KU LEUVEN

ARENBERG DOCTORAL SCHOOL
Faculty of Engineering Science

Cross-Reality

An Investigation into Interaction across Physical, Augmented, and Virtual Realities

Robbe Cools

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor of Engineering Science (PhD): Computer Science

Supervisor:
Prof. dr. Adalberto L. Simeone

February 2025

CROSS-REALITY

AN INVESTIGATION INTO INTERACTION ACROSS PHYSICAL, AUGMENTED, AND VIRTUAL REALITIES

Robbe COOLS

Supervisor:
Prof. dr. Adalberto L. Simeone

Members of the
Examination Committee:
Prof. em. dr. arch. Hilde Heynen, chair
Prof. dr. Katrien Verbert
Prof. dr. arch. Andrew Vande Moere
Prof. dr. Maria Torres Vega
Prof. dr. Augusto Esteves
(University of Lisbon ITI / LARSyS, Instituto Superior Técnico)

Dissertation presented in partial fulfilment of the requirements for the degree of Doctor of Engineering Science (PhD): Computer Science

February 2025

Uitgegeven in eigen beheer, Robbe Cools, Leuven (Belgium)

Acknowledgements

The past years have been exciting and eventful, yet also stressful and challenging. I am therefore grateful to all people that I shared this time with, and that supported me. I want to thank my promotor Adalberto L. Simeone, Adal, thank you for your guidance, support, and trust. Thanks to the members of the jury, Katrien Verbert, Andrew Vande Moere, Maria Torres Vega, Augusto Esteves, and Hilde Heynen.

Many thanks to my colleagues and collaborators. Song, Jihae, and Taiyu from our ARIA group, I look back on our times together with fond memories. I want to thank my fellow PhD researchers, or recent doctors, from other research groups, Jeroen, Alex, Ozan, and Matt. Sharing experiences with all of you is the fuel that kept this research going.

A special thanks goes out to the participants of all the user studies in this thesis, without whom this research would have been impossible.

A PhD is not only about scientific contributions, but also about passing on knowledge to the next generation of students. I want to thank the master students I supervised, Stan, Piet, Robbe D., Jelle, Richard, Laura, Tim, Xander, Natalie, Renée, Elias S., Elias D.C., Jef, Inne, and Noemi. I am proud to have seen all of you learn, grow, defend the thesis, and graduate. I also want to thank Philip Dutré and the 'gegevensstructuren en algoritmen' teaching assistant team, it was my pleasure to contribute to the education of the first bachelor students.

I thank all my family and friends. Dear Mom and Dad, your upbringing is what shaped me, you were always supportive and allowed me to pursue my interests, which is what put me on the path towards computer science and this dissertation.

Kirsten, sharing the ups and downs of pursuing a PhD with you was a very special and comforting experience, and I look forward to what the future has in store for us. I am so proud that I can call you doctor, I am even prouder to call you my wife.

Robbe Cools

This thesis was supported by Internal Funds KU Leuven (C14/20/078).

Popularised Abstract

Mixed reality enables its user to wear a headset and experience alternate realities to place virtual objects in their environment, or to transport them to virtual worlds. So far, research and applications in this area mainly focused on users interacting in one of these artificial realities, with limited consideration of interaction across multiple realities. Therefore, this thesis investigates 'Cross-Reality (CR)' systems, which operate across multiple mixed realities.

First, we investigate how a real-world user can see a fully virtual world without obstructing their real-world view, by filtering parts of this virtual world to only show its most relevant elements. Second, we focus on how people and objects transition between mixed realities. We create methods for transferring an object between fully virtual worlds and virtual representations in the real world, as well as onto computer monitors. Additionally, we created a system based on a mobile vertical surface, that allows a user to transition between varying degrees of real and virtual environments. Third, we imagine how near-future mixed reality could be shaped by contact lenses rather than headsets, and speculate on their potential real-world implications. Furthermore, we present a set of design patterns, which provide blueprints of common solutions found in cross-reality systems, in our own and related works.

Through creating and evaluating these systems, we establish guidelines and patterns to form a foundation for both practical cross-reality applications, as well as the future of ubiquitous mixed reality systems.

Abstract

Mixed reality research mainly focused on intra-reality systems, where users can enter other realities that augment our physical reality, or that present fully virtual environments. However, specific applications that involve multiple manifestations of reality have been explored, such as bystander inclusion, asymmetric virtual reality, environmental awareness, transitional interfaces, or hybrid interfaces. These systems allow users to interact with virtual realities from external perspectives, transition between manifestations, or interact with different realities simultaneously. As these are systems that operate across multiple realities, there is a need to unify them under the term 'Cross-Reality (CR)' systems. Cross-reality supports users in asymmetric roles or fusion of affordances on different points of the reality-virtuality continuum, for which there are applications in domains such as education, research, architectural design, computer-aided design, museums, and virtual stores. Relating cross-reality to other frameworks like ubiquitous computing and cross-device interaction, allows for a future perspective on everyday mixed reality.

The goal of this thesis is to establish a foundation for cross-reality systems by exploring visualisations, transition techniques, context, and design patterns. We achieved this goal by addressing the following three research objectives: 1) Visualising Virtual Environments to Augmented Reality Users; 2) Enabling Object and User Transitions Across the Reality-Virtuality Continuum; and 3) Context and Design of Future Cross-Reality. These objectives were addressed by creating prototype cross-reality systems, which were demonstrated and evaluated with qualitative and quantitative methods. Furthermore, we performed an immersive speculative enactment to gain insights on future contexts of use, and a literature analysis to derive a framework with cross-reality design patterns.

In the first objective, we filtered virtual environments for display to augmented reality users, finding that a static selection of interactive objects and their immediate context preserved recognition of events. We then used these findings to develop a cross-reality user study tool, allowing researchers to observe virtual reality study participants from an augmented reality perspective. In the second objective, we investigated transitions of objects and users between realities. We found that a novel blended space technique was most performant for object transitions between virtual and augmented reality. In contrast, object transition between desktop monitors and augmented reality benefited from minimal modality switches and from objects being in reach of the other modality after transition. In the third objective, we focused on the context and design of cross-reality. We found that a future ubiquitous mixed reality context requires devices to communicate user state to others, intuitive sharing of virtual content, and support for personalisation. Furthermore, we identified an initial set of eleven design patterns, presented in terms of fundamental, origin, display, and interaction properties of cross-reality systems.

