Gestaltungsprinzipien für soziale Interaktion in co-located Mixed-Reality-Tabletop-Spielen am Beispiel eines Prototyps

Julian Haider

BACHELORARBEIT

eingereicht am
Fachhochschul-Bachelorstudiengang
Medientechnik und -design

an der
Fachhochschule Oberösterreich
in Hagenberg

2026

Betreuer*in: Jeremiah Diephuis

© Copyright 2026 Julian Haider

Diese Arbeit wird unter den Bedingungen der Creative Commons Lizenz *Attribution-NonCommercial-NoDerivatives 4.0 International* (CC BY-NC-ND 4.0) veröffentlicht – siehe https://creativecommons.org/licenses/by-nc-nd/4.0/.

Erklärung

Ich erkläre eidesstattlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benutzt und die den benutzten Quellen entnommenen Stellen als solche gekennzeichnet habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt. Die vorliegende, gedruckte Arbeit ist mit dem elektronisch übermittelten Textdokument identisch.

Hagenberg, am 12. Februar 2026

Julian Haider

Inhaltsverzeichnis

Erklärung ii

Kurzfassung vi

Abstract vii

1 Einleitung 1
- 1.1 Einleitung . . . 1
- 1.2 Forschungsfrage und Hypothesen . . . 2
- 1.3 Methodisches Vorgehen . . . 2
- 1.4 Aufbau der Arbeit . . . 2

2 Technische Grundlagen für co-located MR 3
- 2.1 Mixed Reality: Definitionen und Taxonomie . . . 3
  - 2.1.1 Das Reality-Virtuality Continuum . . . 3
  - 2.1.2 Display-Technologien . . . 4
  - 2.1.3 Taxonomie von Multi-User-Systemen . . . 4
- 2.2 Inside-Out SLAM Tracking und Spatial Alignment . . . 5
  - 2.2.1 Grundlagen des visuell-inertialen Trackings . . . 5
  - 2.2.2 Shared Spatial Anchors und Multi-User Alignment . . . 6
- 2.3 Hand- und Controller-Tracking als Interaktionsmodalitäten . . . 9
  - 2.3.1 Hand-Tracking: Natürlichkeit und Embodiment . . . 9
  - 2.3.2 Controller-Input: Präzision und Zuverlässigkeit . . . 10
  - 2.3.3 Integraler Vergleich für co-located Anwendungen . . . 11
- 2.4 Die Meta Quest 3 als Referenzplattform . . . 12

3 Soziale Interaktion in MR-Tabletop-Spielen 13
- 3.1 Charakteristika klassischer Tabletop-Spiele . . . 13
- 3.2 Hybride Tabletop-Systeme: Digitale Transformation . . . 14
- 3.3 Einflussfaktoren der Interaktion in co-located MR . . . 15
  - 3.3.1 Co-Presence und physisch-virtuelle Konsistenz . . . 15
  - 3.3.2 Soziale Präsenz und Qualität der Interaktion . . . 15
  - 3.3.3 Digital Proxemics . . . 16
  - 3.3.4 Gaze Awareness und Joint Attention . . . 16
  - 3.3.5 Räumliche und visuelle Wahrnehmungsbarrieren . . . 17
- 3.4 Ableitung theoretischer Gestaltungsanforderungen . . . 20

3.4.1 Multimodales Feedback . . . 20
3.4.2 Visualisierung sozialer Cues . . . 21
3.4.3 Symmetrische Interaktionsgestaltung . . . 21
3.4.4 Reduktion der kognitiven Belastung . . . 22

4 Stand der Forschung und Technik 24
4.1 Aktuelle Forschungsansätze zu MR-Co-Experience . . . 24
4.2 Marktanalyse: Tabletop in AR und MR . . . 24
4.2.1 Optische AR-Systeme und projektionsbasierte Ansätze . . . 24
4.2.2 HMD-basierte Mixed-Reality-Systeme . . . 25
4.3 Frameworks und Entwicklungsumgebungen für Colocation . . . 26
4.4 Positionierung der Arbeit . . . 26

5 Konzeption und Prototypentwicklung 28
5.1 Spielkonzept und Designentscheidungen . . . 28
5.1.1 Spielmaterial und virtuelle Komponenten . . . 29
5.1.2 Spielablauf und Aktionsmechanik . . . 30
5.1.3 Spielwelt und Narrative . . . 31
5.2 MR-spezifische Funktionen und Interaktionsdesign . . . 32
5.2.1 Avatar-System für soziale Präsenz . . . 32
5.2.2 Input-Modalitäten: Controller und Hand-Tracking . . . 33
5.2.3 Multimodales Feedback . . . 34
5.3 Technische Architektur und Implementierung . . . 34
5.4 Herausforderungen der co-located MR-Entwicklung . . . 37

6 Empirische Evaluation 38
6.1 Studiendesign und Durchführung . . . 38
6.1.1 Experimentelle Bedingungen . . . 38
6.1.2 Durchführung . . . 39
6.2 Datenanalyse und Messinstrumente . . . 39
6.2.1 Operationalisierung der Hypothesen . . . 39
6.2.2 Analysemethoden . . . 40
6.3 Ergebnisse der empirischen Studie . . . 41
6.3.1 Soziale Präsenz . . . 41
6.3.2 Wahrnehmung der Avatare . . . 42
6.3.3 Usability und Eingabemethode . . . 43
6.4 Methodische Reflexion und Limitationen . . . 43

7 Diskussion der Ergebnisse 44
7.1 Interpretation der sozialen Präsenz . . . 44
7.2 Vergleichende Analyse der Eingabemodalitäten . . . 45
7.3 Akzeptanz virtueller Repräsentationen . . . 46
7.4 Akzeptanz und Zukunftspotenzial aus Nutzersicht . . . 46

8 Designleitfaden für co-located MR-Tabletop-Spiele 48
8.1 Strategien für soziale Präsenz und Repräsentation . . . 48
8.2 Interaktionsmodalitäten und Verkörperung . . . 49

8.3 Synchronisation und Umgang mit dem physischen Raum 50
8.4 Kognitive Entlastung und Hybridität 50
8.5 Atmosphäre und Feedback-Design 51
8.6 Narrative Kontextualisierung 51

9 Fazit und Ausblick **53**

A Fragebogen **54**

Quellenverzeichnis **64**
Literatur 64
Medien 69
Software 69
Online-Quellen 69

Kurzfassung

Analoge Brettspiele sind ein zentrales Medium sozialer Interaktion, geprägt durch physische Präsenz und nonverbale Kommunikation. Die Übertragung in co-located Mixed Reality (MR) verbindet physische und virtuelle Elemente, stellt Entwickler*innen jedoch vor die Herausforderung, soziale Isolation durch Head-Mounted Displays (HMDs) zu vermeiden. Die Arbeit untersucht, wie Gestaltungsprinzipien in MR-Tabletop-Spielen soziale Interaktion und soziale Präsenz fördern können. Hierfür wurde der kooperative Prototyp *Depthbound* für die *Meta Quest 3* entwickelt, der Shared Spatial Anchors und Video-Passthrough nutzt. Im Rahmen einer empirischen Studie wurde anschließend evaluiert, wie sich unterschiedliche Interaktionsmodalitäten und Visualisierungen im Vergleich zum analogen Vorbild auf das soziale Erleben auswirken. Basierend auf den gewonnenen Erkenntnissen leitet die Arbeit einen Designleitfaden ab, der Strategien zur technischen und gestalterischen Optimierung des gemeinsamen Spielerlebens in hybriden Umgebungen aufzeigt.

Abstract

Analog board games are a central medium of social interaction, characterized by physical presence and nonverbal communication. Transferring them into co-located Mixed Reality (MR) combines physical and virtual elements but poses challenges for developers to avoid social isolation caused by Head-Mounted Displays (HMDs). This thesis investigates how design principles in MR tabletop games can promote social interaction and social presence. For this purpose, the cooperative prototype *Depthbound* was developed for the *Meta Quest 3*, utilizing shared spatial anchors and video passthrough. In an empirical study, the effects of different interaction modalities and visualizations on the social experience were evaluated in comparison to the analog original. Based on the findings, this work derives a design guide that outlines strategies for the technical and design-oriented optimization of shared gameplay in hybrid environments.

Kapitel 1

Einleitung

1.1 Einleitung

Analoge Brett- und Kartenspiele fungieren seit jeher als zentrale Medien sozialer Interaktion. Im Gegensatz zu vielen rein digitalen Videospielen zeichnen sie sich durch die Unmittelbarkeit des gemeinsamen Erlebens und die physische Präsenz der Spielenden aus [36]. Für fast 60% der Hobby-Spieler*innen stellt die soziale Interaktion den primären Motivationsfaktor dar, noch vor strategischen oder intellektuellen Herausforderungen [52].

Diese besondere soziale Qualität basiert maßgeblich auf der Greifbarkeit der Spielkomponenten sowie der räumlichen Konstellation der Beteiligten. Insbesondere in co-located Tabletop-Spielen konstituiert der Spieltisch einen gemeinsamen, öffentlichen Informationsraum, der eine gegenseitige Wahrnehmung ermöglicht. Die gemeinsame räumliche Situation schafft dabei Voraussetzungen dafür, dass nonverbale Informationskanäle wie Augenkontakt, Gestik und Mimik kontinuierlich wahrgenommen und in die soziale Interpretation des Gegenübers einbezogen werden können [36, 37].

Mit dem Aufkommen der Mixed Reality (MR) eröffnet sich ein neues Feld für hybride Spielerlebnisse, das dieses Prinzip auf co-located Mixed-Reality-Tabletop-Spiele überträgt und die sozialen Qualitäten der analogen Welt mit der Rechenleistung digitaler Systeme verknüpfen möchte. Die Nutzung von Head-Mounted Displays (HMDs) droht die natürliche soziale Interaktion zu beeinträchtigen und eine soziale Isolation zu begünstigen. Während bei klassischen hybriden Ansätzen lediglich der Blick vom Brett abgewendet wird, riskieren HMDs durch die direkte Überlagerung des Sichtfelds eine deutlich stärkere Barriere zwischen den Spielenden. Daraus ergibt sich für die Spieleentwicklung die Aufgabe, virtuelle Inhalte so zu gestalten und einzubetten, dass die verwendete Hardware nicht als trennendes Element wirkt, sondern die soziale Verbundenheit der Spielenden am Tisch möglichst erhält oder unterstützt.

Vor diesem Hintergrund untersucht die vorliegende Arbeit Gestaltungsprinzipien für soziale Interaktion in co-located Mixed-Reality-Tabletop-Spielen. Am Beispiel des entwickelten Prototyps *Depthbound* wird analysiert, wie gezielte Designentscheidungen die technologische Schnittstelle optimieren können. Ziel ist es, Lösungen aufzuzeigen, die den sozialen Austausch fördern und die soziale Präsenz am hybriden Spieltisch aktiv unterstützen.

1.2 Forschungsfrage und Hypothesen

Die vorliegende Arbeit untersucht das Spannungsfeld zwischen technologischer Immersion und zwischenmenschlicher Verbundenheit in co-located MR-Tabletop-Szenarien. Im Zentrum steht dabei die folgende Forschungsfrage:

Welche Gestaltungsprinzipien fördern in co-located Mixed-Reality-Tabletop-Spielen die soziale Interaktion und soziale Präsenz, um die Qualitäten des analogen Miteinanders erfolgreich in den hybriden Raum zu übertragen?

Aus der zentralen Forschungsfrage und den methodischen Zielen ergeben sich zwei Kernhypothesen, die im Rahmen der empirischen Evaluation überprüft werden:

- **H1:** Räumlich verankerte Avatare mit animierten Kopf- und Handbewegungen erhöhen die soziale Präsenz, fördern gemeinsame Aufmerksamkeit und verbessern die Verständlichkeit nonverbaler Signale in co-located MR-Tabletop-Spielen.
- **H2:** Hand-Tracking verstärkt das Erleben verkörperter Interaktion und wird gegenüber Controller-Eingaben als natürlichere Eingabemodalität wahrgenommen.

1.3 Methodisches Vorgehen

Ziel dieser Arbeit ist die Konzeption und anschließende Prototypentwicklung eines kooperativen MR-Tabletop-Spiels für drei Personen. Als Referenzplattform dient die *Meta Quest 3*. Durch die Implementierung gezielter Designprinzipien soll das Spiel die geteilte Präsenz und soziale Interaktion zwischen den Spieler*innen steigern, indem virtuelle Objekte kohärent in den physischen Raum integriert und nonverbale Kommunikationskanäle wie Gesten und gemeinsame Fokussierungen unterstützt werden.

Die Forschungsfrage wird auf Basis von drei methodischen Säulen beantwortet. Zunächst erfolgt eine Literaturanalyse zur systematischen Auswertung aktueller Studien über kollaborative Mixed-Reality-Umgebungen. Darauf aufbauend wird ein technisch funktionsfähiger Prototyp eines MR-Tabletop-Spiels in *Unity* entwickelt. Abschließend wird eine empirische Evaluation mittels einer Studie durchgeführt, die quantitative Leistungsdaten mit qualitativen Daten kombiniert.

1.4 Aufbau der Arbeit

Kapitel 2 erläutert die technologischen Grundlagen von MR-Systemen. Kapitel 3 entwickelt den theoretischen Rahmen zu sozialer Interaktion, sozialer Präsenz, *Embodied Interaction* und Shared Attention in MR-Tabletop-Spielen. Kapitel 4 analysiert den aktuellen Forschungsstand. Kapitel 5 beschreibt Konzeption und Prototypentwicklung. Kapitel 6 präsentiert das Studiendesign sowie die Ergebnisse der empirischen Evaluation. Darauf aufbauend interpretiert Kapitel 7 die Befunde. Kapitel 8 fasst praxisorientierte Designleitlinien zusammen, und Kapitel 9 schließt mit einem Fazit und einem Ausblick auf zukünftige Forschung und Praxis.

Kapitel 2

Technische Grundlagen für co-located MR

2.1 Mixed Reality: Definitionen und Taxonomie

2.1.1 Das Reality-Virtuality Continuum

Das Verständnis von *Mixed Reality* (MR) basiert klassischerweise auf dem von Milgram und Kishino [27] eingeführten *Reality-Virtuality Continuum* (RV-Kontinuum), das ein Spektrum zwischen vollständig realen und vollständig virtuellen Umgebungen beschreibt. Mixed Reality fungiert dabei als generischer Oberbegriff für alle Umgebungen, in denen reale und virtuelle Objekte innerhalb einer Darstellung gemeinsam präsentiert werden.

Entlang dieses Kontinuums werden primär zwei Ausprägungen unterschieden. *Augmented Reality* (AR) ergänzt die reale Welt um virtuelle Objekte, während *Augmented Virtuality* (AV) reale Elemente in eine vorwiegend virtuelle Umgebung integriert [27].

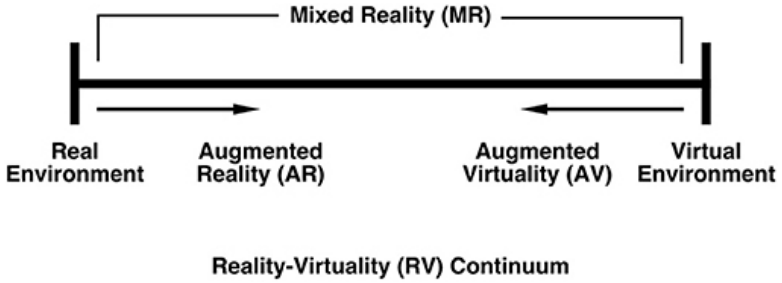


Abbildung 2.1: Reality-Virtuality Continuum nach Milgram und Kishino [27].

Moderne Head-Mounted Displays (HMDs) wie die in dieser Arbeit verwendete *Meta Quest 3* nutzen technologisch bedingt einen *Video-Passthrough*-Ansatz. Im Kontext dieser Arbeit und gemäß der Definition des Kontinuums werden solche hybriden Anwendungen unter dem Begriff Mixed Reality zusammengefasst.

2.1.2 Display-Technologien

Für die Realisierung von Mixed Reality haben sich zwei fundamentale Display-Technologien etabliert, die die Wahrnehmung der physischen Umgebung unterschiedlich lösen:

- **Optical See-Through (OST):** Systeme wie die *Microsoft HoloLens* nutzen halbtransparente Spiegel oder transparente Linsen, um digitale Inhalte direkt in das optische Sichtfeld einzublenden, wobei die nutzende Person die reale Welt unmittelbar durch das Glas sieht. Ein technischer Nachteil ist jedoch die additive Lichtmischung. Da Licht physikalisch nicht subtrahiert werden kann, erscheinen virtuelle Objekte oft halbtransparent und können die reale Welt nicht vollständig verdecken [1, 38].
- **Video-Passthrough (VST):** Geräte wie die *Meta Quest 3* oder *Apple Vision Pro* nehmen die Außenwelt über Kameras auf und stellen sie als Videostrom auf internen Bildschirmen dar. Dies ermöglicht die vollständige Kontrolle über jeden Pixel, sodass virtuelle Objekte die Realität vollständig verdecken können. Kritische Faktoren sind jedoch die Latenz sowie visuelle Verzerrungen, die zu einer Unterschätzung von Distanzen oder durch den mediatisierten Blickkontakt zu einer reduzierten sozialen Präsenz führen können [1, 38, 40].

2.1.3 Taxonomie von Multi-User-Systemen

Während das RV-Kontinuum den Grad der Virtualität beschreibt, erfordert die technische Konzeption kollaborativer Anwendungen eine Klassifizierung anhand der räumlichen Konstellation. Für kollaborative MR-Anwendungen ist die Unterscheidung zwischen remote Kollaboration und co-located Kollaboration von zentraler Bedeutung, da sie fundamentale Auswirkungen auf die technische Architektur sowie die Anforderungen an das Tracking und die Latenzbewältigung hat [13].

Remote vs. co-located Kollaboration

Remote Szenarien beschreiben Situationen, in denen sich die Teilnehmenden an unterschiedlichen physischen Orten befinden. Die primäre technische Herausforderung liegt hier in der Netzwerklatenz und der Replikation von Zustandsdaten, da Verzögerungen die gegenseitige Wahrnehmung und Aufgabenleistung beeinträchtigen können. Da die physischen Umgebungen der Nutzer*innen voneinander getrennt sind, ist eine exakte räumliche Übereinstimmung der realen Räume meist nicht erforderlich [13, 32].

In co-located Szenarien, dem Fokus dieser Arbeit, teilen sich die Nutzenden denselben physischen Raum. Hier verschiebt sich die technische Priorität von der reinen Latenzkompensation hin zur präzisen räumlichen Synchronisation. Die Glaubwürdigkeit der Interaktion und die Vermeidung von *Cybersickness*[1] hängen maßgeblich davon ab, ob virtuelle Objekte für alle Beteiligten an exakt derselben physischen Position verankert erscheinen und räumliche Konsistenz gewahrt bleibt [32, 42].

[1] Unwohlsein durch immersive Displays, ausgelöst durch Diskrepanzen zwischen gesehener Bewegung und dem Gleichgewichtssinn.

Individual Space vs. Shared Space

Eine weitere notwendige technische Unterscheidung betrifft das Koordinatensystem, in dem die Anwendung operiert. In der Literatur wird hierbei, basierend auf der Unterscheidung zwischen lokalen und globalen Koordinatensystemen [42], zwischen individuellen und geteilten Räumen differenziert:

- **Individual Space (Lokales System):** Jedes HMD initialisiert beim Start eine eigene, lokale Karte, die nicht inhärent mit anderen Geräten geteilt wird. Der Ursprung $(0, 0, 0)$ liegt typischerweise an der Position und Ausrichtung, an der das Gerät eingeschaltet wurde. In einem Multi-User-Szenario führt dies dazu, dass identische Koordinaten $P(x, y, z)$ auf verschiedenen Geräten völlig unterschiedlichen physischen Positionen entsprechen. Ohne Abgleich bleiben andere Nutzende im virtuellen Raum daher fehlausgerichtet oder unsichtbar [32].
- **Shared Space (Globales System):** Um eine kohärente Interaktion im selben physischen Bereich zu ermöglichen, müssen diese lokale Koordinatensysteme in ein gemeinsames globales Referenzsystem überführt werden [13]. Alle Geräte müssen sich auf einen gemeinsamen Ursprung und eine gemeinsame Ausrichtung einigen, damit virtuelle Objekte für alle Nutzenden geometrisch korrekt und deckungsgleich mit der physischen Welt erscheinen [32, 42].

Die zentrale technische Herausforderung bei der Entwicklung von co-located MR-Tabletop-Spielen besteht darin, diesen *Shared Space* herzustellen. Da moderne HMDs meist *Inside-Out-Tracking* via VI-SLAM nutzen, ist ihre Lokalisation zunächst lokal und ohne Kenntnis über die relative Position zu anderen Geräten [32]. Um diese isolierten Karten bei Standalone-Geräten zu einer kohärenten Umgebung zu fusionieren, kommen häufig sogenannte *Shared Spatial Anchors* zum Einsatz, welche im folgenden Kapitel detailliert erläutert werden.

2.2 Inside-Out SLAM Tracking und Spatial Alignment

2.2.1 Grundlagen des visuell-inertialen Trackings

Moderne Consumer-MR-Systeme setzen auf *Inside-Out Tracking*, bei dem Kameras und *Inertial Measurement Units* (IMU) direkt am HMD integriert sind [26, 32]. Das System schätzt kontinuierlich seine Pose (Position x, y, z und Orientierung Roll, Pitch, Yaw) relativ zur Umgebung. Dieses *6 Degrees of Freedom (6DoF)*-Tracking basiert auf *Visual-Inertial Simultaneous Localization and Mapping (VI-SLAM)* [32].

Der Algorithmus extrahiert visuelle Features aus Kamerabildern, verfolgt deren Bewegung zwischen Frames und fusioniert diese mit IMU-Daten für hochfrequente Pose-Updates [26, 49]. Das System erstellt dabei kontinuierlich eine Karte der Umgebung und lokalisiert sich darin [7]. Trotz dieser Selbstlokalisation bleibt die Drift-Akkumulation eine Herausforderung, da sich kleine Schätzfehler über die Zeit summieren [8].

Für co-located MR-Anwendungen sind sowohl die absolute Positionsgenauigkeit über mittlere Distanzen als auch die lokale Stabilität bei feinmotorischen Interaktionen als komplementäre Genauigkeitsdimensionen von zentraler Bedeutung. Aktuelle Benchmark-Studien zur *Meta Quest 3* zeigen, dass das System bei geschlossenen Pfaden von ungefähr 50 Metern durchschnittliche *Loop Errors* von 8,4 cm ($\pm$ 4,5 cm) auf-

weist [8]. Dies bestätigt die Effektivität der *Loop-Closure*-Mechanismen[2], die bei wiederholter Erkennung derselben Umgebungsmerkmale die Trajektorien-Drift korrigieren.

Entscheidender für Tabletop-Szenarien ist jedoch die lokale Präzision bei begrenzten Bewegungsradien. Evaluationen von co-located Multi-User-Systemen mit *Meta Quest 2* innerhalb eines 7×7 Meter-Bereichs belegen *Absolute Trajectory Errors* von 3,1 cm bis 4,9 cm bei verschiedenen Bewegungsmustern. Bei direkten Nahbereichs-Interaktionen zwischen mehreren Nutzenden, wie sie beim gemeinsamen Manipulieren von Spielelementen auf einem physischen Tisch auftreten, liegt die durchschnittliche Abweichung bei etwa 5,2 cm. Diese Genauigkeit ermöglicht glaubwürdige soziale Interaktionen, bei denen Nutzende beispielsweise gemeinsam auf virtuelle Karten zeigen oder Spielfiguren auf entsprechend skalierten Feldern platzieren. Alle durchgeführten Fist-Bump-Interaktionen zwischen co-located Nutzenden waren trotz der Abweichungen erfolgreich durchführbar, was die praktische Eignung für kollaborative Tabletop-Szenarien unterstreicht [32].

Die temporale Konsistenz der Pose-Schätzung ist dabei ebenso kritisch wie die absolute Genauigkeit. Visuelle Artefakte wie das scheinbare „Schwimmen“ virtueller Objekte auf Tischoberflächen entstehen durch hochfrequente Mikro-Oszillationen (*Jitter*) in der geschätzten HMD-Pose [32]. Die Fusion von visuellen und inertialen Messungen stabilisiert die Pose-Schätzung zwischen den diskreten Kamera-Frames und reduziert so diese visuellen Inkonsistenzen erheblich. Für stationäre Tabletop-Szenarien bedeutet dies, dass virtuelle Objekte räumlich stabiler erscheinen, selbst wenn die nutzende Person minimale Kopfbewegungen ausführt oder sich leicht vorbeugt, um Spielelemente genauer zu betrachten.

2.2.2 Shared Spatial Anchors und Multi-User Alignment

Für co-located MR-Tabletop-Szenarien müssen mehrere Inside-Out SLAM-Systeme in einen gemeinsamen räumlichen Referenzrahmen ausgerichtet werden. *Meta Quest* realisiert dieses *Multi-User Alignment* über *Spatial Anchors* [82].

Funktionsweise von Spatial Anchors

Ein Spatial Anchor fungiert als ortsfester Referenzrahmen, der virtuellen Objekten eine dauerhafte Position und Orientierung (6DoF-Pose) in der physischen Welt zuweist. Der technische Lebenszyklus eines Anchors lässt sich in fünf fundamentale Operationen unterteilen [78]:

1. **Create:** Zunächst wird ein Anchor an einer spezifischen Position im Raum instanziiert. Das System verknüpft dabei die virtuellen Koordinaten mit markanten visuellen Merkmalen der physischen Umgebung (Feature Points), um den Punkt trackbar zu machen.
2. **Persist:** Um die Nutzung über die aktuelle Laufzeit hinaus zu ermöglichen, kann der Anchor persistiert werden. Dabei werden die räumlichen Daten serialisiert und entweder lokal auf dem Gerät oder in der Cloud gespeichert.
3. **Share:** Für kollaborative Szenarien (Colocation) wird der persistierte Anchor mit anderen Geräten geteilt. Durch den Austausch der Anchor-ID erhalten co-located

[2] Ein Verfahren, das erkennt, wenn das Gerät an einen bereits besuchten Ort zurückkehrt und akkumulierte Fehler in der Karte korrigiert [7].

