichteten, dass sie während der VR-Erfahrung für einen Moment vergessen hätten, dass sie sich eigentlich in einem Pflegeheim befinden, da die virtuelle Welt so realistisch wirkte. Obwohl es zu Beginn einige Probleme mit der Steuerung gab, haben sich die Teilnehmenden nach kurzer Zeit daran gewöhnt und konnten die VR problemlos nutzen.

Die Ergebnisse dieser Studie zeigen, dass die immersive und interaktive Natur der VR-Technologie eine starke Wirkung auf die Senioren und Seniorinnen hatte, sowohl in Bezug auf das soziale Miteinander als auch auf das Gefühl, neue Welten zu entdecken.

Am zweiten Teil der Studie nahmen sieben Paare teil: sieben Seniorinnen (5 weiblich, 3 männlich) im Alter von 71 bis 86 Jahren sowie ihre Angehörigen (6 weiblich, 2 männlich) im Alter von 21 bis 58 Jahren. Die Tests wurden erneut in den Pflegeeinrichtungen der AWO durchgeführt. Die Teilnehmenden haben die Studie aus verschiedenen Räumen heraus durchgeführt, sodass sie sich nur in der virtuellen Welt hören und sehen konnten. Alle verbrachten zwischen 30 und 40 Minuten in der VR-Welt, wobei die Studienleitung bei Problemen unterstützte. Einige Rückmeldungen der Teilnehmenden sind in Abbildung 13 dargestellt.

Das VR-Erlebnis hat nicht nur Senioren und Seniorinnen, sondern auch für Angehörige beeindruckt. Eine Teilnehmerin hat durch die Studie herausgefunden, wie viel Verbindung sie eigentlich zu ihrer Mutter hat: »*… es war auch interessant zu sehen, wie nah wir beide uns eigentlich doch sind. Man lebt irgendwie das Leben, dann denkt nicht nach und dann bei sowas kommt sowas raus, faszinierend. (TN6, w, 55 Jahre alt).*

Während der Studie hatten alle Teilnehmenden das Gefühl, zusammen an einem Ort gewesen zu sein und die andere Person gegenüber in der VR-Welt zu »fühlen«. Eine Teilnehmerin sagte: »*Ich habe ihren Avatar gesehen und sie ist halt auf das eingegangen, worauf ich gezeigt habe. Wenn ich gesagt habe, auch noch ein Stück rechts oder drück noch mal und das schafft so eine Verbindung. Ich habe auch ihre Hand berührt, ich weiß es ist komisch, aber ich hatte das gefühlt, ich hatte sie wirklich berührt.*« (53, Angehörige, w, TN7)

Alle Teilnehmenden äußerten den Wunsch, das Erlebnis erneut zu teilen, zusammen zu spielen und zu reisen.



Abbildung 13: Kommentare der Teilnehmer und Teilnehmerinnen zur Erfahrung mit dem ZEIT-System
Quelle: Eigene Darstellung

3. Fazit und Ausblick

Im beschriebenen Forschungsprojekt wurde eine immersive VR-Technologie entwickelt, welche sich das Ziel setzt, die soziale Teilhabe älterer Menschen zu stärken. Das Projekt zeigt, wie Senioren und Seniorinnen in einer Multi-User-VR-Umgebung Erlebnisse teilen, neue Erfahrungen machen und soziale Bindungen stärken können. Ein partizipativer und nutzerzentrierter Designprozess war dabei der Schlüssel, um sicherzustellen, dass die Technologie auf die Bedürfnisse und Einschränkungen der älteren Zielgruppe abgestimmt ist. Die Nutzer und Nutzerinnen wurden von Anfang an in die Gestaltung der Technologie eingebunden und die Zielgruppe über die Dauer des Projektes genau erforscht. Die Ergebnisse zeigen, dass diese Form der virtuellen Interaktion das Potenzial besitzt, Einsamkeit zu verringern und ein Gefühl von sozialer Nähe zu schaffen.

Die Einbindung von Senioren und Seniorinnen und ihren Angehörigen im Entwicklungsprozess hat gezeigt, dass die VR-Technologie nicht nur akzeptiert werden, sondern auch als bereichernd empfunden werden kann. Insbesondere die Nutzung des Systems zur Reise an virtuelle Orte oder zur Interaktion mit Angehörigen in digitalen Gesellschaftsspielen fand große Zustimmung. Durch die Kombination von VR, taktilem Feedback und Emotionserkennung konnte eine tiefere Immersion erreicht werden, welche für die Zielgruppe von besonderer Bedeutung ist.

Für die Zukunft bieten die Projektergebnisse spannende Perspektiven. Dabei ist es wichtig, die Technologie noch zugänglicher zu machen, beispielsweise durch die Entwicklung von vereinfachten Steuerungselementen und die Integration zusätzlicher sensorischer Feedbacksysteme wie z. B. akustische Signale. Ebenso bietet die Untersuchung der Technologie in unterschiedlichen kulturellen Kontexten einen interessanten Ansatz, um deren Wirksamkeit und Anpassungsfähigkeit in verschiedenen sozialen Umfeldern zu untersuchen.

Insgesamt zeigt das Projekt, dass immersive Technologien eine vielversprechende Lösung zur Förderung der sozialen Teilhabe älterer Menschen darstellen. Für die Zukunft ist es notwendig, diese Ansätze weiter zu optimieren und auf eine breitere Nutzerbasis zu skalieren, damit langfristig zur Verbesserung der Lebensqualität älterer Menschen beigetragen werden kann.

Förderhinweis

Das Projekt ZEIT wurde vom Bundesministerium für Bildung und Forschung (BMBF)[1] gefördert.

Literatur

Barrero, Jose Maria, Bloom, Nicholas & Davis, Steven J. (2021). Why Working From Home Will Stick. In: *SSRN Electronic Journal*, Elsevier Verlag, Amsterdam. https://dx.doi.org/10.2139/ssrn.3741644

Bundesministerium für Familie, Senioren, Frauen und Jugend (Hg.) (2024). *Einsamkeitsbarometer 2024. Langzeitentwicklung von Einsamkeit in Deutschland,* Schobin, Janosch, Arriagada, Céline & Gibson-Kunze, Martin (Autorenschaft), Kompetenzzentrum Einsamkeit, Institut für Sozialarbeit und Sozialpädagogik e.V., BMFSFJ Bundesministerium für Familie, Senioren, Frauen und Jugend, Berlin.

Destatis (2020). *16,9 Millionen Menschen lebten 2018 alleine*, Destatis Statistisches Bundesamt, Wiesbaden. https://www.destatis.de/DE/Presse/Pressemitteilungen/2020/03/PD20_N014_122.html%3Bjsessionid=8CDE48277A5F6FF92EBCC935658AF63C.internet8711 [Zugriff 21.07.2020]

Ekman, Paul (1999). Basic Emotions. In: *Handbook of Cognition and Emotion*, Dalgleich, Tim & Power, Mick J., John Wiley & Sons Ltd., Hobokon. https://doi.org/10.1002/0470013494.ch3

Haslam, Catherine et al. (2015). Social Connectedness and Health. In: *Encyclopedia of Geropsychology*, Springer, Singapore. https://doi.org/10.1007/978-981-287-080-3_46-2

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

Jacob, Louis, Haro, Joseph Maria & Koyanagi, Ai (2019). Relationship between Living Alone and Common Mental Disorders in the 1993, 2000 and 2007 National Psychiatric Morbidity Surveys. In: *PLoS ONE*, Vol. 14(5), e0215182, Public Library of Science (PLOS), San Francisco. https://doi.org/10.1371/journal.pone.0215182

Kimmel, Simon & Heuten, Wilken (2022). Elevating Social Presence in Multi-User VR by Increasing Behavioral Realism. In: *Social Presence in Virtual Event Spaces – A CHI 2022 Workshop*, New Orleans, Association for Computing Machinery (ACM) New York. https://www.offis.de/offis/publikation/elevating-social-presence-in-multi-user-vr-by-increasing-behavioral-realism.html [Zugriff 20.03.2025]

Kimmel, Simon et al. (2023). Let's Face It: Influence of Facial Expressions on Social Presence in Collaborative Virtual Reality. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery (ACM) New York, S. 1–16. https://doi.org/10.1145/3544548.3580707

McAuliffe, M., & Triandafyllidou, A. (2021). *World Migration Report 2022* [White paper], International Organization for Migration (IOM), Geneva.

Mehmet Ozaslan et al. (2020). Practical Measures to Prevent COVID-19: A Mini-Review. In: *Journal of Biological Sciences*, Vol. 20(2), Asian Network for Scientific Information, Dubai, S. 100–102. https://scialert.net/abstract/?doi=jbs.2020.100.102 [Zugriff 20.03.2025]

Statistisches Bundesamt (Destatis), Wissenschaftszentrum Berlin für Sozialforschung (WZB), Bundesinstitut für Bevölkerungsschutz (BiB) (Hg.) (2021). Datenreport 2021. Ein Sozialbericht fur die Bundesrepublik Deutschland, In: *Zeitbilder,* Bundeszentrale für politische Bildung bbp, Bonn. https://www.bib.bund.de/Publikation/2021/pdf/Datenreport-2021-Ein-Sozialbericht-fuer-die-Bundesrepublik-Deutschland.pdf?__blob=publicationFile&v=2 [Zugriff 20.03.2025]

Röhr, Susanne et al. (2020). Psychosoziale Folgen von Quarantänemaßnahmen bei schwerwiegenden Coronavirus-Ausbrüchen: ein Rapid Review. In: *Psychiatrische Praxis*, Vol. 47, Georg Thieme Verlag KG, Stuttgart. S. 179–189. https://doi.org/10.1055/a-1159-5562

Smith, Mathew Lee, Steinman, Lesley E. & Casey, E. A. (2020). Combatting Social Isolation Among Older Adults in a Time of Physical Distancing: The COVID-19 Social Connectivity Paradox. In: *Frontiers in Public Health*, Vol. 8, Frontiers Media S.A., Lausanne. https://doi.org/10.3389/fpubh.2020.00403

United Nations (2013). *International Migration Report 2013* [White paper], United Nations, Department of Economic and Social Affairs, Population Division. New York. https://www.un.org/en/development/desa/population/publications/pdf/migration/migrationreport2013/Full_Document_final.pdf [Zugriff 20.03.2025]

V.
Tranformative Technologien für eine inklusivere Gesellschaft

V/AR-Technologie für die digitale Gesellschaft – Living Lab »XR-Interakt«

Jana Tröge, Heike Wiesner und Christoph Runde

Zusammenfassung

Virtuell und Augmented Reality (V/AR) gewinnen als Technologien in allen Lebens- und Arbeitskontexten zunehmend an Bedeutung. Die Bürgerinnen und Bürger der Zukunft werden nicht mehr nur passiv von diesen transformativen Technologien betroffen sein, sondern müssen diese Angebote auch aktiv mitgestalten, um nicht nur die Akzeptanz zu erhöhen, sondern auch einen gesellschaftlichen Nutzen daraus zu generieren. Neue Anwendungsfelder von V/AR-Technologien sollten daher nicht nur vielfältige Nutzergruppen adressieren, sondern diese auch partizipativ in die Entwicklungsprozesse von Anfang an mit einbinden. Das Bundesministerium für Bildung und Forschung (BMBF)[1] hat dafür das Förderprogramm »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« für die Jahre 2020 bis 2024 ins Leben gerufen. In diesem Rahmen wurden 15 Projekte und das übergreifende Begleitprojekt Living Lab »XR-Interakt« gefördert. Dieses Begleitprojekt, bestehend aus dem Team der Hochschule für Wirtschaft und Recht Berlin (HWR Berlin) und dem Virtual Dimension Center (VDC) Fellbach, förderte die Zusammenarbeit und den Austausch zwischen den einzelnen Projekten. Ziel der geförderten Initiativen war es, immersive und effektive Wissensräume durch die Weiterentwicklung der V/AR-Technologie zu schaffen und soziale sowie kulturelle Teilhabe zu ermöglichen. Es wurde besonders darauf geachtet, die Nutzerinnen und Nutzer sowie die Bürgerinnen und Bürger aktiv einzubeziehen und Gender- sowie Diversity-Aspekte (Initiative D21 e.V. 2022[2]) in die Entwicklung und Gestaltung der V/AR-Technologie und ihrer Anwendungsszenarien zu integrieren. Das Living Lab »XR-Interakt« hat ein gemeinsames Kooperations-

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

2 Das Netzwerk für die digitale Gesellschaft e.V. begleitet seit 25 Jahren die digitale Transformation. Im jährlichen Lagebericht D21 Digital-Index wird das Lagebild zur Digitalen Gesellschaft abgebildet (vgl. Initiative D21 e.V., Berlin).

konzept erarbeitet und die Kollaboration und Vernetzung gefördert sowie den 15 Teilprojekten die Möglichkeit der Demonstration und Erprobung ihrer V/AR-Entwicklungen und -Szenarien sowohl vor einem Fachpublikum als auch der breiten Öffentlichkeit ermöglicht. Eine projektübergreifende partizipative Evaluationsstrategie wurde entwickelt und umgesetzt. Diese Strategie umfasste eine übergreifende partizipative Evaluation der eingesetzten interaktiven Systeme, die Eruierung von Good-Practice-Beispielen, eine Expert:innenbefragung und eine übergreifende Projektevaluation insgesamt. In dem vorliegenden Artikel werden im zweiten und im dritten Abschnitt neben einer kurzen Projektbeschreibung und Erläuterung der zentralen Arbeitspakete des Living Lab »XR-Interakt«-Projekts auch die wichtigsten Ergebnisse aus der begleitenden empirischen Exploration zusammenfassend dargelegt. Der Fokus wurde insbesondere auf den partizipativen Leitgedanken der Programmlinie und auf die anvisierten Zielgruppen gelegt. Die Möglichkeiten und der Mehrwert von V/AR-Projekten für die digitale Gesellschaft werden abschließend im Fazit zusammenfassend abgerundet.

1. Einführung in das Begleitprojekt Living Lab »XR-Interakt«

Das Projekt Living Lab »XR-Interakt« ist als Begleitprojekt im Rahmen des BMBF-Förderschwerpunkts »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« mit dem Ziel entstanden, die Vernetzung und interdisziplinäre Zusammenarbeit zwischen allen 15 beteiligten V/AR-Projekten zu initiieren, zu moderieren und zu fördern, eine nachhaltige übergreifende Nutzung der Projektergebnisse zu ermöglichen und die Ergebnisse in einem gemeinsamen Kontext zusammenzuführen.

In der eingangs beschriebenen BMBF-Förderlinie kommt es vor, dass die Einzelprojekte zeitgleich ähnliche Entwicklungen im V/AR-Bereich initiieren. Den Einzelprojekten fehlen in der Regel die Ressourcen und operativen Möglichkeiten, um umfassendere, projektübergreifende Fragestellungen gemeinsam zu bearbeiten. Auch Standardisierungsfragen können im Rahmen von Einzelprojekten nur schwer angegangen werden, da sie allein nicht die nötige Hebelwirkung entfalten können. Für Standardisierungsaktivitäten sind zudem signifikante zusätzliche Ressourcen erforderlich. Somit besteht die Gefahr, dass in den Einzelprojekten proprietäre Insellösungen entwickelt werden, was wiederum die Akzeptanz und Durchsetzungskraft dieser Lösungen schmälern könnte. Die Lösungen der Einzelprojekte würden nur bei den jeweiligen Projektpartner:innen, also punktuell und verteilt, zu besichtigen sein, was den Gesamtüberblick über das Spektrum

der Projektlösungen aus der Förderlinie erschweren würde (vgl. dazu auch den nachfolgenden Beitrag von Christopher Runde im vorliegenden Sammelband).

Damit »das Rad nicht immer wieder neu erfunden wird«, sollten von Beginn an mögliche Synergieeffekte genutzt werden. Das Living Lab »XR-Interakt«-Projekt fungierte daher als sogenannter »passage point« um die programmatische Zielsetzung der VR/AR-Programmlinie für und mit den 15 Projekten zu erreichen. Dabei handelt es sich um ein Verbundprojekt zwischen der Hochschule für Wirtschaft und Recht Berlin (HWR Berlin) und dem Virtual Dimension Center (VDC) Fellbach.

Ein weiteres sehr wichtiges Ziel der Förderlinie war die Adressierung der digitalen Gesellschaft. Damit dieses Ziel tatsächlich erreicht werden konnte, wurde eine partizipative übergreifende Evaluationsstrategie (gemeinsam) entwickelt und umgesetzt, die im folgenden Abschnitt kurz erläutert werden soll.

1.1 Vorhaben, Zielsetzung und Meilensteine des Living Lab »XR-Interakt«

Im Living Lab »XR-Interakt« flossen die Erfahrungen und Ergebnisse der einzelnen Projekte zusammen. Während der gesamten Laufzeit in den Jahren 2021 bis 2024 wurde den 15 Teilprojekten die Möglichkeit gegeben, aktiv zusammenzuarbeiten und ihre Ergebnisse in einer eigens dafür vorgesehenen Transferplattform zusammenzuführen und zu teilen. In den nachfolgenden Teilabschnitten wird auf das Vorgehen und die Umsetzung eingegangen.