The outcomes of these objectives are 14 guidelines, which we discussed in the broader context of cross-reality and ubiquitous mixed reality, and eleven cross-reality design patterns. These guidelines and patterns form a foundation for both practical cross-reality applications, as well as future ubiquitous mixed reality systems.

Gepopulariseerde samenvatting

Met mixed reality kan de gebruiker een headset dragen en alternatieve realiteiten ervaren om virtuele objecten in hun omgeving te plaatsen of naar virtuele werelden te transporteren. Tot nu toe waren onderzoek en toepassingen in dit gebied vooral gericht op gebruikersinteractie in één van deze kunstmatige realiteiten, waarbij slechts beperkt aandacht werd besteed aan interactie tussen meerdere realiteiten. Daarom onderzoekt dit proefschrift ‘Cross-Reality (CR)’-systemen, die in meerdere kunstmatige realiteiten opereren.

Ten eerste onderzoeken we hoe een gebruiker uit de echte wereld een volledig virtuele wereld kan zien zonder hun zicht op de echte wereld te belemmeren, door delen van deze virtuele wereld te filteren om alleen de meest relevante elementen weer te geven. Ten tweede concentreren we ons op de manier waarop mensen en objecten overgaan tussen gemengde realiteiten. We creëren methoden voor het overbrengen van een object tussen volledig virtuele werelden en virtuele representaties in de echte wereld, maar ook naar computermonitors. Daarnaast hebben we een systeem gemaakt op basis van een mobiel verticaal oppervlak, waarmee een gebruiker kan overstappen tussen reële en virtuele omgevingen in verschillende mate. Ten derde stellen we ons voor hoe mixed reality in de nabije toekomst zou kunnen worden gevormd door contactlenzen in plaats van door headsets, en speculeren we over de mogelijke implicaties ervan in de echte wereld. Verder presenteren we een reeks ontwerppatronen, die blauwdrukken zijn van gemeenschappelijke oplossingen die gevonden worden in cross-realitysystemen, in onze eigen en aanverwante werken.

Door deze systemen te creëren en te evalueren, stellen we richtlijnen en patronen voor die een basis vormen voor zowel praktische cross-reality-toepassingen als de toekomst van alomtegenwoordige mixed reality-systemen.

Samenvatting

Onderzoek naar gemengde realiteit richtte zich voornamelijk op intra-realiteitssystemen, waarbij gebruikers andere realiteiten kunnen betreden die onze fysieke realiteit vergroten, of die volledig virtuele omgevingen presenteren. Er zijn echter specifieke toepassingen onderzocht die meerdere manifestaties van de realiteit omvatten, zoals omstandersinclusie, asymmetrische virtuele realiteit, omgevingsbewustzijn, transitionele interfaces of hybride interfaces. Deze systemen stellen gebruikers in staat om vanuit externe perspectieven met virtuele realiteiten te interageren, tussen manifestaties te wisselen of tegelijkertijd met verschillende realiteiten te interageren. Omdat dit systemen zijn die over meerdere realiteiten heen werken, is er behoefte om ze te verenigen onder de term 'Cross-Reality (CR)'-systemen. Cross-reality ondersteunt gebruikers in asymmetrische rollen of fusie van affordances op verschillende punten van het realiteit-virtualiteitscontinuüm, waarvoor er toepassingen zijn in domeinen zoals onderwijs, onderzoek, architectonisch ontwerp, computerondersteund ontwerp, musea en virtuele winkels. Door cross-reality te relateren aan andere frameworks zoals ubiquitous computing en cross-device interaction, ontstaat er een toekomstig perspectief op alomtegenwoordige mixed reality.

Het doel van dit proefschrift is om een basis te leggen voor cross-reality systemen door visualisaties, transitietechnieken, context en ontwerppatronen te onderzoeken. We bereikten dit doel door de volgende drie onderzoeksdoelstellingen aan te pakken: 1) Virtuele omgevingen visualiseren voor gebruikers van augmented reality; 2) Object- en gebruikerstransities mogelijk maken over het continuüm van realiteit en virtualiteit; en 3) Context en ontwerp van toekomstige cross-reality. Deze doelstellingen werden aangepakt door prototype cross-reality systemen te creëren, die werden gedemonstreerd en geëvalueerd met kwalitatieve en kwantitatieve methoden. Verder voerden we een immersive speculative enactment uit om inzicht te krijgen in toekomstige gebruikscontexten, en een literatuuranalyse om een raamwerk te creëren met cross-reality ontwerppatronen.

In het eerste doel filterden we virtuele omgevingen voor weergave aan augmented reality-gebruikers, waarbij we ontdekten dat een statische selectie van interactieve objecten en hun directe context de herkenning van gebeurtenissen bewaarde. Vervolgens gebruikten we deze bevindingen om een cross-reality gebruikersstudietool te ontwikkelen, waarmee onderzoekers deelnemers aan virtual reality-studies konden observeren vanuit een augmented reality-perspectief. In het tweede doel onderzochten we overgangen van objecten en gebruikers tussen realiteiten. We ontdekten dat een nieuwe blended space-techniek het beste presteerde voor objectovergangen tussen virtuele en augmented reality. Daarentegen profiteerden objectovergangen tussen desktopmonitoren en augmented reality van minimale modaliteitswisselingen en van objecten die binnen bereik waren van de andere modaliteit na de overgang. In het derde doel richtte we ons op de context en het ontwerp van cross-reality. We ontdekten dat een toekomstige alomtegenwoordige mixed reality-context vereist dat apparaten de gebruikersstatus aan anderen communiceren, dat virtuele content intuïtief wordt gedeeld en dat personalisatie wordt ondersteund. Verder identificeerden we een set van elf ontwerppatronen, gepresenteerd in termen van fundamentele, oorsprongs-, weergave- en interactie-eigenschappen van cross-reality-systemen.

De uitkomsten van deze doelstellingen zijn 14 richtlijnen, die we bespraken in de bredere context van cross-reality en alomtegenwoordige mixed reality, en elf cross-reality ontwerppatronen. Deze richtlijnen en patronen vormen een basis voor zowel praktische cross-reality toepassingen, als toekomstige alomtegenwoordige mixed reality systemen.