Nutzer*innen Zugriff auf denselben räumlichen Referenzpunkt, was ein synchronisiertes Erlebnis ermöglicht.

4. **Query:** Um gespeicherte oder geteilte Anchors wiederherzustellen, führt die Anwendung eine Abfrage durch. Das System sucht anhand von Filterkriterien (z. B. UUIDs oder räumliche Nähe) nach den entsprechenden Ankern und lokalisiert diese in der aktuellen Umgebung.
5. **Destroy:** Nicht mehr benötigte Anchors können explizit gelöscht werden, um Speicherressourcen freizugeben und die Liste der abfragbaren Referenzpunkte zu bereinigen.

Genauigkeits-Charakteristik

Die erreichbare Präzision des Multi-User Alignment wird stark von den spezifischen räumlichen Gegebenheiten beeinflusst. Die höchste Zuverlässigkeit wird in Umgebungen erreicht, die reich an visuellen Merkmalen wie Möbel, Bilder oder strukturierte Wände sind und eine gleichmäßige, intensive Beleuchtung aufweisen. Wenn die Umgebung nur mäßig strukturiert ist und die Lichtverhältnisse wechselhaft sind, verringert sich die erreichbare Genauigkeit. In ungünstigen Situationen, wie sie bei leeren, hellen Wänden oder extrem schwacher Beleuchtung vorliegen, kann die Ausrichtung stark beeinträchtigt oder sogar ganz unmöglich werden [82].

Die Zuverlässigkeit der Verankerung hängt vorrangig von den folgenden kritischen Umweltfaktoren ab, die für Spatial-Anchor-Systeme relevant sind [82]:

1. **Dichte visueller Merkmale:** Das Mapping ist in Räumen mit vielen unterschiedlichen Oberflächenmerkmalen präziser. Eine geringe Texturierung erschwert das stabile Tracking erheblich.
2. **Lichtverhältnisse:** Eine unzureichende Lichtstärke beeinträchtigt die Erkennung von Features drastisch. Auch schnelle, signifikante Änderungen der Beleuchtung können den Trackingprozess stören.
3. **Bewegung im Raum:** Sich verlagernde Objekte oder Personen können essenzielle Referenzpunkte temporär verdecken oder überlagern, was zu einem fehlerhaften Korrespondenzaufbau führt.
4. **Strategische Platzierung der Ankerpunkte:** Für eine maximale Stabilität sollten die Anker an Stellen positioniert werden, die über eine besonders hohe und stabile visuelle Textur verfügen.

Die praktische Umsetzung dieser Prinzipien wird in Abbildung 2.2 verdeutlicht. Das Szenario zeigt, wie Spatial Anchors genutzt werden können, um heterogene virtuelle Inhalte von zweidimensionalen Medienpanels wie Wetteranzeigen oder Videoplayern bis hin zu interaktiven dreidimensionalen Objekten wie Spielzeug oder Pflanzen persistent im physischen Raum zu verorten. Ein wesentlicher Vorteil besteht dabei in der Wiederverwendbarkeit der Anker. Wie in der Darstellung ersichtlich, muss nicht für jedes virtuelle Objekt ein eigener Anker erstellt werden. Stattdessen können vier strategisch platzierte Anker als Referenzsystem für eine Vielzahl von Inhalten dienen. Dies optimiert sowohl die Ressourcennutzung als auch das Management der Szene [78].

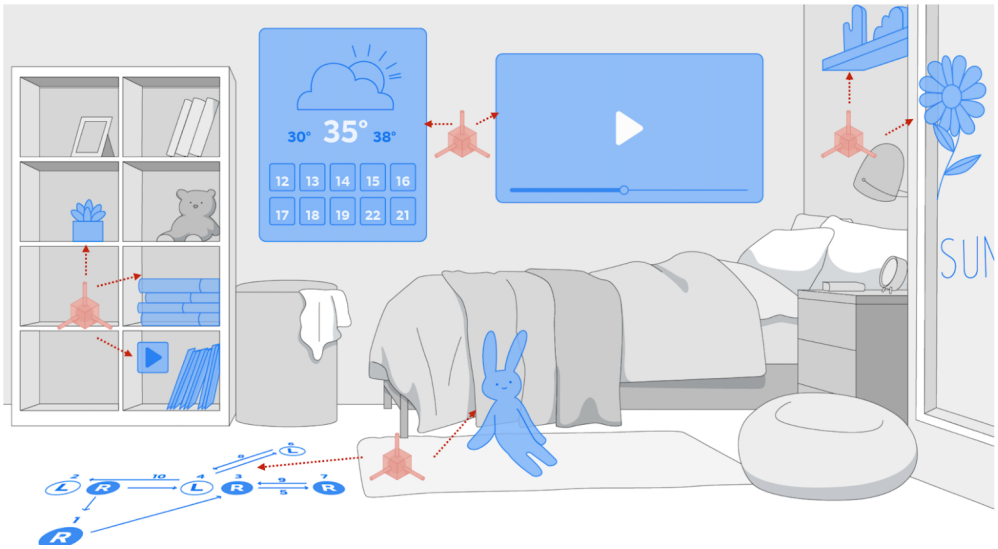


Abbildung 2.2: Exemplarische Anwendung von vier Spatial Anchors zur Verankerung verschiedener virtueller Inhalte in einer Wohnumgebung [78].

Strategien zur Multi-User-Colocation

Die technische Realisierung präziser Multi-User-Synchronisation erfolgt heute primär durch zwei komplementäre Ansätze, die unterschiedliche Anforderungen an Infrastruktur und Umgebungsbedingungen stellen [28]:

1. **Cloud-basierte Spatial Anchors (markerlos):** Kommerzielle Systeme wie Meta Shared Spatial Anchors nutzen SLAM-Algorithmen zur Feature-Erkennung und synchronisieren räumliche Kartendaten über Cloud-Server. Dies ermöglicht plattformübergreifende Persistenz und globale Zugänglichkeit. Der Ansatz erfordert jedoch stabile Internet-Konnektivität und wirft Datenschutzfragen auf. Zudem bleibt die Genauigkeit von der Qualität der Umgebungstextur abhängig, da das zugrunde liegende SLAM-Tracking in strukturarmen Räumen zu Drift-Akkumulation neigt [28, 32].
2. **Fiducial Markers (lokale Referenzpunkte):** Als lokale Alternative dienen gedruckte optische Marker wie AprilTags zur absoluten Positionsbestimmung ohne Cloud-Abhängigkeit. Sie ermöglichen eine schnelle Kalibrierung, wobei sich die Marker-Erkennung typischerweise innerhalb von 1,7 bis 3 Sekunden stabilisiert. Dieser Ansatz ist robust gegen wechselnde Lichtverhältnisse und ermöglicht eine präzise Initialisierung auch in texturarmen Umgebungen, wenngleich das fortlaufende Tracking nach der Kalibrierung weiterhin auf die internen Sensoren angewiesen ist. Er ist kosteneffizient und wahrt die Privatsphäre, da keine Daten an externe Server gesendet werden müssen [28].

2.3 Hand- und Controller-Tracking als Interaktionsmodalitäten

Die Wahl der Eingabemethode ist entscheidend für das Immersionsempfinden und die Usability in Mixed-Reality-Anwendungen. Papadopoulos et al. [31] klassifizieren Interaktionen in MR-Umgebungen anhand verschiedener Modalitäten, darunter visuelle, auditive und haptische Eingaben. Während multimodale Ansätze in der Forschung oft als ideal gelten, dominieren bei aktuellen Consumer-HMDs wie der *Meta Quest 3* zwei primäre Eingabemethoden:

2.3.1 Hand-Tracking: Natürlichkeit und Embodiment

Das kamerabasierte Hand-Tracking nutzt die Infrarot-Kameras des HMDs, um die Hände der nutzenden Person zu erfassen. Durch Machine-Learning-Modelle wird in Echtzeit die Pose eines detaillierten Skeletts mit 26 Gelenken mit einer typischen Bildwiederholrate von 60 Hz geschätzt [74]. Die 26 Gelenke sind in Abbildung 2.3 dargestellt und umfassen jeweils vier Gelenke pro Daumen, fünf Gelenke pro Finger sowie zusätzliche Gelenke für Handgelenk und Handfläche gemäß dem OpenXR-Standard.

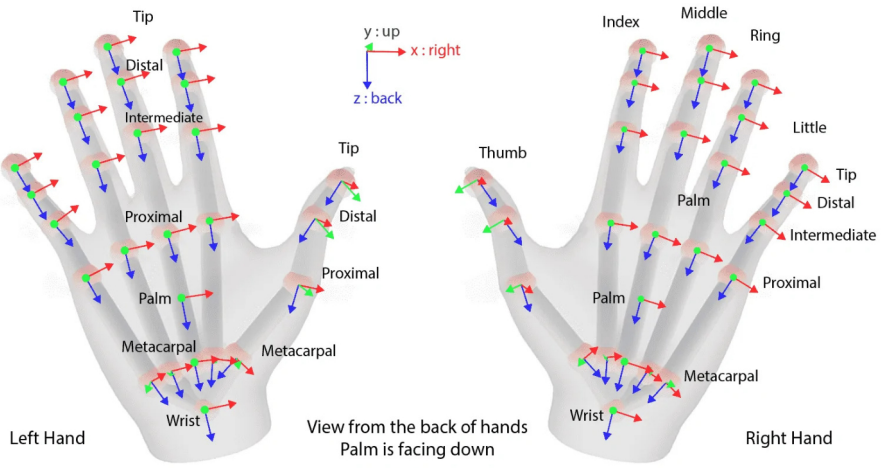


Abbildung 2.3: OpenXR-konformes Hand-Tracking-Skelett mit 26 Joints für Meta Quest [81].

Vorteile: Intuition und soziale Präsenz

Die Interaktion mittels Hand-Tracking gilt als natürlich und intuitiv, da sie direkte Gesten wie Greifen und Zeigen ermöglicht, was die Immersion steigert und die Kommunikation durch Gestik unterstützt. Da die Hände der nutzenden Person ohne störende Hardware visualisiert werden, kann dies das Gefühl des Embodiments fördern [6]. Ein weiterer Vorteil zeigt sich in Mixed-Reality-Szenarien. Die gleichzeitige Sichtbarkeit von realen und virtuellen Inhalten kann dazu führen, dass Nutzer*innen virtuelle Hände eher als ihre eigenen akzeptieren, was die kognitive Verschmelzung von realer und virtueller Welt begünstigt [15].

Aktuelle Studien bestätigen zudem, dass spezifische Manipulationen, wie die dreidimensionale Rotation von Objekten, von Nutzer*innen mit Hand-Tracking oft als einfa-

cher und natürlicher empfunden werden als mit Controllern, da die Bewegungsfreiheit nicht durch die Geräteform eingeschränkt ist [46]. Besonders bei der Manipulation virtueller Objekte spielt die Art der Greifgeste eine wesentliche Rolle für die wahrgenommene Natürlichkeit. Jude et al. untersuchten die Präferenzen der Nutzer*innen bei gestischen Interaktionen und differenzierten dabei zwischen *Grasp*, *Grab* und *Pinch*. Ihre Ergebnisse zeigten eine deutliche Präferenz für den *Grasp*, da dieser Zeigegesten ähnelt und so eine natürliche Interaktion unterstützt [22]. Für filigrane Objekte wird in aktuellen Anwendungen hingegen oft der *Pinch* eingesetzt, da er eine präzise Auswahl kleiner Elemente ermöglicht [46].

Herausforderungen: Robustheit und Stabilität

Trotz technologischer Fortschritte leidet das Hand-Tracking unter spezifischen Limitationen. Grundsätzlich ist es anfällig für Selbstverdeckung oder Tracking-Verluste, sobald die Hände das Sichtfeld der Kameras verlassen [6].

Für co-located Szenarien erweitert Reimer [34] diese Problematik um den Aspekt der gegenseitigen Verdeckung. Hände können sowohl durch physische Objekte und den eigenen Körper als auch durch Mitspielende verdeckt werden, wodurch die Sichtlinie der HMD-Sensoren unterbrochen wird. Besonders in Tabletop-Situationen, in denen Nutzer*innen räumlich nah beieinander agieren, wird die korrekte Unterscheidung und Zuordnung der Hände erschwert. Während die Zuweisung virtueller Hände bei großen Abständen unproblematisch ist, beschreibt Reimer die Handzuordnung in enger räumlicher Nähe, definiert als ein Bereich innerhalb einer Armlänge, als nicht trivial. In diesen Situationen, in denen sich die Hände verschiedener Personen im selben Erfassungsbereich befinden oder überlappen, hat das System Schwierigkeiten, die extrahierten Merkmale korrekt der jeweiligen Person zuzuordnen, was zu einer erhöhten Fehleranfälligkeit führt [34].

Zusätzlich zur räumlichen Konstellation hängt die Genauigkeit stark von Umweltfaktoren wie den Lichtverhältnissen ab, die die Datenerfassung beeinträchtigen können [11]. Im direkten Vergleich wird die optische Handerkennung daher oft als unpräziser empfunden als die Nutzung physischer Controller [6]. Ein spezifisches Problem stellt dabei die Erkennung des Loslassens dar. Nutzer*innen beschreiben virtuelle Objekte beim Hand-Tracking oft als „klebrig“, da das System das Öffnen der Hand erst mit einer leichten Verzögerung als Befehl interpretiert, was die präzise Platzierung von Spielelementen auf dem Tisch erschwert [41].

2.3.2 Controller-Input: Präzision und Zuverlässigkeit

Dedizierte drahtlose Controller bieten eine hochgradig robuste Interaktionsgrundlage. Sie werden mittels Inside-Out-Tracking im Raum lokalisiert und erfordern keine manuelle Kalibrierung [94].

Vorteile: Deterministische Leistung

Der Hauptvorteil von Controllern liegt in der extremen Stabilität und Präzision der Positions- und Rotationserfassung [6]. Controller gewährleisten eine zuverlässige Interaktion auch unter ungünstigen Bedingungen, da sie auf optische Tracking-Marker setzen,

die von den Kameras des Systems erfasst werden [11]. Messungen von Warburton et al. zeigen, dass die Motion-to-Photon-Latenz[3] moderner VR-Controller wie *Oculus Rift*, *Rift S*, *HTC Vive* und *Valve Index* zu Bewegungsbeginn im Bereich von etwa 21–42 ms liegt und nach dem Einsetzen der integrierten Motion-Prediction während kontinuierlicher Bewegungen auf ungefähr 2–13 ms reduziert wird [50].

Markerlose Hand-Tracking-Systeme weisen demgegenüber deutlich variablere Motion-to-Photon-Latenzen auf. Robotergestützte Messungen für die *Meta Quest 3* zeigen Werte zwischen 14,4 ms und 220,5 ms [18]. In vergleichenden Studien zur Greifpräzision schnitten Controller daher signifikant besser ab als verschiedene Hand-Gesten, sowohl in Bezug auf die Geschwindigkeit als auch auf die Platzierungsgenauigkeit [41]. Steed und Lai [46] zeigten, dass Controller eine „Trial-and-Error"-Strategie begünstigen, bei der Nutzer*innen signifikant mehr Objektplatzierungen pro Minute vornehmen als im Hand-Tracking-Modus, da der Greif- und Loslass-Mechanismus per Tastendruck schneller und fehlerfreier ausführbar ist.

Einschränkungen: Indirekte Interaktion und Abstraktion

Die Nutzung von Controllern vermittelt jedoch die Interaktion. Das Drücken von Knöpfen ist weniger intuitiv als direkte Gesten und kann das Gefühl des Embodiments verringern, da keine direkte Korrespondenz zwischen der eigenen Hand und der virtuellen Darstellung besteht [6]. Zwar bieten Controller durch ihre Masse und Vibration haptisches Feedback, dieses ist jedoch oft symbolisch und erreicht nicht die Qualität direkter physischer Interaktion oder echter Kraftrückkopplung, welche für ein tiefes Embodiment förderlich wäre [35].

2.3.3 Integraler Vergleich für co-located Anwendungen

In co-located Mixed-Reality-Szenarien besteht ein Spannungsfeld zwischen Präzision und natürlichem Erleben. Während Hand-Tracking die soziale Präsenz und intuitive Kommunikation fördert [6], ist es anfällig für Umwelteinflüsse wie Beleuchtung [11]. Für Aufgaben, die eine hohe, wiederholbare Präzision und geringste Latenz erfordern, bleiben Controller aufgrund ihrer technischen Reife und der effektiven Latenzkompensation die überlegene Wahl [50].

Über die reine Eingabepräzision hinaus ergibt sich bei co-located Anwendungen eine zusätzliche technische Hürde bei der Datensynchronisation. Reimer [34] betont, dass die Übertragung komplexer Hand-Skelett-Daten deutlich höhere Anforderungen an die Netzwerk-Bandbreite und Latenz stellt als die Übermittlung einfacher Controller-Zustände. Eine unzureichende Synchronisation führt zu visuellem Jitter der virtuellen Hände des Gegenübers, was die soziale Wahrnehmung und die koordinierte Zusammenarbeit am Tisch empfindlich stören kann.

Interessanterweise führt dieser technische Unterschied nicht zwangsläufig zu unterschiedlichen Lösungszeiten bei komplexen Aufgaben, sondern zu unterschiedlichen Lösungsstrategien. Controller fördern schnelles, repetitives Probieren, während Hand-Tracking eher zu bedachteren Handlungen führt [46]. Für das Design von co-located

[3] Die Zeitspanne zwischen einer physischen Bewegung der nutzenden Person und der entsprechenden Aktualisierung des Bildes auf dem Display.

Spielen bedeutet dies, dass die Wahl der Eingabemethode nicht nur die Usability, sondern auch die fundamentale Spielstrategie der Spielenden beeinflusst.

2.4 Die Meta Quest 3 als Referenzplattform

Basierend auf den in den vorangegangenen Abschnitten dargelegten Anforderungen an Video-Passthrough und Tracking-Präzision wird die *Meta Quest 3* als Referenzplattform für diese Arbeit gewählt. Neben ihrer technischen Eignung nimmt sie mit einem Marktanteil von 84 % im vierten Quartal 2024 die dominierende Position im Consumer-MR-Markt ein [62], was die Relevanz der Ergebnisse für eine breite Nutzerbasis sicherstellt.

Die Plattform bietet eine vollständige SDK-Unterstützung für die in Abschnitt 2.1.3 klassifizierten co-located Multi-User-Szenarien. Insbesondere die native Integration der Shared Spatial Anchors ermöglicht die notwendige Synchronisation des gemeinsamen Koordinatensystems [76]. Ergänzend liegen bereits umfangreiche wissenschaftliche Studien zur Tracking-Genauigkeit vor, die eine fundierte Einordnung erlauben.

Hinsichtlich der Display-Technologie stellt die Integration von zwei 4-Megapixel-RGB-Kameras in Kombination mit einem aktiven Tiefensensor einen qualitativen Sprung dar. Dies ermöglicht ein stereoskopisches Vollfarb-Passthrough mit einer Auflösung von 18 Pixeln pro Grad (PPD) [93]. Diese Auflösung ist essenziell, um in co-located Settings feine nonverbale Signale und die Mimik der Mitspielenden trotz der HMD-Okklusion wahrnehmen zu können.

Zudem erlaubt die hardwareseitige Tiefenerfassung über die *Depth API* eine dynamische Verdeckung virtueller Objekte durch reale Elemente der Umgebung in Echtzeit [68]. Diese Funktion ist für Tabletop-Szenarien von zentraler Bedeutung, da sie glaubwürdige visuelle Interaktionen zwischen der physischen Tischoberfläche, den Händen der Nutzenden und den virtuellen Spielelementen gewährleistet.

Kapitel 3

Soziale Interaktion in MR-Tabletop-Spielen

3.1 Charakteristika klassischer Tabletop-Spiele

Analoge Tabletop-Spiele fungieren seit jeher als zentrales Medium sozialer Interaktion. Im Gegensatz zu Videospielen, die häufig durch Bildschirme mediatisiert werden, zeichnet sich das analoge Spiel durch die Unmittelbarkeit des gemeinsamen Erlebens aus [36]. Eine aktuelle Untersuchung mit über 1500 Brettspieler*innen zeigt, dass soziale Interaktion die wichtigste Motivation für das Spielen darstellt, noch vor strategischem Denken oder anderen Herausforderungen. In der Studie bewerteten die Teilnehmenden soziale Interaktion als primären Motivationsfaktor mit einem Median von 80 von 100 Punkten, während strategisches Denken mit 78 Punkten knapp dahinter rangierte [9]. Diese Qualität basiert maßgeblich auf der physischen Beschaffenheit des Spiels und der räumlichen Konstellation der Spielenden.

Ein zentrales Charakteristikum analoger Brettspiele ist die Materialität der Spielkomponenten. Physische Objekte wie Würfel, Karten und Figuren fungieren dabei nicht rein funktional, sondern besitzen einen intrinsischen Wert als haptisches und visuelles Erlebnis. Die taktile Interaktion mit den Materialien sowie das gemeinsame Arrangement der Komponenten auf dem Tisch sind essenzielle Bestandteile des rituellen Spielprozesses [36].

Der Spieltisch konstituiert dabei einen gemeinsamen, öffentlichen Informationsraum. Rogerson et al. beschreiben das physische Brett als zentrales Element, durch das die Spielenden das Spiel erleben. Verlässt der Fokus das Brett, verliert sich oft auch der Bezug zum gemeinsamen Spielgeschehen. Durch die Sitzanordnung rund um dieses Zentrum entsteht eine Situation lückenloser gegenseitiger Wahrnehmung. Anders als bei digital vermittelten Interaktionen, in denen der Blick auf private Bildschirme gerichtet ist, erlaubt der analoge Spieltisch einen geteilten Aufmerksamkeitsfokus, innerhalb dessen sowohl der Spielzustand als auch die sozialen Reaktionen der Mitspielenden simultan wahrgenommen werden können, ohne durch die Vermittlung individueller Bildschirme unterbrochen zu werden. Zugleich bleiben durch diese räumliche Anordnung zentrale Kanäle nonverbaler Kommunikation wie flüchtiger Augenkontakt, subtile Gestik und Mimik erhalten, die für das soziale Lesen der Mitspielenden und die Koordination des

gemeinsamen Spielvollzugs wesentlich sind [37].

3.2 Hybride Tabletop-Systeme: Digitale Transformation

Die Entwicklung hybrider Brettspiele zielt darauf ab, die sozialen Qualitäten der analogen Welt zu bewahren, während gleichzeitig die Rechenleistung digitaler Systeme genutzt wird. Häufig werden dabei kognitive Lasten wie komplexes Punktezählen oder Regelverwaltung an das digitale System ausgelagert, damit sich die Spielenden auf die Interaktion konzentrieren können. Dabei besteht jedoch das Risiko, dass die digitale Technik zur Barriere zwischen den Spielenden wird oder die Aufmerksamkeit vom gemeinsamen Spielbrett abzieht [37].

Designprinzipien für die soziale Interaktion

Um sicherzustellen, dass die technologische Erweiterung das soziale Miteinander nicht stört, entwickelten Rogerson et al. auf Basis von Expert*inneninterviews fünf Designprinzipien für digitale Werkzeuge in Brettspielen [37]. Für die Gestaltung von co-located Mixed-Reality-Szenarien sind insbesondere zwei dieser Prinzipien von zentraler Bedeutung:

- **Integration:** Dieses Prinzip fordert, dass digitale Elemente die Spielenden zum gemeinsamen Brett hinziehen und nicht davon weglenken. Anstatt als isolierende Bildschirme zu fungieren, müssen virtuelle Objekte wie physische Komponenten behandelt werden, die sich nahtlos in das gemeinsame Spielmaterial einfügen. Wenn die Technik die Aufmerksamkeit der Spielenden vereinnahmt, leidet die soziale Verbindung, die den Kern des Brettspiels ausmacht.
- **Nachvollziehbarkeit:** Während Videospiele Berechnungen oft nach dem Black-Box-Prinzip verbergen, ist für Brettspiele die Transparenz der Spielmechanik essenziell. Spielende müssen verstehen, warum ein Zug erlaubt oder verboten ist und woher Ressourcen stammen. Rogerson et al. warnen vor Situationen, in denen Spielende und Elektronik unterschiedliche Vorstellungen der Spielrealität haben, weil der digitale Zustand vom physischen abweicht. Für MR-Systeme bedeutet dies, dass Automatisierung und Regelüberwachung stets visuell transparent und erklärbar bleiben müssen, um das Vertrauen in die gemeinsame Spielrealität zu sichern.

Während weitere Prinzipien wie Vollständigkeit oder Datenschutz für die Produktentwicklung relevant sind, entscheiden primär Integration und Nachvollziehbarkeit über die Qualität der unmittelbaren sozialen Interaktion. Das Design eines MR-Tabletop-Spiels muss sich daran messen lassen, ob es die Technologie nutzt, um Verwaltungsaufgaben zu eliminieren, ohne dabei die Transparenz oder den gemeinsamen Fokus auf dem Tisch zu opfern [37].

Co-located MR-Tabletop-Systeme können in diesem Kontext als eine technologische Evolution klassischer hybrider Ansätze verstanden werden. Die Anwendung dieser Prinzipien ist hierbei von erhöhter Dringlichkeit. Während bei Tablet-gestützten Spielen lediglich der Blick vom Brett abgewendet wird, riskieren HMDs durch die direkte Überlagerung des Sichtfelds eine stärkere soziale Isolation. Die konsequente Verankerung

digitaler Inhalte im gemeinsamen physischen Raum (World Space) statt in privaten Menüs (Screen Space) ist daher keine rein ästhetische Entscheidung, sondern eine fundamentale Bedingung, um die soziale Präsenz am Tisch aufrechtzuerhalten.