Für ein gutes Gelingen einer konstruktiven Zusammenarbeit mit den einzelnen V/AR-Teilprojekten mit- und untereinander und für die Umsetzung des bürgerwissenschaftlichen partizipativen Aspekts in der digitalen Gesellschaft hat das Living Lab »XR-Interakt« die folgenden Meilensteine auf- und umgesetzt:

1. Schaffung eines Kooperationsschemas mit Vertraulichkeitserklärung als Basis für die gemeinsame Zusammenarbeit sowie Analyse und Umsetzung von gemeinsamen Zielstellungen, Vorgehensweisen, Nutzung von Synergie- und Umsetzungspotenzialen in einem gestaltungsorientierten Begleitungs- und Vernetzungsprozess,
2. Entwicklung und Umsetzung einer gemeinsamen Kollaborationsplattform für die Zusammenarbeit und den Informations- und Erfahrungsaustausch,
3. Unterstützung der Projekte bei der Entwicklung interoperabler, standardisierter XR-Lösungen, Aufbau eines globalen XR-Normenregisters, Formulierung und Umsetzung einer Normennutzungs- und -beteiligungsstrategie,

4. Entwicklung und Umsetzung einer gemeinsamen, projektübergreifenden Evaluationsstrategie der V/AR-Szenarien mit dem besonderen Fokus auf die partizipative Beteiligung von diversen Zielgruppen,
5. Umsetzung von bürgerwissenschaftlichen Ansätzen (D21-Digital-Index),
6. Technologietransfers durch ein Demonstrationszentrum,
7. Mitwirkung bei der Organisation der BMBF-Vernetzungstreffen.

Da die Programmlinie während der Corona-Pandemie im Jahr 2021 startete, stand die Vernetzung für die Kooperation zwischen den 15 Teilprojekten vor erheblichen Herausforderungen.

1.2 Kooperation der 15 Teilprojekte der Förderlinie

In den folgenden Abschnitten wird auf das Kooperationsschema, die Kollaborationsplattform und die Arbeitsweise als Grundlage für das Gelingen einer erfolgreichen Zusammenarbeit eingegangen.

1.2.1 Kooperationsschema

Eine nachhaltige, übergreifende Nutzung der Projektergebnisse wurde angestrebt. Dafür wurde ein Kooperationsschema entwickelt. Das Kooperationsschema legt fest, in welchen Bereichen, wann und wie zusammengearbeitet wird mit dem Ziel, für alle Projekte eine verlässliche Arbeitsgrundlage und Arbeitsanweisung in Bezug auf die Zusammenarbeit zu erhalten.

Die Kooperationsbedarfe der Einzelprojekte wurden mittels zweier schriftlicher Umfragen (2022/23) erhoben: Um die Zielstellungen und Vorgehensweise der Projekte zu ermitteln, wurden diese zunächst bei allen Projekten abgefragt. Anhand der Rückmeldungen konnten im Anschluss Projektporträts inklusive Ansprechpartner:innen für alle 15 Projekte erstellt werden, die wiederum an alle Projekte verteilt wurden.

Die nächste Befragung der Einzelprojekte zur Identifizierung möglicher Synergien zeigte verschiedene Kooperationspotenziale. Sämtliche Projekte wurden zu zwölf grundsätzlich möglichen Kooperationspotenzialen (siehe Abbildung 1) in Bezug auf jedes einzelne der anderen 14 XR-Entwicklungsprojekte hin befragt. Diese Antworten wurden numerisch gewertet und aufsummiert. Aus den Projekten mit hoher gegenseitiger Punktevergabe wurden erstens Projektpools gebildet (siehe Farbgruppen in Abbildung 2). Zweitens erhielt jedes Projekt einen individuellen Kooperationsbericht, der ausführte, welche anderen Projekte welches Kooperationsinteresse an ebendiesem Projekt zeigen.

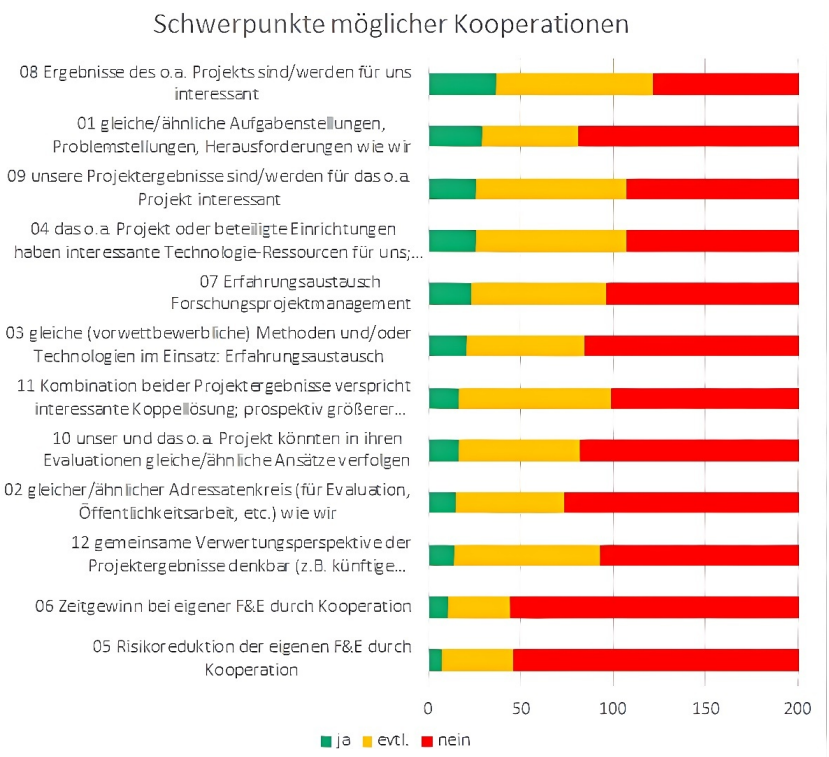
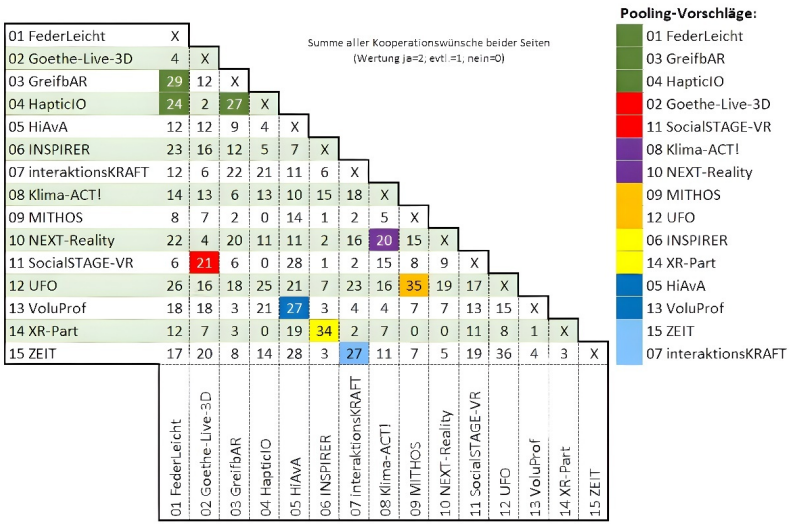


Abbildung 1: Kooperationsinteressen innerhalb der BMBF-Förderlinie
Quelle: Projekteigene Darstellung Living Lab »XR-Interakt«

Ebenfalls war mit den ersten schriftlichen Umfragen festzustellen, dass alle 15 Projekte mit ihren Entwicklungen von V/AR-Anwendungen die digitale Gesellschaft adressieren. Die V/AR-Projekte ließen sich grob in die Bereiche Gesundheit und Pflege (4 Projekte), Bildung (5 Projekte) und Öffentlicher Raum (6 Projekte) einordnen. Zwei Projekte konnten zwei Bereichen gleichzeitig zugeordnet werden. Die Entscheidungsgrundlage wurde in beiden Fällen entlang der adressierten Zielgruppe festgelegt.

Als Basis für die Zusammenarbeit wurden Vertraulichkeitserklärungen mit allen beteiligten Projekten geschlossen. Die einzelnen Projekte haben in mehreren Aktivitäten über die gesamte Projektlaufzeit miteinander kooperiert. Unterstützend dafür wurde eine Kollaborationsplattform vom Living Lab »XR-Interakt« entwickelt und umgesetzt.



	01 FederLeicht	02 Goethe-Live-3D	03 GreifbAR	04 HapticIO	05 HiAvA	06 INSPIRER	07 interaktionsKRAFT	08 Klima-ACT!	09 MITHOS	10 NEXT-Reality	11 SocialSTAGE-VR	12 UFO	13 VoluProf	14 XR-Part	15 ZEIT
01 FederLeicht	X														
02 Goethe-Live-3D	4	X													
03 GreifbAR	29	12	X												
04 HapticIO	24	2	27	X											
05 HiAvA	12	12	9	4	X										
06 INSPIRER	23	16	12	5	7	X									
07 interaktionsKRAFT	12	6	22	21	11	6	X								
08 Klima-ACT!	14	13	6	13	10	15	18	X							
09 MITHOS	8	7	2	0	14	1	2	5	X						
10 NEXT-Reality	22	4	20	11	11	2	16	20	15	X					
11 SocialSTAGE-VR	6	21	6	0	28	1	2	15	8	9	X				
12 UFO	26	16	18	25	21	7	23	16	35	19	17	X			
13 VoluProf	18	18	3	21	27	3	4	4	7	7	13	15	X		
14 XR-Part	12	7	3	0	19	34	2	7	0	0	11	8	1	X	
15 ZEIT	17	20	8	14	28	3	27	11	7	5	19	36	4	3	X

Abbildung 2: Projekt-Pooling
Quelle: Projekteigene Darstellung Living Lab »XR-Interakt«

1.2.2 Die Kollaborationsplattform

Die Implementierung einer Kollaborationsplattform hatte zum Ziel, eine effiziente und effektive Zusammenarbeit der Einzelprojekte zu gewährleisten und auch neue Erfahrungen zu sammeln. Die Hauptziele waren, die Kontaktdaten aller beteiligten Personen für alle verfügbar und aktuell zu halten, Inhalte revisionsfähig abzulegen und notwendige Funktionen wie Dateiablagesysteme, Diskussionsforen, Nachrichtensysteme, Terminkalender und Video-Telekonferenzen bereitzustellen. Neben dem Einsatz von Groupware-Systemen wurde die Nutzung verteilter, kollaborativer V/AR-Umgebungen für die Arbeit des Living Labs »XR-Interakt« ausdrücklich forciert, um hieraus wertvolle sowie für den Transfer mögliche Alltagserfahrungen erlangen zu können.

Für die projektübergreifende Kollaborationsplattform wurde ein Lastenheft mithilfe einer Umfrage zu den Wünschen und Rahmenbedingungen (z. B. rechtliche) über alle Einzelprojekte hinweg als Lösungsansatz erstellt, das Kriterien für die Auswahl der Kollaborationsplattform festlegte. Die Recherche resultierte in 27 Kollaborationslösungen. Anhand des Lastenhefts wurden alle recherchierten Kollaborationslösungen bezogen auf die Erfüllung der Kriterien bewertet. Um die Ergebnisse weiter einzugrenzen bzw. eine Entscheidung treffen zu können, wurden die 15 Einzelprojekte und das Team des Living Lab »XR-Interakt« durch eine Umfrage in den Auswahlprozess eingebunden. Nach der Auswahl hat das

VDC die Kollaborationsplattform »Microsoft SharePoint« als eine übergreifende »Kommunikationswebsite« eingerichtet, über die alle Inhalte visualisiert und abgerufen werden können. Zusätzlich zu der übergeordneten Kommunikationswebsite, auf die alle 15 Einzelprojekte Zugriff haben, hat das VDC auch eine eigene Kollaborationsplattform als untergeordnete Seite für jedes der Einzelprojekte erstellt. Eine Anleitung und Schulung für die Nutzung der Plattform wurde bereitgestellt.

1.2.3 (Vernetzungs-)Workshops und Demonstrationszentrum

Mit dem Kooperationsschema und der Kollaborationsplattform wurden die Grundlagen für eine gemeinsame Zusammenarbeit der im Rahmen der Corona-Pandemie möglichen Aktivitäten gelegt. Mit dem Vernetzungstreffen des BMBF im Oktober 2022 und dem damit verbundenen direkten persönlichen Kennenlernen der 15 beteiligten Projekte untereinander beim Living Lab an der HWR Berlin wurde der Grundstein für die direkte Kooperation gesetzt. Während der gesamten Projektlaufzeit trafen sich die Expertinnen und Experten der Projekte zur gemeinsamen Information, zum Austausch und zur Bearbeitung von projektübergreifenden Fragestellungen.

Das Erleben von V/AR ist häufig noch eine praktische Angelegenheit, bei der die Benutzer:innen V/AR-Systeme (sowohl Software als auch Hardware) anfassen, ausprobieren, anziehen und aufsetzen möchten. Gerade weil es sich um Interaktionstechnik handelt, ist hier häufig die persönliche Präsenz gefragt. Ausgehend von diesen Überlegungen bedarf es der praktischen Möglichkeit, die V/AR-Lösungen der Einzelprojekte der Förderlinie selbst zu erleben und auszuprobieren. Um dies zu ermöglichen, wurden Demonstrationszentren beim VDC Fellbach und der HWR Berlin erstellt und betrieben, sodass die Demonstratoren aller 15 Projekte präsentiert und über die Projektbeteiligten hinaus auch von breiteren Bevölkerungsschichten ausprobiert werden konnten. Dieses Vorgehen unterstützte den Wissenstransfer innerhalb der Förderlinie und das Lessons Learned von den Erfahrungen und Ergebnissen der Projekte untereinander. Neben den jeweils anvisierten *projektspezifischen* Zielgruppen wurden die Einzelprojekte auch mit ausgewählten Bevölkerungsgruppen unter Berücksichtigung des D21-Digital-Index erprobt und formativ (begleitend) evaluiert.

Das Living Lab »XR-Interakt« führte vier aufeinander aufbauende Workshops durch, die stichpunktartig dargestellt werden.

1. Workshop »Vorstellung der Projekte und Evaluationskonzepte« (Oktober 2022 an der HWR Berlin):
   – Vorstellung der Projekte und Personen

- Vorstellung der Evaluationsstrategie der 15 Einzelprojekte
- Gemeinsame Schwerpunktsetzung der übergreifenden Evaluation

Es wurden gemeinsame Evaluationsszenarien entwickelt, die durch alle Einzelprojekte zu berücksichtigen waren. Informationen zur Nutzung (Arbeitsweise, Rechte) einer Kollaborationsplattform wurden evaluiert und dokumentiert.

2. Workshop »Austausch und Evaluationsstrategie« (März 2023 an der HWR Berlin)
   - Projektpräsentation und Vorstellung sowie Erprobung der ersten fünf Demonstratoren inklusive Evaluation und direktem Feedback an das jeweilige Projekt
   - Erarbeitung einer gemeinsamen Evaluationsstrategie der 15 Einzelprojekte der Fördermaßnahme in den Fokusfeldern Technik, ELSI und Akzeptanz mit Blick auf die spezifische Zielgruppe
   - Vorstellung der Living Lab »XR-Interakt«-Kollaborationsplattform
3. Workshop »Austausch, Evaluation, Normung/Standardisierung« (Oktober 2023 im VDC Fellbach)
   - Projektpräsentation (Recap) und Vorstellung der Prototypen-Projekte sowie Erprobung weiterer Demonstratoren inklusive Evaluation und direktes Feedback an das Projekt
   - Projektpräsentation mit Fokus Evaluation Zielgruppe (Wer wurde wann wie evaluiert?)
   - Vortrag Normenarbeit und Workshop Schwerpunkte-Ausarbeitung zu den Normierungsthemen mit dem Ergebnis Kodierung/Streaming, UX/Ergonomie Mensch-System-Interaktion, ELSI/Diversity
4. Workshop im Rahmen der XR EXPO »Präsentation der Projektentwicklungen, Testung durch Bürgerinnen und Bürger und Verbreitung« (April 2024, XR EXPO 2024)
   - Testung und Verbreitung, Qualifizierung, Zielgruppe
   - Präsentation und Testung der 15 Prototypen auf der XR EXPO 2024 sowohl beim Fachpublikum als auch von diversen Bürgerinnen und Bürger sowie Evaluation durch die Zielgruppe
   - Evaluationsworkshop mit Zielgruppen, Elevator Pitches zu den gewünschten Zielgruppen-Feedbacks, Workshop-Gespräche mit den Zielgruppen, Ergebnispräsentation und Diskussion im Plenum
   - Qualifizierungsworkshop »Normung« als Serious Game, organisiert durch das Living Lab »XR-Interakt«, durchgeführt durch das Deutsche Institut für Normung (DIN) e. V.
   - Veröffentlichung der Ergebnisse auf der Living Lab »XR-Interakt«-Plattform

- Testung der weiteren Prototypen durch die Expertinnen und Experten der Förderlinie sowie begleitende Evaluation und Bericht an die Projekte
- Evaluationsbericht #3 – 15 Interviews mit 29 Expertinnen und Experten aus den Projekten der Förderlinie (Juni 2023–Dezember 2023),
- Zielsetzung: Umsetzung der Projektziele der BMBF-Förderlinie im Hinblick auf die digitale Gesellschaft: zwischenmenschliche Kommunikation & Kompetenzerwerb, Partizipation & Inklusion, Schaffung von immersiven und effektiven Wissensräumen, soziale und kulturelle Teilhabe in der digitalen Gesellschaft

Durch die halbjährlichen Workshops und Treffen konnte die Vernetzung und der Austausch unter den Projekten der Förderlinie und die Zusammenarbeit kontinuierlich ausgebaut werden.