Abbreviations

ABS Automatic Blended Space
AE Augmented Environment
AR Augmented Reality
AV Augmented Virtuality
BT Binary Transition
CR Cross-Reality
CReST Cross-Reality user Study Tool
HCI Human-Computer Interaction
HMD Head-Mounted Display
ISE Immersive Speculative Enactment
MBS-B Manual Blended Space with Button transition
MBS-T Manual Blended Space with Touch transition
MR Mixed Reality
OST Optical See-Through
PE Physical Environment
PR Physical Reality
RV Reality-Virtuality
TI Transitional Interface
VE Virtual Environment
VML Virtual Magic Lens
VR Virtual Reality
VST Video See-Through
XR eXtended Reality

Contents

Popularised Abstract iii

Abstract v

Gepopulariseerde samenvatting vii

Samenvatting ix

Abbreviations xi

Contents xiii

1 Introduction 1
- 1.1 A History of Mixed Reality . . . 1
- 1.2 Realities of the Reality-Virtuality Continuum . . . 4
- 1.3 What is Cross-Reality . . . 6
- 1.4 Why Users Need Cross-Reality . . . 9
  - 1.4.1 Application Domains . . . 9
  - 1.4.2 Towards Everyday Mixed Reality: From Cross-Device to Cross-Reality . . . 11

2 Thesis Overview 13
- 2.1 Research Objectives . . . 14
  - 2.1.1 Research Objective 1: Visualising Virtual Environments to Augmented Reality Users . . . 14
  - 2.1.2 Research Objective 2: Enabling Object and User Transitions Across the Reality-Virtuality Continuum . . . 14
  - 2.1.3 Research Objective 3: Context and Design of Future Cross-Reality . . . 15
- 2.2 Methodology & Contributions . . . 15
- 2.3 Thesis Structure . . . 17

I Visualising Virtual Environments to Augmented Reality Users 21

3 The Body in Cross-Reality: A Framework for Selective Augmented Reality Visualisation of Virtual Objects 23
3.1 Introduction . . . 25
3.1.1 Cross-Reality Interaction . . . 25
3.1.2 The Body in Cross-Reality . . . 25
3.2 Related work . . . 26
3.3 A Framework for Selective Visualisation . . . 26
3.3.1 Proximity Threshold . . . 27
3.3.2 Field of View . . . 27
3.3.3 Importance Ranking . . . 28
3.3.4 Technical Implementation . . . 28
3.4 Use cases . . . 28
3.4.1 Vocational Training - Food Hygiene: . . . 28
3.4.2 Virtual Exhibition - Architecture . . . 29
3.5 Evaluation . . . 29
3.6 Conclusion . . . 29

4 SelectVisAR: Selective Visualisation of Virtual Environments in Augmented Reality 31
4.1 Introduction . . . 33
4.2 Related Work . . . 34
4.2.1 Tablet and screen-based VE Visualisation . . . 34
4.2.2 Projection-based VE Visualisation . . . 34
4.2.3 Miniatures . . . 35
4.2.4 AR-based VE Visualisation . . . 35
4.2.5 Selection within VEs . . . 36
4.3 Visualisation and Implementation . . . 36
4.3.1 Pilot Study . . . 37
4.3.2 Selective Visualisation Framework . . . 38
4.3.3 Implementation of System . . . 38
4.3.4 The Virtual Environment . . . 39
4.4 Main User Study . . . 41
4.4.1 Procedure . . . 41
4.4.2 Results . . . 42
4.5 Follow-up User Study . . . 44
4.5.1 Changes to Selective Techniques . . . 44
4.5.2 Procedure . . . 44
4.5.3 Results . . . 44
4.6 Discussion . . . 47
4.7 Conclusion and Future Work . . . 49

5 CReST: Design and Evaluation of the Cross-Reality Study Tool 51
5.1 Introduction . . . 53
5.2 Related Work . . . 53
5.2.1 User Study Tools . . . 54
5.2.2 Cross-Reality . . . 55
5.3 Developing the Cross-Reality Study Tool . . . 56
5.3.1 Participant Observation . . . 56
5.3.2 Study Control . . . 57
5.4 Example cases . . . 59
5.4.1 RE1 Distractor Interactivity . . . 59
5.4.2 RE2 Immersive Speculative Enactment (ISE) . . . 60
5.4.3 RE3 Foldable Spaces . . . 60
5.4.4 Adapting Examples to CReST . . . 61
5.5 Expert Evaluation . . . 62
5.5.1 Verifying Participant Activity . . . 62
5.5.2 Monitoring Participant Progress . . . 63
5.5.3 Augmented Reality Participant Observation . . . 63
5.5.4 Expert Feedback Iteration . . . 64
5.6 Case Study: Graphic and Interaction Fidelity . . . 65
5.6.1 Study Procedure . . . 65
5.6.2 Integration with CReST . . . 67
5.6.3 CReST Functionality . . . 67
5.7 Discussion . . . 68
5.7.1 Added Value of CReST . . . 68
5.7.2 CReST for Different Study Types . . . 69
5.7.3 Limitations & Future Work . . . 70
5.8 Conclusion . . . 71

II Enabling Object and User Transitions Across the Reality-Virtuality Continuum 73

6 Blending Spaces: Cross-Reality Interaction Techniques for Object Transitions Between Distinct Virtual and Augmented Realities 75
6.1 Introduction . . . 77
6.2 Related Work . . . 79
6.2.1 Cross-Reality Interaction . . . 79
6.2.2 Lenses Between Virtual Environments . . . 80
6.2.3 Transitioning Between Virtual Environments . . . 80
6.2.4 Blending Space . . . 81
6.3 Cross-Reality Transition Techniques & Implementation . . . 82
6.3.1 Binary Transition (BT) . . . 82
6.3.2 Virtual Magic Lens (VML) . . . 82