3.3 Einflussfaktoren der Interaktion in co-located MR

Die Übertragung dieser Prinzipien in Mixed Reality erfordert ein tiefes Verständnis der psychologischen und räumlichen Wirkungsweisen von MR-Systemen.

3.3.1 Co-Presence und physisch-virtuelle Konsistenz

Co-Presence beschreibt nach Goffman [19] eine Situation, in der Individuen unmittelbar aufeinander reagieren können, indem sie sich gegenseitig zugänglich, verfügbar und unterworfen sind. Zhao [56] entwickelte darauf aufbauend eine Taxonomie der menschlichen Co-Location, welche die technologischen Bedingungen der Anwesenheit und der Nähe klassifiziert und deren Einfluss auf das subjektive Erleben untersucht.

In co-located MR-Tabletop-Spielen liegt zwischen den Spielenden objektiv Zhaos *Corporeal Copresence* vor, da diese sich physisch im selben Raum befinden und face-to-face interagieren. Die MR-Technologie erzeugt jedoch eine hybride Form der Interaktion, in der virtuelle Elemente die physische Realität überlagern. Studien zeigen hierzu, dass die Fähigkeit, physischen Einfluss auf diese geteilte Umgebung zu nehmen, das subjektive Gefühl von Co-Presence signifikant steigert. Obwohl Lee et al. [24] dies spezifisch im Kontext von virtuellen Menschen untersuchten, unterstreicht es die Bedeutung physisch-virtueller Konsistenz für das gemeinsame Erleben. Yin et al. [54] stützen diese Beobachtung in einer aktuellen Untersuchung, indem sie aufzeigen, dass handlungsbasierte Interaktionen und die direkte Manipulation virtueller Objekte die soziale Präsenz signifikant fördern und die Kommunikation intuitiver gestalten, was das Gefühl der Zusammenarbeit stärkt.

3.3.2 Soziale Präsenz und Qualität der Interaktion

Die soziale Präsenz, ein zentrales psychologisches Konstrukt im Kontext der virtuellen Kommunikation und Interaktion, wird maßgeblich durch die räumliche Präsenz und die wahrgenommene zwischenmenschliche Wärme bestimmt. Son et al. [44] konnten nachweisen, dass insbesondere die Wahrnehmung des Gegenübers als „warmherzig“ und nahbar einen stärkeren Einfluss auf die soziale Präsenz hat als dessen wahrgenommene Kompetenz oder die visuelle Detailtreue des Avatars. Dies spricht für das Design von co-located Szenarien, da die physische Anwesenheit am gemeinsamen Tisch die räumliche Präsenz maximiert und die direkte Face-to-Face-Kommunikation feine soziale Nuancen wie Lächeln und Zuwendung transportiert, die die wahrgenommene Wärme steigern.

Für eine erfolgreiche Zusammenarbeit in VR/MR-Umgebungen ist zudem die Gestaltung der Avatare und der Kommunikation essenziell. Yousefdeh et al. [55] heben hervor, dass Faktoren wie Avatar-Realismus und eine funktionierende verbale Kommunikation entscheidend sind, um Co-Presence und effektive Kollaboration in geteilten virtuellen Räumen zu gewährleisten. Ergänzend dazu untersuchte Srivastava [45] die Qualität sozialer Interaktionen im direkten Vergleich verschiedener Medien. In dieser

Studie erzielte MR zwar erwartungsgemäß geringere Werte bei der sozialen Präsenz als die physische Begegnung, förderte jedoch signifikant tiefere Gespräche und ein stärkeres Verbundenheitsgefühl als herkömmliche mobile Anwendungen. Dies unterstreicht die Eignung von MR als Medium für qualitativ hochwertige soziale Interaktionen, die über reine Informationsübermittlung hinausgehen [45].

Diesen positiven Aspekten stehen jedoch Befunde gegenüber, die auf eine potenzielle Minderung der sozialen Wahrnehmung durch die Hardware hindeuten. Santoso und Bailenson [40] konnten quantitativ nachweisen, dass die Nutzung von Video-Passthrough das Phänomen der sogenannten *Social Absence* auslösen kann. Dabei werden physisch im Raum anwesende Personen als weniger real und weniger präsent wahrgenommen als in einer direkten Face-to-Face-Interaktion ohne HMD. Die Autoren vergleichen diesen Effekt mit der Mediatisierung eines Videotelefonats, bei dem reale Personen durch das Video vermittelt werden und die soziale Präsenz im Vergleich zur Interaktion ohne Headset um etwa 30 % sinkt.

Während Santoso und Bailenson [40] die Minderung der sozialen Präsenz durch die Hardware-Barriere quantifizieren, zeigt Detto [10] einen möglichen Lösungsansatz für die lokale MR-Kollaboration auf. Er untersuchte die Überlagerung der durch das Headset verdeckten Kopfpartien mittels virtueller Avatarköpfe. Zwar zeigten die Fragebogendaten keine generelle Steigerung der subjektiven sozialen Präsenz allein durch das Hinzufügen von Köpfen, jedoch führte die Nutzung von Cartoon-Avataren zu signifikant längeren Blickkontakten und wurde als attraktiver empfunden als realistische Avatare. Dies deutet darauf hin, dass stilisierte Avatare in Kombination mit Lippensynchronisation helfen können, die Aufmerksamkeit auf das Gegenüber zu erhöhen und die Interaktionsqualität zu verbessern.

3.3.3 Digital Proxemics

Während die physische Positionierung am Tisch weitgehend statisch bleibt, folgt die räumliche Wahrnehmung im virtuellen Raum eigenen Gesetzmäßigkeiten, die Williamson et al. [51] unter dem Begriff *Digital Proxemics* beschreiben. In ihrer Untersuchung zeigen sie, dass für Nutzer*innen von Head-Mounted Displays insbesondere die verwendete Eingabemodalität das räumliche Verhalten maßgeblich prägt. So weisen HMD-Nutzer*innen im Vergleich zu Desktop-Nutzer*innen ein ausgeprägteres Bewusstsein für ihren persönlichen Raum auf und vermeiden Verletzungen der Intimdistanz häufiger. Die Autoren führen dies unter anderem auf die Sichtbarkeit der Hände und die Möglichkeit zurück, soziale Signale wie Kopfbewegungen aktiv zur Signalisierung von Aufmerksamkeit einzusetzen. Das durch das HMD unterstützte Embodiment führt demnach dazu, dass physische soziale Normen stärker in den virtuellen Raum übertragen werden.

3.3.4 Gaze Awareness und Joint Attention

Die Wahrnehmung der Blickrichtung des Gegenübers (*Gaze Awareness*) und die daraus resultierende *Joint Attention* fungieren als zentrale Mechanismen für koordinierte Interaktionen in co-located Szenarien. Während Tomasello [48] *Joint Attention* als fundamentalen Schritt der sozialen Kognition definiert, bei dem zwei Individuen ihre Aufmerksamkeit gemeinsam auf ein Objekt richten, zeigen Piumsomboon et al. [33] im Kontext von Mixed Reality auf, dass dieser Prozess durch visuelle Hilfsmittel technisch

augmentiert werden kann. Ihre Untersuchungen zu sogenannten „Awareness Cues“ wie Frustums oder Blickstrahlen legen die Basis für MR-Kollaboration. Ergänzend demonstrieren Jing et al. [21] spezifisch für co-located Szenarien, dass bi-direktionale Visualisierungen der Blickrichtung nicht nur die Herstellung von Joint Attention unterstützen, sondern auch das Verständnis der Intentionen des Gegenübers signifikant verbessern.

Die Nutzung von HMDs kann diesen natürlichen Kommunikationskanal jedoch beeinträchtigen. In VR-Umgebungen wird die Gesichtsmimik durch das Headset verdeckt, was oft den Einsatz von Avataren notwendig macht [5]. Ähnlich führen AR-Brillen dazu, dass Blickrichtungen schwerer lesbar sind und Nutzer*innen ihre Aufmerksamkeit expliziter zwischen dem*der Partner*in und der Aufgabe aufteilen müssen, da implizite Hinweise fehlen [21]. Besonders der *Mutual Gaze*, also der wechselseitige Blickkontakt als zentrale soziale Bindungskomponente, wird durch die Hardware physisch blockiert, was die nonverbale Kommunikation erschwert [10].

Untersuchungen bestätigen, dass fehlende explizite Blickhinweise in MR-Umgebungen die Kollaboration behindern, da die Identifikation des gemeinsamen Fokusobjekts ohne verbale oder gestische Zusatzinformationen ungenau bleibt [21]. Jing et al. heben hervor, dass insbesondere räumlich eingebettete Cues wie „Laser Eyes“ helfen, diese Lücke zu schließen, da sie durch ihre Richtungsvektoren auch im peripheren Sichtfeld wahrgenommen werden können und so eine intuitive *Mutual Awareness* schaffen. Dieser Aspekt ist besonders in co-located Settings relevant, da physische Nähe häufig ein vorsichtigeres räumliches Verhalten sowie größere Distanzen zwischen den Akteur*innen impliziert und technologische Barrieren andernfalls die intuitive Deutung von Bewegungen beeinträchtigen können [5].

Diesen Befund untermauert Detto [10] durch Blickregistrierungsdaten. Er stellte fest, dass Teilnehmende ohne virtuelle Hilfsmittel das verdeckte Gesicht des Gegenübers deutlich seltener fixieren. Durch die Integration von Avatarköpfen, welche die Blickrichtung visualisieren, stieg die Zeit, in der sich die Partner*innen gegenseitig ansehen, signifikant an. Interessanterweise erwiesen sich hierbei stilisierte „Cartoon-Avatare“ als effektiver und attraktiver als realistische Darstellungen, da sie das *Uncanny-Valley*-Phänomen[1] vermeiden und so die Wiederherstellung des Mutual Gaze technisch optimal unterstützen.

3.3.5 Räumliche und visuelle Wahrnehmungsbarrieren

Trotz dieser potenziellen Vorteile ergibt sich eine wesentliche technische Einschränkung für die soziale Präsenz in Co-located-MR-Szenarien aus den begrenzten Sichtfeldern aktueller HMDs.

Eingeschränktes Sichtfeld (FOV)

Die *Meta Quest 3* bietet einen horizontalen Sichtwinkel von 110° und einen vertikalen von 96° [75], während das natürliche binokulare menschliche Sichtfeld etwa 214° horizontal sowie bis zu 150° vertikal umfasst [47]. Diese Diskrepanz führt zu einer erheblichen Einschränkung des Wahrnehmungsbereichs. Während Menschen ihre Umgebung über

[1] Bezeichnet das Unbehagen, das entsteht, wenn künstliche Figuren fast, aber nicht perfekt wie echte Menschen wirken [88].

ein breites peripheres Sichtfeld erfassen, nehmen HMD-Nutzer*innen lediglich einen zentral fokussierten Ausschnitt wahr. Laut den Gesundheits- und Sicherheitsrichtlinien von Meta kann diese Begrenzung insbesondere bei lokalen Multiplayer-Erlebnissen problematisch sein, da die Sicht auf die physische Umgebung und die Mitspieler*innen eingeschränkt wird [77]. Dies beeinträchtigt die gegenseitige räumliche Wahrnehmung und mindert somit die soziale Präsenz.

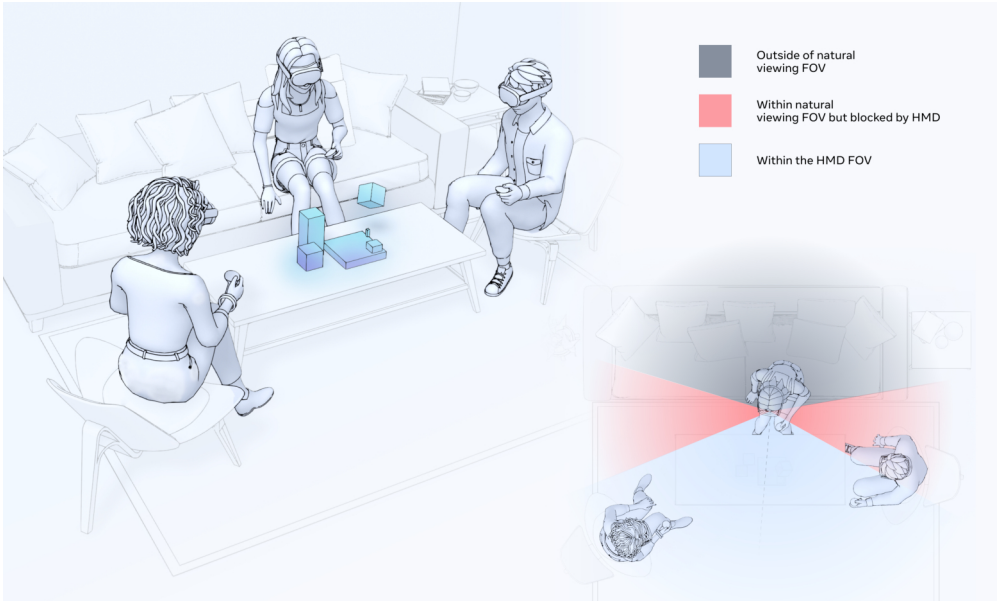


Abbildung 3.1: FOV-Limitierungen in Co-located-MR-Tabletop-Spielen [77].

Wie Abbildung 3.1 verdeutlicht, führt dieser Effekt dazu, dass physisch ko-präsente Mitspieler*innen außerhalb des HMD-FOV schlicht nicht wahrgenommen werden. Diese fundamentale Beeinträchtigung der peripheren Awareness reduziert die Fähigkeit, intuitiv auf nonverbale Signale wie Gesten oder Körperbewegungen zu reagieren.

Hinsichtlich der Sicherheit bewertet Meta solche stationären Tabletop-Szenarien zwar als risikoärmer im Vergleich zu bewegungsintensiven Anwendungen, da die Nutzer*innen an festen Positionen verbleiben [77]. Dennoch bleibt die Einschränkung der sozialen Wahrnehmung bestehen, da die natürliche Erfassung der Gruppe durch den „Tunnelblick“ der Hardware unterbrochen wird.

Perzeptive Verzerrungen und Distanzfehler

Zusätzlich zu den rein geometrischen Einschränkungen des Sichtfeldes beeinflusst die technische Vermittlung durch Video-Passthrough die perzeptive Genauigkeit der Nutzerwahrnehmung. Untersuchungen zur *Meta Quest 3* zeigen, dass Distanzen unter den Bedingungen des Passthroughs im Vergleich zur natürlichen Sicht systematisch unterschätzt werden, was unter anderem auf die physische Positionierung der Kameras sowie auf die algorithmische Fusion der Videostreams zurückgeführt wird, da diese nicht exakt der natürlichen Augenposition entsprechen.

Darüber hinaus treten Veränderungen in der Wahrnehmung des eigenen Körper-

schemas sowie subjektiv empfundene Körperverzerrungen auf, die insbesondere für stationäre Tabletop-Anwendungen relevant sind, da sie die Präzision manueller Interaktionen mit virtuellen Objekten auf physischen Oberflächen erheblich beeinträchtigen können [40].

Perspektiven-Diskrepanz

Ein fundamentales Problem kollaborativer MR ist die *Perspektiven-Diskrepanz*. Obwohl sich die Nutzer*innen physisch am gleichen Tisch befinden, sehen sie die virtuelle Szene aus unterschiedlichen Blickwinkeln. Dadurch entstehen Inkonsistenzen im gemeinsamen Wahrnehmungsraum. Ein Objekt, das Person A frontal sieht, kann für Person B verdeckt sein oder auf dem Kopf stehen. Solche Differenzen erschweren deiktische Referenzen erheblich, da das räumliche Verständnis zwischen den Teilnehmenden asymmetrisch ist [20].

Traditionell wird versucht, dieses Problem durch eine starre räumliche Synchronisation zu lösen, bei der alle virtuellen Objekte fest an der physischen Welt verankert sind. Gottsacker et al. [20] argumentieren jedoch, dass diese starre Kopplung bei komplexen Objekten hinderlich sein kann. Wenn Person B die Sicht von Person A einnehmen muss, würde eine erzwungene virtuelle Übernahme der Sicht oft zu Cybersickness oder eine physische Bewegung zu Ineffizienz führen.

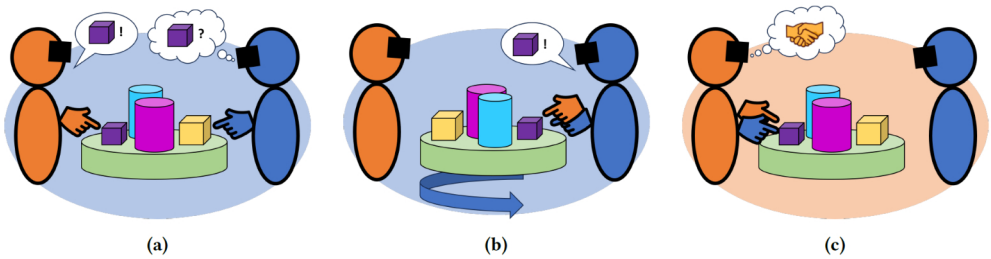


Abbildung 3.2: Schematische Darstellung der Perspektiven-Diskrepanz in co-located MR-Kollaboration. (a) Zwei Nutzer*innen sehen dasselbe virtuelle Objekt aus unterschiedlichen Blickwinkeln, was zu inkonsistenten Referenzen führt. (b) Prozess der perspektivischen Ausrichtung durch dynamische Transformation lokaler Kopien. (c) Aufgelöste gemeinsame Sicht mit konsistenter räumlicher Referenz trotz unterschiedlicher physischer Positionen [20].

Als technischer Lösungsansatz wurde die „Decoupled Hands“-Technik entwickelt. Dabei erhält jede*r Nutzer*in eine lokale Kopie der virtuellen Inhalte. Um die Perspektiven anzugleichen, wird diese Kopie dynamisch rotiert, sodass die exakte Sicht des Gegenübers eingenommen wird, ohne die physische Sitzposition verlassen zu müssen. Entscheidend für die soziale Interaktion ist, dass die virtuellen Repräsentationen der Hände des*der Leader*in mit den Objekten transformiert werden. Dadurch zeigen die virtuellen Hände im Sichtfeld des*der Follower*in präzise auf die korrekte Stelle des rotierten Objekts. Dieser Ansatz bewahrt die Face-to-Face-Interaktion und ermöglicht das Deuten auf Details, ohne dass die Nutzer*innen ihre physische Position verlassen oder auf die Wahrnehmung von Mimik und Körpersprache verzichten müssen.

Für die Gestaltung sozialer Interaktion in co-located Mixed-Reality-Tabletop-Spielen

bietet diese Technik zwar eine Lösung für visuelle Okklusionsprobleme bei komplexen 3D-Modellen, steht jedoch in einem Spannungsfeld zur territorialen Konsistenz klassischer Tabletop-Szenarien. Während Gottsacker et al. die Unabhängigkeit von der physischen Sitzposition als Vorteil für die Referenzierung betonen, erfordern Spiele oft eine strikte Verbindung zwischen virtuellem Raum und physischer Präsenz. Eine individuelle Rotation der Spielszene könnte die intuitive Zuordnung von Spielfiguren oder Zonen zu den physischen Mitspielenden stören. Für das Ziel einer glaubwürdigen *Corporeal Copresence* im Spielkontext muss daher abgewogen werden, ob die visuelle Klarheit durch Perspektivteilung den Verlust der räumlichen Orientierung und der physischen Verankerung von Spielbesitz aufwiegt.

3.4 Ableitung theoretischer Gestaltungsanforderungen

Basierend auf den theoretischen Grundlagen zu Co-Presence und Joint Attention lassen sich konkrete Anforderungen für die Gestaltung von co-located MR-Tabletop-Spielen definieren. Diese Anforderungen bilden die Grundlage für die Konzeption und technische Umsetzung des Prototyps.

3.4.1 Multimodales Feedback

Um trotz der virtuellen Natur der Objekte ein starkes Gefühl von Embodiment und Handlungskontrolle zu vermitteln, ist multimodales Feedback entscheidend. Studien belegen konsistent, dass das Zusammenspiel mehrerer Modalitäten einen stärkeren Sinn für Präsenz vermittelt als einzelne Modalitäten allein. Kombinierte Rückmeldungen erleichtern die kognitive Verarbeitung und erhöhen die wahrgenommene Natürlichkeit der Interaktion [12, 17].

Hinsichtlich der räumlichen und physikalischen Wahrnehmung zeigen Untersuchungen, dass sich die Modalitäten funktionell ergänzen. Während visuelles Feedback signifikant die Lokalisierung von Ereignissen verbessert (z. B. wo ein Objekt auftrifft), ermöglicht haptisches Feedback eine direktere Wahrnehmung, *dass* ein Ereignis überhaupt stattgefunden hat. Die haptische Komponente ist dabei besonders effektiv, um die Existenz eines Ereignisses zu bestätigen, selbst wenn die visuelle Zuordnung unsicher ist [17].

Im Kontext von sozialen Interaktionen mit Embodied Conversational Agents[2] erfüllt multimodales Feedback zusätzlich eine regulierende Funktion. Hier konnte gezeigt werden, dass die Integration akustischer Signale (wie gefüllte Pausen) in Kombination mit visuellen Indikatoren (wie Warte-Icons oder nonverbalen Gesten) die negative Wahrnehmung von Latenzen verringert. Diese multimodale Strategie steigert nicht nur die soziale Präsenz, sondern erhöht signifikant die wahrgenommene Belebtheit und Anthropomorphisierung des Gegenübers [12].

[2] Computergenerierte Gesprächspartner*innen, die Sprache und nonverbale Signale für interaktive Konversationen nutzen [12].

3.4.2 Visualisierung sozialer Cues

Da HMDs einen Großteil der nonverbalen Kommunikation verdecken, müssen MR-Tabletop-Spiele explizite Mechanismen bereitstellen, die soziale Hinweise sichtbar machen. Awareness-Mechanismen schließen diese Lücke, indem sie Informationen über Aufmerksamkeit, Aktivität und Position der anderen Spieler*innen repräsentieren [21].

Gaze-Indikatoren sind ein zentrales Werkzeug, um *Shared Gaze* technisch wiederherzustellen. Der „Laser Eye"-Ansatz, bei dem ein virtueller Strahl die Blickrichtung vom HMD-Ursprung zur fokussierten Stelle auf dem Tisch anzeigt, hat sich als besonders effektiv erwiesen. Er verbessert die Koordination und das gemeinsame Verständnis der Situation signifikant, da er durch die räumliche Verkörperung des Strahls auch im peripheren Sichtfeld wahrgenommen werden kann [21]. Erickson et al. [14] bestätigen, dass solche Cues die Leistung steigern, weisen jedoch darauf hin, dass technische Faktoren wie Latenz (ab ca. 260 ms) und Tracking-Ungenauigkeit die Fehlerrate bei der Zielidentifikation erhöhen können.

Zusätzlich können *Proximity Cues* wie das Field-of-View-Frustum genutzt werden, um anzuzeigen, was Mitspieler*innen gerade sehen können. Piumsomboon et al. [33] identifizierten die Kombination aus Gaze-Rays und Frustum-Visualisierung als effektivstes Hilfsmittel, um Nutzer*innen bewusst zu machen, welche Bereiche für ihr Gegenüber sichtbar sind. Dies ist besonders in co-located Szenarien relevant, da Nutzer*innen hier ein vorsichtigeres räumliches Verhalten zeigen und größere physische Abstände wahren als in rein virtuellen Umgebungen [5].

Neben Blickrichtungen sind Aktivitätsindikatoren wichtig. Piumsomboon et al. [33] zeigten, dass farbliche Hervorhebungen der virtuellen Hände (z. B. Rot für Greifen, Grün für *Thumbs-up*) Missverständnisse reduzieren. Jing et al. [21] ergänzen, dass solche visuellen Gaze-Cues die Notwendigkeit für explizite Handgesten verringern, da deiktische Referenzen durch den Blickstrahl bereits eindeutig geklärt sind und somit die physische Kommunikation effizienter gestaltet wird.

Avatar-basierte Repräsentationen dienen dazu, Körperhaltung und Gesten trotz HMDs sichtbar zu machen. Bereits abstrahierte Avatare ermöglichen die Einschätzung der räumlichen Position und erhöhen die soziale Präsenz [30]. Hinsichtlich des Gestaltungsstils liefert Detto [10] wichtige Evidenz für die Vermeidung des *Uncanny-Valley*-Effekts. Seine Studie zeigt, dass stilisierte „Cartoon-Avatare" im Vergleich zu realistischen Darstellungen als attraktiver wahrgenommen werden und die gegenseitige Blickdauer erhöhen. Zudem erwies sich ein einfaches Audio-basiertes *Lip-Sync* für die Mundanimation als effektiver für die soziale Bindung als komplexes Facial-Tracking, da letzteres bei realistischen Avataren oft befremdlich wirkte [10].

3.4.3 Symmetrische Interaktionsgestaltung

Forschung zu kooperativen Brettspielen zeigt, dass ein kontinuierlicher verbaler Austausch sowie das Teilen von Informationen grundlegende Voraussetzungen für eine erfolgreiche gemeinsame Strategiebildung darstellen [3]. In hybriden Spielsystemen besteht jedoch die Gefahr, dass technische Unterschiede, etwa durch unterschiedliche Hardwareausstattungen wie Augmented Reality und Virtual Reality oder durch die exklusive Nutzung eines Eingabegeräts, unbeabsichtigt asymmetrische Rollenstrukturen hervorbringen. In der Folge können einzelne Spielende vorwiegend die Rolle technischer

Operator*innen einnehmen oder Führungsrollen allein aufgrund der Art der Informationsdarstellung entstehen. Dies schränkt die gleichberechtigte Teilhabe am kooperativen Spielgeschehen ein [5, 37].