2. Die projektübergreifende partizipative Evaluation

Um eine gemeinsame – übergreifende – Evaluationsstrategie aller geförderten Einzelprojekte erfolgreich umsetzen zu können, mussten die vielfältigen Technologien »an sich« evaluiert werden, um die jeweilige Trag- und Reichweite jeder (neu) entwickelten Technologie zu testen. Andererseits mussten die eingesetzten Technologien »in Aktion« eruiert und weiterentwickelt werden. Darüber hinaus mussten die Evaluationsergebnisse, die aus den Einzelprojekten abgeleitet wurden, aufeinander bezogen ausgewertet werden, um eine übergreifende Evaluationsstrategie zu entwickeln und umsetzen zu können.

Das Living Lab »XR-Interakt« hat vor diesem Hintergrund einen *partizipativen Evaluationsprozess* entwickelt, der qualitativ und explorativ ausgerichtet war. Im Folgenden wird zunächst unter Punkt 2.1. auf die methodische Vorgehensweise eingegangen. Die Ergebnisse werden anschließend im Abschnitt 2.2 zusammenfassend dargestellt.

2.1 Der partizipative Evaluationsprozess vom Living Lab »XR-Interakt«

Das partizipative Evaluationsmodell vom Living Lab »XR-Interakt« umfasste vier Analysebereiche, die gezielt auf die definierten Programmziele ausgerichtet wurden: 1. Partizipative Evaluation der interaktiven Systeme, 2. Partizipative Evaluation von Best-Practice-Beispielen aus den Projekten (einschließlich Materialbefragungen), 3. Partizipative Evaluation von Best-Practice-Praxisbeispielen in der

digitalen Gesellschaft (einschließlich Materialbefragungen) und 4. Partizipative übergreifende Projektevaluation:

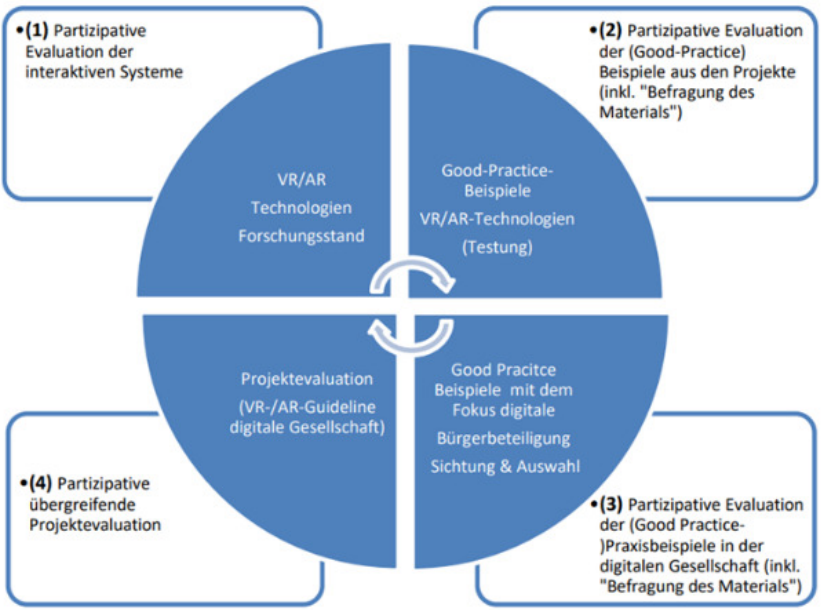


Abbildung 3: Die vier Analysebereiche des Evaluationsmodells vom Living Lab »XR-Interakt«
Quelle: Eigene Darstellung

Mit Blick auf das Evaluationsmodell (Abbildung 3) lag in den ersten Monaten nach dem Projektstart im Februar 2022 der erste Evaluationsfokus auf der (1) *partizipativen Evaluation der interaktiven Systeme* (V/AR-Technologien). Neben dem Einsatz eines halbstandarisierten Fragebogens wurde im ersten übergreifenden Workshop (Oktober 2022) auch auf die eingesetzten Systeme eingegangen, um die Möglichkeiten zur Zusammenarbeit und die Integration aller Projektbeteiligten in den partizipativen Prozess voranzutreiben. Ein zusätzlicher Schwerpunkt lag auf der Entwicklung und Abstimmung einer gemeinsamen Evaluationsstrategie.

Der zweite Evaluationsfokus (Abbildung 3) lag auf der (2) *partizipativen Evaluation der (Good-Practice-)Beispiele*. Die Projekte erhielten die Möglichkeit, ihre entwickelten Prototypen »in Aktion« von den Expertinnen und Experten der V/AR-Programmlinie direkt (gegenseitig) testen und bewerten zu lassen. Auch dafür wurde ein Fragebogen entwickelt, der neben den technischen Aspekten auch auf ethische, rechtliche und soziale Aspekte abhob. Zusätzlich enthielt er Fragen zur

Akzeptanz und zu Aspekten der digitalen Repräsentanz. Die Expertinnen und Experten konnten sich gegenseitig ein direktes Feedback zum Gesamteindruck der V/AR-Anwendung geben sowie dezidierte Verbesserungsvorschläge formulieren. Diese wurden anonymisiert ausgewertet und als direktes Feedback aus Expert:innensicht dem jeweiligen Einzelprojekt direkt über die Transferplattform zur Verfügung gestellt. Ebenfalls erhielten die Projekte durch Befragung einen Eindruck darüber, ob sich der selbst eingeschätzte Entwicklungsstand des Prototyps mit der Einschätzung der Expertinnen und Experten deckt oder wo es Abweichungen gab. Dies konnte ebenfalls direkt in die Weiterentwicklung der V/AR-Entwicklungen und Szenarien einfließen. Der zweite Evaluationsfokus wurde im Rahmen dreier Workshopsreihen umgesetzt (Abbildung 4):



Abbildung 4: Die partizipative Evaluation der (Good-Practice-)Beispiele mit dem Fokus V/AR-Prototypentestung aller Projekte im Zeitraum März 2023 bis April 2024
Quelle: Eigene Darstellung

Im dritten Schwerpunkt des Evaluationsmodells (Abbildung 3) handelt es sich um die (3) *partizipative Evaluation der (Good-Practice-)Praxisbeispiele in der digitalen Gesellschaft.* In der Förderlinie des BMBF ging es um transformative Technologien, die eine neue Qualität der zwischenmenschlichen Kommunikation und des Kompetenzerwerbs ermöglichen sollen. An dieser Stelle wurde daher methodisch auf eine Expert:innenbefragung gesetzt, um alle 15 Projekte auf ihre praktischen Einsatzszenarien hin prüfen zu können. Neben der anvisierten Zielgruppe wurden auch Fragen zur Reichweite, zum Mehrwert und der Nachhaltigkeit gestellt. Auch der Einbindungsgrad der Zielgruppe mit Blick auf den D21-Digital-Index wurde dabei im Besonderen evaluiert. Methodisch wurde als qualitatives Erhebungsinstrument das leitfadengestützte Interview gewählt. Die Interviews mit den Expertinnen und Experten aus den Projekten der Programmlinie wurden in Form »halbstandardisierter Interviews« geführt und vollständig anonymisiert und aus-

gewertet. Die relativ offenen Fragen boten den Befragten die Möglichkeit, auch ihr informelles Wissen zu entfalten.

Die 15 Teilprojekte der Programmlinie wurden im Zeitraum vom Juni 2023 bis Dezember 2023 befragt.[3] Insgesamt nahmen an den 15 Interviews insgesamt 29 Beteiligte (24 männlich/5 weiblich) aus unterschiedlichen Statusgruppen (9 Projektleitung/20 Wissenschaftliche Mitarbeiter:innen) aus den jeweiligen Teilprojekten der Programmlinie teil. Der fachliche Hintergrund der Befragten war entsprechend der Projektbezüge sehr breit gefächert. Als Expertinnen und Experten auf ihrem Gebiet wurden Teilnehmende aus Technik, Informatik und natur- und ingenieurwissenschaftlichen Bereichen und aus Fachgebieten der Arbeits-, Sozial- und Bildungswissenschaften, der Psychologie und weiterer Fachgebiete befragt. Für die qualitative Erhebung und Auswertung wurden alle geführten Interviews herangezogen. Um die Interviews »als Ganzes« mit ihrer Vieldeutigkeit und ihren eventuellen Widersprüchen zu berücksichtigen, wurde die Methode der Qualitativen Inhaltsanalyse nach Kuckartz und Rädicker (2022: 132 ff.) gewählt. Die Datenaufbereitung und -auswertung wurde mithilfe der QDA-Software MAXQDA (Kuckartz) vorgenommen. Die Arbeitsschritte erfolgten nach den Gütekriterien der qualitativen Forschung unter Berücksichtigung der Transparenz und Nachvollziehbarkeit in Anlehnung an Kuckartz und Rädicker (2022: 132 ff.).

Mit Blick auf das Evaluationsmodell (Abbildung 3) wurde der vierte und letzte Schwerpunkt auf die (4) *partizipative projektübergreifende Evaluation* gelegt. Auf die Partizipation und Inklusion sowie die Schaffung von immersiven und effektiven Wissensräumen sowie die soziale und kulturelle Teilhabe in der digitalen Gesellschaft wurde in diesem Kontext ein besonderes Augenmerk gerichtet. Als methodische Vorgehensweise wurde hier ein bürgerwissenschaftlicher Ansatz verfolgt und umgesetzt:

Im Rahmen der XR EXPO 2024 am 03.–04. April 2024 in Stuttgart wurde den 15 Projekten in der BMBF-Programmlinie »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« die Möglichkeit zur Präsentation und Demonstration ihrer entwickelten V/AR-Anwendungen vor Bürgerinnen und Bürger und dem Fachpublikum ermöglicht. Die Testerinnen und Tester waren mit dem Ziel der Einbindung der adressierten digitalen Gesellschaft durch das Begleitprojekt »XR-Interakt« organisiert und geführt. Zusätzlich konnte auch das »Laufpublikum« (gemeint ist das Messepu-

3 Die Interviewdauer umfasste eine Länge von 30 min bis zu 1,5 h – abhängig von der Anzahl der teilnehmenden Interviewten. Die Befragungen wurden sowohl in Einzel- als auch in Gruppengesprächen mit maximal vier Projektbefragten durchgeführt.

blikum) die Projekte testen und ein Feedback geben. Der Fokus wurde – wie es die Programmlinie es auch vorsah – auf die breite digitale Gesellschaft gelegt.

Die vom Living Lab »XR-Interakt« im Vorfeld organisierten 18 Testpersonen wurden mit Blick auf den D21-Digital-Index »rekrutiert«. Neben der Kategorie Gender (7 weiblich/11 männlich) wurde auch auf die Kategorie Alter und die sozialen und kulturellen Hintergründe geachtet, was sich als große Herausforderung gerade mit Blick auf V/AR-Szenarien herausstellte. Auffällig häufig wurden die Demonstratoren der 15 Projekte vorwiegend nur von männlichen Studierenden getestet. Es fehlten häufig jüngere, aber auch ältere (insbesondere weibliche) Testpersonen. Somit wurden insgesamt fünf Schüler:innen, zwei Student:innen, zwei Fachpersonen aus dem medizinischen Bereich sowie neun weitere (ältere) Bürger:innen als Testgruppe gewonnen. Die Proband:innen beurteilten die Anwendungen in den Projektclustern »Bildung«, »Gesundheit« und »Öffentlicher Raum« mit einem Feedbackbogen. Zusätzlich enthielt der Feedbackbogen Fragen, die von den 15 Projekten formuliert wurden und spezifische eigene Projektinteressen abbildeten. Die V/AR-Anwendung jedes der Projekte wurde von mehreren Personen getestet, ausprobiert und evaluiert.

Neben den Auswertungsergebnissen aus den Feedbackbögen kam es zu einem direkten Austausch zwischen den 18 Tester:innen in den jeweiligen Bereichen (Bildung, Gesundheit und Öffentlicher Raum), die im Rahmen eines dreigeteilten Workshops am 05.04.2024 durch das »XR-Interakt«-Team moderiert und durchgeführt wurden. Die eruierten Ergebnisse wurden zusammengefasst über die Transferplattform allen 15 Teilprojekten der Förderlinie zur Verfügung gestellt.

2.2 Ausgewählte Evaluationsergebnisse des Living Lab »XR-Interakt«

Im Folgenden werden die wichtigsten Erkenntnisse und Ergebnisse aus der partizipativen Evaluation des Begleitprojekts Living Lab »XR-Interakt« zusammenfassend dargestellt. Das partizipative Evaluationsmodell vom Living Lab »XR-Interakt« umfasste vier Analysebereiche (vgl. Abbildung 3). Zu jeder Analysedimension wurde ein umfangreicher Evaluationsbericht verfasst und allen 15 Projekten zugänglich gemacht. Der vorliegende Abschnitt beinhaltet nur die wichtigsten Ergebnisse aus allen vier Evaluationsberichten über den Gesamtförderbereich des Begleitprojekts entlang den besonderen Herausforderungen der Programmlinie insgesamt. Alle Aussagen wurden vollständig anonymisiert und die Einzelzitate wurden in der Regel zusammengefasst.

2.2.1 Projektmanagement und Projektarbeit in V/AR-Kontexten

Die Forschungsprojekte der BMBF-Förderlinie hatten alle eine begrenzte Laufzeit (in der Regel drei Jahre). Somit lastete der Zeitdruck auf allen 15 Projekten der BMBF-Förderlinie.

Das *Zeitmanagement* war somit besonders herausfordernd. Unvorhergesehene Ereignisse wie Mitarbeiter:innenwechsel oder Lieferschwierigkeiten – gerade während der Corona Pandemie – wirkten sich strukturverwerfend auf das Projektgeschehen und die gesamte Projektplanung und -entwicklung aus. Aber auch andere Aspekte wie die Teilnahme an Workshops und Vernetzungstreffen als auch die BMBF-Vorgabe einer Umsetzung einer spezifischen Evaluationsstrategie waren zu berücksichtigen.

Im Rahmen der Projektarbeit standen die Projektteams einerseits vor der Aufgabe, die Projekt(teil)ziele im standortübergreifenden Verbundprojekt gemeinsam umzusetzen, andererseits standen sie vor der Herausforderung, eine neue innovative V/A/XR-Technologie für einen sehr spezifischen gesellschaftlichen Anwendungskontext mit einer spezifisch anvisierten Zielgruppe äußerst passgenau zu verwirklichen:

Bei den 15 geförderten Projekten handelte es sich in der Mehrheit um Verbundprojekte, deren *Beteiligte aus den unterschiedlichsten Fachgebieten* (Technik, Natur- und Ingenieurwissenschaften, Medizin, Arbeits-, Sozial-, Geistes- und Erziehungswissenschaften und vieles mehr) deutschlandweit an dem jeweiligen Projekt mitgewirkt haben, d. h., es musste mit einem teilweise *sehr verteilten, multidisziplinären Konsortium, die Koordination des Projektteams und der Arbeitsprozesse* umgesetzt werden. Die Koordination der Arbeitspakete, insbesondere an den Schnittstellen zwischen verschiedenen Projektpartner:innen und Projektbereichen war eine immer wiederkehrende Herausforderung, die sich keinesfalls nur auf die projektinterne komplexe technische Abstimmung begrenzte. Dabei konnte beobachtet werden, dass die Komplexität der Koordination durch die Größe des Projekts und die Vielzahl der verschiedenen (sozio-technischen) Beteiligten und der daran gekoppelten Arbeitspakete stetig wuchs.