6.3.3 Manual Blended Space - Button Transition (MBS-B) 83
6.3.4 Auto Blended Space (ABS) 84
6.3.5 Implementation 85
6.4 Transitions User Study 85
6.4.1 Environments & Task 85
6.4.2 Experimental Setup & Design 86
6.4.3 Metrics 86
6.5 Transitions Results 87
6.5.1 Task completion times 87
6.5.2 NASA-TLX 89
6.5.3 Preference & Interviews 89
6.6 Mixed Manipulations User Study 90
6.6.1 Manual Blended Space - Touch Transition (MBS-T) 91
6.6.2 Environments 91
6.6.3 Experimental Setup & Design 91
6.7 Mixed Manipulations Results 93
6.7.1 Task completion times 94
6.7.2 NASA-TLX 94
6.7.3 Preference & Interviews 94
6.8 Discussion 95
6.8.1 Design Guidelines on Transition Techniques 95
6.8.2 Future Design of Blended Spaces 96
6.8.3 Limitations 97
6.9 Conclusion 98

7 Towards a Desktop–AR Prototyping Framework: Prototyping Cross-Reality Between Desktops and Augmented Reality 99
7.1 Introduction 101
7.2 Related Work 102
7.3 Proposed Desktop–AR Prototyping Framework 104
7.3.1 Input Modalities 104
7.3.2 Desktop–AR CR Interactions 105
7.3.3 Desktop–AR CR Transitions 106
7.4 Proof-of-Concept Desktop–AR Workspace 107
7.5 Use Cases 109
7.6 Conclusion 109

8 Comparison of Cross-Reality Transition Techniques Between 3D and 2D Display Spaces in Desktop–AR Systems 113
8.1 Introduction 115
8.2 Related Work 116
8.3 Desktop–AR Design Rationale 118
8.4 Design & Implementation of Transition Techniques 119

8.5 User Study 122
8.5.1 Procedure 122
8.5.2 Task 123
8.5.3 Metrics 123
8.6 Results 124
8.7 Discussion 127
8.7.1 Design Guidelines 128
8.7.2 Limitations & Future Work 129
8.8 Conclusion 129

9 How a Vertical Surface Supports Cross-Reality Transitional Interface Tasks at Different Virtuality Levels 131
9.1 Introduction 133
9.2 Related Work 134
9.2.1 Mixed Reality & Robots 134
9.2.2 Cross-Reality 134
9.2.3 Transitional Interfaces 135
9.3 Vertical Surface Interactions Across Realities 136
9.3.1 Dimensions of the Transitional Interface 136
9.3.2 Characteristics of Possible Interactions 138
9.3.3 Implementation 139
9.4 User Study 140
9.4.1 Interactive Surface Tasks 140
9.4.2 Procedure 142
9.4.3 Metrics 142
9.5 Results 142
9.5.1 Quantitative Data: Time, Surface Movement, and Questionnaire 143
9.5.2 Qualitative Data: Interview 147
9.6 Discussion 149
9.6.1 Physical Environment Preferred for User Study Tasks (RQ1) . 149
9.6.2 Virtual Surface is More Convenient (RQ2), Physical Surface is Used More (RQ3) 150
9.6.3 Limitations & Future Work 151
9.7 Conclusion 152

III Context and Design of Future Cross-Reality 153

10 The Impact of Near-Future Mixed Reality Contact Lenses on Users' Lives via an Immersive Speculative Enactment and Focus Groups 155
10.1 Introduction 157
10.2 Related Work 158
10.2.1 Immersive Speculative Enactments 159

10.2.2 Towards Mixed Reality Contact Lenses 160
10.2.3 Social Acceptability & Practical Concerns of Public Mixed Reality 161
10.2.4 Privacy & Ethical Concerns 162
10.3 The Immersive Speculative Enactment 162
10.4 User Study 165
10.4.1 Participants 166
10.4.2 Procedure 167
10.4.3 Focus Group 168
10.5 Results 169
10.5.1 Privacy, Security, and Perceptual Agency 169
10.5.2 Social Acceptability of Increased Virtuality 171
10.5.3 Excessive use Taking Away From Real Life 172
10.5.4 Future Existence for Practical Application 173
10.6 Discussion 174
10.6.1 Perceptual Agency as new Issue 175
10.6.2 Social Interaction and Acceptability 175
10.6.3 Towards a Desirable Future 176
10.6.4 Limitations 177
10.7 Conclusion 178

11 From Display to Interaction: Design Patterns for Cross-Reality Systems 179
11.1 Introduction 181
11.2 Related Work 182
11.2.1 Mixed Reality & Cross-Reality 182
11.2.2 HCI Theory & Design Patterns 183
11.3 Methodology 184
11.4 Fundamental Patterns 185
11.4.1 Scope 185
11.4.2 Directionality 187
11.5 Origin Patterns 188
11.5.1 Live Copy 189
11.5.2 Replay 189
11.5.3 Notification 191
11.6 Display Patterns 192
11.6.1 Miniature 192
11.6.2 Windows 193
11.6.3 Transformed Environments 196
11.6.4 Combined Visual Modalities 197
11.7 Interaction Patterns 198
11.7.1 Inclusion of Physical Objects 198
11.7.2 Different Interaction Modalities 198

11.8 Example Applications of CR Design Patterns 200
11.8.1 Example 1: Story-based Transitional Interface 200
11.8.2 Example 2: Seamless Bystander Inclusion 201
11.9 Discussion 202
11.9.1 Fundamental Pattern Opportunities 203
11.9.2 Limitations & Future Work 203
11.10Conclusion 204

IV Conclusions 205

12 Conclusions 207
12.1 Summary of Contributions 207
12.1.1 Research Objective 1: Visualising Virtual Environments to Augmented Reality Users 207
12.1.2 Research Objective 2: Enabling Object and User Transitions Across the Reality-Virtuality Continuum 209
12.1.3 Research Objective 3: Context and Design of Future Cross-Reality 211
12.2 Reflection 214
12.2.1 Directionality, Scope, and Content Origin as Aspects of Cross-Reality 214
12.2.2 Generalising Findings 215
12.2.3 From Cross-Reality to Everyday Mixed Reality 216
12.2.4 When Should a Mixed Reality System be Cross-Reality 217
12.3 Limitations 218
12.4 Future Research Directions 220
12.4.1 Cross-Reality Social Presence 220
12.4.2 Cross-Reality State Visualisation 220
12.4.3 Cross-Reality Agents 221
12.5 Closing Remarks 221

Bibliography 223

Chapter 1

Introduction

The availability of Mixed Reality (MR) devices gives rise to challenges and opportunities that need to be addressed and researched, which were primarily focused on intra-reality systems. However, specific scenarios that involve multiple manifestations of reality have been explored, such as bystander inclusion, asymmetric virtual reality, environmental awareness, transitional, or hybrid interfaces. There is a need to unify these disparate research areas with systems that operate across multiple realities, under the term 'Cross-Reality (CR)' systems. CR has applications in a variety of domains like education, architectural design, user studies, computer-aided design, game development, museums, and virtual shopping, and also contributes to adoption of MR in everyday contexts.