Um dieser Entwicklung entgegenzuwirken, ist das Prinzip der Integration (vgl. Abschnitt 3.2) konsequent zu berücksichtigen. Ziel ist eine Umsetzung, bei der digitale Werkzeuge lediglich eine von mehreren Komponenten darstellen und die Aufmerksamkeit nicht vom gemeinsamen physischen Raum ablenken. Der Beitrag zum Erreichen des Spielziels sollte primär durch die Interaktion mit dem geteilten physischen Spielfeld sowie mit den Mitspielenden erfolgen, sodass die eingesetzte Technik den sozialen Austausch unterstützt und nicht als Barriere im kooperativen Spielgeschehen wirkt [3, 37].

3.4.4 Reduktion der kognitiven Belastung

MR-Systeme können durch visuelle Überladung, eingeschränkte Sichtfelder sowie die Notwendigkeit, die Aufmerksamkeit zwischen Interaktionspartner*innen und -partnern und virtuellen Objekten aufzuteilen, eine erhöhte kognitive Belastung verursachen. Nutzende müssen gleichzeitig die physische Umgebung, virtuelle Inhalte und die Handlungen anderer Spielender berücksichtigen. Ohne unterstützende Gestaltungsmechanismen führt diese geteilte Aufmerksamkeit zu einem erhöhten Koordinationsaufwand, da Intentionen und Aufmerksamkeitsfokus des Gegenübers explizit kommuniziert oder inferiert werden müssen [21].

Um kognitive Ressourcen für die eigentliche Aufgabe freizusetzen, ist die Gestaltung konsistenter Feedbackmechanismen von zentraler Bedeutung. Eine einheitliche Ausgestaltung des Feedbacks über unterschiedliche technische Setups hinweg, etwa in Augmented-Reality- und Virtual-Reality-Konfigurationen, erleichtert die Synchronisation gemeinsamer Aktionen und reduziert Unsicherheiten hinsichtlich des geteilten Referenzraums [5]. Diese Entlastung begünstigt kollaborative Prozesse, in denen Spielende Regeln gemeinsam aushandeln und Denkprozesse externalisieren. In Anlehnung an Beobachtungen aus strategischen Tabletop-Spielen fungiert die Gruppe dabei als verteilter kognitiver Prozessor, wodurch die individuelle mentale Belastung reduziert wird [3].

Ein weiterer wesentlicher Einflussfaktor auf die kognitive Belastung ist die Wahl der Visualisierungsform. Räumlich eingebettete Hinweise, wie das sogenannte *Laser Eye*, nutzen das periphere Sehen und erfordern gemäß Messungen signifikant weniger mentale Anstrengung als abstrakte visuelle Indikatoren [21]. Demgegenüber können Visualisierungen mit hoher Informationsdichte, etwa Blickverlaufshistorien in Form von *Trail Paths*, insbesondere in räumlich eingeschränkten Arbeitsumgebungen kontraproduktiv wirken. Sie überlagern relevante Inhalte und beeinträchtigen dadurch den Aufmerksamkeitsfokus [21]. Ergänzend zeigen Messungen der elektrodermalen Aktivität, dass eine geringe Realitätstreue von Avataren oder das Fehlen verbaler Kommunikation emotionalen Stress begünstigen kann, was wiederum die für kollaborative Prozesse verfügbare kognitive Kapazität reduziert [55].

Dabei steht eine erhöhte technische Komplexität nicht zwangsläufig in einem positiven Verhältnis zu einer geringeren kognitiven Belastung oder einem verbesserten Nutzungserleben. Detto [10] konnte zeigen, dass einfaches audiobasiertes Lip-Syncing für Mundanimationen häufig als effektiver wahrgenommen wird als rechenintensives

kamerabasiertes Facial Tracking. Letzteres kann bei wahrgenommenen Inkonsistenzen irritierend wirken und durch den Uncanny-Valley-Effekt von der eigentlichen Interaktion ablenken. Für die Aufrechterhaltung sozialer Präsenz und die damit verbundene kognitive Entlastung durch intuitive Kommunikation sind daher Konsistenz und ein stimmiger sozialer Ausdruck, etwa durch stilisierte oder cartoonhafte Avatare, bedeutsamer als eine hohe technische Detailtiefe [10].

Kapitel 4

Stand der Forschung und Technik

4.1 Aktuelle Forschungsansätze zu MR-Co-Experience

Ghoshal et al. [16] untersuchten den Einfluss unterschiedlicher visueller Umgebungen auf das Spielerleben in co-located Szenarien, indem Teilnehmende dasselbe Spiel auf der *Meta Quest 2* sowohl in einer Virtual-Reality-Umgebung als auch im Video-Passthrough-basierten Augmented-Reality-Modus spielten.

Die Ergebnisse zeigten eine deutliche Differenzierung in der Wahrnehmung der beiden Modi. Während die VR-Umgebung aufgrund der visuellen Abschottung ein signifikant stärkeres Präsenzerleben erzeugte, führte der AR-Modus zu einem ausgeprägteren Gefühl von Co-Presence. Dieser Effekt wird von den Autor*innen darauf zurückgeführt, dass Mitspielende im AR-Modus nicht ausschließlich als abstrakte Avatare wahrgenommen werden, sondern weiterhin als physische Körper im realen Raum präsent bleiben. Dadurch wird die Wahrnehmung eines geteilten räumlichen Kontexts gestärkt.

Hinsichtlich des Spielspaßes konnten keine signifikanten Unterschiede zwischen den beiden Plattformen festgestellt werden. Dies ist insofern bemerkenswert, als ein höheres Maß an Immersion, wie es in VR gemessen wurde, in der Regel mit einem gesteigerten Spielerleben assoziiert wird. Die Autorinnen und Autoren argumentieren jedoch, dass die in der VR-Bedingung deutlich stärker ausgeprägte Cybersickness diesen positiven Effekt möglicherweise kompensiert hat. Insgesamt unterstreichen diese Ergebnisse das Potenzial von Mixed-Reality-Ansätzen für Anwendungsszenarien, in denen soziale Interaktion und physisches Wohlbefinden gegenüber maximaler virtueller Immersion priorisiert werden.

4.2 Marktanalyse: Tabletop in AR und MR

4.2.1 Optische AR-Systeme und projektionsbasierte Ansätze

Es existieren mehrere kommerzielle AR-basierte Tabletop-Systeme, die unterschiedliche technische Ansätze verfolgen. *Tilt Five* (Tilt Five, 2020) nutzt eine retroreflektierende Spielfläche, spezielle AR-Brillen und projektionsbasierte Verfahren, um virtuelle Inhalte über dem Tisch darzustellen [91]. Das System ist explizit für co-located Szenarien konzipiert, bei denen Spieler*innen physisch am selben Tisch sitzen. Es basiert auf ei-

ner externen optischen Tracking-Methode mittels Markierungspunkten, die präzises Alignment ermöglicht, jedoch eine räumliche Beschränkung auf das retroreflektive Material aufweist. Zudem treten distanzabhängige Helligkeitsprobleme auf, da die Projektoren ihre Helligkeit nicht automatisch an die Entfernung anpassen [90].

Mirrorscape (Mirrorscape Inc., 2023) verfolgt einen anderen Ansatz und nutzt mobile AR-Geräte wie Smartphones, Tablets sowie die *Apple Vision Pro* [87]. *Mirrorscape* (Mirrorscape Inc., 2023) verfolgt einen anderen Ansatz und nutzt mobile AR-Geräte wie Smartphones, Tablets sowie die *Apple Vision Pro* [87]. Die Plattform richtet sich primär an Tabletop-Rollenspiele wie *Dungeons & Dragons* (TSR, 1974) und *Pathfinder* (Paizo Publishing, 2009) und ermöglicht das Platzieren virtueller Miniaturen auf realen Tischflächen, wobei sowohl co-located als auch remote Multiplayer-Szenarien unterstützt werden [60]. Im Unterschied zu *Tilt Five* benötigt *Mirrorscape* keine spezielle Hardware, sondern nutzt die in modernen Mobilgeräten integrierten AR-Funktionen. Dies erhöht zwar die Zugänglichkeit, führt jedoch zugleich zu Einschränkungen hinsichtlich Immersion und Tracking-Präzision.

4.2.2 HMD-basierte Mixed-Reality-Systeme

Während AR-basierte Systeme wie *Tilt Five* oder *Mirrorscape* bereits etablierte Lösungen darstellen, befinden sich Mixed-Reality-Ansätze mit Head-Mounted Displays weiterhin in einer frühen Entwicklungsphase. Der aktuelle Markt für co-located MR-Anwendungen lässt sich dabei grob in zwei Kategorien differenzieren: bewegungsintensive Action-Titel und strategische Tabletop-Umsetzungen.

Ein Großteil des Angebots, darunter Titel wie *Spatial Ops* (Resolution Games, 2024), *Does It Stack?* (Joy Way, 2024) oder *Home Sports* (Resolution Games, 2024), setzt auf physische Aktivität und großräumige Interaktionen [63, 83, 85, 86]. Diese Konzepte nutzen MR als „physischen Spielplatz“, liefern jedoch nur begrenzte Erkenntnisse für ruhige, strategische Szenarien, die maßgeblich von stationärer Face-to-Face-Kommunikation leben.

Ein aktuelles Beispiel für die spezifische Ausrichtung auf familiäre Co-Location ist der Titel *Game Night* (Fantail Games, 2025). Der Entwickler Stephen Linton verfolgt hierbei den Ansatz, Mixed Reality explizit als digitales Brettspiel zu inszenieren, das die soziale Dynamik des gemeinsamen Raums in den Mittelpunkt stellt, anstatt sie durch komplexe Spielmechaniken zu überlagern [58]. Der Titel zeigt die praktische Umsetzung dieses Konzepts auf aktueller Consumer-Hardware.

Im Bereich der strategischen Tabletop-Spiele gilt *Demeo* (Resolution Games, 2021) als technologischer Vorreiter. Seit der Integration von Colocation-Support via Shared Spatial Anchors und Hand-Tracking im Jahr 2023 demonstriert der Titel, wie komplexe Brettspiele technisch in MR funktionieren können [84]. Auch Nachfolger wie *Battlemarked* (Brink XR, 2024) setzen auf diese Technologien [61]. Obwohl diese Titel beweisen, dass die technische Basis für Tabletop in MR existiert, liegt ihr Fokus primär auf der Spielmechanik von Dungeon-Crawlern und weniger auf der expliziten Unterstützung zwischenmenschlicher Kommunikation durch augmentierte soziale Cues.

Hier zeigt sich eine Forschungslücke im Design von co-located Tabletop-Erlebnissen. Während Sandbox-Ansätze wie *Table Space* (Boxtree Inc., 2024) zwar Trading Card Games mit Hand-Tracking ermöglichen [79], fehlt es an systematischen Untersuchungen

zur Gestaltung der sozialen Interaktion unter Video-Passthrough-Bedingungen. Insbesondere die Frage, wie sich unterschiedliche Eingabemodalitäten und die Visualisierung von Awareness-Cues auf das Gefühl der gemeinsamen Präsenz in ruhigen, strategischen Spielsituationen auswirken, ist durch die aktuellen kommerziellen Titel noch nicht hinreichend beantwortet. Andere Plattformen wie *All On Board!* (The Game Kitchen, 2023) konzentrieren sich hingegen fast vollständig auf rein virtuelle Umgebungen und Remote-Multiplayer [89], wodurch die Relevanz von co-located MR-Tabletop-Forschung weiter unterstrichen wird.

4.3 Frameworks und Entwicklungsumgebungen für Colocation

Die Implementierung von co-located MR-Erlebnissen basiert maßgeblich auf etablierten Game Engines wie *Unity* und *Unreal Engine*, die Rendering, Logik und die Integration von spezialisierten SDKs übernehmen. Für co-located MR-Szenarien mit Shared Spatial Anchors stellt Meta mehrere Referenzprojekte bereit.

Metas *Multiplayer Building Blocks* bieten vorgefertigte Unity-Komponenten, die häufige Multiplayer-Anforderungen abstrahieren und die Entwicklung beschleunigen [72]. Zu den zentralen Komponenten gehören `Networked Grabbable Object` für synchronisierte Objektmanipulation, `Networked Avatar` für Spieler*innen-Repräsentationen mit automatischer Pose-Synchronisation und `Colocation` für die Integration von Shared Spatial Anchors [72]. Die Building Blocks reduzieren die Implementierungskomplexität erheblich, setzen jedoch spezifische Networking-Frameworks voraus. Sie bieten zudem begrenzte Anpassungsmöglichkeiten für spezialisierte Interaktionsmuster wie kartenspezifische Gesten, was manuelle Erweiterungen erforderlich macht [72].

Unreal Engine 5 bietet durch das Meta XR Plugin grundlegende MR-Funktionalität, erfordert für Shared Spatial Anchors jedoch eine manuelle Integration des Austauschs der Anker. Im Unterschied zu Unity fehlen dedizierte Colocation-Templates.

Godot 4 unterstützt zwar OpenXR-basiertes MR, bietet jedoch keine nativen Lösungen für die Colocation, sodass Mehrspielerfunktionen manuell implementiert werden müssen [66].

Das umfassendste Referenzprojekt ist *Unity MR Motifs* [59], das die Implementierung von co-located Multiplayer-Erlebnissen mit der `Colocation Discovery API` demonstriert, welche das automatische Auffinden von Nutzer*innen ohne manuelle Raum-Codes ermöglicht [71]. Unitys eigenes MR-Multiplayer-Tabletop-Template, das im Februar 2025 veröffentlicht wurde, ist hingegen für Remote-Multiplayer konzipiert [92]. Die technische Infrastruktur ist somit grundsätzlich vorhanden, erfordert für komplexe Szenarien jedoch spezifische Implementierungen jenseits der Standard-Templates.

4.4 Positionierung der Arbeit

Diese Arbeit adressiert die identifizierten Forschungslücken durch die Entwicklung eines funktionalen co-located MR-Tabletop-Spiels für drei Spieler*innen auf *Meta Quest*-HMDs. Anders als bestehende Arbeiten verfolgt sie einen holistischen Ansatz, der technische Implementierung, empirische Evaluation und soziale Gestaltungsprinzipien miteinander verknüpft. Der Schwerpunkt liegt auf der empirischen Untersuchung symme-

trischer Designlösungen, die eine gleichberechtigte soziale Interaktion in einem kontrollierten co-located Szenario ermöglichen.

Kapitel 5

Konzeption und Prototypentwicklung

5.1 Spielkonzept und Designentscheidungen

Depthbound ist ein kooperatives, co-located Mixed Reality Tabletop-Spiel für drei Spieler*innen. Das Spiel adaptiert die Kernmechanik des Kartenspiels *Hanabi* und erweitert sie um räumliche und multimodale Interaktionsmöglichkeiten. Die Wahl von *Hanabi* als konzeptionelle Grundlage erfolgte aufgrund mehrerer strategischer Überlegungen. Das Spiel implementiert asymmetrische Informationsverteilung als zentrales Designprinzip, da alle Spieler*innen die Karten der Mitspielenden sehen, niemals jedoch die eigenen. Diese fundamentale Asymmetrie schafft Abhängigkeit zwischen den Spielenden und macht Kommunikation essentiell für den Spielerfolg, wodurch soziale Präsenz und Joint Attention natürlich gefördert werden. Einen visuellen Eindruck der praktischen Umsetzung sowie Ausschnitte des Prototyps zeigt das begleitende Projektvideo [57].

Für die Umsetzung in Mixed Reality erfolgten signifikante Anpassungen. Um die Spieldauer an MR-Sessions anzupassen und die kognitive Belastung durch HMDs zu berücksichtigen, wurde die Gesamtkartenzahl reduziert und die Ressourcen-Ökonomie verschärft. Ein zentrales Gestaltungsprinzip ist die Kongruenz von physischem und virtuellem Raum. Anders als in Remote-VR-Szenarien, in denen die Umgebung beliebig ist, teilen sich alle Teilnehmenden hier einen realen Ort. Der virtuelle Spieltisch fungiert dabei als synchronisierter Ankerpunkt, der mittels Shared Spatial Anchors für alle HMDs an exakt derselben Position im Raum fixiert wird.

Dabei wurde bewusst auf die zwingende Notwendigkeit einer physischen Entsprechung verzichtet. Das virtuelle Spielfeld ist frei im Raum positionierbar, um Unabhängigkeit von der vorhandenen Möblierung zu gewährleisten und haptische Konflikte mit unpassenden realen Objekten zu vermeiden. Dennoch entsteht durch die präzise Kalibrierung ein gemeinsamer Referenzrahmen, der es den Spielenden erlaubt, natürliche soziale Verhaltensweisen zu nutzen. Ein Zeigen auf eine virtuelle Karte ist für das Gegenüber räumlich eindeutig interpretierbar, da sich alle Spielenden innerhalb desselben Koordinatensystems bewegen. Das zugrunde liegende Design verfolgt das Ziel, die Aufmerksamkeit nicht in private Menüstrukturen zu lenken, sondern sie konsequent im *World Space* zu bündeln, der als gemeinsamer Interaktionsraum in der Mitte der Gruppe verortet ist.

5.1.1 Spielmaterial und virtuelle Komponenten

Das virtuelle Kartendeck wurde in Anlehnung an das Originalregelwerk [2] von ursprünglich 50 auf 28 Karten reduziert. Das Deck setzt sich somit aus vier Farbsätzen zusammen, die jeweils drei Karten des Wertes 1, drei Karten des Wertes 2 und eine Karte des Wertes 3 umfassen. Diese Verteilung macht höhere Zahlen strategisch wertvoller. Durch die Erhöhung auf drei Karten des Wertes 2 wurde die mittlere Spielphase stabilisiert, während die Einzigartigkeit der Zahl 3 die Komplexität der Entscheidung am Ende einer Farbreihe wahrt. Jede*r Spieler*in hält vier Karten, deren Vorderseiten durch das System nur für die Mitspielenden gerendert werden. Dieser Ansatz nutzt die selektive Render-Pipeline der MR-Technologie zur systemseitigen Regeldurchsetzung. Da die Informationen auf den eigenen Karten für das Headset der besitzenden Person gar nicht erst generiert werden, ist sowohl ein versehentliches als auch ein absichtliches Spicken, das eine Schwachstelle des analogen Vorbilds darstellt, technisch ausgeschlossen.

Da das dauerhafte Halten virtueller Objekte ergonomisch ermüdend ist, schweben die Handkarten leicht erhöht vor den Sitzplätzen der Spielenden. Diese Positionierung garantiert, dass die Karten jederzeit gut sichtbar und zugänglich bleiben, ohne komplexe Greifinteraktionen zu erfordern oder die Sicht auf den Tisch zu versperren. Abbildung 5.1 illustriert diesen räumlichen Aufbau sowie die selektive Informationsdarstellung aus der First-Person-Perspektive.

Abbildung 5.1: First-Person-View in *Depthbound*: Eigene Handkarten (Rückseiten), sichtbare Mitspielerkarten, farbkodierte Spielzonen

Die virtuelle Spielfläche nutzt räumliche Affordanzen, um Interaktionen intuitiv zu gestalten. Nach Norman bezeichnen Affordanzen die wahrgenommenen Handlungsmöglichkeiten, die ein Objekt seinen Nutzenden bietet [29]. In *Depthbound* wird dieses Prinzip durch sogenannte *Snap-Zones* umgesetzt. Vier farbkodierte Ablagebereiche umgeben das Zentrum und korrespondieren visuell mit den Kartenfarben. Diese Gestaltung fun-

giert als intuitiver Signifikant, der den Spielenden ohne textuelle Erklärung signalisiert, wo eine Karte abgelegt werden kann. Eine zentrale Spielzone dient der Kartenvalidierung. Spieler*innen platzieren dort Karten, woraufhin das System korrekte Karten automatisch in die jeweilige Farbzone animiert und fehlerhafte auf den Ablagestapel verschiebt.

Ergänzend dazu verwaltet das System die gemeinsamen Ressourcen. Zum Spielstart stehen dem Team fünf Hinweis-Tokens für die Kommunikation sowie drei Fehler-Tokens als Lebensanzeige zur Verfügung. Diese werden als zentral sichtbare UI-Elemente auf der virtuellen Schriftrolle dargestellt.

Die UI-Architektur folgt dabei einem hybriden Ansatz, der zwischen globalem Spielzustand und individueller Steuerung differenziert. Während die Aktionsauswahl über ein privates, nur für den*die aktive*n Spieler*in sichtbares Interface erfolgt, werden kollektive Informationen bewusst in den gemeinsamen *World Space* ausgelagert. Die Platzierung der Statusanzeigen auf der zentralen Schriftrolle statt in einem persönlichen Head-Up-Display (HUD) ist eine gezielte Maßnahme zur Stärkung der Co-Location. Da der Blick auf den Tisch notwendig ist, um den Ressourcenstand zu prüfen, wird die Joint Attention gebündelt und die soziale Isolation durch rein private Menüstrukturen minimiert.

5.1.2 Spielablauf und Aktionsmechanik

Der Spielablauf folgt einer festen Rundenstruktur. Der Host, also Spieler*in 1, beginnt. Danach sind die weiteren Spielenden im Uhrzeigersinn an der Reihe. In jedem Zug wird genau eine von drei möglichen Aktionen ausgeführt. Die Auswahl erfolgt über UI-Buttons auf einem Canvas, das am `LeftHandAnchor` des*der aktiven Spieler*in erscheint. Die Buttons stehen für die Aktionen „Karte spielen“, „Karte abwerfen“ und „Hinweis geben“. Nicht verfügbare Aktionen werden visuell ausgegraut, um eine klare Bedienbarkeit zu gewährleisten.

Bei der Aktion „Karte spielen“ wählt die spielende Person eine Karte aus der eigenen Hand und legt sie in die passende Farbreihe. Die Karte muss der nächsten erwarteten Zahl dieser Reihe entsprechen. Jede Farbreihe beginnt mit der Zahl eins, gefolgt von zwei und drei. Ist die Karte korrekt, wird sie in die Farbzone animiert. Ist sie falsch, verliert das Team ein Fehler-Token, und die Karte wird auf den Ablagestapel gelegt.

Bei der Aktion „Karte abwerfen“ legt der Spielende eine Karte auf den Ablagestapel und erhält, falls möglich, ein Hinweis-Token zurück. Diese Aktion dient der Regeneration von Ressourcen und spielt vor allem in längeren Spielphasen eine wichtige Rolle.

Bei der Aktion „Hinweis geben“ informiert der*die aktive Spieler*in ein Teammitglied über eine Eigenschaft der eigenen Handkarten. Der Hinweis bezieht sich auf entweder die Farbe oder die Zahl der Karten. Ein Beispiel wäre die Aussage „Diese beiden Karten sind blau“ oder „Diese Karte ist eine eins“. Das Geben eines Hinweises kostet ein Hinweis-Token. Die betroffenen Karten werden während des Hinweises visuell hervorgehoben, um Missverständnisse zu vermeiden und die gemeinsame Aufmerksamkeit zu unterstützen.

Nach den Aktionen „Karte spielen“ oder „Karte abwerfen“ wird automatisch eine neue Karte gezogen. Ist der Nachziehstapel leer, wird nicht mehr nachgezogen. Nach Abschluss einer Aktion endet der Zug der aktiven Person, und die Aktionsbuttons er-

scheinen vor dem*der nächsten Spieler*in.

Sieg- und Niederlagebedingungen

Das Spiel kann auf drei unterschiedliche Arten enden:

- **Sieg:** Alle vier Farbreihen werden in aufsteigender Reihenfolge von eins bis drei vervollständigt.
- **Niederlage:** Alle drei Fehler-Tokens sind aufgebraucht.
- **Erschöpfung des Nachziehstapels:** Nach Ziehen der letzten Karte erhält jede*r Spieler*in noch einen Zug. Danach wird geprüft, ob alle Farbreihen vollständig sind. Ist dies der Fall, gilt das Team als siegreich, andernfalls endet das Spiel mit einer Niederlage.

5.1.3 Spielwelt und Narrative

Während die Kernmechanik adaptiert wurde, erhielt *Depthbound* eine eigenständige narrative und visuelle Identität, um Immersion und soziale Präsenz zu stärken. Narrative Kohärenz steigert nachweislich Motivation und Engagement in kooperativen Spielen [39].

Narrative Rahmung

Die Spielhandlung verortet die Spieler*innen in einem maritimen Fantasy-Szenario. Die Zahlen symbolisieren die Machtstufen der Kreaturen, während die vier Farben unterschiedliche Ozeanzonen repräsentieren. Das Auslegen der Karten wird als Befreiungsritual inszeniert, das dem abstrakten Sortierprozess eine thematische Bedeutung verleiht.

Abbildung 5.2: Visuelles Mockup des Kartendesigns mit repräsentativen Karten, das Stil, Farbgebung und Gestaltung des Sets zeigt.

Visuelle und auditive Gestaltung

Die visuelle Gestaltung orientiert sich an einer maritimen Dark-Fantasy-Ästhetik. Die Karten weisen gezielt Alterungsspuren auf, um eine authentische Materialwirkung zu erzielen, während vier gesättigte Primärfarben die Unterscheidbarkeit der Spielobjekte sichern. Die Umgebungsgestaltung ist bewusst minimalistisch konzipiert. Ein Holztisch mit ausgewählten Objekten fungiert als zentrale Spielfläche, auf der eine Schriftrolle wesentliche Spielinformationen anzeigt. Der umliegende Raum bleibt leer, um visuelle Ablenkungen zu minimieren und die Aufmerksamkeit auf die relevanten Spielinhalte zu lenken.

Das Sounddesign nutzt ein mehrschichtiges Ambient-System aus Schiffsgeräuschen, Wasserklängen und Wettereffekten. Wie Kern und Ellermeier [23] in ihrer empirischen Untersuchung zeigen, erhöhen kohärente Ambient-Soundscapes das Präsenzgefühl und die wahrgenommene Realitätsnähe virtueller Umgebungen signifikant, wobei die Kombination aus diffusen Hintergrundgeräuschen und ereignisbasierten Sounds die stärksten Effekte erzielt.