Für die Mehrheit der Projekte war es daher wichtig, bereits in der Planungsphase realistische Zeitpläne (neu) zu erstellen und Methoden für effizientes Zeitmanagement zu implementieren (wie beispielsweise eine parallele Abwicklung von Prozessen oder eine zyklische Projektplanung). Die Besonderheit in den Projekten lag aufgrund der *Neuartigkeit der V/AR-Technologie* neben der einzelnen Zielsetzung auch in einer *starken Verortung der Projekte in der Grundlagenforschung*. Dies spielte eine evidente Rolle bei der Konzeptionierung, Projektentwicklung, -umsetzung, aber auch in den angestrebten Zielsetzungen (vgl. dazu auch Punkt 2.2.2 Zielgruppenorientierte Technologie).

Die Entwicklung einer *»gemeinsamen Sprache«* zwischen den verschiedenen Projektbeteiligten war eine sehr zentrale Herausforderung, die in allen untersuchten Projekten festgestellt werden konnte. Das betraf insbesondere die Schwierigkeit des gemeinsamen Verständnisses zwischen Beteiligten aus verschiedenen Berufsfeldern innerhalb der Verbundprojekte. Das wurde vor allem in den Experten:inneninterviews von vielen Projekten herausgestellt. Dies erforderte eine längere Zeit. Ein effektiver Verständnisaustausch, die Entwicklung einer gemeinsamen Sprache und Sichtweise auf die Projektentwicklung zwischen allen Projektbeteiligten ist grundlegend, um Informationen und Ergebnisse effektiv nutzen zu können. Der *regelmäßige Austausch* sowie klare Kommunikationskanäle und Tools unterstützten dabei die Erhaltung eines effizienten Verständnisaustauschs.

Eine weitere Problematik war jene, die aus der dynamischen Natur der Projektentwicklung entsteht. Von den Befragten der Projekte wurde dabei betont, dass sich die technischen und konzeptionellen Anforderungen im Laufe der Projektlaufzeit ändern konnten, wodurch eine ständige Anpassung und Flexibilität erforderlich war. Fast aus allen Projekten wurde bestätigt, dass die *rasante Entwicklung der V/AR-Technologie* wie beispielsweise die Entwicklung der VR-Brille innerhalb der letzten fünf Jahre durch immer wieder neu zu bewertende Anpassungen mit im Projekt berücksichtigt werden mussten. Schnittstellen und Systeme im Laufe der Projektentwicklung mussten immer wieder neu angepasst werden. Komplikationen zwischen technischen Aspekten und (neuen) Anforderungen an die Projekt(ziel)entwicklung erforderten, dass häufig auch *»Abstriche«* in der Funktionalität oder dem eigentlichen Anwendungsszenario gemacht werden mussten (vgl. dazu auch Einzelbeiträge der Projekte im vorliegenden Sammelband).

Diese dynamische und häufig auch unvorhergesehene Entwicklung erforderte für alle Forschungsprojekte eine ständige Überprüfung und Anpassung der Projektziele, Methoden und Anforderungen. Je höher und transparenter der Reifegrad der Kommunikation innerhalb der Projektteams, umso erfolgreicher konnten die Projekte auf diese querlaufenden und ineinander verstrickten soziotechnischen Artefakte und Akteur:innen (un)intendiert reagieren.

2.2.2 Zielgruppenorientierte V/AR-Technologie

Die Weiterentwicklung der V/AR-Technologie war für alle Projekte herausfordernd. Das wurde von allen Projekten immer wieder herausgestellt. Die Auswahl der Soft- und Hardware wurde mit äußerster Sorgfalt getätigt. Ein wesentliches Entscheidungskriterium für den Einsatz der Soft- und Hardware war grundlegend die *Usability*.

Die häufig eingesetzte Game Engine »Unity3D« bietet viele Schnittstellen für eine leichte Integration und Entwicklung von V/AR-Programmen, welche von den Teilprojekten als logische Schlussfolgerung genutzt wurde. Die Spannbreite der Nutzung verschiedener Software und Programmiersprachen ist für die V/AR-Entwicklungen sehr groß (z. B. wie C+, Python, Eigenprogrammierungen, Open CV, Frameworks, mobile Applikationen, Künstliche Intelligenz, Avatare etc.). Darüber hinaus können an dieser Stelle keine vereinheitlichten Aussagen getroffen werden, da die VR- und AR-Hardware-Komponenten in unterschiedlichen Maßen entsprechend der ausgewählten V/AR-Szenarien eingesetzt wurden. Faktoren wie beschränkte Verfügbarkeit, technische Anforderungen und Ressourcen wie Datenverbindung, Datenverarbeitung und Datenschutz beeinflussten die Auswahl maßgeblich. Virtual, Augmented und Mixed Reality sind Technologien, die als immersive Erfahrungslösung in den unterschiedlichen Bereichen durch die Teilprojekte verwendet wurden. Die Mensch-Technik-Interaktion der Projekte erforderte die Berücksichtigung der Usability und Benutzerfreundlichkeit, aber auch die Sicherheit der Anwender:innen und Nutzenden. Verschiedene Technikanwendungen wurden im Projektverlauf geprüft, genutzt, verworfen und auch mal gänzlich neu entschieden, z. B. bei der Auswahl der XR-Brillen.

Zusammenfassend lässt sich feststellen, dass die Auswahl der Soft- und Hardware gezielt getroffen wurde, um eine möglichst reibungslose, zielorientierte und effiziente Umsetzung der Szenarien für die jeweils spezifische Zielgruppe entwickeln und umsetzen zu können. Je nach Anforderungen und Schwerpunktsetzung des Projekts wurden die passenden Technologien ausgewählt. Gerade die technischen Entscheidungen beeinflussten jedoch viele Faktoren wie beispielsweise den Grad der Partizipation und die verschiedenen Anwendungsmöglichkeiten. Immersion, Präsenzerleben und Interaktivität sollten gezielt für und mit den Anwender:innen umgesetzt werden (Radianti et al., 2020).

In allen Lösungen ging es darum, alltagstaugliche und erfolgreiche Angebote zu entwickeln. Nur langsam etablieren sich Quasi-Standards der 3D-Interaktion, die es einfach machen würden, zwischen 3D-Anwendungen zu wechseln. Gute Usability, ein niedrigschwelliger Zugang und positive User Experience (UX) werden aber neben der Erfüllung sachlich-technischer Anforderungen durch die in den verschiedenen Projekten erarbeiteten Lösungen entscheidend sein, wenn es um die Akzeptanz geht.

2.2.3 Zielgruppe und Partizipation aus Sicht der Projektgruppe

In der BMBF-Förderlinie wurde explizit die digitale Gesellschaft als Zielgruppe adressiert. Die Bereitschaft diverser Bürger:innengruppen, sich mit V/AR-Technologien auseinanderzusetzen, wird nur mit einer signifikanten Einbeziehung

einer diversen Zielgruppe gelingen. Citizen Science ist – durch den hohen Grad an Partizipation im Sinne von Teilhabe – ein bewährtes Konzept, um (diverse) Bürgerinnen und Bürger an gesellschaftlichen Transformationsprozessen aktiv zu beteiligen (Bonn et al., 2016). Der Beteiligungsgrad der Bürgerinnen und Bürger wurde bei den verschiedenen Vorhaben durchaus sehr unterschiedlich gehandhabt.

Citizen Science heißt: Bürger:innen wirken an Forschungsprojekten mit. Bei der Förderlinie »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« stehen verschiedene Herausforderungen im Mittelpunkt: Es geht um neue Formen des informellen Wissens und Lernens im digitalen Zeitalter und darum, Interesse und Kompetenzen im Umgang mit komplexen transformativen Technologien zu entwickeln. Es stellen sich Fragen danach, wie Sachkenntnisse aus der Wissenschaft in die Gesellschaft gelangen, wie die Gesellschaft an Erkenntnisprozessen teilnehmen kann und wie sie die Methoden der Wissenschaft besser verstehen und nachvollziehen kann. Aber es geht auch umgekehrt darum, innovative Technologien in gesellschaftliche Kontexte so zu transferieren, dass auch ein Nutzen für die Bürger:innen erkennbar wird (Wiesner & Stepczynski, 2020).

Durch die Einbindung von (diversen) Bürger:innengruppen werden die Anwendungsfälle der Projekte näher an die Bedürfnisse der breiten Öffentlichkeit herangeführt. Die Zielgruppenvielfältigkeit der *Bürgerschaft* als Handelnde sowohl als Privatperson als auch im Beruf und in der Gesellschaft musste daher von den Projekten gezielt inkludiert werden (vgl. dazu u. a. auch die Projektbeiträge im Sammelband von »Klima-ACT!«, »Goethe-Live-3D« und »INSPIRER«). Da die Mehrheit der Teilprojekte an Hochschulen und Universitäten angesiedelt war, wurde jedoch überdurchschnittlich oft auf (männliche) Studierende als Proband:innen zurückgegriffen, auch wenn der Projektfokus nicht explizit im Bildungsbereich angesiedelt war.

Durch das Living Lab »XR-Interakt« wurde die breite Adressierung der anvisierten Zielgruppen der 15 Projekte (in den Bereichen »Bildung«, »Öffentlicher Raum« und »Gesundheit«) nicht nur unter dem D21-Digital-Index abgefragt, sondern auch begleitend analysiert und beratend begleitet. Aus den Studien rund um das Thema »I-Methodology« ist bekannt, dass insbesondere sehr homogene Entwicklungsteams dazu neigen, ihre eigenen Attribute zur Ausgangstellung ihrer Entwicklungsziele zu setzen. Gemeint ist, dass in erster Linie Produkte entwickelt werden, die aus der verengten Perspektive des Entwicklungsteams stammen. Die seither weiterentwickelten Ansätze partizipativer Softwaregestaltung bieten Möglichkeiten an, dieser »I-Methodology«-Problematik (z. B. Rommes, 2014; Vines et al., 2015; Floyd et al., 1989 sowie für den deutschsprachigen Kontext Bath, 2002) entgegenzuwirken. Somit sollte ein userzentrierter

partizipativer Softwaregestaltungsansatz bevorzugt eingesetzt werden, um nicht nur geschlechtskonnotierte Stereotypisierungen zu vermeiden, sondern auch Stereotypisierungen unter den Aspekten Alter und kulturellem Kontext von Beginn an vorzubeugen. Entwicklungsprojekte im Bereich transformativer Technologien sollten somit intersektionale Aspekte berücksichtigen und einen partizipative Softwaregestaltungsansatz durch eine D21-Digital-indexierte Perspektive erweitern und nutzbar machen (vgl. dazu auch Wiesner et al., 2019). Geht man davon aus, dass insbesondere die Entwicklungsteams ihren Fokus zu Beginn nur auf die V/AR-Technologien »an sich« setzen, ist es der Arbeit des Living Lab »XR-Interakt« geschuldet, dass die Projektmitarbeitenden im Projektverlauf ihre Priorität verstärkt auf die Erhöhung des Integrationsgrads ihrer jeweils spezifischen Zielgruppe gelegt haben. Jedes Projekt hat sich im Rahmen der Programmlinie verstärkt darüber Gedanken gemacht, wie, wo und in welcher Entwicklungsphase die jeweilige Zielgruppe wie oft gezielt eingebunden werden sollte.

War bei der Initiierung und Ideenentwicklung bereits die Zielgruppe klar im Visier, wurden die Nutzenden bei diesen Projekten von Anfang an involviert, indem sie z.B. bei der Szenarien- und/oder Softwareentwicklung selbst aktiv mitgewirkt haben. In den Teilprojekten ging es um Inklusion und Teilhabe für die Einbindung der digitalen Gesellschaft, aber auch um Sorgfalt mit Blick auf die gesellschaftliche Verantwortung im Handeln während der Projektarbeit und darüber hinaus. Ebenso ging es um Kommunikation und dem Umgang mit Konflikten, der durch einen partizipativen Ansatz nicht selten als produktiver Lernprozess wahrgenommen wurde.

Mit der Förderlinie wurden ebenfalls die Inklusion und Unterstützung spezifischer Zielgruppen und Zielgruppen mit besonderen Bedürfnissen mit der Entwicklung von V/AR-Anwendungen gezielt adressiert. In einigen Projekten wurden Zielgruppen direkt in die Projektentwicklung von den jeweiligen Projekten absichtsvoll und kontinuierlich integriert, vgl. dazu u. a. die Beiträge im Sammelband von »MITHOS« (Training von Konfliktsituationen), »NEXT Reality« (Simulation und Training in der Fachausbildung) und »UFO« (Training für neurodiverse Personen). Entsprechend der Projektideen und -inhalte der Teilprojekte wurden die Bürgerinnen und Bürger in ihrer Rolle und Funktion als Schüler:innen, Auszubildende, Studierende, Fachkräfte, Eltern- und Familienmitglieder oder in ihrer beruflichen Funktion als Anwender:innen und Fachpersonal wie z. B. Lehrende, Mitarbeitende in Kommunen/der Stadtplanung und im Gesundheitswesen adressiert.

Zusammenfassend kann festgestellt werden, dass die 15 Teilprojekte der Förderlinie des BMBF die digitale Gesellschaft als Ganzes zumindest in der Konzeptionsphase adressiert haben. Auch Bürgerinnen und Bürger wurden sowohl als

Privatpersonen mit ihren Bedürfnissen und Besonderheiten angesprochen und auch in ihren jeweiligen Bildungs- und Entwicklungsphasen, Berufen und Funktionen innerhalb der demokratischen Gesellschaft eingebunden. Alle Zielgruppen wurden in der *Projektkonzeption* berücksichtigt. Es konnte weiter festgestellt werden, dass ein Teil der Projekte ihre Zielgruppe auch in der *Entwicklungs- und Umsetzungsphase* einbezogen hatten, um die Anwendung der V/AR-Technologie zielgruppenorientiert mit ihren Anwendungsfällen testen und optimieren zu können. In einigen wenigen Fällen wurde die spezifische Zielgruppe sogar in das Entwicklungsteam direkt als Entwickler:in integriert. Insgesamt ist ein echtes Bemühen zu erkennen, die Zielgruppe und die Zielsetzung des Projekts zumindest in Teilen miteinander zu verzahnen, um die Demonstratoren durch (eine oder auch mehrere) Feedbackschleifen zu optimieren und an den Bedürfnissen der Zielgruppe partizipativ anzupassen.

Besonders erfolgreich war die zielgruppengerechte Realisierung der Anwendung, wenn die Projekte ihr Augenmerk auf die vielfältigen Merkmale ihrer Zielgruppe wie die demografischen, geografischen, psychografischen und verhaltensbezogenen Faktoren, aber auch Alter, Geschlecht, Wohnort, Bildungsgrad, Affinität zu neuen Technologien usw. legten (Initiative D21 e.V. 2022). Den D21-Digital-Index ernst zu nehmen, hat sich im Projektverlauf als eines der wichtigsten Kriterien herausgestellt, allein um den ELSI-Anforderungen (u. a. mentale und ethische Fragen wie Überforderung, Einsatz von Strom am Menschen, rechtliche Aspekte wie digitale Souveränität) zu entsprechen, und sei es für die Genehmigung des Ethikrats für das Ausrollen des Projekts.

Eine intendiert breitgefächerte Einbeziehung von diversen Bürger:innen erfolgte in erster Linie nur bei den Projekten, die sich dem »Öffentlichen Raum« zuordnen ließen, auch wenn dieser hohe Anspruch in der Förderlinie als ein wichtiges Förderkriterium ausformuliert wurde. Beispielhaft wurde dies in den Stadtplanungsformaten der Projekte »INSPIRER« und »XR-Part« umgesetzt. Dort fungierten die beteiligten Bürger:innen als sogenannte »kritische Instanz« sowie als Multiplikator:innen, die sowohl die Vorteile als auch die Herausforderungen der V/AR-Technologie über die gesamten Projektphasen hinweg kommunizierten (vgl. dazu auch Straßburger & Rieger, 2019: 230–234).

Die Lücke der geringen Bürgerbeteiligung wurde durch das zusätzliche bürgerwissenschaftliche Angebot vom Living Lab »XR-Interakt« geschlossen, in dem diverse Bürger:innen alle 15 Demonstratoren getestet haben. Das Feedback wurde systematisch erfasst und an die Einzelprojekte zurückgespiegelt. Auf diese Weise konnten die Projekte in der letzten Projektphase für ihre finalen Prototypen ein letztes Feedback einholen und konstruktiv verwerten.

2.2.4 Zielgruppenorientierte Herausforderungen bei der Entwicklung von V/AR-Anwendungsszenarien

Dieser Abschnitt soll über die Besonderheiten und Herausforderungen bei der Entwicklung der Anwendungsszenarien für die anvisierte Zielgruppe Aufschluss geben. Dies wird anhand der Szenarien in den Bereichen Bildung, Gesundheit und dem Öffentlichen Raum beschrieben.