1.1 A History of Mixed Reality

Throughout history, humanity has striven to bring to life imaginary realities and share them with others through storytelling and interactive experiences. From books and paintings to cinema and video games, people have consistently explored new ways to envision and create fictional realities. This ongoing fascination has led to the emergence of MR as a groundbreaking immersive medium. The history of MR starts with stereoscopes in the 1800s, and leads up to the commercially available head-mounted displays (HMDs) of the past decade. These devices enable experiences across a continuum that ranges from the Physical Reality (PR) to entire Virtual Realities (VR).

In 1832 Charles Wheatstone invented the stereoscope (Figure 1.1), which is a device that used mirrors to present two different images to the eyes, creating the illusion of depth. Images were generated by capturing two photographs from a slightly different

Figure 1.1: Wheatstone Stereoscope from 1832, which used angled mirrors that provide the user with a different image in each eye. Dr. Brian May, who is a collector of stereoscopic photographs, demonstrates how the stereoscope is used. (Image credit: David Trett; https://www.kcl.ac.uk/charles-wheatstone-the-father-of-3d-and-virtual-reality-technology-2)

angle, which the stereoscope could then display to each eye separately. In 1851, the first commercially available stereoscope for immersive entertainment was introduced, using lenses instead of mirrors, leading to a more compact design. [125]. The concept introduced by the stereoscope was further developed with advances in electronics, computing, and video.

In the 1950s and 1960s Morton Heilig and Ivan Sutherland laid the foundation for a new generation of electronic MR devices, by introducing the Telesphere mask, Sensorama, and 'Sword of Damocles' (Figure 1.2). Heilig's contributions with the Telesphere mask and Sensorama centred around stereoscopic video and multisensory stimulation, such as haptic and olfactory, which immersed users in a virtual scene [110]. Additionally, the Telesphere mask introduced the HMD form-factor that became common for current MR devices [109]. In 1965, Sutherland described the ultimate display as 'a room within which the computer can control the existence of matter' [269], which he built further upon with the 'Sword of Damocles' prototype. The prototype augments the environment with optical see-through displays that render virtual content on top of users' views, and tracks head movements to allow users to look around [270]. Even at this early stage of research, there was a distinction in virtuality between Sensorama, which provides a fully recorded view, and 'Sword of Damocles', which overlaid digital

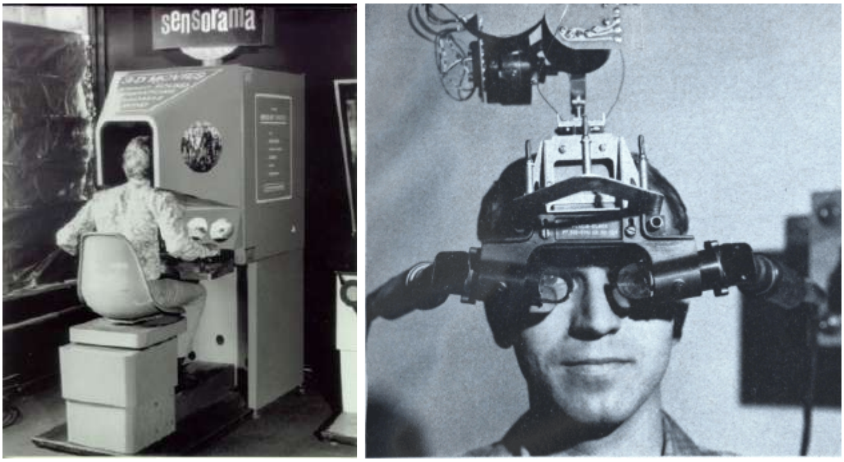


Figure 1.2: Early prototypes of MR devices. Sensorama (left) used stereoscopic video recordings to give users the feeling of being in a different environment. The 'Sword of Damocles' (right) used 3D wireframe graphics overlaid on top of the real world, to give users the illusion that virtual objects exist in their physical space.

content onto the real world. Both prototypes were limited by the computing capabilities of their time and required further advancements in graphics and rendering technology.

The 1990s mark a period of renewed interest in MR following advancements in real-time 3D rendering, and commercial availability of hardware such as Nintendo Virtual Boy [27], Virtuality Group, and CAVE systems [125]. At the time, it was not feasible to develop devices that were both capable and affordable, for example, the Nintendo Virtual Boy was affordable but had a monochrome display and was uncomfortable to use. Other more capable devices, such as those from Virtuality Group, were too expensive for everyday users and remained limited to arcades or amusement parks. CAVE systems were also developed, which abandoned the HMD form-factor and required a room to be equipped with 3D projectors to depict a Virtual Environment (VE) to the user on its walls and floor [55]. Nevertheless, research conducted in this era is fundamental to this thesis, such as systems combining physical displays and AR [70], multi-user AR collaboration [272, 21], handheld miniatures [203], and the 'MagicBook' Transitional Interface (TI) [20].

In 1994 Milgram et al.'s theoretical paper 'Augmented Reality: A class of displays on the reality-virtuality continuum' [173] was published. This work lays the foundation for the characterisation of CR used in this thesis by introducing the Reality-Virtuality (RV) continuum and its three components (Extent of World Knowledge, Reproduction

Fidelity, and Extent of Presence Metaphor). In section 1.2, the RV continuum introduces the terminology that is used to denote different realities.