Die narrative Begleitung wird durch zwei komplementäre Charaktere realisiert, die unterschiedliche Feedback-Ebenen bedienen. Kapitän Mortis tritt rein akustisch als verfremdete Geisterstimme aus dem Hintergrund in Erscheinung. Seine Kommentare sind allgemein gehalten und dienen der Kontextualisierung der Handlung, indem sie Hintergründe zum Ritual und der Spielwelt erklären.

Als visuelles und humoristisches Gegenstück fungiert der animierte Papagei Flint, der zu Spielbeginn auf den virtuellen Tisch fliegt, dort als animierter NPC verweilt und direkt auf das Spielgeschehen reagiert. Er kommentiert konkrete Aktionen wie erfolgreiche Spielzüge oder Fehler mit frechen und humorvollen Einwürfen. Während der Kapitän somit die narrative Kohärenz sichert, übernimmt Flint die Funktion eines diegetischen Feedback-Systems, das den Spielzustand auf unterhaltsame Weise spiegelt.

Die narrative Integration dient nicht primär ästhetischen Zwecken, sondern unterstützt die Forschungsziele durch Vermeidung ludonarrativer Dissonanz. Nach Seraphine entsteht dieses Phänomen, wenn die Anreize der Spielmechanik (Ludik) und die Vorgaben der narrativen Ebene miteinander in Konflikt geraten. Ein solches Spannungsverhältnis führt zu Emersion, also dem Gegenteil von Immersion [43]. Die kohärente thematische Einbettung in *Depthbound* stellt sicher, dass Spielmechanik und maritimes Narrativ harmonieren, wodurch ein immersives Spielerlebnis entsteht, das natürlichere soziale Interaktionen zwischen den Spielenden begünstigt.

5.2 MR-spezifische Funktionen und Interaktionsdesign

Mixed Reality erweitert das klassische Spielkonzept um einzigartige Funktionalitäten, die in analoger Form nicht möglich wären.

5.2.1 Avatar-System für soziale Präsenz

Zur Untersuchung von H1 implementiert Bedingung C ein Avatar-System basierend auf dem Meta Avatars SDK. Die Avatare nutzen eine Ganzkörperdarstellung und verwenden das Movement SDK, um aus dem 3-Point-Tracking (HMD und Hände) mittels „Inverser

Kinematik“[1] eine Ganzkörper-Pose zu generieren [69].

Zusätzlich kommt die *Oculus-Lipsync*-Technologie zum Einsatz, um die verbale Kommunikation visuell zu unterstützen. Dabei analysiert das System die Audio-Wellenform des Mikrofons in Echtzeit und interpoliert daraus passende Mundbewegungen. Diese basieren auf sogenannten Visemen, wobei das System das Audiosignal auf 15 definierte Mundstellungen abbildet [73].

Abbildung 5.3: Full-Body-Avatare in Bedingung C. Die Körperhaltung wird mittels 3-Point-Tracking generiert.

5.2.2 Input-Modalitäten: Controller und Hand-Tracking

Zur Untersuchung von H2 implementiert der Prototyp zwei unterschiedliche Eingabemodalitäten. Die Controller-basierte Interaktion nutzt die Touch Controller der *Meta Quest 3*, die durch dedizierte Infrarot-LEDs räumlich getrackt werden und eine präzise, deterministische Eingabe mit niedriger Latenz ermöglichen. Die Implementierung erfolgt über das *Meta XR Interaction SDK* mit Ray Interactors für distanzbasierte Selektion und Grabbable-Komponenten für physikbasierte Manipulation virtueller Karten [70]. Als alternative Modalität wurde Hand-Tracking integriert, bei dem die Infrarot-Kameras des HMDs ein 26-Joint-Skelett der Hände in Echtzeit erfassen [81]. Die Hand-Tracking-Implementierung verwendet Pinch-Gesten für Kartenselektion und natürliche Greifbewegungen für Objektmanipulation.

[1] Ein Verfahren der Computeranimation, bei dem aus der Position des Endeffektors (z. B. der Hand) die notwendigen Gelenkwinkel der gesamten kinematischen Kette (z. B. Arm und Schulter) berechnet werden.

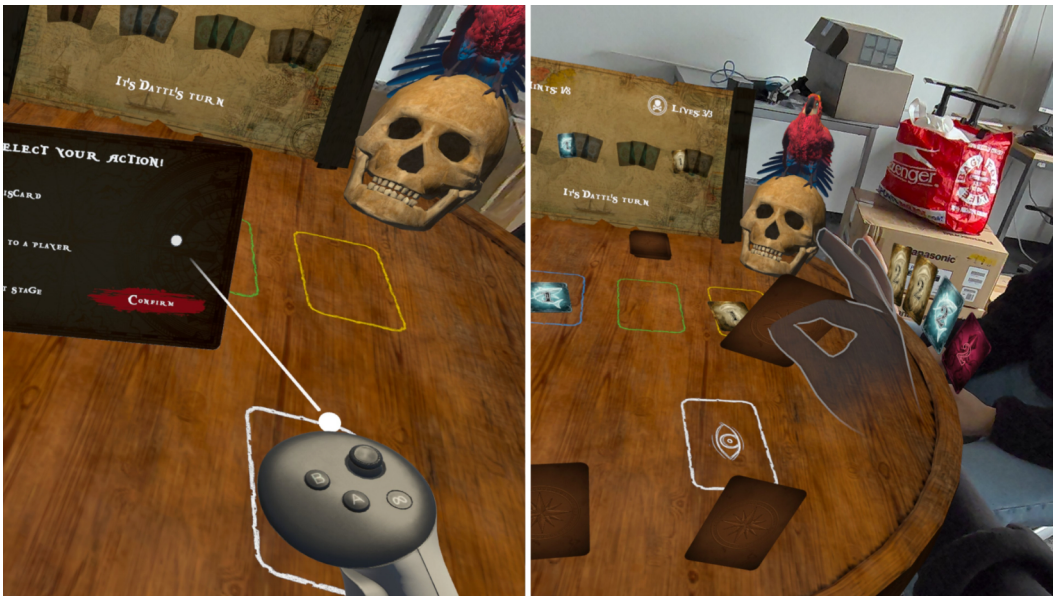


Abbildung 5.4: Vergleich der Interaktionstechniken: Controller-basierte Selektion (links) gegenüber Hand-Tracking mittels Gestik (rechts).

5.2.3 Multimodales Feedback

Basierend auf Erkenntnissen zu multimodalem Feedback in VR ist in den Bedingungen B und C multimodales Feedback integriert. Multimodale Systeme, die visuelle, auditive und haptische Rückmeldungen kombinieren, haben sich als überlegen gegenüber rein audiovisuellen Systemen erwiesen, insbesondere hinsichtlich Präsenzgefühl und Nutzerpräferenz [25]. Dies umfasst in Depthbound visuelle Animationen, auditive Sound-Effekte sowie haptische Controller-Vibrationen bei wichtigen Ereignissen wie dem Schweben über Snap-Zones. Letzteres entfällt naturgemäß, wenn Nutzer*innen Hand-Tracking verwenden. Aufgrund des Ziels, eine generell gute User Experience zu schaffen, wird der Effekt unterschiedlicher Feedback-Level nicht isoliert als Hypothese getestet, sondern multimodales Feedback als feste Design-Entscheidung implementiert.

5.3 Technische Architektur und Implementierung

Die Implementierung des Prototyps basiert auf dem offiziellen *MR Motifs Template* von *Meta*, das in *Unity 6* realisiert wurde. In das Projekt wurden mehrere spezialisierte Software Development Kits integriert. Das *Meta XR Core SDK* unterstützt SLAM-basiertes Tracking sowie Passthrough-Funktionen, während das *Meta XR Interaction SDK* die Interaktion über Hände und Controller ermöglicht. Für co-located Multiplayer-Networking kommt *Photon Fusion 2* zum Einsatz. Die Darstellung der virtuellen Präsenz mehrerer Nutzer*innen wird durch das *Meta Avatars SDK* realisiert. Insgesamt entsteht dadurch eine modulare technologische Architektur, welche die Grundlage für alle im Prototyp implementierten Funktionen bildet.

Die Architektur basiert auf einer funktionalen Trennung zwischen der Präsentations-

schicht, welche die visuelle Darstellung und Nutzerinteraktion via *Meta XR SDKs* handhabt, und der Geschäftslogikschicht, die den Spielzustand verwaltet. *Photon Fusion 2* fungiert dabei als zentrales Netzwerk-Framework und bildet das technische Fundament, auf dem die zentralen Management-Komponenten aufbauen. Die räumliche Synchronisation zwischen co-located Spielenden wird durch *Shared Spatial Anchors* realisiert, die eine konsistente Verankerung virtueller Objekte im gemeinsamen physischen Raum ermöglichen [76]. Dies garantiert, dass virtuelle Objekte wie Kartenstapel für alle Beteiligten an exakt derselben physischen Position auf dem Tisch erscheinen. Ohne diese präzise Kalibrierung würde die räumliche Konsistenz brechen, etwa wenn eine spielende Person auf eine Karte zeigt, ihr Finger für eine andere spielende Person jedoch ins Leere deutet. Die Architektur stellt sicher, dass die virtuelle Welt physikalisch plausibel in den geteilten realen Raum eingebettet ist.

In *Photon Fusion 2* wird die Synchronisation des Spielzustands durch den `NetworkRunner` und die integrierte *State Replication Pipeline* gesteuert [64]. Dabei werden Zustandsdaten deklarativ über `NetworkBehaviour`-Komponenten mittels *Networked Properties* und *NetworkArrays* automatisch repliziert. Diskrete Spieleraktionen hingegen werden via *Remote Procedure Calls* (RPCs) vom Client an den Host übermittelt [65]. Dieser Mechanismus ermöglicht die Ausführung von Funktionen auf einem entfernten System analog zu lokalen Methodenaufrufen. Der Host fungiert dabei als *State Authority*, die die alleinige Entscheidungsgewalt über den validen Spielzustand besitzt und die verbindliche Spiellogik ausführt, um Inkonsistenzen zwischen den Clients zu vermeiden.

Ein zentrales Element der serverseitigen Autorität ist die Validierung von Spieleraktionen. Programm 5.1 zeigt exemplarisch die Methode `HandleCardDrop` im `NetworkManager`, welche die Logik für das Ausspielen von Karten steuert. Hier werden Validierung und Zustandssynchronisation in einer autoritativen Methode zusammengeführt.

Programm 5.1: Autoritative Validierung eines Spielzugs im NetworkManager

```
1   // Auszug aus NetworkManager.cs
2   public void HandleCardDrop(NetworkedCard card, PlayerRef droppingPlayer,
3   SnapZoneType zoneType, byte colorIndex)
4   {
5     // Only State Authority (Server) is allowed to change state
6     if (!Object.HasStateAuthority) return;
7
8     // [...] (Validation of PlayerIndex and TurnManager removed for brevity)
9
10    CardData cardData = card.Data;
11
12    switch (zoneType)
13    {
14      case SnapZoneType.PlayZone:
15      // Validate against game rules (Color & Sequence)
16      // Create copy of current state for validation logic
17      int[] colorProgress = new int[gameConfig.colorCount];
18      // [...] Fill array
19
20      var playValidation = ValidationService.ValidatePlayCard(cardData, ...,
    colorProgress);
21
22      if (playValidation.IsValid)
23      {
24        // 1. Update card network state
25        card.Location = CardLocation.ColorStack;
26        card.TargetIndex = cardData.Color;
27
28        // 2. Update global game state
29        gameStateManager.ApplyCardPlay(cardData.Color, cardData.Value);
30      }
31      else
32      {
33        // Error case: Discard card & remove life token
34        cardManager.DiscardCount++;
35        gameStateManager.ApplyCardPlayError(cardData.Color, cardData.Value);
36      }
37      break;
38
39      case SnapZoneType.DiscardPile:
40      // Logic for discarding (regenerate hint token)
41      gameStateManager.ApplyCardDiscard(cardData.Color, cardData.Value);
42      break;
43    }
44
45    // End turn and activate next player
46    turnManager.CompleteAction();
47    card.RPC_ConfirmDrop();
48  }
49
```

5.4 Herausforderungen der co-located MR-Entwicklung

Die Entwicklung von co-located Mixed-Reality-Anwendungen ist mit spezifischen Hürden verbunden, die den aktuellen Reifegrad der Technologie widerspiegeln.

Während für VR etablierte Standards existieren, stellt co-located MR-Gaming noch eine wenig standardisierte Domäne dar. Wissenschaftliche Literatur und Gestaltungsrichtlinien für diesen spezifischen Anwendungsfall sind begrenzt. Die Entwicklung erforderte daher oft experimentelle Iterationen, insbesondere bei der Gestaltung sozialer Interaktionen, die nicht durch bestehende Frameworks abgedeckt waren.

Die Einrichtung einer Multiplayer-Umgebung mit Shared Spatial Anchors erfordert eine enge Verzahnung der Unity-Projektkonfiguration mit der zugehörigen Cloud-Infrastruktur im *Meta Developer Dashboard*. Die Abhängigkeit von korrekten Data Use Checkups und Organisations-IDs im Backend stellt eine signifikante Eintrittsbarriere dar, da lokale Fehler oft auf fehlende Berechtigungen in der Cloud-Infrastruktur zurückzuführen sind [72].

Eine zentrale Herausforderung im Workflow ist die Validierung der räumlichen Synchronisation. Während Spiellogik teilweise simuliert werden kann, erfordert das Testen von Shared Spatial Anchors und der relativen Positionierung von Avataren zwingend physische Co-Presence und vollständige Builds. Funktionen wie Remote Avatare und SLAM-Kalibrierung lassen sich nicht via *Quest Link* im Editor testen [80]. Dies erzwingt zeitintensive Build- und Deploy-Zyklen auf alle beteiligten HMDs, was den Iterationsprozess im Vergleich zur klassischen Spieleentwicklung deutlich verlangsamt.

Kapitel 6

Empirische Evaluation

6.1 Studiendesign und Durchführung

Das Studiendesign folgt einem Mixed-Methods-Ansatz innerhalb eines *Within-Subject-Designs*. Dabei werden quantitative psychometrische Daten mit qualitativem Feedback verknüpft, um die Auswirkungen unterschiedlicher Technologie-Konfigurationen systematisch zu untersuchen. Bei diesem Versuchsplan durchläuft jede teilnehmende Person alle Stufen der unabhängigen Variable, wodurch die Versuchspersonen als ihre eigene Kontrolle dienen [67].

6.1.1 Experimentelle Bedingungen

Um die in Kapitel 1 formulierten Kernhypothesen zur sozialen Interaktion zu überprüfen, wurde das Studiendesign auf drei fokussierte experimentelle Bedingungen reduziert (Tabelle 6.1).

Tabelle 6.1: Übersicht der experimentellen Bedingungen

Bedingung	Beschreibung
A (Analog)	Das Spiel wird klassisch mit physischen Karten am Tisch gespielt. Diese Bedingung bildet die grundlegende Spiellogik ab und fungiert als analoge Referenz ohne technologische Vermittlung.
B (MR–Controller)	Das Spiel wird mittels des MR-Prototyps mit integriertem multimodalen Feedback durchgeführt. Die Eingabe erfolgt ausschließlich über Controller. Avatar-Repräsentationen sind in dieser Bedingung deaktiviert, sodass nur die spielrelevanten UI-Elemente und virtuelle Spielobjekte dargestellt werden.
C (MR–Avatare)	Diese Bedingung nutzt denselben MR-Prototyp wie Bedingung B, visualisiert jedoch zusätzlich räumlich verankerte Avatare mit animierten Kopf- und Handbewegungen. Die Eingabe erfolgt ausschließlich über Hand-Tracking. Die Avatare stellen die Blickrichtung mittels Head-Tracking sowie die Gestik der Mitspielenden virtuell dar.

6.1.2 Durchführung

Die Evaluation erfolgt mit Dreiergruppen. Jede Gruppe durchläuft alle drei Bedingungen. Bedingung A (Analog) wird von allen Gruppen an erster Stelle durchgeführt, da sie als strukturierte Einführung in die Spielmechanik dient. Nach dieser Einführungsphase folgen die Bedingungen B (MR–Controller) und C (MR–Avatare) in randomisierter Reihenfolge. Dieses Vorgehen dient dazu, potenzielle Reihenfolgeeffekte[1] zu kontrollieren. Nach jeder Bedingung sowie zum Abschluss der Studie füllen die Teilnehmenden umfangreiche Fragebögen aus, die sowohl standardisierte Skalen als auch offene Feedbackfelder enthalten.

6.2 Datenanalyse und Messinstrumente

Die Datenerhebung erfolgt über ein digitales Erfassungstool, das eine präzise Zuordnung der Daten ermöglicht. Die Analyse basiert auf der methodischen Triangulation objektiver demografischer Hintergrunddaten, quantitativer psychometrischer Skalen und qualitativer Selbstauskünfte.

6.2.1 Operationalisierung der Hypothesen

Die Operationalisierung der Hypothesen erfolgt durch die Zuordnung der theoretischen Konstrukte zu messbaren Variablen und entsprechenden Erhebungsinstrumenten, wie in Tabelle 6.2 dargestellt. Im Zentrum steht die konzeptionelle Abgrenzung sozialer

[1] Effekte, die durch die Reihenfolge der Bedingungen entstehen, z. B. Übungseffekte, Ermüdung oder Übertragungseffekte zwischen Bedingungen [53].

Verbundenheit. Während die Corporeal Copresence durch die physische Anwesenheit am Tisch objektiv gegeben ist, richtet diese Arbeit den Fokus auf die soziale Präsenz. Hierfür wird eine adaptierte Version des *Networked Minds Social Presence Inventory* (NMSPI) [4] eingesetzt, ergänzt durch spezifische Items zur Wahrnehmung der Avatare und der Joint Attention. Zur Überprüfung der Eingabemodalität wird neben der *System Usability Scale* (SUS) eine direkte Präferenzabfrage (Controller vs. Hand-Tracking) eingesetzt und um eine qualitative Begründung ergänzt. Dadurch wird ermöglicht, die Präferenzentscheidung primär in Bezug auf ihrer Bedienpräzision oder aufgrund ihrer Immersion einzuordnen. Zur Kontrolle potenzieller Verzerrungen werden VR-Vorerfahrung, der Bekanntheitsgrad innerhalb der Triade sowie die Affinität zu Brett- und Videospielen als Kovariablen berücksichtigt.

Tabelle 6.2: Operationalisierungs-Matrix: Zuordnung von Hypothesen, Variablen und Messinstrumenten.

Bereich	**Variable**	**Instrument**
Allgemein	Demografie	Alter, VR-Erfahrung, Bekanntheit
	Zukunftspotenzial	Subjektive Markteinschätzung
H1	Soziale Präsenz	NMSPI (Skala 1–7)
	Avatar-Wahrnehmung	Items zu Blick & Mimik
	Qualitative Bewertung	Offenes Feedback (Vergleich)
H2	Usability & UX	SUS (Score 0–100)
	Embodiment	Eigene Items (Teil 3)
	Präferenz	Präferenzwahl + Begründung

6.2.2 Analysemethoden

Die Auswertung folgt einem kombinierten Ansatz, der quantitative Tendenzen mit qualitativen Erklärungen verknüpft. Aufgrund der geringen Stichprobengröße ($N = 9$) und des explorativen Charakters der Studie liegt der Schwerpunkt der quantitativen Analyse auf der deskriptiven Statistik. Für die psychometrischen Skalen (NMSPI, SUS) werden Mittelwerte (M) und Standardabweichungen (SD) berechnet, um zentrale Tendenzen und die Streuung der Antworten zu erfassen. Zur visuellen Darstellung der Unterschiede zwischen den experimentellen Bedingungen werden Boxplots herangezogen.

Die Antworten auf die offenen Fragen werden im Rahmen einer qualitativen Inhaltsanalyse thematisch kategorisiert. Ziel ist es, die deskriptiven Ergebnisse zu kontextualisieren und erklärende Muster für beobachtete Effekte zu identifizieren.

6.3 Ergebnisse der empirischen Studie

In diesem Abschnitt werden die quantitativen und qualitativen Ergebnisse der durchgeführten Studie ($N = 9$) vorgestellt. Die statistische Auswertung berücksichtigte sowohl Gruppenvergleiche über alle drei Bedingungen als auch direkte Paarvergleiche zweier Bedingungen. Das Signifikanzniveau wurde auf $\alpha = 0,05$ festgelegt.

6.3.1 Soziale Präsenz

Zur Überprüfung von Hypothese 1 wurden die Global Scores des NMSPI analysiert, deren Skala von 1 (sehr niedrige Präsenz) bis 7 (sehr hohe Präsenz) reicht. Die deskriptive Statistik in Tabelle 6.3 zeigt, dass die analoge Bedingung A den höchsten Mittelwert erzielte ($M = 6,49$), während die MR-Bedingungen B ($M = 5,72$) und C ($M = 5,82$) auf einem vergleichbaren Niveau liegen. Die Verteilung der Daten wird in Abbildung 6.1 illustriert.

Tabelle 6.3: Deskriptive Statistik der NMSPI Global Scores ($N = 9$).

Bedingung	**Mittelwert (M)**	**Standardabw. (SD)**	**Median**
A: Analog	6,49	0,44	6,56
B: MR-Controller	5,72	0,79	5,67
C: MR-Avatare	5,82	0,53	5,94

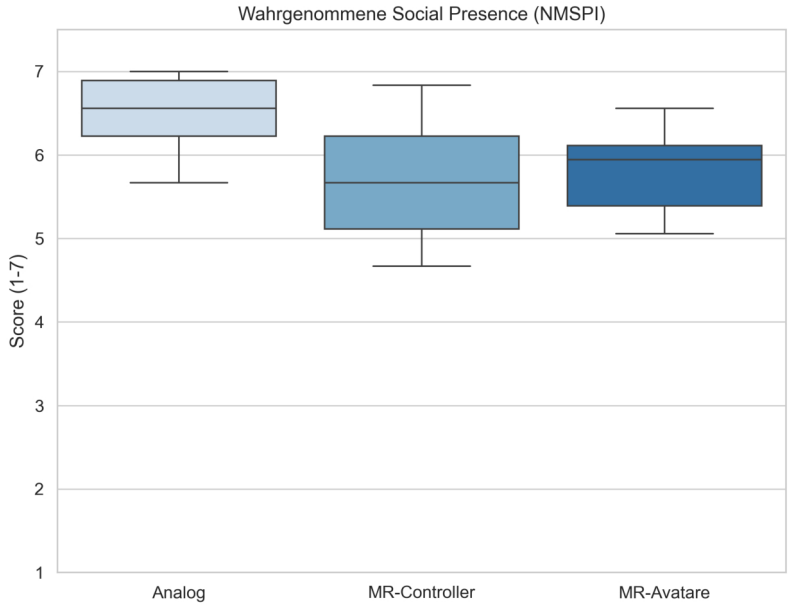


Abbildung 6.1: Boxplot der NMSPI Global Scores in den drei Bedingungen

Die statistische Analyse bestätigte, dass ein signifikanter Unterschied zwischen den drei Bedingungen vorliegt ($p = 0,005$).

Die detaillierte Analyse der Paarvergleiche lieferte folgende Ergebnisse:

- **Bedingung B vs. C:** Es konnte kein signifikanter Unterschied zwischen der Bedingung ohne Avatare und der Bedingung mit Avataren festgestellt werden ($p = 0,910$).
- **Bedingung A vs. C:** Zwischen der analogen Bedingung und der MR-Umgebung mit Avataren liegt ein signifikanter Unterschied vor ($p = 0,008$), wobei die soziale Präsenz im analogen Setting höher bewertet wurde.

Abbildung 6.2 zeigt die detaillierte Aufschlüsselung der NMSPI-Subskalen. In der Dimension Co-Presence erzielten Bedingung A ($M = 6,46$) und Bedingung B ($M = 6,37$) höhere Werte als Bedingung C ($M = 5,89$).

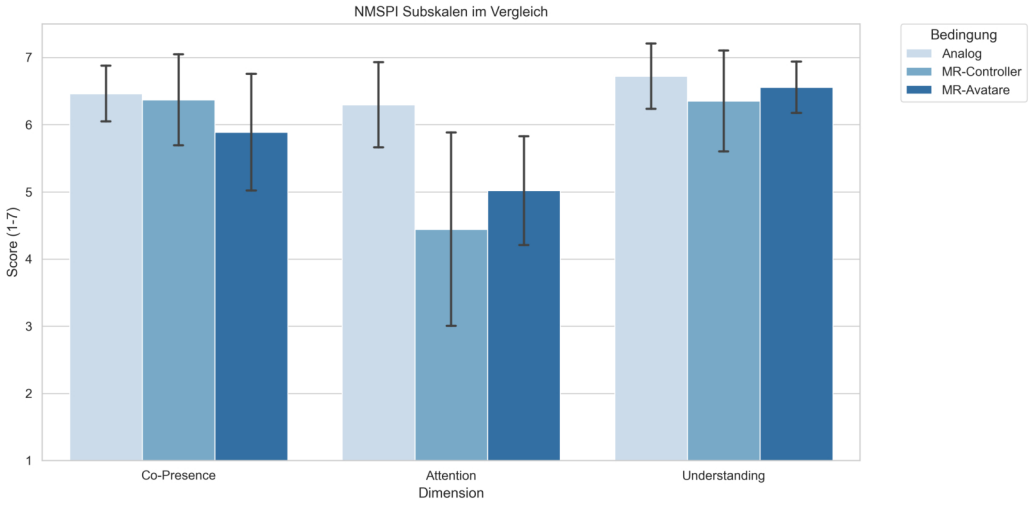


Abbildung 6.2: Detaillierter Vergleich der NMSPI-Subskalen: Co-Presence, Attention Allocation und Perceived Understanding

6.3.2 Wahrnehmung der Avatare

Die Bewertung der Avatare in Bedingung C (Skala 1–7) ergab differenzierte Ergebnisse. Die Unterstützung bei der Erkennung der Blickrichtung wurde mit $M = 4,56$ neutral bis leicht positiv bewertet. Niedrigere Werte zeigten sich bei der Wahrnehmung von Kopfbewegungen ($M = 3,33$) sowie bei der Einschätzung, ob Avatare die Präsenz der Mitspielenden verstärken ($M = 4,00$).