Mithilfe von V/AR-Szenarien kann der *Bildungsbereich* modernisiert und erweitert werden. Durch Simulationen sind spezielle Anwendungsszenarien für das Training von besonderen Situationen, das Gruppenlernen, lernunterstützende Formate im Unterricht, Studium oder Nachhilfe wie z. B. XR-Vorlesungen möglich. Das kann das Gruppen- und auch das Einzellernen betreffen. Für die Umsetzung mussten daher u. a. situative Herausforderungen berücksichtigt werden, wie die eines komplexen Unterrichtsszenarios oder einer Klassensituation, die nur in begrenzter Größe dargestellt werden konnten. In weiteren Projekten musste die Anwendergruppe erstmal auf das bevorstehende Anwendungsszenario »gebrieft« werden, damit sich die Testperson in der virtuellen Umgebung überhaupt zurechtfindet. Bei einer klar umrissenen Trainingsumgebung spielten Konstitution und Demografie der Zielgruppe wie Größe, Gewicht und Mobilitätsgrad eine große Rolle. Zu berücksichtigen war – insbesondere bei der Anwendung im Grundschulkontext – auch das Gewicht eines größeren Tablets. Es konnten aber auch die (eigenen) Vorstellungen über die Möglichkeiten der Interaktion in der virtuellen Realität mit den tatsächlichen Gegebenheiten abweichen. Räumliche, aber auch sicherheitstechnische Anforderungen fanden Eingang in die Anwendungsszenarien.

Im *Gesundheitswesen* entwickelten die entsprechenden Teilprojekte Szenarien u. a. für den Therapieeinsatz und das Training für Fachkräfte und Auszubildende wie beispielsweise das Knotenbinden und das Legen medizinischer Instrumente. Spezielle Herausforderungen im Gesundheitswesen waren im Besonderen die Berücksichtigung diverser ethischer Fragen. In vielen Teilen musste zunächst die Ethikkommission der medizinischen Fakultäten oder Institutionen für die Einschätzung der sorgenfreien Anwendung am Menschen eingeschaltet werden. Das betraf z. B. Fragen, die mit dem kritischen Einsatz von Strom einhergingen. Themen wie wenig vorhandene Vorarbeiten/-entwicklungen, das betretene Neuland der Thematik sowie die Operabilität und Standards mussten ebenfalls gelöst werden.

Die Teilprojekte im *Bereich Öffentlicher Raum* entwickelten in der Regel mehrere Szenarien zur partizipativen Gestaltung der Gesellschaft. Es wurden Szenarien wie

- Museumsbesuche für Familien, Freund:innen und Schulklassen, Integration in den Unterricht,
- Klimahandeln für Bürger:innen,
- soziale Treffen und Veranstaltungen für Bürger:innen, Familien und Freund:innen,
- Beteiligungsszenarien in der Stadtplanung und Stadtteilplanungen wie Gestaltung eines Kreisels, Stadtparks oder Stadtteilviertels

entwickelt. In diesen Bereichen hatte die Inklusion der Bürger:innen in den Prozessen eine evidente Stellung. Die intensive Beteiligung der Bürger:innen hatte für all diese Projekte eine zentrale Bedeutung. Das Potenzial für mehr Bürgerengagement bei der Evaluation konnte allein schon anspruchsvoll sein. Neben der Akquise *diverser* Bürger:innen stellten sich auch unvorhergesehene Fragestellungen als große Herausforderung heraus. Wie viele Feedbackschleifen konnten tatsächlich berücksichtigt werden, damit das Projekt noch innerhalb der Laufzeit der Förderung erfolgreich umgesetzt werden konnte? Die Projektbeteiligten stießen dabei immer wieder auf unerwartete Erwartungen und Reaktionen wie z. B. die Erwartungshaltung, dass das Szenario 1:1 umgesetzt werden sollte. Der Realisierbarkeitsgrad musste immer wieder mit den Erwartungen neu angepasst, aber auch mit den beteiligten Bürger:innen (erneut) kommuniziert werden.

3. Mehrwert der V/AR-Technologie für die digitale Gesellschaft

Aus den Befragungen der Bürgerinnen und Bürger auf der einen Seite und der Befragung der Expertinnen und Experten der Projekte auf der anderen Seite konnte das Living Lab »XR-Interakt« das Potenzial und die Möglichkeiten der V/AR-Einsatzszenarien herausarbeiten. Diese positiven Effekte sind in Abbildung 5 zusammenfasst.

Mit Blick auf die Abbildung lassen sich viele positive Aspekte herausarbeiten. V/AR-Anwendungsszenarien sind orts-, zeit- und anwenderunabhängig einsetzbar. Dadurch lassen sich – auf mittlere Sicht – sicherlich auch die Kosten reduzieren. Ebenso wird auf die Möglichkeit der Anwendung in der digitalen Gesellschaft durch vielfältige Szenarien hingewiesen. Der Aspekt der Sozialisierung mit der Technologie trifft besonders auf die Projekte im Bildungskontext zu. Gerade im Gesundheitswesen, aber auch im Öffentlichen Raum lässt sich ein sehr hoher Innovationsindex durch die vielfältigen Anwendungen vermuten. Anwender:innen können zu Multiplikator:innen für den Einsatz der V/AR-Technologie werden, d. h., nutzt man V/AR-Anwendungen bereits als technische und soziale Innovation und als digitales Werkzeug in der Bildung, kann der Zugang, die Sozialisie-

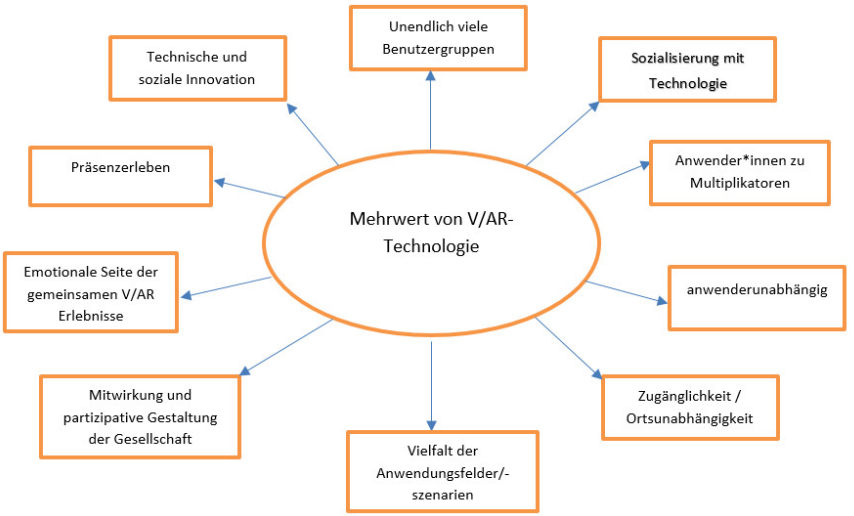


Abbildung 5: Mehrwert von V/AR-Technologie – Beispiel
Quelle: Eigene Darstellung

rung, die Verbreitung und der Umgang sinnhaft vollzogen werden. Im sozialen Bereich sind besonders die emotionale Seite der gemeinsamen V/AR-Erlebnisse und das Zusammenführen von Menschen von verschiedenen Orten hervorzuheben. Das gemeinsame Erleben von bestimmten V/AR-Umgebungen und/oder das gemeinsame Durchführen von Aktivitäten kann den sozialen Zusammenhalt der Gesellschaft mit Blick auf Gemeinschaft und soziale Teilhabe stärken. Das Präsenzerleben als technische und soziale Innovation als Paradigmenwechsel ist an dieser Stelle ebenfalls zu betonen. Da die Technologie an sich anwenderunabhängig ist, könnten unendlich viele Benutzergruppen davon partizipieren und eine Vielfalt von Anwendungsfelder zum Tragen kommen.

Die spezifischen Anforderungen variierten von Projekt zu Projekt, wobei zwischen ein bis fünf Anwendungsszenarien zumeist entwickelt und auch erprobt wurden. Die V/AR-Technologien wurden sowohl von den Expertinnen und Experten als auch den beteiligten Bürgerinnen und Bürgern in ihrer Anwendung als zukunftsweisende soziale und technische Innovation angesehen, die eine breite Palette von Einsatzmöglichkeiten für unterschiedlichste Anwendungsgruppen in verschiedenen Bereichen bietet. Dies schafft einen erheblichen Mehrwert für die digitale Gesellschaft, sowohl in persönlichen, sozialen, beruflichen als auch in gesellschaftlichen partizipativen Kontexten.

4. Fazit

Die 15 geförderten Projekte in der Förderlinie des BMBF »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« haben vielfältige V/A/XR-Anwendungsszenarien entwickelt und umgesetzt. Dabei wurden die Projekte durch das Living Lab »XR-Interakt« zur Förderung der Kooperation, des Austausches und der Interaktion begleitet.

In allen Projekten wurde die anvisierte Zielgruppe mit den entwickelten V/ARSzenarien und der eingesetzten Technologie gezielt angesprochen. Die Integration der Zielgruppen erfolgte zu verschiedenen Zeitpunkten: War die Zielgruppe von Anfang an hinreichend bekannt, konnte sie bei der Ideenentwicklung und in der Konzeptions- und Beantragungsphase bereits direkt involviert werden. Bei anderen Projekten wurde der Integrationsprozess der Zielgruppe erst nach der Prototypentwicklung angeschoben. In wenigen Ausnahmefällen war die Zielgruppe sogar Teil des Entwicklungsteams. Bei einigen Projekten wurden die anvisierten Zielgruppen durch die Initiierung des Living Lab »XR-Interakt« in die Entwicklungs- und Umsetzungsphase aktiv integriert.

Eine verstärkte Orientierung an den D21-Zielstellungen (Initiative D21 e.V. 2022) wurde durch spezifische bürgerwissenschaftliche Angebote durch das Living Lab »XR-Interakt« explizit befördert.

Die Projekte sind sich bewusst, dass sie sowohl Grundlagenforschung im Bereich der Technologien Virtual Reality, Augmented Reality und Mixed Reality betreiben, aber auch zielgruppenorientierte Anwendungsgebiete für die digitale Gesellschaft erobern. Angesichts des begrenzten Budgets und der dreijährigen Projektlaufzeit wurden die Projektgruppen vor besondere Herausforderungen gestellt. Die interdisziplinäre Zusammenarbeit, ein vielfältiges Konsortium, schnelle technologische Entwicklungen und eine gemeinsame Sprache wurden besonders häufig genannt. Aber auch die Integration der Zielgruppe und die Entwicklung der spezifischen Anwendungsszenarien haben die Projekte stark gefordert.

Die Mehrheit der Projekte strebt an, die gewonnenen Erkenntnisse und Ergebnisse nachhaltig zu implementieren, z. B. durch die Integration in Ausbildungs-Curricula. Projektergebnisse und Erkenntnisse wurden durch Publikationen, Handlungsempfehlungen, Leitfäden und Open-Source-Bereitstellungen verbreitet. Dabei können jedoch Einschränkungen durch Patente, den Umgang mit großen Datenmengen und -sicherheit, kostenpflichtige Frameworks sowie spezielle Interessen der wirtschaftlichen Kooperationspartner:innen auftreten. Zudem sind weitere Projekte und Folgeprojekte zur Verstetigung, Fortführung und Weiterentwicklung in Planung.

Die geförderten Projekte sind sich ihrer gesellschaftlichen Verantwortung bewusst, die mit der Einführung einer neuen Technologie in die Gesellschaft einhergeht. Die Mehrheit der Projekte entwickelte Zukunftsstrategien, um ihre Vorhaben und Lösungen in der Zukunft weiterzuentwickeln und nachhaltig zu verstetigen.

Bei allen Einzelprojekten ging es darum, alltagstaugliche und erfolgreiche Angebote zu entwickeln. Nur langsam etablieren sich in dem V/AR-Sektor Quasi-Standards der 3D-Interaktion, die es einfach machen würden, zwischen 3D-Anwendungen zu wechseln. Gute Usability, ein niedrigschwelliger Zugang und positive UX werden aber neben der Erfüllung sachlich-technischer Anforderungen durch die in den verschiedenen Projekten erarbeiteten Lösungen entscheidend sein, wenn es um die Akzeptanz geht. Die Digitalisierung bietet an dieser Stelle ganz neue Möglichkeiten der Partizipation, erfordert angesichts der daran gekoppelten Herausforderungen zugleich auch einen sehr reflektierten Blick auf die jeweiligen Einsatzszenarien von transformativen Technologien wie V/AR-Technologien.

Förderhinweis

Das Projekt Living Lab – XR-Interakt wurde vom Bundesministerium für Bildung und Forschung (BMBF)[4] gefördert.

Literatur

Bath, Corinna (2002). Genderforschung in der Informatik: 10 Jahre zurück – 10 Jahre voraus? In: *FIfF – Kommunikation, Frauen in der Informationsgesellschaft*, Vol. 3, Teil 2, Forum InformaterInnen für Frieden und gesellschaftliche Verantwortung e.V., Bremen, S. 41–46.

Bonn, Aletta, et al. (2016). *Grünbuch Citizen Science Strategie 2020 für Deutschland*, Helmholtz-Zentrum für Umweltforschung (UFZ), Deutsches Zentrum für integrative Biodiversitätsforschung (iDiv) Halle-Jena-Leipzig, Leipzig, Museum für Naturkunde Berlin, Leibniz-Institut für Evolutions- und Biodiversitätsforschung (MfN), Berlin-Brandenburgisches Institut für Biodiversitätsforschung (BBIB), Berlin. https://www.buergerschaffenwissen.de/sites/default/files/assets/dokumente/gewiss-gruen-buch_citizen_science_strategie.pdf [Zugriff 06.02.2025]

Bundesministerium für Bildung und Forschung (2020). Bekanntmachungen. Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft

4 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

(VAR2), Berlin. https://www.interaktive-technologien.de/foerderung/bekanntmachungen/var2 [Zugriff 06.02.2025]

Bundesministerium für Bildung und Forschung (2020). Bekanntmachungen. Richtlinie zur Förderung und Forschung und Entwicklung auf dem Gebiet »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft«, Bundesanzeiger vom 02.04.2024, Berlin. https://www.bmbf.de/bmbf/shareddocs/bekanntmachungen/de/2020/04/2912_bekanntmachung.html#searchFacets [Zugriff 06.02.2025]

Bundesministerium für Bildung und Forschung (2020). Bekanntmachungen. Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft (VAR2), Projekte, XR-Interakt, Berlin. https://www.interaktive-technologien.de/projekte/xr-interakt [Zugriff 06.02.2025]

Floyd, Christiane, Reisin, Fanny-Michaela & Schmidt, Gerhard (1989). STEPS to Software Development with Users. In: Ghezzi, Carlo & McDermid, John A. (Hg.), ESEC '89, Lecture Notes in Computer Science No. 387. Springer, Berlin, Heidelberg. S. 48–64. https://doi.org/10.1007/3-540-51635-2

Kuckartz, Udo & Rädicker, Stefan (2022). Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung (Grundlagentexte Methoden), 5. Aufl., Beltz Juventa, Weinheim.

Flick, Uwe (2014). Gütekriterien qualitativer Sozialforschung. In: Baur, Nina & Blasius, Jörg (Hg.), *Handbuch Methoden der empirischen Sozialforschung*, Springer VS, Wiesenbaden, S. 411–423.

Initiative D21 e.V. (2022). *D21-Digitalindex 2023/24*. https://initiatived21.de/themen/digitalisierung-und-nachhaltigkeit [Zugriff 06.02.2025]

Radianti, Jaziar et al. (2020). A systematic review of immersive virtual reality applications for higher education. Design elements, lessons learned, and research agenda. In: *Computers & Education*, Vol. 147, Art. 103778, Elsevier B.V., Amsterdam. https://doi.org/10.1016/j.compedu.2019.103778

Rommes, Els (2014): Feminist Interventions in the Design Process. In: Ernst, Waltraud & Horwath, Ilona (Hg.), *Gender in Science and Technology*, transcript Verlag, Bielefeld, S. 41–55.

Schenk, Johannes et al. (2022): Wissensmanagement in virtuellen Welten: Wissensgenerierung in Virtual Reality-Umgebungen. In: *HMD Praxis der Wirtschaftsinformatik* Vol. 59, S. 159–176. https://doi.org/10.1365/s40702-021-00814-z

Straßburger, Gaby & Rieger, Judith (2019). *Partizipation kompakt*, Beltz Verlag, Weinheim, S. 230–234.

Vines, John et al. (2015). The beginnings, middles and endings of participatory research in HCI: An introduction to the special issue on ›perspectives on participation‹. In: *International Journal of Human-Computer Studies*, Vol. 74(0), S. 77–80.

Wiesner, Heike et al. (2019). Digitale Souveränität und Geschlecht – das geht uns alle an! In: *Ksoe Dossier digital, digital + geschlechtergerecht*, Vol. 01, S. 10–15. https://www.ksoe.at/dl/qMrlJmoJmMOJqx4KJKJKJKkmoLl/Dossier_01_2019_digital.pdf [Zugriff 06.02.2025]

Wiesner, Heike & Stepczynski, Jan (2020). MEHR MEDIENVIELFALT, Digitale Akademie Pflege 4.0. In: *Die Schwester | Der Pfleger*, Vol. 06/20, S. 17–19.