The release of Oculus Rift DK1 in 2013, marks the beginning of a new era of relatively affordable MR devices, which continued in 2016 with other notable releases such as HTC Vive, Oculus Rift CV1, HoloLens 1, and Google Daydream, leading to a rapid increase in VR and AR research (in all domains, as seen in Figure 1.3). Vive and Rift CV1 mark the first widely adopted tethered HMDs, aimed primarily at VR gaming on platforms like SteamVR. HoloLens 1 is a standalone optical see-through (OST) AR device, whereas Google Daydream marks the first standalone VR HMDs with positional tracking, a trend which would continue with later devices. Vive Pro was released in 2018, accompanied by a wireless adapter. In 2019 the trend towards standalone devices continued with the Oculus Quest 1. Furthermore, 2019 also marks the release of HoloLens 2 and Valve Index, which alongside the aforementioned Vive Pro, are the devices available when this thesis was initiated. VR-centred HMDs such as Vive Pro and Valve Index were already equipped with outward-facing cameras that enabled video see-through (VST) AR, though with limited quality and developer support. Meanwhile, other manufacturers, such as Varjo with its VR-1 HMD, already included support for higher quality VST.

Integration of VST as one of the main features of HMDs would continue, such as with the most recent devices at the time of writing: Meta Quest 3 and Apple Vision Pro. This marks a shift from HMDs being purpose-built for VR or AR, towards a single device that can smoothly transition users on the continuum. Over these years, various terms have emerged, such as virtual, augmented, mixed, and extended reality, to describe the realities of the RV continuum.

1.2 Realities of the Reality-Virtuality Continuum

Mixed Reality (MR) and eXtended Reality (XR) MR, also known as XR, is an umbrella term that encompasses all points on the RV continuum (Figure 1.4). Other MR characterisations exist that align more with AR [257], but MR is most commonly adopted as an umbrella term in research, which aligns with its original meaning [173, 251]. (Figure 1.3). Within MR four distinct 'realities' are identified (VR, AR, AV, and PR, which are explained in the following paragraphs), which allow for a more precise description of the virtuality of the user experience.

Virtual Reality (VR) VR refers to the most virtual mediated reality that hardware can currently achieve, creating an experience that makes its users feel as if they are in a VE detached from PR. As technology advances, the rightmost part of the RV continuum extends further than what is currently referred to as VR, called a 'Matrix-like' VE by

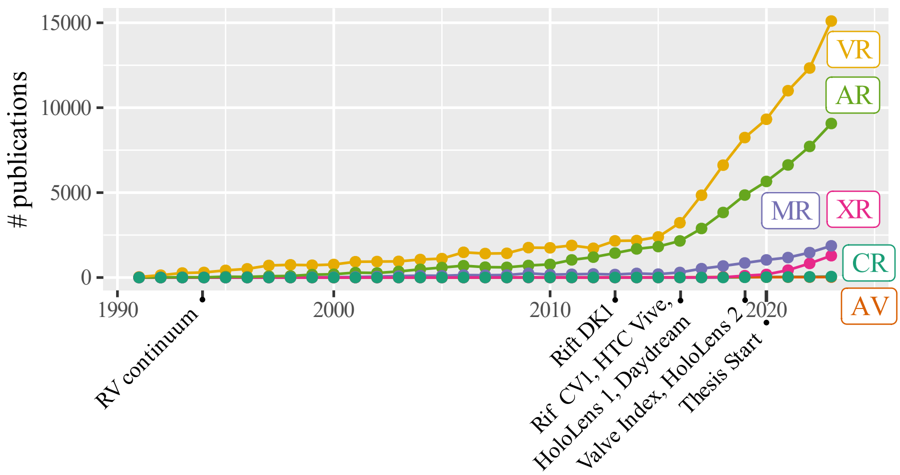


Figure 1.3: Prevalence of terms in scientific publications, based on a query for occurrence of the term in title or abstract (data collected from https://www.dimensions.ai/ and visualised with ggplot).

Skarbezet al. [251]. With 'Matrix-like', they refer to the 1999 film in which a VR is accessed through a brain implant, depicting a higher degree of immersion than can currently be achieved. For now VR refers to an experience mediated by HMDs or CAVE systems that artificially stimulates primarily sight and hearing. Conversely to VR, which presents a fully artificial reality, AR focuses on adding virtual augmentations to the real world.

Augmented Reality (AR) AR refers to a mediated reality that is based in PR, but augmented with virtual elements, thus situated towards the left-hand side of the RV continuum. This thesis uses AR to describe an immersive reality achieved through HMD or Spatial Augmented Reality (SAR) systems [13], while users of non-immersive mobile AR do not experience presence in the augmented environment. Typically, HMD AR is achieved via one of two technologies, video see-through (VST) or optical see-through (OST). VST uses external facing cameras that capture the PR and combine it with virtual content to render on an opaque display. OST uses translucent display technology that allows light from the real world to travel through it, while also superimposing virtual content on top. VST typically offers superior field of view, allowing it to show virtual content in the user's peripheral vision, and ensures that virtual content is displayed fully opaque. OST offers a superior, unaltered view of the real world, at the cost of a reduced field of view and translucency of virtual content. Moreover, with OST it is currently sub-optimal to create fully immersive experiences due to translucency of

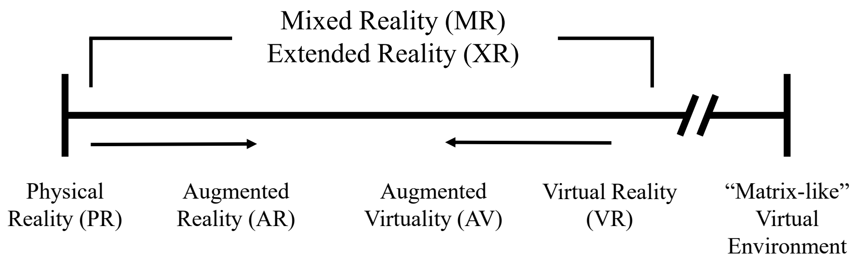


Figure 1.4: The Reality-Virtuality Continuum, adapted from Milgram et al. [173], and Skarbez et al. [251]

displays, which cannot fully occlude external light, including light entering from the sides. In contrast to AR, which augments the real world with virtual elements, there is also augmented virtuality, that augments VR experiences with real-world elements.

Augmented Virtuality (AV) AV refers to a reality in which most of the content is virtual, but there is some inclusion of real-world elements [173]. What constitutes these real-world elements can differ, such as physical objects, but also images or videos of the real world. Hence, AV can be described as a VR experience that is expanded upon by including physical props [250] or portals into the real world [290]. Because a VR experience becomes AV when these real-world elements are introduced, the distinction between VR and AV is not always clear.