Hinsichtlich der Notwendigkeit von Avataren sprachen sich alle 9 Teilnehmer*innen (100 %) für die Option „Optional einschaltbar“ aus. Die Antwortmöglichkeiten „Pflicht“ oder „Nicht verwenden“ erhielten keine Zustimmungen.

6.3.3 Usability und Eingabemethode

Zur Überprüfung von Hypothese 2 wurden die Scores der SUS sowie die explizite Präferenzabfrage ausgewertet.

System Usability Scale (SUS)

Die SUS-Scores (Skala 0–100) zeigten für beide Eingabemethoden gute bis sehr gute Werte:

- **Bedingung B (Controller):** $M = 77,2$
- **Bedingung C (Hand-Tracking):** $M = 78,9$

Der statistische Vergleich ergab keinen signifikanten Unterschied ($p = 0,496$). Beide Systeme wurden von den Teilnehmenden als gleichermaßen gebrauchstauglich eingestuft.

Subjektive Präferenz

Am Ende der Studie wurden die Teilnehmenden gefragt, welche Eingabemethode sie für das Spiel *Depthbound* bevorzugen würden. Trotz der ähnlichen Usability-Bewertungen sprach sich die Mehrheit für das Hand-Tracking aus:

- **7 von 9 Teilnehmer*innen (77,8 %)** bevorzugten das Hand-Tracking.
- **2 von 9 Teilnehmer*innen (22,2 %)** bevorzugten den Controller.

6.4 Methodische Reflexion und Limitationen

Das gewählte Studiendesign erfordert eine kritische Einordnung hinsichtlich der Stichprobengröße, der Teilnehmerauswahl und der technologischen Rahmenbedingungen.

Eine zentrale Limitation der Studie stellt die Stichprobengröße von $N = 9$ Teilnehmern dar. Diese Begrenzung resultiert aus der hohen zeitlichen und logistischen Komplexität des Versuchsaufbaus. Um trotz der kleinen Stichprobe valide Aussagen treffen zu können, wurde ein Within-Subject-Design gewählt, das die Varianz zwischen den Subjekten eliminiert [67]. Dennoch sind die statistischen Ergebnisse primär als explorative Tendenzen zu interpretieren.

Ein vollfaktorielles Design wurde zugunsten der Studiendauer und physiologischer Belastungsgrenzen verworfen. Der Vergenz-Akkommodations-Konflikt[2] sowie das Gewicht des HMDs führen bei längerer Nutzung zu Ermüdung, wodurch die Datenqualität abnehmen würde. Außerdem würde eine höhere Anzahl an Bedingungen die kognitive Last erhöhen und das soziale Erleben überlagern. Das gestraffte Design ermöglicht valide Daten innerhalb der Belastungsgrenzen.

Die Rekrutierung erfolgte mittels Gelegenheitsstichprobe im Umfeld der Fachhochschule, was zu einer vergleichsweise technikaffinen Stichprobe führte. Um einer möglichen Verzerrung entgegenzuwirken, wurde die individuelle Vorerfahrung mit VR/MR-Systemen im demografischen Fragebogen explizit erhoben und bei der Auswertung qualitativ berücksichtigt.

[2] Bezeichnet die physiologische Entkopplung von Akkommodation (Fokus auf Displayebene) und Vergenz (Blick auf virtuelles Objekt), die visuelle Ermüdung begünstigt.

Kapitel 7

Diskussion der Ergebnisse

In diesem Kapitel werden die in Kapitel 6.3 dargestellten Ergebnisse interpretiert und in Bezug zu den formulierten Hypothesen sowie zum aktuellen Forschungsstand gesetzt. Die Diskussion gliedert sich in die Interpretation der quantitativen und qualitativen Befunde, die Einordnung in den theoretischen Rahmen und eine kritische Reflexion der Grenzen dieser Studie.

7.1 Interpretation der sozialen Präsenz

Die quantitative Auswertung von Hypothese 1 konnte die Annahme, dass Avatare die soziale Präsenz signifikant steigern, in dieser Studie nicht bestätigen. Es fand sich kein statistisch signifikanter Unterschied zwischen der Bedingung mit Avataren (C) und der Bedingung ohne Avatare (B). Zudem blieb die analoge Bedingung (A) unerreicht und wurde signifikant höher bewertet ($M = 6{,}49$).

Für dieses Ergebnis lassen sich durch die Synthese der qualitativen Rückmeldungen und der quantitativen Daten vier zentrale Erklärungsansätze identifizieren:

1. **Prädominanz der auditiven Kommunikation:** Da sich die Spieler*innen im selben physischen Raum befanden und direkt miteinander sprechen konnten, war der akustische Kanal bereits sehr stark ausgeprägt. Die visuelle Repräsentation durch abstrakte Avatare lieferte in diesem co-located Setting zu wenig zusätzlichen Informationsgehalt, um einen messbaren Unterschied in der empfundenen Präsenz zu bewirken.
2. **Räumliche Inkongruenz:** Ein wesentlicher Störfaktor war die Diskrepanz zwischen dem physischen Körper und dem überlagerten Avatar. Im Gegensatz zu reiner VR fehlt in MR die visuelle Konsistenz, da das Gehirn permanent den realen Arm mit dem virtuellen Avatar-Arm abgleicht. Da diese aufgrund technischer Limitierungen nie perfekt deckungsgleich sind, entsteht eine kognitive Dissonanz. Das Gehirn muss kontinuierlich verarbeiten, welche Position als Referenz dient, was die kognitive Belastung erhöht und die Immersion bricht.
3. **Latenz als „Störfaktor“:** In einem co-located Setting erfolgt die reale Interaktion ohne Zeitverzögerung. Die Bewegungen des Avatars unterliegen jedoch der Netzwerklatenz der Synchronisation. Dies führt dazu, dass deiktische Gesten, die für Tabletop-Spiele essenziell sind, unpräzise werden. Der physische Finger deutet

bereits auf das Ziel, während der Avatar noch in der Bewegung ist. Diese Asynchronizität führt zu Missverständnissen und entwertet den Avatar als zuverlässiges Kommunikationswerkzeug.

4. **Fehlende mimische Expressivität:** Während analoge Spieler*innen feine Nuancen im Gesichtsausdruck ihres Gegenübers lesen konnten, verdeckte das HMD diese Informationen. Da die *Meta Quest 3* im Gegensatz zur *Quest Pro* über keine internen Kameras für Face-Tracking verfügt, konnten stille nonverbale Signale (wie skeptisches Stirnrunzeln oder ein Lächeln ohne Worte) nicht übertragen werden. Softwareseitige Lösungen wie *Audio-to-Expression* wurden als unzureichend betrachtet, da sie Mimik nur während des Sprechens simulieren, nicht jedoch in den für das Spiel kritischen Momenten der stillen Beobachtung (8 von 9 Nennungen im qualitativen Feedback).

Dennoch berichteten Teilnehmer*innen qualitativ, dass Avatare zumindest als grober Anker für Joint Attention fungierten, wie etwa bei der Frage „Schaut er*sie nach links oder rechts?“, auch wenn dies den globalen NMSPI-Score nicht signifikant anhob.

7.2 Vergleichende Analyse der Eingabemodalitäten

Bezüglich Hypothese 2 offenbarte die Studie ein interessantes Spannungsfeld zwischen objektiver Messung und subjektivem Erleben. Obwohl die quantitativen SUS-Scores für beide Eingabemethoden nahezu identisch waren (beide im Bereich „Gut“ bis „Exzellent“), bevorzugte eine deutliche Mehrheit der Teilnehmer*innen (77,8 %) das Hand-Tracking.

Spannungsfeld zwischen Intuitivität und Präzision

Diese Diskrepanz legt nahe, dass Hand-Tracking zwar nicht effizienter oder fehlerfreier ist als die Nutzung von Controllern, jedoch als hedonisch wertvoller empfunden wird. Das Greifen von Spielkarten stellt eine physisch vertraute Metapher dar und reduzierte die mentale Hürde der Mediennutzung.

Besonders deutlich zeigte sich dies darin, dass Probanden ohne Vorkenntnisse des Spiels *Hanabi* sehr hohe Usability-Werte vergaben, was die intuitive Natur gestengesteuerter Interfaces bestätigt.

Die Minderheit der Teilnehmer*innen, die Controller bevorzugte, argumentierte hingegen leistungsorientiert. Für diese Nutzergruppe waren haptisches Feedback und die Präzision der Tasten ausschlaggebend. Daraus lässt sich ableiten, dass mit zunehmendem kompetitiven Anspruch die Verlässlichkeit der Eingabemethode an Bedeutung gewinnt. Für das soziale Erlebnis und die Geselligkeit scheint jedoch die Barrierefreiheit des Hand-Trackings vorteilhafter zu sein.

Technische Limitationen und visuelle Interferenzen

Im qualitativen Feedback wurde ein spezifisches Phänomen der Mixed Reality besonders hervorgehoben, das sogenannte „Ghosting“ der eigenen Hände. Da das System die virtuellen Avatar-Hände über das Videobild der realen Hände legte, kam es insbesondere bei auftretenden Latenzen zu störenden visuellen Dopplungen.

Dieses Problem verdeutlicht die Notwendigkeit präziserer Okklusionsalgorithmen sowie einer verbesserten Segmentierung in zukünftigen MR-Systemen, da nur durch eine genauere visuelle Trennung von realen und virtuellen Elementen die Immersion nachhaltig aufrechterhalten werden kann.

7.3 Akzeptanz virtueller Repräsentationen

Ein besonders eindeutiges Ergebnis ergab sich aus der Frage nach der Notwendigkeit von Avataren. Die Gesamtheit der Teilnehmer*innen sprach sich dafür aus, diese optional ein- und ausschaltbar zu gestalten. Dies verdeutlicht, dass visuelle Repräsentationen in co-located MR-Settings nicht als fundamentale Notwendigkeit wie in reinen Remote-VR-Umgebungen wahrgenommen werden, sondern vielmehr als optionales Zusatzmerkmal.

Die Analyse der Feature-Wichtigkeit zeigte zudem, dass die Kopfrotation und die Bewegungs-Smoothness als wichtiger erachtet wurden als die Blickrichtung. Dies deutet darauf hin, dass grobe Aufmerksamkeitsrichtungen (z. B. ob der Fokus auf den eigenen oder fremden Spielobjekten liegt) für die Spieler*innen relevanter sind als detaillierter Augenkontakt, solange keine Mimik übertragen wird.

Diese Ergebnisse legen zudem nahe, dass für co-located Szenarien eine Unterscheidung zwischen einer vollständig virtuellen Präsenz und einer erweiterten Präsenz notwendig ist. Während Remote-Szenarien eine vollständige virtuelle Repräsentation erfordern, um das Gegenüber sichtbar zu machen, wird diese in der Co-Location oft als redundantes „Over-Engineering“ wahrgenommen. Vollständige Avatare verdecken den echten Menschen und erzeugen durch die genannten Tracking-Ungenauigkeiten eher Distanz. Minimalistische Hinweise, die die physische Präsenz im Passthrough lediglich ergänzen statt sie zu ersetzen, scheinen hier der zielführendere Ansatz zu sein, um das „Uncanny Valley“ der asynchronen Überlagerung zu vermeiden.

7.4 Akzeptanz und Zukunftspotenzial aus Nutzersicht

Zum Abschluss der Studie wurden die Teilnehmer*innen gebeten, das Potenzial von co-located Mixed Reality subjektiv zu bewerten. Da die Proband*innengruppe über technologische Vorerfahrung verfügte, spiegeln ihre Aussagen die Einschätzung einer technologieaffinen Zielgruppe wider, lassen jedoch keine generalisierbaren Rückschlüsse auf die allgemeine Marktentwicklung zu.

Subjektive Barrieren der Adoption

Auf die Frage, ob derartige Systeme innerhalb der nächsten fünf Jahre den Massenmarkt erreichen werden, reagierten die Teilnehmer*innen zurückhaltend. Der errechnete Mittelwert von $M = 2{,}33$ (auf einer Skala von 1 = „Unwahrscheinlich“ bis 5 = „Sehr wahrscheinlich“) deutet darauf hin, dass die Technologie von den Teilnehmenden im aktuellen Entwicklungsstadium noch primär als Nischenprodukt wahrgenommen wird.

Die qualitative Analyse der Rückmeldungen identifiziert die zentralen Gründe für diese zurückhaltende Nutzungsmotivation:

- **Hardwareergonomie (44 %):** Als am häufigsten genanntes Hemmnis erwies sich

der Formfaktor der Head-Mounted Displays. Die Rückmeldungen der Teilnehmenden bezogen sich überwiegend auf Aspekte wie die wahrgenommene Sperrigkeit, das Gewicht sowie den eingeschränkten Tragekomfort der Geräte. Solange HMDs aufgrund ihrer physischen Präsenz das unmittelbare soziale Zusammensein beeinträchtigen, wird die Nutzung im Kontext eines informellen, gemeinschaftlichen Brettspielabends von den Nutzenden als wenig attraktiv eingeschätzt.

- **Kostenbarriere (33 %):** Ein weiteres zentrales Hindernis stellt die ökonomische Zugangsschwelle dar. Insbesondere die Voraussetzung, dass jede teilnehmende Person über ein eigenes, kostenintensives Headset verfügen muss, was hohe Anschaffungskosten für mehrere Geräte verursacht, wurde als wesentlicher limitierender Faktor für die private Organisation lokaler Mehrspieler-Szenarien identifiziert.
- **Mangel an Inhalten (22 %):** Demgegenüber verortete lediglich eine Minderheit der Befragten die geringe Verbreitung primär in einem unzureichenden Angebot an Software oder spielerischen Inhalten. Fehlende oder qualitativ unzureichende Anwendungen wurden somit als vergleichsweise nachrangige Ursache eingeschätzt.

Zusammenfassend lässt sich festhalten, dass der operative Aufwand derzeit noch als zu hoch eingeschätzt wird, um MR-Lösungen in geselligen Runden klassischen physischen Spielen vorzuziehen. Diese zurückhaltende Einschätzung der Studienteilnehmenden steht in einem interessanten Spannungsfeld zur Perspektive von Branchenexpert*innen. Stephen Linton identifiziert die *Meta Quest 3* aufgrund ihrer Passthrough-Qualität und des Preispunkts als ersten echten Wendepunkt hin zur Massenmarkt-Tauglichkeit für co-located MR. Während die Proband*innen dieser Studie primär die aktuellen ergonomischen Hürden betonten, sehen Entwickler*innen in der Hardware erstmals die notwendige Basis, um Konzepte umzusetzen, die gänzlich ohne klassische Benutzeroberflächen auskommen und somit auch Personen ohne Gaming-Erfahrung erreichen [58].

Potenzial trotz Hardware-Grenzen

Wesentlich ist die Differenzierung zwischen der Kritik an der Hardware und der Akzeptanz des zugrundeliegenden Spielkonzepts. Während die identifizierten Hürden überwiegend extrinsischer Natur waren (z. B. Preis und Gewicht), wurde die intrinsische Motivation zur Nutzung von MR-Spielen insgesamt positiv bewertet.

Die qualitativen Daten zeigen, dass das hybride Erlebnis aus digitalen Effekten und einem physischen Spieltisch grundsätzlich als bereichernd wahrgenommen wird, sofern die Einstiegshürden ausreichend niedrig sind. Daraus ergibt sich die Schlussfolgerung, dass die Akzeptanz von co-located Mixed Reality in naher Zukunft weniger von der Komplexität der Avatarsysteme abhängen wird, sondern vielmehr davon, wie reibungslos sich die Technologie in bestehende soziale Situationen integrieren lässt.

Da Spieleentwickler*innen keinen direkten Einfluss auf den Hardware-Formfaktor haben, kommt Software-Designentscheidungen eine besondere Bedeutung zu, um diese hardwareseitigen Reibungspunkte zu kompensieren. Die aus dieser Studie abgeleiteten Design-Implikationen werden im folgenden Kapitel systematisch zusammengefasst.

Kapitel 8

Designleitfaden für co-located MR-Tabletop-Spiele

Basierend auf den theoretischen Grundlagen, den technischen Erfahrungen der Implementierung sowie der empirischen Evaluation von *Depthbound* werden im Folgenden systematische Gestaltungsprinzipien formuliert. Der Leitfaden verknüpft Designtheorie, technische Rahmenbedingungen und Nutzerwahrnehmung. Er richtet sich an XR-Entwickler*innen und Game Designer*innen, die soziale Interaktion in co-located Mixed-Reality-Szenarien gestalten.

8.1 Strategien für soziale Präsenz und Repräsentation

Die Studie zeigte, dass virtuelle Avatare die soziale Präsenz im Vergleich zur physischen Realität nicht zwangsläufig erhöhen, sondern durch technische Unzulänglichkeiten wie Latenz oder fehlende Mimik sogar störend wirken können. Daraus ergeben sich folgende Gestaltungsrichtlinien:

Prinzip der erweiterten Präsenz statt virtueller Präsenz

In co-located Szenarien mit Video-Passthrough ist der physische Körper der Mitspielenden bereits sichtbar. Eine vollständige Überlagerung durch Avatare führt häufig zu Diskrepanzen zwischen realer und virtueller Wahrnehmung und kann subtile nonverbale Signale wie Blickrichtung oder Mimik verdecken [10, 40].

Empfehlung: Verzichten Sie auf vollständige Avatare, wenn kein präzises Face- und Eye-Tracking verfügbar ist. Nutzen Sie stattdessen minimalistische, akzentuierende Elemente, die die physische Mimik nicht verdecken, sondern visuell unterstützen.

Vermeidung des „Uncanny Valley“ durch Abstraktion

Da aktuelle Consumer-Hardware wie die *Meta Quest 3* keine native Gesichtserfassung bietet, wirken statische Avatar-Gesichter häufig wie emotionslose Masken und können Irritationen im Sinne des „Uncanny Valley“ hervorrufen [6].

Empfehlung: Wählen Sie stilisierte oder abstrakte Darstellungsformen, da diese bei fehlender Gesichtsanimation oft eine höhere soziale Akzeptanz und Attraktivität

aufweisen als realistische Darstellungen [10]. Wenn Avatare eingesetzt werden, sollten sie optional ein- und ausschaltbar sein, wie die einhellige Zustimmung der Proband*innen zu diesem Feature in der Studie zeigte.

Priorisierung der Latenzminimierung bei Gestik

Latenzen bei der Synchronisation von Avataren führen dazu, dass deiktische Gesten unpräzise oder zeitlich versetzt wahrgenommen werden, da die reale Handbewegung bereits ausgeführt wird, während die entsprechende Avatar-Geste zeitlich verzögert dargestellt wird.

Empfehlung: Optimieren Sie den Netzwerkcode gezielt auf minimale Latenz bei Positions- und Rotationsupdates der Hände, selbst auf Kosten der visuellen Qualität, um Missverständnisse in der Kommunikation zu vermeiden.

8.2 Interaktionsmodalitäten und Verkörperung

Die Evaluation offenbarte ein Spannungsfeld zwischen objektiver Gebrauchstauglichkeit und subjektivem Erleben. Obwohl Controller- und Hand-Tracking ähnliche Leistungswerte erzielten, wurde Hand-Tracking von den Teilnehmenden signifikant bevorzugt.

Natürlichkeit als Einstiegsmechanismus

Hand-Tracking senkt die Einstiegshürde, da die zugrunde liegende Interaktionsmetapher aus der physischen Welt vertraut ist. Die Visualisierung der eigenen Hände ohne störende Hardware fördert dabei nachweislich das Gefühl des Embodiment [6].

Empfehlung: Nutzen Sie Hand-Tracking als primäre Eingabemethode für gesellige, strategische Spiele, bei denen soziale Interaktion wichtiger ist als wettbewerbsorientierte Präzision.

Vermeidung von visuellen Doppelungen

Die visuelle Überlagerung der realen Hand im Passthrough und einer virtuellen Handdarstellung kann die Immersion erheblich beeinträchtigen.

Empfehlung: Verwenden Sie präzise Handmaskierung zur Okklusion oder blenden Sie virtuelle Hände nur dann ein, wenn sie aktiv mit virtuellen Objekten interagieren, um visuelle Interferenzen zu minimieren.

Multimodales Feedback als Ersatz für Haptik

Bei der Nutzung von Hand-Tracking fehlt das physische Feedback, das bei der Interaktion mit realen Objekten selbstverständlich ist. Untersuchungen belegen jedoch, dass multimodale Systeme, die visuelle und auditive Rückmeldungen kombinieren, die wahrgenommene Natürlichkeit der Interaktion erhöhen [25]. Zudem gilt es, dem identifizierten Problem der „klebrigen“ Hände beim Loslassen virtueller Objekte [41] durch gezieltes Feedback entgegenzuwirken.

Empfehlung: Kompensieren Sie fehlende Haptik durch bewusst überzeichnetes audiovisuelles Feedback (z. B. Soundeffekte beim Berühren, visuelles Hervorheben von Snap-Zones), um Handlungssicherheit und das Gefühl von Kontrolle zu gewährleisten.

8.3 Synchronisation und Umgang mit dem physischen Raum

Das Fundament eines jeden Tabletop-Spiels ist der gemeinsame Tisch als zentraler Informations- und Interaktionsraum. Dieser wird in der Literatur als öffentlicher Raum beschrieben, der eine lückenlose gegenseitige Wahrnehmung ermöglicht [36].

Konsistenz durch geteilte räumliche Anker

Eine geteilte Realität entsteht nur dann, wenn virtuelle Objekte für alle Beteiligten an exakt derselben physischen Position verankert sind [32].

Empfehlung: Nutzen Sie Shared Spatial Anchors als technische Basis. Das virtuelle Spielfeld muss physikalisch plausibel im Raum verortet sein und darf nicht an die Kopfbewegung einzelner Spieler*innen gekoppelt werden.

Zentrierung der Aufmerksamkeit

Head-Mounted Displays schränken das Sichtfeld ein und begünstigen einen „Tunnelblick“. Meta warnt in den Guidelines explizit davor, dass dies die Wahrnehmung von Mitspielern im peripheren Bereich einschränkt [77].

Empfehlung: Platzieren Sie alle spielrelevanten Informationen zentral auf dem virtuellen Tisch zwischen den Spielenden. Vermeiden Sie schwebende Menüs direkt vor den Gesichtern, da diese den Blickkontakt blockieren und die soziale Dynamik der Gruppe stören.

8.4 Kognitive Entlastung und Hybridität

Während die Nutzung von Head-Mounted Displays durch visuelle Reize und geteilte Aufmerksamkeit potenziell eine erhöhte kognitive Belastung verursachen kann [21], liegt der Mehrwert hybrider Systeme in der gezielten Entlastung der Spielenden.

Automatisierung von Verwaltungsaufwänden

Digitale Systeme eignen sich besonders für repetitive und regelbasierte Aufgaben, die in analogen Spielen häufig den Spielfluss unterbrechen. Dies ermöglicht es der Gruppe, als „verteilter kognitiver Prozessor“ zu fungieren und sich auf die Strategie zu konzentrieren [3].

Empfehlung: Nutzen Sie die Technik für die Automatisierung von Verwaltungsaufgaben wie Ressourcentracking, Spielaufbau und Regelprüfung, damit sich die Spielenden primär auf Strategie, Kommunikation und soziale Interaktion konzentrieren können.

Transparenz des Spielzustands

In hybriden Spielsystemen ist für Nutzer*innen oft nicht nachvollziehbar, welche Regeln das System intern anwendet oder warum bestimmte Aktionen nicht erlaubt sind. In diesem Kontext wird vor „Black-Box"-Entscheidungen gewarnt, die das Vertrauen untergraben [37].

Empfehlung: Visualisieren Sie Regelverstöße und Spielzüge sofort und eindeutig im gemeinsamen Raum. Das System muss nachvollziehbar bleiben, um Vertrauen in die digitale Spielleitung und Akzeptanz für automatisierte Entscheidungen zu schaffen.

8.5 Atmosphäre und Feedback-Design

Die visuelle Dominanz von Mixed Reality verleitet dazu, die atmosphärische Ebene zu vernachlässigen. Die Studie zeigte jedoch, dass Audio und Feedback-Gestaltung eine tragende Rolle für die soziale Präsenz und die Wahrnehmung der Umgebung spielen.

Einsatz von Spatial Audio und Soundscapes

Während das Sichtfeld technologisch begrenzt ist, umgibt der Klang die Spielenden vollständig. Kohärente Soundscapes können die physische Umgebung emotional aufladen und das Gefühl des gemeinsamen „Eintauchens" verstärken, ohne die visuelle Aufmerksamkeit vom Tisch abzulenken [23].

Empfehlung: Nutzen Sie räumliches Audio, um virtuelle Ereignisse im physischen Raum zu verorten. Ergänzen Sie funktionale Sounds durch atmosphärische Hintergrundgeräusche, um technische Stille zu vermeiden und die Immersion zu stabilisieren.

Diegetisches Feedback und Interface

Rein abstrakte UI-Sounds oder visuelle Overlays können den Immersionsbruch verstärken. Feedback, das aus der Spielwelt selbst kommt und damit diegetisch ist, wirkt natürlicher und weniger störend.

Empfehlung: Integrieren Sie Feedback-Systeme in die Spielwelt, anstatt sie nur als 2D-Overlay einzublenden. Personifizierte Feedback-Geber (wie z. B. ein begleitender NPC oder Erzähler) können technische Hinweise oder Spielstatus-Updates liefern, ohne die soziale Illusion zu zerstören.