Normung für XR und Metaverse – Living Lab »XR-Interakt«

Christoph Runde

Zusammenfassung

Dieser Fachbeitrag thematisiert die Normung im Bereich Extended Reality (XR) und Metaverse. Ziel dieses Teilprojekts der Begleitforschung der Programmlinie »Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft« (Bundesministerium für Bildung und Forschung [BMBF][1], 2020) war die Erstellung eines umfassenden Verzeichnisses von XR-Normen, -Standards und -Normungsorganisationen, um die Fragmentierung des Wissens zu überwinden und interoperable Lösungen der Teilprojekte zu fördern. Das im Rahmen dieses Projekts entwickelte Verzeichnis ist heute das weltweit größte seiner Art. Es unterstützt Unternehmen und Forschungseinrichtungen durch die Kategorisierung nach Themen wie Grundlagen, Hardware, Anwendungen und Usability. Normen bieten Vorteile wie Kostenreduktion, Markttransparenz und Rechtssicherheit, fördern aber auch Wettbewerb und Innovation. Das Verzeichnis wird öffentlich bereitgestellt, um Unternehmen und Forschungseinrichtungen den Zugang zu relevanten Informationen zu erleichtern und den technologischen Fortschritt zu unterstützen. Der Beitrag betont die Bedeutung der aktiven Mitgestaltung von Normen, um Zukunftsmärkte erfolgreich mitzugestalten.

1. Herausforderung des Begleitprojekts in der Normung

Eine der Aufgaben des Teilprojekts Living Lab »XR-Interakt« bestand darin, die parallellaufenden 15 XR-Entwicklungsprojekte bei der gemeinsamen Nutzung von Normen und der Entwicklung interoperabler Lösungen – die auf Standards beruhen – zu unterstützen. Dazu wurde es notwendig, zu ermitteln, welches

1 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

der 15 BMBF-Entwicklungsprojekte welche Normen und Standards einzusetzen plante. Gleichzeitig sollten die Projekte über sinnvoll einsetzbare Normen informiert und zur Normenarbeit befähigt werden. Zur Förderung der Kooperation zwischen den Projekten war es ferner sinnvoll, zu eruieren, ob sich Projektpartner:innen bereits in der Normenarbeit engagierten und ob andere Projekte davon ebenfalls profitieren könnten.

Die Herausforderung bei der Normennutzung bestand zunächst darin, dass das Wissen zur XR-Normung sehr zersplittert war. Ohne eine umfassende Recherche war es aber unmöglich, zu bestimmen, welche XR-Normen ein bestimmtes Thema tangierten oder welche Normungsarbeitsgruppen ein definiertes Sachgebiet bearbeiteten. Dies ist u. a. auf die Vielzahl an Normungsorganisationen (sogenannte Standard Development Organisations [SDOs]) im XR-Kontext zurückzuführen, die – teils konkurrierend – Normen und Standards definieren (Runde, 2023: 74). Die Zielsetzung dieser Arbeit war daher u. a. die Schaffung eines Verzeichnisses, das eine qualifizierte Übersicht über alle XR- und Metaverse-relevante Normen, Industriestandards, Richtlinien, Empfehlungen, SDOs und Arbeitsgruppen bietet.

2. Normung: Hintergrund

2.1 Die Bedeutung der Normung

Die Bedeutung der Normung für die moderne Wirtschaft ist nicht zu unterschätzen. Sie eröffnet enorme ökonomische Potenziale und trägt maßgeblich zur wirtschaftlichen Stärke Deutschlands bei. Deutschland hat hier immer eine führende Rolle eingenommen. Diejenigen, die die Normen definieren, beeinflussen maßgeblich die Märkte der Zukunft, was auch für den Bereich der XR gilt. Es ist daher von entscheidender Bedeutung, diese Zukunftsmärkte aktiv mitzugestalten. Auch ohne aktive Beteiligung ist es jedoch ratsam, die bestehenden Normen zu kennen und anzuwenden. Es empfiehlt sich, bewährte Lösungen zu nutzen und die Ressourcen auf die wirklich innovativen Aspekte der Entwicklungen zu fokussieren. Normen definieren den Stand der Technik und sollten daher beachtet werden.

2.2 Standardisierung und Normung im Überblick

Die Standardisierung ist allgemein die Vereinheitlichung von Produkten, Bauteilen oder Verfahren auf eine oder wenige Varianten. Es ist eine Vereinheitlichung

von Objekten nach gewissen Vorlagen. Die Vorgehensweisen zur Standardisierung sind Normung und Typisierung. Der Begriff kann auf verschiedene Gebiete angewandt werden. Beispiele in der Wirtschaft sind etwa in der Fertigung die Normung und Typisierung von Teilen, Zwischen- oder Endprodukten sowie im Rechnungswesen die Standardisierung von Kostenarten (Springer Fachmedien Wiesbaden, 2014: 519).

Normung bezeichnet die Formulierung, Herausgabe und Anwendung von Regeln, Leitlinien oder Merkmalen durch eine anerkannte Organisation und deren Normengremien (DKE, 2020: 1). Sie sollen auf den gesicherten Ergebnissen von Wissenschaft, Technik und Erfahrung basieren und auf die Förderung optimaler Vorteile für die Gesellschaft abzielen. Die Festlegungen werden mit Konsens erstellt und von einer anerkannten Institution angenommen. Anerkannte Normungsinstitutionen sind etwa die International Organization for Standardization (ISO), die International Electrotechnical Commission (IEC) oder das Deutsche Institut für Normung (DIN) (DKE, 2020: 1).

2.2.1 Argumente für die Normung und Standardisierung

Der Hauptzweck der Standardisierung ist in der Kostenersparnis und der Arbeitsvereinfachung zu sehen. Die Standardisierung führt zur Erhöhung der Markttransparenz und zur Kostensenkung (bei Herstellungskosten, Informationskosten, Transaktionskosten, Versandkosten, Vertriebskosten, Wechselkosten) (Springer Fachmedien Wiesbaden, 2014: 519). Durch die Verwendung von Standards können sich Hersteller auf die wirklich innovativen Aspekte ihrer Produkte konzentrieren.

Aufseiten der Käufer:innen standardisierter Produkte, Services und Dienstleistungen sind besonders die reduzierten Wechselkosten sowie die Kompatibilität wichtig: Eingekaufte Produkte und Services werden deutlich leichter austauschbar sowie technisch anschlussfähig und dadurch damit integrierbar. Dadurch intensiviert sich der Wettbewerb, was zu Preisdegression und Leistungssteigerung führen muss.

Ein ganz wesentlicher Vorteil der Normung besteht in der Erlangung gewisser Rechtssicherheit von auf den Markt gebrachten Produkten. Juristisch relevant ist hier vor allem die sogenannte Beweislastumkehr: Kommt es zu einem Schadensfall mit einem nichtnormierten Produkt, so muss der Hersteller nachweisen, das Produkt fehlerfrei entwickelt zu haben. Kommt es jedoch zu einem Schadensfall mit einem normierten Produkt, so wird dem Hersteller zugestanden, auf der Höhe des Stands der Technik gehandelt zu haben. In diesem Fall muss der/die Kund:in nachweisen, dass der Hersteller fehlerhaft gehandelt hat. Hier wird also die Beweislast umgekehrt.

Das DIN, die Deutsche Gesellschaft für Qualität (DGQ) und die Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE (DKE) stellten die Vorteile der aktiven Normungsarbeit deutlich heraus:

So formuliert das DIN: »Normen entwickeln diejenigen, die sie später anwenden. Damit der Markt die Normen akzeptiert, sind eine breite Beteiligung, Transparenz und Konsens Grundprinzipien bei allen beteiligten Organisationen. Durch die Mitarbeit in der Normung bringt ein Unternehmen einerseits seine eigenen Interessen ein, andererseits kann es den Markt und mögliche Wettbewerber beobachten. Wer sich aktiv an der Normung beteiligt, erwirbt einen Wissensvorsprung, kann seine Kosten senken und somit Marktöffnungsstrategien verfolgen.« (DIN, 2023: 1)

Die DGQ führt aus: »Aktive Normungsarbeit ist eine Zukunftsinvestition – sowohl für die Normungsexpertinnen und -experten als auch für die entsendenden Organisationen. Durch aktive Mitgestaltung der Zukunft über Normen haben die mitwirkenden Personen und Organisationen direkten Einfluss auf die spätere Anwendung. Sie erhalten tiefere Einblicke in Hintergründe und Zusammenhänge sowie Logiken in Bezug auf das Normungsthema bzw. die Normungsinhalte. Vor allem profitieren die Beteiligten davon, dass sämtliche Normungsaktivitäten im Themenfeld des jeweiligen Ausschusses transparent für alle Experten kommuniziert werden: Dies gibt ihnen die Möglichkeit, sich über Stellungnahmen einzubringen und über Änderungen und Vorschläge zu Normen und -entwürfen abzustimmen. Nicht zu unterschätzen sind auch die Netzwerkaspekte: Bei aktiver Mitwirkung entstehen zahlreiche fachliche Kontakte mit Kolleginnen und Kollegen maßgeblicher Organisationen – und potenziell auch Karriereoptionen. Der direkte Austausch mit Vertretern anderer Interessengruppen, sich einen Wissensvorsprung verschaffen, Einfluss nehmen, Respekt und Akzeptanz erfahren und schließlich den Bekanntheitsgrad von Person und Organisation zu erhöhen, sind darüber hinaus weitere positive Effekte.« (DGQ, 2023: 2)

Die DKE führte 2019 Kurzinterviews mit einigen wenigen Unternehmern zum Nutzen des Engagements in der Normung durch. Die Befragten nannten daraufhin (DKE, 2023: 2):

- Frühzeitig auf relevante Themen einstellen und reagieren können
- Eigene Produkte sicherer machen und hierdurch am Markt positionieren
- Die Wirtschaft international stärken und hierdurch Unternehmenserfolg steigern
- Normen durch persönliches Wissen und eigene Interessen mitgestalten.

Es gibt also zahlreiche Argumente dafür, Normen, Standards und Richtlinien nicht nur passiv, also durch deren Anwendung, zu nutzen, sondern diese aktiv selbst mitzugestalten.

2.2.2 Argumente wider die Normung und Standardisierung

Den positiven Aspekten der Normierung und Standardisierung stehen allerdings auch mögliche Nachteile gegenüber. Für den Erfolg eines Produktes, einer Dienstleistung und somit des Gesamtunternehmens sollte eine »unique selling proposition« (USP), also ein Alleinstellungsmerkmal, gegeben sein. Damit kann es problematisch erscheinen, dass eine Standardlösung Basis einer guten USP sein kann, da sie sich von der Masse absetzen muss (Runde, 2023: 73). Das heißt zwangsläufig, dass der besondere Mehrwert des Produktes nicht durch eine standardisierte Wertschöpfungskette, sondern nur durch eine Abweichung vom Standard entstehen kann.

Analysiert man bestehende Standards, so wird deutlich, dass es sich bei den formulierten Kriterien häufig um Minimalanforderungen handelt (also das, was Kund:innen sowieso erwarten können sollten). Das konsequente Verfolgen von Standards kann also dazu führen, dass Unternehmen sich immer am untersten Rand der Möglichkeiten befinden und ihre Alleinstellungsmerkmale sogar aufgeben.

Blue-Ocean-Strategien beinhalten – etwas verkürzt gesagt – insbesondere das Weglassen erlernter Features eines Produkts oder einer Leistung, wenn dadurch möglich wird, den Kund:innen an anderer Stelle einen erheblichen Mehrwert zu schaffen (Gabler Wirtschaftslexikon, 2025: 1). Hier ist es sinnvoll, Standards nicht zu erfüllen, wenn dies dem Alleinstellungsmerkmal und dem Erfolg des Produktes dient.

Standardisierung begründet die Gefahren der Schematisierung und des Flexibilitätsverlustes. Es kann vorkommen, dass Aspekte, die nicht standardisiert werden können, vernachlässigt werden. Standardisierung ist ein abstimmungsintensiver Prozess, der hohe Kosten verursacht und im Normalfall eine Dauer von einigen Jahren benötigt. Daraus ergibt sich ein zu statischer Rahmen. Standardisierung kann dazu führen, dass der kreative Freiraum eingeschränkt wird. Dieses Abwürgen von Innovationsfeldern ist insbesondere im Umfeld der Entwicklung innovativer Technologien nicht erwünscht.

2.2.3 Fazit

Normung und Standardisierung bieten enorme Mehrwerte und sind daher ökonomisch grundsätzlich als sehr sinnvoll zu bewerten. Bewusst auf Normung zu verzichten, kann allenfalls eine mit einem gewissen Risiko behaftete Innovationsstrategie sein, die auf komplett neue, wettbewerbsarme Märkte ausgerichtet ist.

3. Vorgehensweise: Arbeitsmethode

Die Zielsetzungen dieses Teilprojekts (Überwindung des zersplitterten Normenwissens und Befähigung der Teilprojekte zur Normenarbeit) erforderten zwei Arbeitsansätze. Der erste beschäftigte sich mit Qualifizierungsfragestellungen, der zweite mit der Schaffung eines umfassenden Registers.

3.1 Qualifizierung zur Normennutzung und zur Normenarbeit

Im Projekt wurden bestehende Angebote zur Normenlehre recherchiert, dokumentiert und Nutzungsrechte identifiziert. In Deutschland bieten insbesondere die deutschen Normungsorganisationen DIN und DKE Qualifizierungsangebote zur Normenarbeit. Dazu zählen Schulungsliteratur, Vorträge, Planspiele und Veranstaltungen bei DIN und DKE. Auch die Danish Standards Foundation (DSF) (Bøgh, 2015) und das European Telecommunications Standards Institute (ETSI) (Abdelkafi, 2021) bieten hervorragende digitale Lehrbücher zur Normung. Die ISO (ISO, 2014) und der Verein Deutscher Ingenieure (VDI) (VDI, 2017) geben Normungs-Lehrenden Hilfestellung.

3.2 Normenrecherche

Die Identifikation sämtlicher relevanter XR-Normungsorganisationen, XR-Normen und Normungsprojekte impliziert umfangreiche, langwierige und iterative Rechercheaufgaben. Diese wurden mittels folgender Maßnahmen umgesetzt: Es erfolgte der aktive, persönliche Austausch mit Vertreter:innen von Normierungsorganisationen. Projektmitarbeiter:innen nahmen an einschlägigen Veranstaltungen teil, bei denen die XR-Normung thematisiert wurde. Aufgrund der Breite, Streuung und Unübersichtlichkeit des Themas wurde die Recherche mit unterschiedlichen Startpunkten durchgeführt. Normungsorganisationen, aber auch Verbände von XR-Anwenderbranchen wurden ermittelt und nach XR-Arbeitsgruppen und XR-Normungsprojekten durchsucht. Ausgehend von weiteren bekannten XR-Normen konnte wiederum auf Arbeitsgruppen, Normungsorganisationen und Verbände geschlossen werden. Die Suche nach Dokumenten wurde ebenfalls über (Hochschul-)Bibliotheken und eine freie WWW-Recherche initiiert. Dieser Prozess wurde viele Male zyklisch durchlaufen. Als Ergebnis liegen komplettierte Listen der Normungsorganisationen, der Arbeitsgruppen, der laufenden Aktivitäten und der XR-Normen vor (rote Kästchen in Abbildung 1).