Physical Reality (PR) PR differentiates the unmediated ('real') reality from the mediated realities discussed before (AR, AV and VR), and is also referred to as the 'real world'. PR includes experiences that are not immersive, such as users with phones, tablets, projection, etc. This distinction between the physical and the virtual, and specifically the differences between the realities encompassed by MR, give rise to systems that operate across multiple realities.

1.3 What is Cross-Reality

CR refers to a type of MR system that operates across multiple 'realities', with a reality defined by the RV continuum [173]. Research into these types of systems has gained increased relevance by recent advancements in MR technology, which make not only VR but also AR available to consumers. Being a nascent field, variations exist on how the term is spelled, I chose to write it with hyphen (-) and use CR as abbreviation.

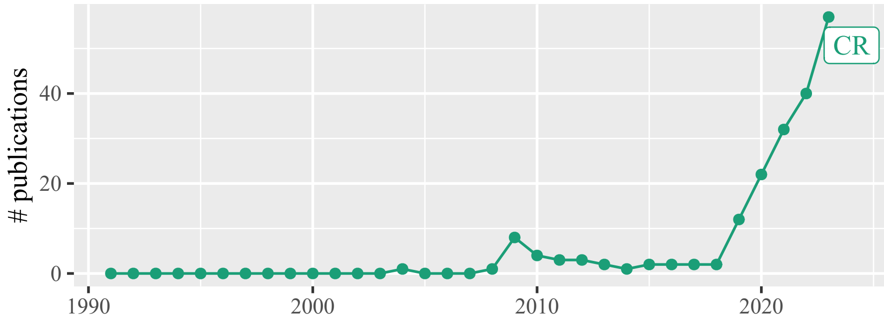


Figure 1.5: Prevalence of the term Cross-Reality in scientific publications, based on a query for occurrence of the term in title or abstract (data collected from `https://www.dimensions.ai/` and visualised with ggplot). The spike in 2009 is a special issue by Paradiso et al. [202]. From 2019 onwards there is an increased research interest in CR.

'Cross-reality' is not used to describe the nature of a reality, i.e. a user is not immersed in *a* cross-reality, rather cross-reality refers to the attribute that allows a system to operate across multiple realities. Confusion is caused by other researchers abbreviating cross-reality as XR, or even using it as a synonym for extended reality. Consistently writing cross-reality (CR) ensures it is distinct from the other terms and carries the meaning conveyed in this thesis.

The first peak of popularity for the term CR stems from a special issue by Paradiso et al. [202] in 2009, who define CR as 'the union between ubiquitous sensor/actuator networks and shared online virtual worlds - a place where collective human perception meets the machine view of pervasive computing.' This characterisation differs from the one in this thesis, as it emphasizes the integration of virtual worlds with pervasive computing. However, there are similarities, including the connection between real and virtual worlds, and the focus on enabling access to VEs through different devices. From 2020 onwards, CR saw a rise in related publications and workshops (Figure 1.5), starting with a workshop by Simeone et al. [244] who characterise it as '1) a smooth transition between systems using different degrees of virtuality; or 2) collaboration between users using different systems with different degrees of virtuality'.

Before the term CR gained popularity, research had already explored topics that now fall under this category. This research primarily focuses on specific scenarios within CR, such as asymmetric VR, bystander exclusion, awareness of surroundings (like bystanders and obstacles), substitutional reality, Transitional Interfaces (TIs), and hybrid interfaces.

Asymmetric VR [198] allows multiple users to interact with the shared VE through asymmetric interfaces. Interface asymmetry, such as equipping external users with tablets and desktops, is related to user role asymmetry, such as spectating, dueling, cooperation, 'navigator and pilot', 'boss vs horde', 'hide and seek', game mastering, and teleguidance [199]. Asymmetric VR's main motivation lies in *device cost*, *accessibility*, and *asymmetric applications* in education, productivity, and entertainment. However, with the trend towards less expensive standalone HMDs, the cost of equipping multiple users with VR devices has become less of an issue. Asymmetric VR makes the VE accessible to bystanders without equipping them with HMDs, through hardware setups involving projectors [99, 292] or displays [100]. Applications include pair-learning between tablet and VR users [62], or allowing VR and desktop users to engage in games together [134]. More generally, **bystander exclusion** refers to the disconnect between HMD and non-HMD users, which asymmetric VR solves by externalising parts of the VE for bystanders to see and interact with. Similarly, AR HMDs also exclude bystanders, for which solutions such as projection were explored [124].

VR users are **unaware of their surroundings**, making it difficult for bystanders to **interrupt** the VR user, and can lead to **collisions**. Interruptions are facilitated by representations of the bystander in the VR experience, through avatars [87], widgets [246], or external facing cameras [288]. Similarly, VR users are unaware of any obstacles in their environment which can lead to collisions [315]. As such, VR users should be made aware of their surroundings through guardian systems, which can be configured by external users [315], or by generating the VE to match physical obstacles [40]. The integration of physical objects into the VE allows users to avoid colliding with them, as well as improving the experience and interaction possibilities in the VE, such as with **substitutional reality** [250].

In contrast to the previous problems, which mainly centred around a VR user, a **Transitional Interface (TI)** is another type of CR system that 'enables users to freely move along the reality-virtuality continuum' [233]. The first TI, the MagicBook, focused on storytelling and allowed transitions between reading a physical book, an augmented version of the book, and an immersive scene [20]. User transitions are essential to allow immersed users to return to the PR without taking off the HMD [81], to smoothly immerse users in VEs [284], and to offer users different perspectives during collaboration [233].

Complementary to TIs, **hybrid user interfaces** combine devices to provide a mutual display space, for example between AR and desktop monitors [70]. This combination of interfaces allows users to benefit from the strengths of each display space: perform precise, familiar manipulations on a 2D screen, and use 3D AR for better spatial understanding and larger working areas, for example, in the context of data visualisation [157].

Previous research is divided into aforementioned sub-fields, centred around immediate VR-related issues such as bystanders or collisions, or specific types of systems such as TI or hybrid interfaces. Thus, there is an opportunity for CR to unify and develop a deeper understanding of interactions across realities to facilitate application development in various domains.