8.6 Narrative Kontextualisierung

Die eingesetzte Mixed-Reality-Technologie sollte nicht als Selbstzweck verstanden werden. Um soziale Interaktionen als kohärent und natürlich erfahrbar zu machen, bedarf die ungewohnte Situation, wie das Tragen von Head-Mounted Displays sowie die Einbindung virtueller Objekte, einer narrativen Rahmung.

Vermeidung ludonarrativer Dissonanz

Wenn die Spielhandlung und die technische Ausführung im Widerspruch stehen, entsteht eine Distanz zum Erlebnis. Eine thematische Rahmung, die erklärt, warum virtuelle

Technologie in den physischen Raum eindringt, hilft den Spielenden, die Hardware zu akzeptieren.

Empfehlung: Entwickeln Sie ein Narrativ, das die Mixed-Reality-Situation rechtfertigt (z. B. durch ein Ritual oder ein Sci-Fi-Hologramm). Eine starke thematische Einbettung fördert die intrinsische Motivation [39] und lässt technische Hürden wie das Tragen eines HMDs eher in den Hintergrund treten.

Kapitel 9

Fazit und Ausblick

Die vorliegende Arbeit untersuchte das Potenzial von Mixed Reality, die soziale Interaktion analoger Tabletop-Spiele technologisch zu augmentieren, ohne deren inhärente Qualitäten zu beeinträchtigen. Die Entwicklung und Evaluation des Prototyps *Depthbound* verdeutlicht, dass eine erfolgreiche Übertragung in den hybriden Raum an spezifische gestalterische Rahmenbedingungen geknüpft ist.

Die empirische Studie zeigte, dass das analoge Original bezüglich der sozialen Präsenz weiterhin unerreicht blieb. Dies unterstreicht, dass die physische Co-Presence der primäre Referenzrahmen ist. Virtuelle Repräsentationen sollten daher die soziale Bandbreite der physischen Mimik unterstützen, statt sie durch starre Avatare zu verdecken. Gleichzeitig erwies sich das Hand-Tracking gegenüber Controllern als die präferierte Eingabemethode, da die natürliche Gestik die Einstiegshürde senkt und das soziale Erleben trotz technischer Limitationen fördert. Da Barrieren wie Latenzen oder ergonomische Einschränkungen aktueller HMDs den Interaktionsfluss stören können, erweist sich eine erweiterte Präsenz häufig als effizienter gegenüber einer vollständigen Virtualisierung.

Blickend auf die Zukunft steht das Feld der co-located Mixed Reality vor einer technologischen Transition. Während aktuelle Headsets durch ihren Formfaktor und das eingeschränkte Sichtfeld den direkten Blickkontakt oft noch erschweren, könnten kommende Optical See-Through-Systeme und die Etablierung von Facial-Tracking diese Barriere auflösen. Ein weiterführendes Potenzial liegt zudem im Einsatz generativer KI als kontextbewusste Spielleitung, die Gespräche und Handlungen versteht, um die Spielwelt dynamisch anzupassen. Zusammenfassend ist co-located MR nicht als Ersatz für das analoge Spiel zu verstehen, sondern als komplementäres Medium, das soziale Wärme mit digitaler Immersion verbindet.

Anhang A

Fragebogen

Der im Rahmen der Evaluation verwendete Fragebogen ist im Folgenden vollständig dargestellt und wurde als externes PDF-Dokument in den Anhang eingebunden.

Depthbound Studie - Datenerhebung

Gesamtdokumentation des Fragebogens (Anhang)

Teil 1: Demografische Angaben

Bitte füllen Sie vor Beginn der ersten Runde diesen demografischen Fragebogen aus.

Teilnehmer-Name: _____

Triade (Gruppen-Nr): _____

Datum: _____

Persönliche Angaben

1. Alter:

_____ Jahre

2. Geschlecht:

○ Männlich ○ Weiblich ○ Divers ○ Keine Angabe

VR/MR-Erfahrung

3. Wie oft haben Sie in den letzten 12 Monaten Virtual Reality (VR) oder Mixed Reality (MR) Geräte genutzt?

○ Nie
○ Einmal
○ 2-5 Mal
○ 6-10 Mal
○ Mehr als 10 Mal
○ Regelmäßig (mindestens 1x pro Monat)

4. Welche VR/MR-Systeme haben Sie bereits genutzt? (Mehrfachauswahl möglich)

☐ Meta Quest (1/2/3/Pro)
☐ PlayStation VR
☐ HTC Vive
☐ Valve Index
☐ Microsoft HoloLens
☐ Andere

Gaming-Erfahrung

5. Wie viele Stunden pro Woche spielen Sie durchschnittlich digitale Spiele? (PC, Konsole, Mobile)

○ 0 Stunden (Ich spiele keine digitalen Spiele)
○ 1-3 Stunden
○ 4-7 Stunden
○ 8-14 Stunden
○ Mehr als 14 Stunden

6. Wie oft spielen Sie Brettspiele oder Kartenspiele?

○ Nie
○ Selten (weniger als 1x pro Monat)
○ Monatlich (1-3x pro Monat)
○ Wöchentlich
○ Mehrmals pro Woche

7. Kennen Sie das Kartenspiel "Hanabi"?

○ Ja, ich habe es bereits gespielt
○ Ja, ich kenne die Regeln, aber habe es noch nie gespielt
○ Ich habe davon gehört, kenne aber die Regeln nicht
○ Nein, noch nie davon gehört

8. Wie häufig spielen Sie kooperative Multiplayer-Spiele? (Spiele, in denen Sie MIT anderen gegen das Spiel spielen)

○ Nie ○ Selten ○ Manchmal ○ Häufig ○ Sehr häufig

Bekanntheit mit Mitspielern

9. Wie gut kennen Sie die beiden anderen Teilnehmer in Ihrer Gruppe?

Mitspieler A: _____
(Skala: Sehr gut, Gut, Oberflächlich, Gar nicht)

Mitspieler B: _____
(Skala: Sehr gut, Gut, Oberflächlich, Gar nicht)

Fragebogen nach Bedingung A (Analog)

Bitte bewerten Sie die folgenden Aussagen über Ihre Erfahrung während des analogen Spiels.

Frage	*Stimme überhaupt nicht zu* 1	2	3	4	5	6	7 *Stimme voll und ganz zu*
1. Ich hatte das Gefühl, dass die anderen Spieler mit mir im selben Raum waren.	○	○	○	○	○	○	○
2. Die anderen Spieler schienen für mich präsent zu sein.	○	○	○	○	○	○	○
3. Ich fühlte mich mit den anderen Spielern verbunden.	○	○	○	○	○	○	○
4. Ich hatte das Gefühl, dass ich wirklich mit anderen Menschen zusammen war.	○	○	○	○	○	○	○
5. Ich konnte die Absichten der anderen Spieler verstehen.	○	○	○	○	○	○	○
6. Die Kommunikation mit den anderen Spielern war klar und verständlich.	○	○	○	○	○	○	○
7. Ich konnte leicht nachvollziehen, worauf die anderen Spieler aufmerksam waren.	○	○	○	○	○	○	○
8. Die nonverbalen Signale (Gesten, Blicke) der anderen waren für mich gut erkennbar.	○	○	○	○	○	○	○
9. Ich konnte die Emotionen der anderen Spieler gut wahrnehmen.	○	○	○	○	○	○	○
10. Ich hatte das Gefühl, dass die anderen Spieler meine Gefühle verstehen konnten.	○	○	○	○	○	○	○
11. Es war einfach, die Stimmung der anderen Spieler einzuschätzen.	○	○	○	○	○	○	○
12. Die Handlungen der anderen Spieler hatten direkten Einfluss auf meine Entscheidungen.	○	○	○	○	○	○	○
13. Ich musste meine Aktionen mit den anderen Spielern koordinieren.	○	○	○	○	○	○	○
14. Mein Verhalten war eng mit dem Verhalten der anderen Spieler verknüpft.	○	○	○	○	○	○	○
15. Wir haben als Team zusammengearbeitet.	○	○	○	○	○	○	○
16. Ich fühlte mich während des Spiels in die Interaktion eingebunden.	○	○	○	○	○	○	○

Frage	*Stimme überhaupt nicht zu*	1	2	3	4	5	6	7	*Stimme voll und ganz zu*
17. Die Interaktion mit den anderen Spielern fühlte sich natürlich an.		○	○	○	○	○	○	○	
18. Ich war mental in die gemeinsame Aktivität involviert.		○	○	○	○	○	○	○	

Fragebogen nach Bedingung B (MR-Basic)

Teil 1: Bitte bewerten Sie die Usability des Systems.

Frage	*Stimme überhaupt nicht zu*	1	2	3	4	5	*Stimme voll und ganz zu*
1. Ich denke, dass ich das System gerne häufiger nutzen würde.		○	○	○	○	○	
2. Ich fand das System unnötig komplex.		○	○	○	○	○	
3. Ich fand das System einfach zu nutzen.		○	○	○	○	○	
4. Ich glaube, ich würde die Hilfe einer technisch versierten Person benötigen, um das System nutzen zu können.		○	○	○	○	○	
5. Ich fand, die verschiedenen Funktionen in diesem System waren gut integriert.		○	○	○	○	○	
6. Ich denke, das System enthielt zu viele Inkonsistenzen.		○	○	○	○	○	
7. Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen.		○	○	○	○	○	
8. Ich fand das System sehr umständlich zu nutzen.		○	○	○	○	○	
9. Ich fühlte mich bei der Nutzung des Systems sehr sicher.		○	○	○	○	○	
10. Ich musste eine Menge lernen, bevor ich mit dem System zurechtkam.		○	○	○	○	○	

Teil 2: Bitte bewerten Sie das soziale Erleben.

Frage	*Stimme überhaupt nicht zu*	1	2	3	4	5	6	7	*Stimme voll und ganz zu*
1. Ich hatte das Gefühl, dass die anderen Spieler mit mir im selben Raum waren.		○	○	○	○	○	○	○	
2. Die anderen Spieler schienen für mich präsent zu sein.		○	○	○	○	○	○	○	
3. Ich fühlte mich mit den anderen Spielern verbunden.		○	○	○	○	○	○	○	
4. Ich hatte das Gefühl, dass ich wirklich mit anderen Menschen zusammen war.		○	○	○	○	○	○	○	
5. Ich konnte die Absichten der anderen Spieler verstehen.		○	○	○	○	○	○	○	
6. Die Kommunikation mit den anderen Spielern war klar und verständlich.		○	○	○	○	○	○	○	
7. Ich konnte leicht nachvollziehen, worauf die anderen Spieler aufmerksam waren.		○	○	○	○	○	○	○	

Frage	*Stimme überhaupt nicht zu*	1	2	3	4	5	6	7	*Stimme voll und ganz zu*
8. Die nonverbalen Signale (Gesten, Blicke) der anderen waren für mich gut erkennbar.		○	○	○	○	○	○	○	
9. Ich konnte die Emotionen der anderen Spieler gut wahrnehmen.		○	○	○	○	○	○	○	
10. Ich hatte das Gefühl, dass die anderen Spieler meine Gefühle verstehen konnten.		○	○	○	○	○	○	○	
11. Es war einfach, die Stimmung der anderen Spieler einzuschätzen.		○	○	○	○	○	○	○	
12. Die Handlungen der anderen Spieler hatten direkten Einfluss auf meine Entscheidungen.		○	○	○	○	○	○	○	
13. Ich musste meine Aktionen mit den anderen Spielern koordinieren.		○	○	○	○	○	○	○	
14. Mein Verhalten war eng mit dem Verhalten der anderen Spieler verknüpft.		○	○	○	○	○	○	○	
15. Wir haben als Team zusammengearbeitet.		○	○	○	○	○	○	○	
16. Ich fühlte mich während des Spiels in die Interaktion eingebunden.		○	○	○	○	○	○	○	
17. Die Interaktion mit den anderen Spielern fühlte sich natürlich an.		○	○	○	○	○	○	○	
18. Ich war mental in die gemeinsame Aktivität involviert.		○	○	○	○	○	○	○	

Fragebogen nach Bedingung C (MR-Sozial)

Teil 1: Bitte bewerten Sie die Usability des Systems.

Frage	*Stimme überhaupt nicht zu*	1	2	3	4	5	*Stimme voll und ganz zu*
1. Ich denke, dass ich das System gerne häufiger nutzen würde.		○	○	○	○	○	
2. Ich fand das System unnötig komplex.		○	○	○	○	○	
3. Ich fand das System einfach zu nutzen.		○	○	○	○	○	
4. Ich glaube, ich würde die Hilfe einer technisch versierten Person benötigen, um das System nutzen zu können.		○	○	○	○	○	
5. Ich fand, die verschiedenen Funktionen in diesem System waren gut integriert.		○	○	○	○	○	
6. Ich denke, das System enthielt zu viele Inkonsistenzen.		○	○	○	○	○	
7. Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen.		○	○	○	○	○	
8. Ich fand das System sehr umständlich zu nutzen.		○	○	○	○	○	
9. Ich fühlte mich bei der Nutzung des Systems sehr sicher.		○	○	○	○	○	

Frage	*Stimme überhaupt nicht zu*	1	2	3	4	5	*Stimme voll und ganz zu*
10. Ich musste eine Menge lernen, bevor ich mit dem System zurechtkam.		○	○	○	○	○	

Teil 2: Bitte bewerten Sie das soziale Erleben.

Frage	*Stimme überhaupt nicht zu*	1	2	3	4	5	6	7	*Stimme voll und ganz zu*
1. Ich hatte das Gefühl, dass die anderen Spieler mit mir im selben Raum waren.		○	○	○	○	○	○	○	
2. Die anderen Spieler schienen für mich präsent zu sein.		○	○	○	○	○	○	○	
3. Ich fühlte mich mit den anderen Spielern verbunden.		○	○	○	○	○	○	○	
4. Ich hatte das Gefühl, dass ich wirklich mit anderen Menschen zusammen war.		○	○	○	○	○	○	○	
5. Ich konnte die Absichten der anderen Spieler verstehen.		○	○	○	○	○	○	○	
6. Die Kommunikation mit den anderen Spielern war klar und verständlich.		○	○	○	○	○	○	○	
7. Ich konnte leicht nachvollziehen, worauf die anderen Spieler aufmerksam waren.		○	○	○	○	○	○	○	
8. Die nonverbalen Signale (Gesten, Blicke) der anderen waren für mich gut erkennbar.		○	○	○	○	○	○	○	
9. Ich konnte die Emotionen der anderen Spieler gut wahrnehmen.		○	○	○	○	○	○	○	
10. Ich hatte das Gefühl, dass die anderen Spieler meine Gefühle verstehen konnten.		○	○	○	○	○	○	○	
11. Es war einfach, die Stimmung der anderen Spieler einzuschätzen.		○	○	○	○	○	○	○	
12. Die Handlungen der anderen Spieler hatten direkten Einfluss auf meine Entscheidungen.		○	○	○	○	○	○	○	
13. Ich musste meine Aktionen mit den anderen Spielern koordinieren.		○	○	○	○	○	○	○	
14. Mein Verhalten war eng mit dem Verhalten der anderen Spieler verknüpft.		○	○	○	○	○	○	○	
15. Wir haben als Team zusammengearbeitet.		○	○	○	○	○	○	○	
16. Ich fühlte mich während des Spiels in die Interaktion eingebunden.		○	○	○	○	○	○	○	
17. Die Interaktion mit den anderen Spielern fühlte sich natürlich an.		○	○	○	○	○	○	○	
18. Ich war mental in die gemeinsame Aktivität involviert.		○	○	○	○	○	○	○	

Fragebogen nach Bedingung C (MR-Sozial) - Details

Teil 3: Avatare und Nonverbale Kommunikation

In dieser Bedingung waren Ihre Mitspieler durch Avatare repräsentiert.

Frage	*Sehr selten / fast nie*	1	2	3	4	5	*Sehr häufig / ständig*
1. Wie oft haben Sie die Avatare Ihrer Mitspieler bewusst wahrgenommen?		○	○	○	○	○	

Frage	*Stimme überhaupt nicht zu*	1	2	3	4	5	6	7	*Stimme voll und ganz zu*
2. Die Avatare haben sich realistisch bewegt.		○	○	○	○	○	○	○	
3. Ich konnte die Avatare eindeutig meinen Mitspielern zuordnen.		○	○	○	○	○	○	○	
4. Die Avatare haben mich vom Spielgeschehen abgelenkt.		○	○	○	○	○	○	○	
5. Durch die Avatare konnte ich Kopfbewegungen meiner Mitspieler (z.B. Nicken, Kopfschütteln) besser wahrnehmen.		○	○	○	○	○	○	○	
6. Durch die Avatare konnte ich erkennen, in welche Richtung meine Mitspieler schauten.		○	○	○	○	○	○	○	
7. Die Avatare haben mir geholfen, nonverbale Hinweise meiner Mitspieler zu verstehen.		○	○	○	○	○	○	○	
8. Ich habe selbst bewusst Kopfbewegungen eingesetzt, um mit meinen Mitspielern zu kommunizieren (z.B. Nicken zur Zustimmung).		○	○	○	○	○	○	○	
9. Die Körpersprache meiner Mitspieler war durch die Avatare gut erkennbar.		○	○	○	○	○	○	○	
10. Die Avatare haben das Gefühl verstärkt, dass meine Mitspieler wirklich 'da' sind.		○	○	○	○	○	○	○	
11. Durch die Avatare fühlte sich die Interaktion persönlicher an.		○	○	○	○	○	○	○	
12. Die Avatare haben mir geholfen, eine emotionale Verbindung zu meinen Mitspielern aufzubauen.		○	○	○	○	○	○	○	
13. Ich habe die Avatare eher als 'virtuelle Repräsentationen' meiner Mitspieler wahrgenommen statt als abstrakte Platzhalter.		○	○	○	○	○	○	○	

VERGLEICHSBEWERTUNG

Frage	*Viel schlechter*	1	2	3	4	5	6	7	*Viel besser*
14. Im Vergleich zu Bedingung B (MR OHNE Avatare): Das Gefühl von sozialer Präsenz war in Bedingung C (MIT Avataren)...		○	○	○	○	○	○	○	
15. Im Vergleich zu Bedingung B: Die nonverbale Kommunikation war in Bedingung C...		○	○	○	○	○	○	○	

Frage	*Viel schlechter*	1	2	3	4	5	6	7	*Viel besser*
16. Im Vergleich zu Bedingung B: Das Gefühl von Zusammengehörigkeit im Team war in Bedingung C...		○	○	○	○	○	○	○	
17. Im Vergleich zum analogen Spiel (Bedingung A): Die soziale Interaktion in Bedingung C (MR mit Avataren) war...		○	○	○	○	○	○	○	
18. Im Vergleich zum analogen Spiel: Die Wahrnehmung nonverbaler Signale (Blicke, Gesten, Körpersprache) war in Bedingung C...		○	○	○	○	○	○	○	

VERGLEICH MIT BEDINGUNG A (ANALOG)

19. Was war im analogen Spiel besser für die soziale Interaktion als in MR mit Avataren?

20. Was war in MR mit Avataren besser für die soziale Interaktion als im analogen Spiel?

AVATAR-DESIGN

21. Welche Avatar-Features waren für Sie am wichtigsten? (Wählen Sie die Top 3)

- ☐ Kopfbewegungen (z.B. Nicken, Drehen)
- ☐ Blickrichtung
- ☐ Position im Raum
- ☐ Visuelle Gestaltung (Aussehen des Avatars)
- ☐ Bewegungsflüssigkeit (keine Ruckler)

22. Was hat Ihnen am Avatar-Design gefehlt?

- ☐ Gesichtsausdrücke / Mimik
- ☐ Handbewegungen / Gesten
- ☐ Körperhaltung (Sitzen/Stehen)
- ☐ Individualisierung (verschiedene Avatare pro Person)
- ☐ Nichts - die Avatare waren ausreichend

Sonstiges:

Frage	*Nie*	1	2	3	4	5	*Die ganze Zeit über*
23. Haben Sie während des Spiels vergessen, dass Sie ein VR-Headset tragen und sich 'wirklich' mit Ihren Mitspielern am Tisch gefühlt?		○	○	○	○	○	

Frage	*Stark verschlechtert*	1	2	3	4	5	6	7	*Stark verbessert*
24. Die Avatare haben die soziale Interaktion im Spiel insgesamt...		○	○	○	○	○	○	○	

Frage	*Sehr schlecht*	1	2	3	4	5	6	7	*Sehr gut*
26. Wie würden Sie die Avatar-Implementierung insgesamt bewerten?		○	○	○	○	○	○	○	

25. Avatare sollten in MR-Multiplayer-Spielen wie Depthbound...

○ Pflicht sein ○ Optional sein ○ Nicht genutzt werden ○ Weiß nicht

WEITERE FAKTOREN

Frage	*Sehr unsicher*	1	2	3	4	5	6	7	*Sehr sicher*
27. Wie sicher haben Sie sich gefühlt, sich im physischen Raum zu bewegen/zu agieren? (Räumliche Sicherheit)		○	○	○	○	○	○	○	

Frage	*Stimme gar nicht zu*	1	2	3	4	5	6	7	*Stimme voll zu*
28. Ich konnte jederzeit nachvollziehen, was meine Mitspieler gerade fokussieren/planen. (Social Transparency)		○	○	○	○	○	○	○	

Frage	*Gar nicht*	1	2	3	4	5	6	7	*Sehr stark*
29. Hat das eingeschränkte Sichtfeld (FOV) des Headsets Ihre soziale Interaktion beeinträchtigt?		○	○	○	○	○	○	○	

Abschluss: Eingabemodalität und Embodiment

In Bedingung B haben Sie Controller genutzt, in Bedingung C Ihre eigenen Hände (Hand-Tracking).

Frage	*Controller (Bedingung B)*	1	2	3	4	5	6	7	*Hand-Tracking (Bedingung C)*
1. Welche Eingabemethode empfanden Sie als natürlicher für dieses Spiel?		○	○	○	○	○	○	○	
2. Mit welcher Methode konnten Sie Karten präziser greifen und ablegen?		○	○	○	○	○	○	○	
3. Bei welcher Methode mussten Sie weniger über die Steuerung nachdenken?		○	○	○	○	○	○	○	
4. Bei welcher Methode hatten Sie stärker das Gefühl, dass die virtuelle Repräsentation (Hände bzw. Controller) Teil Ihres eigenen Körpers war?		○	○	○	○	○	○	○	
5. Bei welcher Methode stimmten die visuellen Bewegungen gefühlt besser mit Ihren tatsächlichen körperlichen Bewegungen überein?		○	○	○	○	○	○	○	
6. Bei welcher Methode hatten Sie ein stärkeres Gefühl von direkter Kontrolle über das Spielgeschehen (Sense of Agency)?		○	○	○	○	○	○	○	

7. Wenn Sie das Spiel noch einmal spielen müssten, welche Methode würden Sie wählen?

○ Controller (Bedingung B)
○ Hand-Tracking (Bedingung C)

Warum?

Abschluss: Expertenmeinung & Ausblick

ZUKUNFTSPOTENZIAL

Frage	*Sehr unwahrscheinlich*	1	2	3	4	5	*Sehr wahrscheinlich*
1. Glauben Sie, dass lokale Multiplayer-MR-Spiele in den nächsten 5 Jahren den Massenmarkt erreichen?		○	○	○	○	○	

2. Was ist Ihrer Meinung nach aktuell das größte Hindernis für den Durchbruch dieser Technologie?

- ○ Hardware noch zu unkomfortabel / schwer
- ○ Fehlende gute Spiele/Inhalte
- ○ Einrichtung / Kalibrierung zu kompliziert
- ○ Anschaffungskosten für mehrere Brillen zu hoch
- ○ Soziale Hemmungen ("Man sieht albern aus")

Anderes Hindernis...

SPIELDESIGN & GENRE

3. Welche Art von Spielen eignet sich Ihrer Meinung nach am besten für lokales MR?

- ○ **Tabletop-Fokus:** Ruhiges Sitzen am Tisch (wie Brettspiele), virtuelle Elemente auf dem Tisch.
- ○ **Bewegungs-Fokus:** Aktives Bewegen durch den ganzen Raum, Ausweichen, Sport.
- ○ **Hybrid:** Mischung aus beidem.

4. Haben Sie eine konkrete Idee für ein klassisches Brett- oder Kartenspiel, das als MR-Version besonders viel Spaß machen würde?

z.B. 'Dungeons & Dragons, weil man die Monster sehen könnte' oder 'Schach mit animierten Figuren'...

5. Raum für sonstige Anmerkungen, Feedback zur Studie oder Wünsche:

Quellenverzeichnis

Literatur

[1] Haley Adams u. a. „Depth Perception in Augmented Reality: The Effects of Display, Shadow, and Position“. In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Christchurch, New Zealand, 12.–16. März 2022). New York, NY: IEEE, März 2022, S. 1–10. DOI: 10.1109/VR51125.2022.00011 (siehe S. 4).

[2] Antoine Bauza. *Spielanleitung Hanabi*. ABACUSSPIELE Verlags GmbH & Co. KG. Dreieich, 2012. URL: https://www.brettspiele-report.de/images/h/hanabi/Spielanleitung-Hanabi.pdf (siehe S. 29).

[3] Matthew Berland und Victor R. Lee. „Collaborative strategic board games as a site for distributed computational thinking“. *International Journal of Game-Based Learning* 1.2 (2011), S. 65–81. DOI: 10.4018/ijgbl.2011040105 (siehe S. 21, 22, 50).

[4] Frank Biocca und Chad Harms. *Guide to the Networked Minds Social Presence Inventory (Version 1.2). Measures of co-presence, social presence, subjective symmetry, and intersubjective symmetry.* Techn. Ber. ML-RR-2000-xx. East Lansing, MI: Media Interface und Network Design (M.I.N.D.) Labs, Michigan State University, 2003. URL: https://web-archive.southampton.ac.uk/cogprints.org/6743/1/2002_guide_netminds_measure.pdf (besucht am 12. 02. 2026) (siehe S. 40).