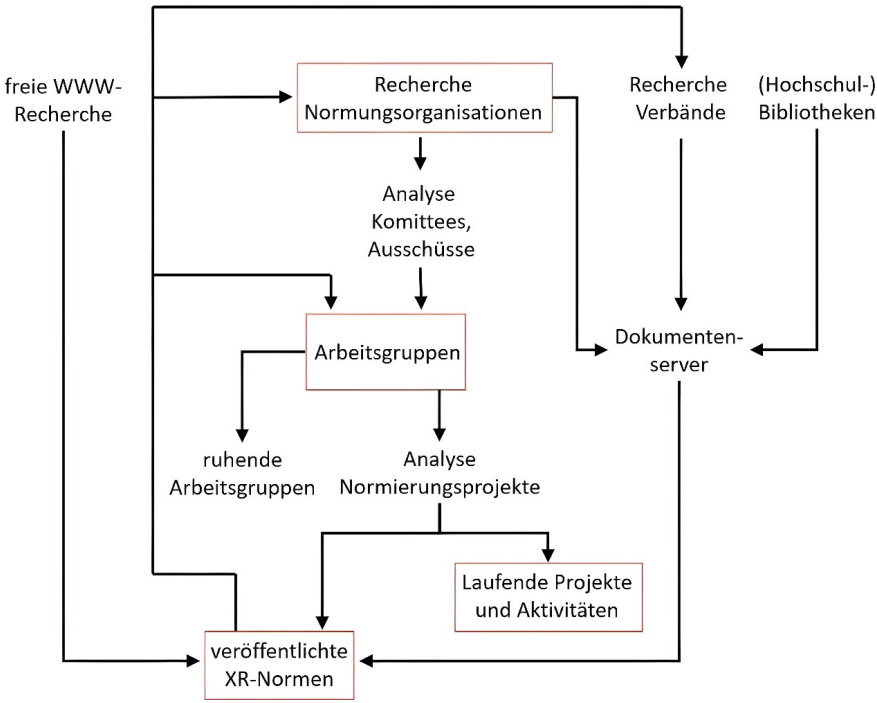


Abbildung 1: Pfade der Recherche zu Normen, Arbeitsgruppen und Organisationen
Quelle: Eigene Darstellung

Die Einträge des Registers wurden zunächst jeweils einem Typus (eigene Typografie) zugeordnet:

- standard (industry)
- standard (official)
- specification
- recommendation/ guideline
- technical report
- whitepaper
- working group
- SDO
- other organisation

Tabelle 1: Dokumententypen im Normenregister.
Quelle: Eigene Darstellung

Die recherchierten XR-Normen und -Arbeitsgruppen wurden inhaltlich analysiert und daraus Kategorien gebildet. Maßgabe dabei war, eine Anzahl von wenigen Dutzend nicht zu überschreiten, dennoch aber für hinreichende Trennschärfe zu sorgen. Aus den Vorarbeiten war bekannt, dass diese mindestens in die sechs Themenfelder Grundlagen, Hardware, Datenformate, Content & Anwendungen, Usability & Ergonomie sowie Interoperabilität zu gliedern sind. Weiterhin waren die wichtigsten Hardwarekategorien (Grafik, Akustik, Haptik, Tracking, mobile XR) zu berücksichtigen. Als zentrale Anwendungsfelder

wurden Training, Engineering/Design, Gesundheit, Assistenz, Kollaboration, Marketing, Unterhaltung/Kultur und Geo/Bauwesen ermittelt. Diese Anwendungsfelder sind als branchenübergreifend zu verstehen: So kann es etwa im Maschinenbau XR-Anwendungen im Engineering (hier: Produktentwicklung), in der Assistenz (hier: AR-Service-Support), in der Kollaboration (hier: verteiltes Entwickeln in XR) oder im Marketing (hier: XR-Vertriebskonfiguratoren) geben. Aus der fortwährenden inhaltlichen Analyse der XR-Normen ergaben sich einige weitere Themen (etwa: human modelling, content creation/modelling), die hinzugefügt wurden. Damit ergaben sich insgesamt 31 Unterkategorien:

- basics/terms
- XR management
- regulation/ethics/governance
- evaluation/conformity
- API application programming interface
- formats/coding/compression
- communication/interoperability
- mapping
- virtual reality
- augmented reality
- user experience (UX)/user interface design (UID)
- ergonomics/usability
- interaction/patterns
- graphics software/algorithms
- graphics hardware/optics/displays
- haptics
- auralisation/audio/ acoustics
- tracking/sensing/motion capturing
- mobile XR
- other XR hardware
- human modelling
- content creation/modelling
- applications: education/training
- applications: design/engineering
- applications: health
- applications: assistance
- applications: collaboration
- applications: marketing
- applications: entertainment/culture
- applications: geo & construction
- applications: others

Tabelle 2: Inhaltliche Unterkategorien im Normenregister.
Quelle: Eigene Darstellung

Werden in der Zukunft neue Themencluster erkannt, die bislang nicht hinreichend durch Kategorien abgedeckt sind, so werden durch das Virtual Dimension Center (VDC) Fellbach neue Kategorien angelegt werden.

Neben der inhaltlichen Verortung der XR-Normen, -Arbeitsgruppen und -Organisationen durch die zuvor genannten Kategorien wurde ebenfalls noch der Status erfasst und den folgenden drei Gruppen zugeordnet:

- published/active
- withdrawn/closed
- work in progress

Tabelle 3: Statusangaben im Normenregister.
Quelle: Eigene Darstellung

Insgesamt ergibt sich durch die vorgeschlagene Arbeitsmethode ein strukturierter Prozess der Recherche und inhaltlichen Einordnung von XR-Normen, -Arbeitsgruppen und -Organisationen. Beides ist notwendig, um die weitestgehende

Vollständigkeit der Sammlung zu gewährleisten als auch trotz der zu erwartenden großen Menge Übersichtlichkeit zu garantieren.

4. Projektergebnisse

Die Projektergebnisse teilen sich auf in die umgesetzten Maßnahmen zur direkten Unterstützung der 15 XR-Entwicklungsprojekte durch Qualifizierung und Vernetzung in der Normenarbeit auf der einen Seite und auf der anderen Seite in die Zur-Verfügung-Stellung umfassenden Normenwissens im neu entwickelten Normenregister.

4.1 Befähigung zur Normenarbeit

4.1.1 Qualifizierung

Normungslehre wurde in sämtliche Teilprojekte transferiert: Das Living Lab »XR-Interakt« hat sich selbst auf einem DIN-Fortbildungsworkshop am 28. September 2023 in Berlin in der Normenlehre weiterqualifizieren lassen; sämtlichen Teilprojekten wurden die ermittelten digitalen Bildungsangebote zur Verfügung gestellt und die Teilprojekte hierüber informiert. Qualifizierungsmaßnahmen wurden geplant und umgesetzt.[2]

4.1.2 Ergebnisse aus der Befragung der 15 XR-Entwicklungsprojekte

Die 15 Entwicklungsprojekte wurden durch das Living Lab »XR-Interakt« zur Normenarbeit und -nutzung befragt. Dazu kam ein Fragebogen zum Einsatz, der bereits Hunderte mögliche einzusetzende XR-Normen beinhaltete. Das Living Lab »XR-Interakt« erstellte aus den Befragungsergebnissen 15 individuelle Projekltauswertungsberichte, die darlegen, welche anderen Projekte die gleichen Normen nutzen, in gleichen Normungsarbeitsgruppen mitwirken oder an gleichen Normungsprojekten mitarbeiten. Diese für jedes Projekt individuellen Berichte konnten durch die Projekte somit genutzt werden, um im

2 Das Living Lab »XR-Interakt« hat einen Qualifizierungsvortrag des DIN beim 3. Living-Lab-Workshop »Austausch und Evaluation« vom 6.–8. Oktober 2023 beim Projektpartner Virtual Dimension Center (VDC) in Fellbach organisiert; für den 4. April 2024 organisierte das Living Lab »XR-Interakt« den DIN-Workshop »Qualifizierung zur Normenarbeit mit Serious Games« im Rahmen des 4. Living-Lab-Workshops »Austausch und Evaluation« im Haus der Wirtschaft in Stuttgart.

Normungskontext zu kooperieren und ihre Entwicklungsarbeiten aufeinander abzustimmen.

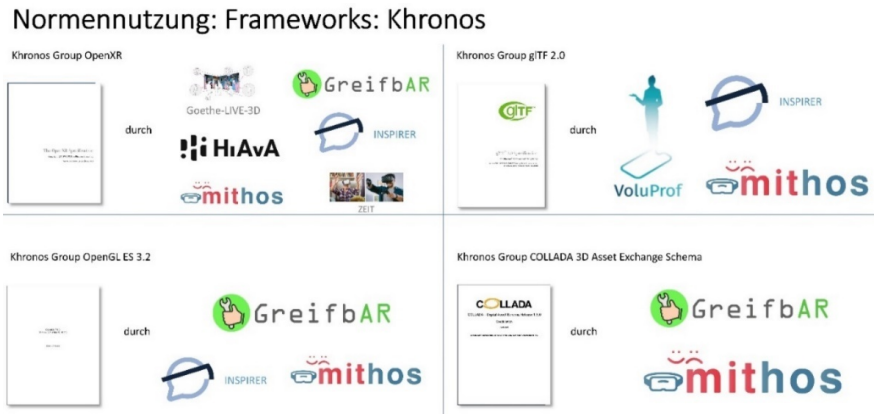


Abbildung 2: Auszug aus dem Ergebnis der projektübergreifenden Normenbefragung. Hier: Nutzung verschiedener Khronos-Frameworks durch die Teilprojekte
Quelle: Eigene Darstellung

Abbildung 2 zeigt einen Auszug aus dem Ergebnis der projektübergreifenden Normenbefragung. Hier haben Projekte angegeben, ob sie Frameworks der Khronos-Gruppe nutzen. Es zeigte sich, dass sechs Projekte OpenXR einsetzen, drei Projekte je glTF und OpenGL ES und zwei Projekte COLLADA 3D. Die Projekte wurden über die Befragungsergebnisse informiert, sodass sich auch hieraus Kooperation motivierte.

4.2 Normenverzeichnis

4.2.1 Vernetzung

Das VDC-Living Lab ist der Khronos Group, dem Metaverse Standards Forum, dem DIN, der IEEE Standards Association und der Simulation Interoperability Standards Organisation (SISO) als Mitglied beigetreten, um auf diese Weise Zugang zu den VAR-Normungsaktivitäten dieser Organisationen zu erhalten. Weiterhin hat das VDC eine Liaison mit dem ISO-IEC JTC1/SC 24 »Computer graphics, image processing and environmental data representation« vereinbart.

Mitglieder in der ISO sind ausschließlich nationale Normungsorganisationen (wie etwa das DIN für Deutschland), die fachliche und nationale Interessen in der ISO vertreten. Zu diesem Zweck organisieren die nationalen Normungsorga-

nisationen sogenannte Spiegelgremien, welche die Arbeitsgruppen der ISO bei sich lokal abbilden. In diesen Spiegelgremien findet die jeweils nationale Willensbildung statt, die dann in die korrespondierende ISO-Arbeitsgruppe eingebracht wird. In der Vergangenheit fehlte allerdings das deutsche Spiegelgremium zur o. a. ISO-Arbeitsgruppe ISO-IEC JTC1/SC 24. Das Living Lab »XR-Interakt« gründete daher am 24. Oktober 2023 gemeinsam mit dem DIN das DIN-Normungsgremium »Metaverse und Extended Reality« (NA 043–01-24 AA). Dieses DIN-Normungsgremium wird künftig die deutsche Stimme im o. a. ISO-IEC JTC 1/SC 24 sein.

4.2.2 Recherche-Ergebnis

Aktuell (Stand Anfang des Jahres 2025) sind im XR- und Metaverse-Normenverzeichnis mehr als 100 Organisationen, mehr als 250 Arbeitsgruppen und mehr als 1.000 Dokumente registriert. Damit ist dieses Verzeichnis heute das weltweit größte seiner Art (Runde, 2024a; Runde, 2024b; Runde, 2024d).

4.2.3 Online-Normenverzeichnis

Das Living Lab hat die so identifizierten relevanten Normen, Arbeitsgruppen und Normungsorganisationen zu den Themenfeldern XR und Metaverse in einem Verzeichnis geordnet und stellt dieses Verzeichnis über ein Web-Interface der Allgemeinheit zur Verfügung. Ziel dieses Angebots ist es, XR-Normenwissen zu verbreiten. Wurde in einem ersten Versuch noch eine Literatur-Datenbanklösung eingesetzt, so wurde im nächsten Schritt eine Webshop-Lösung verwendet, deren Inhalte aus einer Tabelle generiert werden. Die genannte Tabelle enthält die grundlegenden Informationen zu den eingetragenen Normen wie etwa Bezeichnung, Veröffentlichungsjahr, Status, veröffentlichende Organisation, Kurzzusammenfassung und die oben genannten Kategorien. Mit diesem Verzeichnis lassen sich heute vielfältige Suchaufgaben bearbeiten. Es besteht die Möglichkeit, nach Normenveröffentlichungen bestimmter Organisationen (etwa DIN, ISO, ANSI, Khronos Group) zu suchen; man kann Arbeitsgruppen und Normungsorganisationen finden, die sich spezifischen Themenfeldern (etwa Haptik, XR-Sicherheit, Displaytechnologie; Beispiel in u.a. Abbildung 3: Organisationen und Arbeitsgruppen im Kontext Design / Engineering) widmen; es können Normen im Bearbeitungsstand »work in progress« ermittelt werden, sodass ein Blick in die Normungszukunft möglich wird; schließlich können die Nutzer:innen das Gesamtregister schlicht nach eigens gewählten Schlagworten durchforsten.

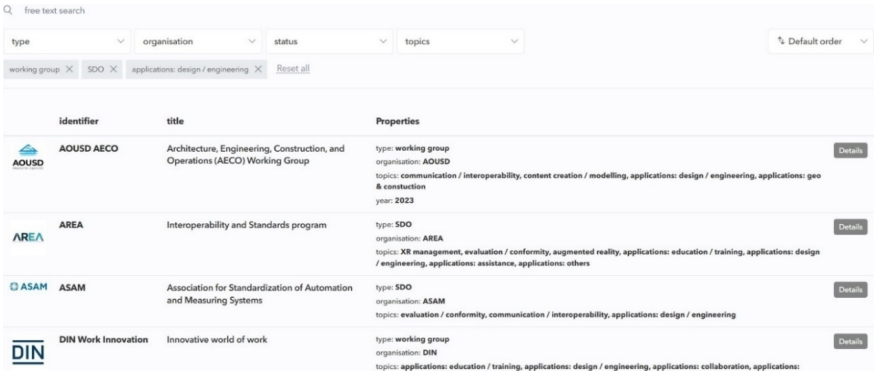


Abbildung 3: Anwendungsbeispiel: Anzeige Arbeitsgruppen und Organisationen, die sich mit der Normung von XR im Kontext Design/Engineering beschäftigen.
Quelle: Eigene Darstellung

Neben der Darstellung auf der eigenen Website lässt sich das Verzeichnis auch als iFrame in andere Websites einbauen, und einige Organisationen machen bereits Gebrauch von dieser Option.[3] Es wird daran gearbeitet, das Verzeichnis weiter zu befüllen und es weiter zu verbreiten. Damit hat das BMBF einen wesentlichen Beitrag dazu geleistet, der breiten XR- und Metaverse-Community eine sehr nachhaltige und langfristige Unterstützung in Form des Verzeichnisses zur Verfügung zu stellen.

5. Fazit

Auch transformative Technologien werden sich in der digitalen Gesellschaft auf breiter Front nur als Standards durchsetzen können. Die Normung ist jedoch ein in ihrer Wichtigkeit häufig unterschätztes Thema. Nicht nur dient die Berücksichtigung von Normen und Standards der Entwicklung interoperabler und marktgängiger Lösungen, sondern lenkt auch Entwicklungsressourcen auf die wirklich innovativen Aspekte einer Idee und bildet die Grundlage technologischer Synergien. Die aktive Gestaltung von Normen kann sogar Zukunftsmärkte mitdefinieren, was von strategischer Bedeutung sein kann (Runde, 2024c: 9). Lei-

3 Das Register wird aktuell auf der Seite des VDC (https://xr-metaverse-standards-register.vdc-fellbach.de/) ausgegeben, ebenso auf den Seiten der baden-württembergischen Metaverse-Initiative CyberLänd (https://cyberlaend.eu), auf der Seite des europäischen XR-Verbands EuroXR Association (https://www.euroxr.org) und auf der Seite der international ausgerichteten X Reality Safety Intelligence Initiative (XRSI) (https://xrsi.eu/xr-standards-register).

der ist aber das Wissen um die Normung – insbesondere in der Forschung und bei kleinen und mittelständischen Unternehmen – nicht in dem Umfang vorhanden, wie es wünschenswert wäre. Das Living Lab »XR-Interakt« hat mit diesem konzentrierten Ansatz zur Schaffung eines XR- und Metaverse-Verzeichnisses eine sehr signifikante Grundlagenarbeit für die gesamt XR- und Metaverse-Fachgemeinschaft erarbeitet. Das Verzeichnis findet in der Fachgemeinschaft aktuell eine sehr positive Resonanz, sodass das Verzeichnis in seiner offenen Form für die Allgemeinheit weiter betrieben und ausgebaut werden wird.

Förderhinweis

Das Projekt Living Lab – XR-Interakt wurde vom Bundesministerium für Bildung und Forschung (BMBF)[4] gefördert.