[5] Hugo Brument, Francesco De Pace und Iana Podkosova. „Does mixed reality influence joint action? Impact of the mixed reality setup on users' behavior and spatial interaction“. *Journal on Multimodal User Interfaces* 19 (2024), S. 119–138. DOI: 10.1007/s12193-024-00445-w (siehe S. 17, 21, 22).

[6] Gavin Buckingham. „Hand Tracking for Immersive Virtual Reality: Opportunities and Challenges“. *Frontiers in Virtual Reality* 2 (2021). DOI: 10.3389/frvir.2021.728461 (siehe S. 9–11, 48, 49).

[7] Cesar Cadena u. a. „Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age“. *IEEE Transactions on Robotics* 32.6 (2016), S. 1309–1332. DOI: 10.1109/TRO.2016.2624754 (siehe S. 5, 6).

[8] Long Cheng, Michael Schreiner und Andreas Kunz. „Comparing Tracking Accuracy in Standalone MR-HMDs: Apple Vision Pro, HoloLens 2, Meta Quest 3, and Pico 4 Pro“. In: *Proceedings of the 30th ACM Symposium on Virtual Reality Soft-*

ware and Technology (VRST '24) (Trier, Germany, 9.–11. Okt. 2024). New York, NY: ACM, Okt. 2024, S. 1–2. DOI: 10.1145/3641825.3689518 (siehe S. 5, 6).

[9] Liam Cross u. a. „Your Move: An Open Access Dataset of Over 1500 Board Gamer's Demographics, Preferences and Motivations". *Simulation & Gaming* 54.5 (2023), S. 1–22. DOI: 10.1177/10468781231189493 (siehe S. 13).

[10] Lucas Detto. „Local collaboration in a Mixed Reality environment: Adding virtual heads to improve social presence". Magisterarb. Stockholm, Sweden: KTH Royal Institute of Technology, 2023 (siehe S. 16, 17, 21–23, 48, 49).

[11] Shahin Doroudian. „Collaboration in Immersive Environments: Challenges and Solutions". *arXiv* (2025). DOI: 10.48550/arXiv.2311.00689 (siehe S. 10, 11).

[12] Morad Elfleet und Mathieu Chollet. „Investigating the Impact of Multimodal Feedback on User-Perceived Latency and Immersion with LLM-Powered Embodied Conversational Agents in Virtual Reality". In: *Proceedings of the 24th ACM International Conference on Intelligent Virtual Agents (IVA '24)* (Glasgow, Scotland, 16.–19. Sep. 2024). New York, NY: ACM, Sep. 2024. DOI: 10.1145/3652988.3673965 (siehe S. 20).

[13] Barrett Ens u. a. „Revisiting Collaboration through Mixed Reality: The Evolution of Groupware". *International Journal of Human-Computer Studies* 131 (2019), S. 81–98. DOI: 10.1016/j.ijhcs.2019.05.011 (siehe S. 4, 5).

[14] Austin Erickson u. a. „Sharing Gaze Rays for Visual Target Identification Tasks in Collaborative Augmented Reality". In: *Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Virtual Event, Brazil, 9.–13. Nov. 2020). Piscataway, NJ: IEEE, Nov. 2020. DOI: 10.1109/ISMAR50242.2020.00083 (siehe S. 21).

[15] Adélaïde Genay, Anatole Lécuyer und Martin Hachet. „Virtual, Real or Mixed: How Surrounding Objects Influence the Sense of Embodiment in Optical See-Through Experiences?" *Frontiers in Virtual Reality* 2 (2021). DOI: 10.3389/frvir.2021.679902 (siehe S. 9).

[16] Moinak Ghoshal u. a. „Co-located Immersive Gaming: A Comparison between Augmented and Virtual Reality". In: *Proceedings of the 2022 IEEE Conference on Games (CoG)* (Virtual Event, China, 21.–24. Aug. 2022). Piscataway, NJ: IEEE, Aug. 2022. DOI: 10.1109/CoG51982.2022.9893708 (siehe S. 24).

[17] Janet K. Gibbs, Marco Gillies und Xueni Pan. „A comparison of the effects of haptic and visual feedback on presence in virtual reality". *International Journal of Human-Computer Studies* 157 (2022). DOI: 10.1016/j.ijhcs.2021.102717 (siehe S. 20).

[18] Eric Godden, William Steedman und Matthew K. X. J. Pan. „Robotic Characterization of Markerless Hand-Tracking on Meta Quest Pro and Quest 3 Virtual Reality Headsets". *IEEE Transactions on Visualization and Computer Graphics* 31.5 (2025), S. 3025–3034. DOI: 10.1109/TVCG.2025.3549182 (siehe S. 11).

[19] Erving Goffman. *Behavior in Public Places: Notes on the Social Organization of Gatherings*. New York, NY: Free Press, 1963, S. 248 (siehe S. 15).

[20] Matt Gottsacker u. a. „Decoupled Hands: An Approach for Aligning Perspectives in Collaborative Mixed Reality“. In: *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)* (Yokohama, Japan, 26. Apr.–1. Mai 2025). New York, NY: ACM, Apr. 2025. DOI: 10.1145/3706599.3720219 (siehe S. 19).

[21] Allison Jing u. a. „Eye See What You See: Exploring How Bi-Directional Augmented Reality Gaze Visualisation Influences Co-Located Symmetric Collaboration“. *Frontiers in Virtual Reality* 2 (2021). DOI: 10.3389/frvir.2021.697367 (siehe S. 17, 21, 22, 50).

[22] Alvin Jude, G. Michael Poor und Darren Guinness. „Grasp, Grab or Pinch? Identifying User Preference for In-Air Gestural Manipulation“. In: *Proceedings of the 2016 ACM Symposium on Spatial User Interaction (SUI '16)* (Tokyo, Japan, 15.–16. Okt. 2016). New York, NY: ACM, Okt. 2016. DOI: 10.1145/2983310.2989209 (siehe S. 10).

[23] Angelika C. Kern und Wolfgang Ellermeier. „Audio in VR: Effects of a Soundscape and Movement-Triggered Step Sounds on Presence“. *Frontiers in Robotics and AI* 7 (2020), S. 20. DOI: 10.3389/frobt.2020.00020 (siehe S. 32, 51).

[24] Myungho Lee u. a. „Mixed Reality Tabletop Gameplay: Social Interaction With a Virtual Human Capable of Physical Influence“. *IEEE Transactions on Visualization and Computer Graphics* 24.8 (Aug. 2019). DOI: 10.1109/TVCG.2019.2959575 (siehe S. 15).

[25] Daniel Martin u. a. „Multimodality in VR: A Survey“. *ACM Computing Surveys* 54.8 (2022). DOI: 10.1145/3508361 (siehe S. 34, 49).

[26] Mark McGill, Jan Gugenheimer und Euan Freeman. „A Quest for Co-Located Mixed Reality: Aligning and Assessing SLAM Tracking for Same-Space Multi-User Experiences“. In: *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology (VRST '20)* (Virtual Event, Canada, 1.–4. Nov. 2020). New York, NY: ACM, Nov. 2020, S. 1–10. DOI: 10.1145/3385956.3418968 (siehe S. 5).

[27] Paul Milgram u. a. „Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum“. In: *Telemanipulator and Telepresence Technologies* (Boston, MA, 31. Okt.–1. Nov. 1994). Bd. 2351. Bellingham, WA: SPIE, 1994, S. 282–292. DOI: 10.1117/12.197321 (siehe S. 3).

[28] Amitabh Mishra und Kevin Foster Carff. „AprilTags in Unity: A Local Alternative to Shared Spatial Anchors for Synergistic Shared Space Applications Involving Extended Reality and the Internet of Things“. *Sensors* 25.14 (2025). DOI: 10.3390/s25144408 (siehe S. 8).

[29] Donald A. Norman. *The Design of Everyday Things*. Revised and expanded edition. Basic Books, 2013 (siehe S. 29).

[30] Niklas Osmers u. a. „The Role of Social Presence for Cooperation in Augmented Reality on Head Mounted Devices: A Literature Review“. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)* (Virtual Event, Japan, 8.–13. Mai 2021). New York, NY: ACM, Mai 2021. DOI: 10.1145/3411764.3445633 (siehe S. 21).

[31] Theofilos Papadopoulos u. a. „Interactions in Augmented and Mixed Reality: An Overview“. *Applied Sciences* 11.18 (2021), S. 8752. DOI: 10.3390/app11188752 (siehe S. 9).

[32] Carlos A. Pinheiro de Sousa u. a. „Co-Located VR with Hybrid SLAM-based HMD Tracking and Motion Capture Synchronization“. In: *Proceedings of the GI VR/AR Workshop 2025* (Chemnitz, Germany, 16.–17. Sep. 2025). Bonn, Germany: Gesellschaft für Informatik, Sep. 2025 (siehe S. 4–6, 8, 50).

[33] Thammathip Piumsomboon u. a. „The Effects of Sharing Awareness Cues in Collaborative Mixed Reality“. *Frontiers in Robotics and AI* 6 (2019). DOI: 10.3389/frobt.2019.00005 (siehe S. 16, 21).

[34] Dennis Reimer. „Hand Tracking in Colocated Multi-User Virtual Reality“. Diss. Wien, Österreich: Technische Universität Wien, 2024. URL: https://repositum.tuwien.at/handle/20.500.12708/198755 (siehe S. 10, 11).

[35] Grégoire Richard u. a. „Studying the Role of Haptic Feedback on Virtual Embodiment in a Drawing Task“. *Frontiers in Virtual Reality* 1 (2021). DOI: 10.3389/frvir.2020.573167 (siehe S. 11).

[36] Melissa J. Rogerson, Martin Gibbs und Wally Smith. „"I love all the bits": The Materiality of Boardgames“. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)* (San Jose, CA, USA, 7.–12. Mai 2016). New York, NY: ACM, Mai 2016. DOI: 10.1145/2858036.2858433 (siehe S. 1, 13, 50).

[37] Melissa J. Rogerson, Laura A. Sparrow und Martin R. Gibbs. „More Than a Gimmick: Digital Tools for Boardgame Play“. *Proceedings of the ACM on Human-Computer Interaction* 5.CHI PLAY (2021), S. 1–23. DOI: 10.1145/3474688 (siehe S. 1, 14, 22, 51).

[38] Jannick P. Rolland und Henry Fuchs. „Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization“. *Presence: Teleoperators and Virtual Environments* 9.3 (2000), S. 287–309. DOI: 10.1162/105474600566808 (siehe S. 4).

[39] Richard M. Ryan, C. Scott Rigby und Andrew Przybylski. „The Motivational Pull of Video Games: A Self-Determination Theory Approach“. *Motivation and Emotion* 30.4 (2006). DOI: 10.1007/s11031-006-9051-8 (siehe S. 31, 52).

[40] Monique Santoso und Jeremy Bailenson. „How Video Passthrough Headsets Influence Perception of Self and Others“. *Cyberpsychology, Behavior, and Social Networking* (2024). DOI: 10.1089/cyber.2024.0398 (siehe S. 4, 16, 19, 48).

[41] Alexander Schäfer, Gerd Reis und Didier Stricker. „Comparing Controller with the Hand Gestures Pinch and Grab for Picking Up and Placing Virtual Objects“. In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (Christchurch, New Zealand, 12.–16. März 2022). New York, NY: IEEE, März 2022, S. 738–739. DOI: 10.1109/VRW55335.2022.00220 (siehe S. 10, 11, 49).

[42] Dieter Schmalstieg und Tobias Höllerer. *Augmented Reality: Principles and Practice*. Boston, MA: Addison-Wesley Professional, 2016 (siehe S. 4, 5).

[43] Frédéric Seraphine. „Ludonarrative Dissonance: Is Storytelling About Reaching Harmony?“ (2016). URL: https://www.fredericseraphine.com/index.php/2016/09/02/ludonarrative-dissonance-is-storytelling-about-reaching-harmony/ (siehe S. 32).

[44] Gayoung Son, Aline Tiemann und Marius Rubo. „I am here with you: an examination of factors relating to social presence in social VR“. *Frontiers in Virtual Reality* 6 (2025). DOI: 10.3389/frvir.2025.1558233 (siehe S. 15).

[45] Sparsh Srivastava. *Quantifying Social Presence in Mixed Reality: A Contemporary Review of Techniques and Innovations*. arXiv preprint arXiv:2404.15325. 2024. DOI: 10.48550/arXiv.2404.15325 (siehe S. 15, 16).

[46] Anthony Steed und Jiaqi Lai. „Comparison of hand tracking-based and controller-based interaction in a consumer virtual reality game“. *Virtual Reality* 29.3 (2025). DOI: 10.1007/s10055-025-01190-5 (siehe S. 10, 11).

[47] Hans Strasburger. „Seven Myths on Crowding and Peripheral Vision“. *i-Perception* 11.3 (2020), S. 1–46. DOI: 10.1177/2041669520913052 (siehe S. 17).

[48] Michael Tomasello. „Joint Attention as Social Cognition“. In: *Joint Attention: Its Origins and Role in Development*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1995, S. 103–130 (siehe S. 16).

[49] Ali Tourani u. a. „Visual SLAM: What Are the Current Trends and What to Expect?“ *Sensors* 22.23 (2022), S. 9297. DOI: 10.3390/s22239297 (siehe S. 5).

[50] Matthew Warburton u. a. „Measuring motion-to-photon latency for sensorimotor experiments with virtual reality systems“. *Behavior Research Methods* 55 (2022), S. 3658–3678. DOI: 10.3758/s13428-022-01983-5 (siehe S. 11).

[51] Julie R. Williamson u. a. „Digital Proxemics: Designing Social and Collaborative Interaction in Virtual Environments“. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)* (New Orleans, LA, USA, 1.–5. Mai 2022). New York, NY: ACM, Mai 2022. DOI: 10.1145/3491102.3517594 (siehe S. 16).

[52] Stewart Woods. *Eurogames. The Design, Culture and Play of Modern European Board Games*. Jefferson, NC: McFarland & Company, 2012 (siehe S. 1).

[53] Lydia Yang. „Within-Subjects Design“. In: *EDCEP 917: Experimental Design*. Kansas State University, 2025. Kap. 5, S. 50–61. URL: https://stats.libretexts.org/Courses/Kansas_State_University/EDCEP_917:_Experimental_Design_(Yang) (besucht am 12. 02. 2026) (siehe S. 39).

[54] Jianing Yin u. a. „A Comparison Study Understanding the Impact of Mixed Reality Collaboration on Sense of Co-Presence“. In: *Proceedings of the 2025 IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR 2025)* (Saint-Malo, France, 8.–12. März 2025). Piscataway, NJ: IEEE, März 2025 (siehe S. 15).

[55] Seyed Alireza Ghasempour Yousefdeh und Solomon Sunday Oyelere. „Investigating co-presence and collaboration dynamics in realtime virtual reality user interactions“. *Frontiers in Virtual Reality* 5 (2024). DOI: 10.3389/frvir.2024.1478481 (siehe S. 15, 22).

[56] Shanyang Zhao. „Toward a Taxonomy of Copresence“. *Presence: Teleoperators and Virtual Environments* 12.5 (2003). DOI: 10.1162/105474603322761261 (siehe S. 15).

Medien

[57] *Depthbound: Demonstration des co-located Mixed Reality Prototyps*. 23. Jan. 2026. URL: https://youtu.be/Yg420FuoetM (siehe S. 28).

[58] XR AI Spotlight. *How to Build Co-Located MR Games for Quest*. 12. Feb. 2025. URL: https://www.youtube.com/watch?v=7nRUkSl6Vq4 (siehe S. 25, 47).

Software

[59] Meta Platforms. *Unity MR Motif (Open-Source Reference Project)*. 2025. URL: https://github.com/oculus-samples/Unity-MRMotifs (besucht am 12. 02. 2026) (siehe S. 26).

Online-Quellen

[60] Grant Anderson. *Mirrorscape launches beta AR tabletop gaming platform for mobile and Apple Vision Pro*. GamesBeat. 2025. URL: https://gamesbeat.com/mirrorscape-launches-beta-ar-tabletop-gaming-platform-for-mobile-and-apple-vision-pro/ (besucht am 12. 02. 2026) (siehe S. 25).

[61] Kent Bye. *Resolution Games' "Battlemarked" Blends Mixed Reality Social Features with Demeo and D&D Gameplay*. Voices of VR. 29. Jan. 2025. URL: https://voicesofvr.com/1665-resolution-games-battlemarked-blends-mixed-reality-social-features-with-demeo-and-dd-gameplay/ (besucht am 12. 02. 2026) (siehe S. 25).

[62] Counterpoint Research. *Global VR Market Declines 12% YoY in 2024; 'AR+AI' Smart Glasses to Take Centre Stage in 2025*. 18. März 2025. URL: https://counterpointresearch.com/en/insights/global-xr-arvr-headsets-market-2024 (besucht am 12. 02. 2026) (siehe S. 12).

[63] Cybernetic Walrus. *Does It Stack?* Meta Horizon Store. 2024. URL: https://www.meta.com/en-gb/experiences/does-it-stack/7641276975930797/ (besucht am 12. 02. 2026) (siehe S. 25).

[64] Exit Games GmbH. *Fusion 2 Introduction*. Photon Engine Documentation. 2025. URL: https://doc.photonengine.com/fusion/current/fusion-2-intro (besucht am 12. 02. 2026) (siehe S. 35).

[65] Exit Games GmbH. *Remote Procedure Calls - Photon Fusion*. Photon Engine Documentation. 2024. URL: https://doc.photonengine.com/fusion/current/manual/data-transfer/rpcs (besucht am 12. 02. 2026) (siehe S. 35).

[66] Godot XR Contributors. *Meta Spatial Anchors*. Godot OpenXR Vendors Plugin Documentation. 2022. URL: https://godotvr.github.io/godot_openxr_vendors/manual/meta/spatial_anchors.html (besucht am 12. 02. 2026) (siehe S. 26).

[67] W. Hemmerich. *Within-Subjects-Design, Messwiederholung*. StatistikGuru. 2022. URL: https://statistikguru.de/lexikon/within-subjects-design.html (besucht am 12. 02. 2026) (siehe S. 38, 43).

[68] Meta Platforms. *Depth API Overview*. Meta for Developers. URL: https://developers.meta.com/horizon/documentation/unity/unity-depthapi-overview (besucht am 12. 02. 2026) (siehe S. 12).

[69] Meta Platforms. *Discover the Power of Movement SDK with a New Sample on Github*. Meta for Developers Blog. 2022. URL: https://developers.meta.com/horizon/blog/movement-sdk-sample-github-eye-tracking/ (besucht am 12. 02. 2026) (siehe S. 33).

[70] Meta Platforms. *Meta XR Interaction SDK for Unity*. Meta for Developers. 2025. URL: https://developers.meta.com/horizon/documentation/unity/unity-isdk-interaction-sdk-overview/ (besucht am 12. 02. 2026) (siehe S. 33).

[71] Meta Platforms. *MR Motifs Overview – Mixed Reality Best Practices*. Meta for Developers. 29. Jan. 2025. URL: https://developers.meta.com/horizon/documentation/unity/unity-mrmotifs-overview/ (besucht am 12. 02. 2026) (siehe S. 26).

[72] Meta Platforms. *Multiplayer Building Blocks Setup Guide*. Meta for Developers. 2025. URL: https://developers.meta.com/horizon/documentation/unity/bb-multiplayer-blocks/ (besucht am 12. 02. 2026) (siehe S. 26, 37).

[73] Meta Platforms. *Oculus Lipsync for Unity*. Meta for Developers. 2026. URL: https://developer.oculus.com/documentation/unity/audio-ovrlipsync-unity/ (besucht am 12. 02. 2026) (siehe S. 33).

[74] Meta Platforms. *OpenXR Hand Skeleton in Interaction SDK*. Meta for Developers. URL: https://developers.meta.com/horizon/documentation/unity/unity-isdk-openxr-hand (besucht am 12. 02. 2026) (siehe S. 9).

[75] Meta Platforms. *Quest-VR-Headsets im Vergleich: Quest 2 vs. Quest Pro vs. Quest 3*. Meta Store. 2025. URL: https://www.meta.com/at/quest/compare/ (besucht am 12. 02. 2026) (siehe S. 17).

[76] Meta Platforms. *Shared Spatial Anchors*. Meta for Developers. 2025. URL: https://developers.meta.com/horizon/documentation/unity/unity-shared-spatial-anchors/ (besucht am 12. 02. 2026) (siehe S. 12, 35).

[77] Meta Platforms. *Shared Spatial Anchors – Health and Safety Guidelines*. Meta for Developers. 2025. URL: https://developers.meta.com/horizon/design/mr-health-ssa/ (besucht am 12. 02. 2026) (siehe S. 18, 50).

[78] Meta Platforms. *Spatial Anchors Overview*. Meta for Developers. URL: https://developers.meta.com/horizon/documentation/unity/unity-spatial-anchors-overview/ (besucht am 12. 02. 2026) (siehe S. 6–8).

[79] Meta Platforms. *Table Space: Board and Card Game Sandbox*. Meta Quest Store. 2024. URL: https://www.meta.com/experiences/table-space/7704631686261190/ (besucht am 12. 02. 2026) (siehe S. 25).

[80] Meta Platforms. *Use Link for App Development*. Meta for Developers. 2024. URL: https://developers.meta.com/horizon/documentation/unity/unity-link/ (besucht am 12. 02. 2026) (siehe S. 37).

[81] Plain Concepts. *HandTracking on Quest*. Evergine. 28. Sep. 2022. URL: https://evergine.com/handtracking-quest/ (besucht am 12. 02. 2026) (siehe S. 9, 33).

[82] Qualium Systems. *What Are Spatial Anchors and Why They Matter*. Qualium Systems. URL: https://www.qualium-systems.com/blog/what-are-spatial-anchors-and-why-they-matter/ (besucht am 12. 02. 2026) (siehe S. 6, 7).

[83] Resolution Games. *Behind the Scenes of Spatial Ops: Building a Seamless MR Experience*. Resolution Games Blog. 20. Juni 2024. URL: https://www.resolutiongames.com/blog/behind-the-scenes-of-spatial-ops-building-a-seamless-mr-experience (besucht am 12. 02. 2026) (siehe S. 25).

[84] Resolution Games. *Demeo*. Resolution Games. 2025. URL: https://www.resolutiongames.com/demeo (besucht am 12. 02. 2026) (siehe S. 25).

[85] Resolution Games. *Home Sports*. Resolution Games. 2024. URL: https://www.resolutiongames.com/homesports (besucht am 12. 02. 2026) (siehe S. 25).

[86] Resolution Games. *Spatial Computing & Mixed Reality Games*. Resolution Games. 2024. URL: https://www.resolutiongames.com/mixed-reality-games (besucht am 12. 02. 2026) (siehe S. 25).

[87] Seal Systems. *Client Spotlight: Mirrorscape is Building for Vision Pro*. Seal Systems. 2025. URL: https://sealsystems.net/go/client-spotlight-mirrorscape-is-building-for-vision-pro/ (besucht am 12. 02. 2026) (siehe S. 25).

[88] Studyflix. *Uncanny Valley — einfach erklärt*. Studyflix. 2026. URL: https://studyflix.de/biologie/uncanny-valley-7316 (besucht am 12. 02. 2026) (siehe S. 17).

[89] The Game Kitchen. *All On Board! – Enjoy Board Games in XR!* The Game Kitchen. 2025. URL: https://thegamekitchen.com/all-on-board-vr/ (besucht am 12. 02. 2026) (siehe S. 26).

[90] The Ghost Howls. *Tilt-Five hands-on – very interesting AR glasses for board games*. The Ghost Howls. 2022. URL: https://skarredghost.com/2022/02/16/tilt-five-hands-on-review/ (besucht am 12. 02. 2026) (siehe S. 25).

[91] Tilt Five Inc. *Tilt Five – AR Made To Crowd Around*. Tilt Five. 2025. URL: https://www.tiltfive.com (besucht am 12. 02. 2026) (siehe S. 24).

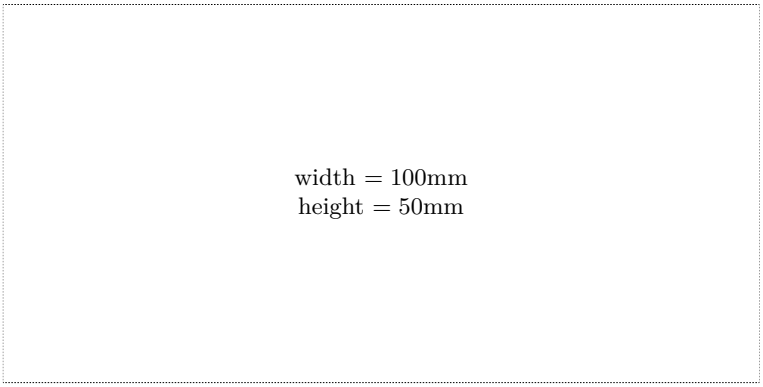
[92] Unity Technologies. *Mixed Reality multiplayer tabletop template*. Unity Documentation. 2025. URL: https://docs.unity3d.com/Packages/com.unity.template.mr-multiplayer@1.0/manual/index.html (besucht am 12. 02. 2026) (siehe S. 26).

[93] VR Compare. *Meta Quest 3*. URL: https://vr-compare.com/headset/metaquest3 (besucht am 12. 02. 2026) (siehe S. 12).

[94] ZyberVR. *How Meta Quest 3 Controller Tracking Works*. 3. Nov. 2023. URL: https://www.zybervr.com/blogs/news/how-meta-quest-3-controller-tracking-works (besucht am 12. 02. 2026) (siehe S. 10).

Messbox zur Druckkontrolle

— Druckgröße kontrollieren! —



— Diese Seite nach dem Druck entfernen! —