Literatur

Abdelkafi, Nizar et al. (2021). *Understanding ICT Standardization: Principles and Practice,* European Telecommunications Standards Institute ETZSI, Sophia Antipolis. https://www.etsi.org/images/files/Education/Slideset_Understanding_ICT_Standardization.pdf [Zugriff 20.03.2025]

Bøgh, Signe Annette (Hg.) (2015). *A World Built on Standards – A Textbook for Higher Education*, Danish Standards Foundation, Göteborg/Dänemark. https://www.ds.dk/media/px5jhney/a-world-built-on-standards.pdf?srsltid=AfmBOop6TjpbdIaMSCSQ4zYx-MAMr7Qlp55d6jYg7h5GXjtyvjCj7vGz [Zugriff 20.03.2025]

Bundesministerium für Bildung und Forschung (2020). Bekanntmachungen. Interaktive Systeme in virtuellen und realen Räumen – Innovative Technologien für die digitale Gesellschaft (VAR2). https://www.interaktive-technologien.de/foerderung/bekanntmachungen/var2 [Zugriff 03.02.2025]

DIN (2023). *Mitarbeit in der Normung. Mitmachen lohnt sich!* Deutsches Institut für Normung e.V., Berlin. https://www.din.de/de/mitwirken/mitarbeit-in-der-normung [Zugriff 03.02.2025]

DKE (2023). *Welche Vorteile bringt das Engagement in der Normung?* Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE, Offenbach am Main. https://www.dke.de/de/normen-standards/bedeutung-der-normung [Zugriff 06.04.2023]

DKE (2020). *Alles rund um den Normungsprozess*, Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE. https://www.dke.de/de/normen-standards/grundlagen-der-normung/normungsprozess [Zugriff 04.02.2025]

4 Zum Zeitpunkt des Drucks: Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR).

DGQ (2023). *Normungsarbeit – wie und warum beteiligen?* Deutsche Gesellschaft für Qualität, Blog, Frankfurt am Main. https://blog.dgq.de/normungsarbeit-wie-und-warum-beteiligen/ [06.04.2023]

Gabler Wirtschaftslexikon (2025). *Blue-Ocean-Strategie,* Springer Gabler×Springer Fachmedien Wiesbaden GmbH, Wiesbaden. https://wirtschaftslexikon.gabler.de/definition/blue-ocean-strategie-120549 [Zugriff 04.02.2025]

ISO (2014). *Teaching Standards*, International Organization for Standardization, ISO Central Secretariat, Genf.

Runde, Christoph (2023). XR Standardisation – Global Overview. In: Helin, K.; Noël, F.; Schäfer, W. (Hg.): *EuroXR 2023, Proceedings of the 20th EuroXR International Conference.* Rotterdam, The Netherlands, 29. November – 1. December 2023, VTT Technology, Finland: 2024.

Runde, Christoph (2024a). The eXtended Reality & Metaverse Standards Register and Regulation Issues. In: oneM2M Partnership Project, *oneM2M Technical Plenary 63*, Bundesnetzagentur, Mainz, 28.2.2024.

Runde, Christoph (2024b). Global Monitor for XR and metaverse standards, specifications, guidelines and recommendations. In: Virtual Dimension Center (VDC), Fellbach, *XR EXPO*, Stuttgart, 4. April 2024.

Runde, Christoph (2024c). Normung frühzeitig mitdenken: Beispiel Metaverse. Wirtschaftliche Ökosysteme und Selbstpositionierung. In: Deutsches Institut für Normung DIN e.V., *Durchstarten mit Normung – zweite interaktive Konferenz zur Normungslehre*, Berlin, 25.9.2024.

Runde, Christoph (2024d). XR & Virtual Worlds Standardisation. A Global Overview. In: *EU VR/AR Industrial Coalition: Online workshop of the VR/AR Industrial Coalition*, European Commission, Brüssel, 27.9.2024.

Springer Fachmedien Wiesbaden [Hg.] (2014). Kompakt-Lexikon Wirtschaft, Springer×Gabler Wiesbaden, S. 519.

VDI (2017). *VDI-Richtlinienarbeit, Grundsätze und Anleitungen, VDI 1000,* VDI Verein Deutscher Ingenieure e.V., Düsseldorf, https://www.vdi.de/fileadmin/pages/vdi_de/redakteure/richtlinien/inhaltsverzeichnisse/2609019.pdf [Zugriff 20.03.2025]

Autorinnen und Autoren

Die genannten Institutionen beziehen sich auf den Projektzeitraum.

Amft, Oliver, Universität Freiburg, Projekt FederLeicht

André, Elisabeth, Universität Augsburg, Projekt MITHOS

Anglet, Manuel, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt MITHOS

Auchtor, Sinah, Universität des Saarlandes, Projekt MITHOS

Bareiß, Laura, Universität Stuttgart, Projekt UFO

Bhuvaneshwara, Chirag, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt MITHOS

Biermann, Antje, Universität des Saarlandes, Projekt MITHOS

Bosse, Sebastian, Fraunhofer Heinrich-Hertz-Institut, Projekt VoluProf

Botsch, Mario, Technische Universität Dortmund, Projekt HiAvA

Brandenburger, Yvonne, Fachhochschule Erfurt, Projekt XR-Part

Brandl, Christopher, Rheinisch-Westfälische Technische Hochschule Aachen/ Fraunhofer-Institut für Kommunikation, Informationsverarbeitung und Ergonomie, Projekt interaktionsKRAFT

Buchholz, Thomas, Deutsche Telekom GmbH, Projekt VoluProf

Bullinger, Angelika C., Technische Universität Chemnitz, Projekt SocialSTAGE-VR

Chehayeb, Lara, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt MITHOS

Damek, Sebastian, Fachhochschule Erfurt, Projekt XR-Part

Dhar, Arya, Laromed GmbH, Projekt ZEIT

Dietrich, Michael, Deutsches Forschungszentrum für Künstliche Intelligenz, Projekt VoluProf

Dittrich, Frank, Technische Universität Chemnitz, Projekt SocialSTAGE-VR

Drüeke, Nicola, Fraunhofer-Institut für Graphische Datenverarbeitung, Projekt NEXT-Reality

Dübel, Steve, Fraunhofer-Institut für Graphische Datenverarbeitung, Projekt NEXT-Reality

Eisert, Peter, Fraunhofer Heinrich-Hertz-Institut, Projekt VoluProf

Evers-Wölk, Michaela, IZT – Institut für Zukunftsstudien und Technologiebewertung, Projekt FederLeicht

Gärtner, Sabine, Hochschule für Angewandte Wissenschaften Hamburg, Projekt Klima-ACT!

Gebhard, Patrick, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt MITHOS

Gröschke, Sophia, Klassik Stiftung Weimar, Projekt Goethe-Live-3D

Halm, Tatjana, Referatsleiterin Recht und Digitales und Rechtsanwältin, Verbraucherzentrale Bayern

Härtel, Anja, AWO Bezirksverband Weser-Ems e. V., Projekt ZEIT

Helfrich, Maik, Hochschule für Angewandte Wissenschaften Hamburg, Projekt Klima-ACT!

Hellge, Cornelius, Fraunhofer Heinricht-Hertz-Institut, Projekt VoluProf

Hemmert, Fabian, Bergische Universität Wuppertal, Projekt ZEIT

Henze, Niels, Universität Regensburg, Projekt HapticIO

Heuten, Wilko, OFFIS – Institut für Informatik, Projekt ZEIT

Hilpert, Bernhard, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt MITHOS

Hilsmann, Anna, Fraunhofer Heinrich-Hertz-Institut, Projekt VoluProf

Holtmann, Christoph, Hochschule für Technik und Wirtschaft Berlin, Projekt INSPIRER

Irmisch, Dorothea, Semvox GmbH, Projekt MITHOS

Israel, Johann Habakuk, Hochschule für Technik und Wirtschaft Berlin, Projekt INSPIRER

Jahn, Thomas, Meilenstein Digital GmbH, Projekt NEXT-Reality

Jaschke, Tobias, Fraunhofer-Institut für Integrierte Schaltungen, Projekt SocialSTAGE-VR

Johannes, Kristina, Klassik Stiftung Weimar, Projekt Goethe-Live-3D

Josef, Andree, Mobile-Etage GmbH, Projekt SocialSTAGE-VR

Jung, Thomas, Hochschule für Technik und Wirtschaft Berlin, Projekt INSPIRER

Jürgens, Rolf, Technische Universität Dortmund, Projekt FederLeicht

Kaiser, Andreas, TriCat GmbH, Projekt MITHOS

Kalus, Alexander, Universität Regensburg, Projekt HapticIO

Kantzow, Boris, Weltenmacher GmbH, Projekt FederLeicht

Karren, Kai, Semvox GmbH, Projekt MITHOS

Keinert, Joachim, Fraunhofer-Institut für Integrierte Schaltungen, Projekt Social-STAGE-VR

Keller, Armin, Universität Rostock, Projekt VoluProf

Keppler, Sebastian, Hochschule für Technik und Wirtschaft Berlin, Projekt INSPI-RER

Kersting, Fabian, TriCat GmbH, Projekt MITHOS

Kiefer, Bernd, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt VoluProf

Kimmel, Simon, OFFIS – Institut für Informatik, Projekt ZEIT

Klasen, Martin, Uniklinik Rheinisch-Westfälische Technische Hochschule Aachen/ AIXTRA – Kompetenzzentrum für Training und Patientensicherheit, Rheinisch-Westfälische Technische Hochschule Aachen, Projekt interaktionsKRAFT

Knopf, Julia, FoBiD -– Forschungsinstitut Bildung Digital, Projekt MITHOS

Kopp, Stefan, Universität Bielefeld, Projekt HiAvA

Körner, Benjamin, Aristech GmbH, Projekt VoluProf

Krabbe, Martin, YOUSE GmbH, Projekt SocialSTAGE-VR

Krome, Niklas, Universität Bielefeld, Projekt HiAvA

Krooß, Felix, Rheinisch-Westfälische Technische Hochschule Aachen, Projekt ZEIT

Krügel, Sebastian, Technische Hochschule Ingolstadt, Projekt HapticIO

Kruse, Rolf, Fachhochschule Erfurt, Projekt XR-Part

Kulik, Alexander, Consensive GmbH, Projekt Goethe-Live-3D

Künsken, Marvin, Christian-Albrechts-Universität zu Kiel, Projekt NEXT-Reality

Lambert, Sophie, Uniklinik Rheinisch-Westfälische Technische Hochschule Aachen/AIXTRA – Kompetenzzentrum für Training und Patientensicherheit, Rheinisch-Westfälische Technische Hochschule Aachen, Projekt interaktionsKRAFT

Landwehr, Eric, OFFIS – Institut für Informatik, Projekt ZEIT

Latoschik, Marc Erich, Universität Würzburg, Projekt HiAvA

Lauinger, Nils, FoBiD – Forschungsinstitut Bildung Digital, Projekt MITHOS

Legler, Franziska, Technische Universität Chemnitz, Projekt SocialSTAGE-VR

Lingenfelser, Florian, Universität Augsburg, Projekt MITHOS

Luzsa, Robert, Universität Passau/Hochschule Landshut, Projekt GreifbAR

Mayr, Susanne, Universität Passau, Projekt GreifbAR

Mehlmann, Gregor, TriCat GmbH, Projekt MITHOS

Mertens, Alexander, Rheinisch-Westfälische Technische Hochschule Aachen, Projekt interaktionsKRAFT

Merz, Christian, Univerität Würzburg, Projekt HiAvA

Meyer, Ann-Kristin, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt MITHOS

Moghadam, Mohammadreza Maleki, Fachhochschule Erfurt, Projekt XR-Part

Nierula, Birgit, Fraunhofer Heinrich-Hertz-Institut, Projekt VoluProf

Olaya-Figueroa, Juan F., Berliner Hochschule für Technik, Projekt HapticIO

Palige, Selina, Technische Universität Chemnitz, Projekt SocialSTAGE-VR

Põder, Johann-Christian, Universität Rostock, Projekt VoluProf

Prison, Christian, Weltenmacher GmbH, Projekt FederLeicht

Probst, Andreas, Meilenstein Digital GmbH, Projekt NEXT-Reality

Queisner, Moritz, Charité – Universitätsmedizin Berlin, Projekt GreifbAR

Quenett, Sibylle, Leiterin des Referats »Interaktive Technologien für Gesundheit und Lebensqualität; Medizintechnik«, Bundesministerium für Bildung und Forschung

Rauschelbach, Anna Larissa, OffTEC Base GmbH, Projekt NEXT-Reality

Räder, Andy, Universität Rostock, Projekt VoluProf

Reipschläger, Patrick, TriCat GmbH, Projekt XR-Part

Remde, Christopher, Charité – Universitätsmedizin Berlin, Projekt GreifbAR

Ristau, Marc, auticon Deutschland GmbH, Projekt UFO

Robertini, Nadia, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt GreifbAR

Rogoll, Svenja, ISP – Institut für Stadtforschung, Planung und Kommunikation der Fachhochschule Erfurt, Projekt XR-Part

Rohden, May-Britt, AWO Bezirksverband Weser-Ems e. V., Projekt ZEIT

Romanelli, Massimo, Semvox GmbH, Projekt MITHOS

Röseler, Jasmin, Hochschule für Angewandte Wissenschaften Hamburg, Projekt Klima-ACT!

Runde, Christoph, Virtual Dimension Center w. V. (VDC) Fellbach, Projekt Living Lab – »XR-Interakt«/Projekt INSPIRER

Saphala, Addythia, Universität Freiburg, Projekt FederLeicht

Sauer, Igor, Charité – Universitätsmedizin Berlin, Projekt GreifbAR

Schimmer, Torsten, Volucap GmbH, Projekt VoluProf

Schlimm, Annika, Klassik Stiftung Weimar, Projekt Goethe-Live-3D

Schlorke, Noro, Technische Universität Dortmund, Projekt FederLeicht

Schmitz, Miriam, Technische Universität Dortmund, Projekt FederLeicht

Schneegass, Stefan, Universität Duisburg-Essen, Projekt FederLeicht

Scholz, Barbro, Hochschule für Angewandte Wissenschaften Hamburg, Projekt Klima-ACT!

Schwarten, Alina, Universität Stuttgart, Projekt UFO

Segundo Garcia Ucharima, Ruben, Semvox GmbH, Projekt MITHOS

Sinning, Heidi, ISP – Institut für Stadtforschung, Planung und Kommunikation der Fachhochschule Erfurt, Projekt XR-Part

Sobol, Yuliya, Bergische Universität Wuppertal, Projekt ZEIT

Son, Jangwoo, Fraunhofer Heinrich-Hertz-Institut, Projekt VoluProf

Sopka, Saša, Uniklinik Rheinisch-Westfälische Technische Hochschule Aachen/ AIXTRA – Kompetenzzentrum für Training und Patientensicherheit, Rheinisch-Westfälische Technische Hochschule Aachen, Projekt interaktionsKRAFT

Sowa, Timo, Semvox GmbH, Projekt MITHOS

Steinicke, Frank, Human-Computer Interaction/Fachbereich Informatik, Universität Hamburg

Steinmayr, Ricarda, Technische Universität Dortmund, Projekt FederLeicht

Stepczynski, Jan, Hochschule für Wirtschaft und Recht Berlin, Projekt Living Lab – »XR-Interakt«/Projekt INSPIRER

Stöber, Michael, Christian-Albrechts-Universität zu Kiel, Projekt NEXT-Reality

Streicher, Ferdinand, Konstruktiv GmbH, Projekt HapticIO

Tejeda Pastor, Sergio, Fraunhofer Heinrich-Hertz-Institut, Projekt VoluProf

Tevfik Lafci, Mustafa, Fraunhofer Heinrich-Hertz-Institut, Projekt VoluProf

Tröge, Jana, Hochschule für Wirtschaft und Recht Berlin, Projekt Living Lab – »XR-Interakt«

Tschanter, Jonathan, Universität Würzburg, Projekt HiAvA

Tsovaltzi, Dimitra, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Projekt MITHOS

Uhl, André, IZT – Institut für Zukunftsstudien und Technologiebewertung, Projekt FederLeicht

Uhl, Matthias, Technische Hochschule Ingolstadt, Projekt HapticIO

Vogt, Lina, Uniklinik Rheinisch-Westfälische Technische Hochschule Aachen/ AIXTRA – Kompetenzzentrum für Training und Patientensicherheit, Projekt interaktionsKRAFT

Vrontos, Apostolos, Rheinisch-Westfälische Technische Hochschule Aachen, Projekt interaktionsKRAFT

Wagner, Nicolas, Technische Universität Dortmund, Projekt HiAvA

Wansitler, Laura, Universität Stuttgart, Projekt UFO

Wegner, Juliane, Universität Rostock, Projekt VoluProf

Weiss, Thomas, Intuity Media Lab GmbH, Projekt HapticIO

Welker, Anna, Semvox GmbH, Projekt MITHOS

Wienrich, Carolin, Universität Würzburg, Projekt HiAvA

Wiesner, Heike, Hochschule für Wirtschaft und Recht Berlin, Projekt Living Lab – »XR-Interakt«/Projekt INSPIRER

Winter, Clemens, Laromed GmbH, Projekt ZEIT

Wintergrün, Dirk, Klassik Stiftung Weimar, Projekt Goethe-Live-3D

Wirzberger, Maria, Universität Stuttgart, Projekt UFO

Wischnowsky, Sven, Deutsche Telekom GmbH, Projekt VoluProf

Wittig, Nick, Universität Duisburg-Essen, Projekt FederLeicht

Wolf, Katrin, Berliner Hochschule für Technik, Projekt HapticIO

Wolter, Anja, ISP – Institut für Stadtforschung, Planung und Kommunikation der Fachhochschule Erfurt, Projekt XR-Part

Zeiler, Anne, YOUSE GmbH, Projekt SocialSTAGE-VR

Zhang, Danchen, Rheinisch-Westfälische Technische Hochschule Aachen, Projekt ZEIT