1.4 Why Users Need Cross-Reality

CR supports the development of applications that require users in asymmetric roles or that require a fusion of affordances available at different points of the RV continuum, which are present in education, research, architectural design, computer-aided design, museums, and virtual stores. Moreover, relating the unified field of CR to other frameworks like ubiquitous computing and cross-device interaction, allows for creating a future perspective on what everyday MR should look like.

1.4.1 Application Domains

In an MR **training simulation** there is a *trainee* learning a skill related to a domain, under the guidance of a *supervisor*. Training simulations are common in domains such as first responder training, medical training, military training, workforce training, or education [314]. It allows the *trainee* to practice vital skills in a safe and controlled environment. In this context, the *trainee* is learning, while the *supervisor* is guiding them. In this way, the *supervisor* benefits from CR by having access to a copy of the object the *trainee* is interacting with to allow them to spectate at a distance without getting in the *trainee*'s way [319]. Additionally, *teachers* should be provided with information and affordances that are distinct from that of the *student*, such as different rights in the VE [213], or an overview environment combining multiple *student* VEs [291, 274].

A **VR user study** involves a *researcher* and one or more *participants*. The *researcher* controls the study, and seeks to gain insight into the behaviour exhibited by the participant. The *participant* takes part in the study, and engages with the VR application created by the *researcher*. Chapter 5 presents CReST, which is an application of selective visualisation to allow researchers in AR to observe co-located *participants* in VR. When *participants* are remote rather than co-located, recording their environments allows *researchers* to observe it to guide the study from a different location [159]. Moreover, in a similar use case Immersive Virtual Reality Evaluations (IVREs) [320] allow designers to evaluate virtual prototypes, and involve the designer and user of the prototype [138].

Architectural design involves the *architect* and the *occupant* of the space. The *architect* employs computer-assisted architectural design tools to create a 3D model of the space, or to visualise architectural interventions for existing spaces [184]. *Occupants* are the ones who will inhabit the space and may have requirements or feedback. In this context, *occupants* benefit from an immersive view to allow them to experience the space, while *architects* benefit from an external view to give them an overview of the space, for example, via a miniature tabletop interface [268]. Furthermore, *architects* could also guide design interventions remotely [184].

Computer-Aided Design requires a combination of precise inputs and spatial reasoning, thus benefiting from a hybrid interface [175]. Moreover, transitions of objects designed on desktop computer-aided design software to VR provide the designer with more opportunities to preview the object, i.e. if they are modelling a machine for deployment on the factory floor, they can preview it in a VE that resembles the factory in which it will be placed. **Game development** consists of multiple activities, such as *creating artefacts, programming, level design*, and *testing* [136]. Considering these four activities, which present a simplified view of the real process, CR supports the transitions between them. Both creating artefacts and programming benefit from a hybrid interface, artists can preview the artefacts they are creating in AR to assess their in-game scale, whereas programmers could test interactions with those artefacts. Then, level design requires an overview of the entire level, which can be achieved with a miniature in which the artefacts are arranged.

Museums already use VR as a medium to provide visitors with immersive and interactive experiences that support their exhibits, or AR to provide visitors with additional information on the exhibits [131]. CR provides opportunities to make these distinct experiences more seamless, allowing visitors to smoothly transition between physical exhibits, augmented exhibits, and VEs depicting their context of use. Users could view artefacts in the augmented museum, but also reach into a VE of a historical site from which they could collect more artefacts. This allows for an interactive experience in which users interact and create new exhibits, which is typically not allowed in physical museums.

On the one hand, VR allows users to shop in **virtual stores**, or even natural VEs, such as through an apartment metaphor [260]. On the other hand, AR apps are already available to allow users to **preview furniture** in their rooms [118], or virtually **try on clothes** [194]. However, CR provides opportunities to connect these applications, where the user transitions the object from VR to AR for preview, to then order it, and ultimately replace the virtual with the physical. The summary in this section is not exhaustive and the list of applications for CR will continue to grow as technology progresses to everyday MR.

1.4.2 Towards Everyday Mixed Reality: From Cross-Device to Cross-Reality

Smooth transitions and users with different degrees of virtuality are central to MR becoming a ubiquitous [97] and everyday technology [245]. Ubiquitous computing, or 'ubicomp' [301], refers to the seamless integration of computers into the world with the following five key components: embeddedness, connectivity, context awareness, adaptability, and transparency [289].

Ubiquitous availability of devices such as smartphones and tablets lead to cross-device interaction, concerned with 'interfaces or applications that move beyond the bounds of a single device screen' [32]. Cross-device interactions may consist of users sharing content from personal mobile devices, such as phones, on a shared tabletop display [236]. CR is an extension of cross-device into immersive realities, thus some CR works are also cross-device, i.e. when HMDs are used together with other devices. However, not all CR works are cross-device, as an HMD user can switch between immersive realities without switching HMDs. Furthermore, CR focuses on the human experience for its characterisation, centring around the subjective notion of reality, rather than being centred around the hardware that is involved, as with cross-device interaction.

In a near-future, as MR becomes ubiquitous, everyday CR scenarios can arise: people on public transport could transition to VEs instead of using noise-cancelling earphones, in the office they could repurpose any surface in the environment for collaborative interaction instead of relying on dedicated whiteboards and projection screens, and at home they could play video games in VEs rather than on physical displays. Users will have multiple virtual layers at their fingertips, necessitating a content transition between them. For example, using a smartphone's precise text input to look up a recipe, and then intuitively pulling it out of the phone and placing it in the environment to reference while cooking.

Ubiquitous and everyday MR envision MR as the dominant computing interface [245, 97]. Ubiquitous MR [97] centres around 'being able to fluidly access and switch between states on the reality-virtuality continuum.' It argues that, similar to ubicomp, MR should seamlessly blend with the environment, and be available anywhere and anytime. Closely integrating MR into users' daily lives, allows them to interact with spatial interfaces without thinking about devices, or where their experience is situated on the RV continuum. Furthermore, everyday MR is concerned with the challenges of adopting MR in everyday contexts, where one of the main challenges is CR [245]. Since MR is not a dominant interface used ubiquitously and in everyday contexts, researchers have to speculate on this possible future [126]. CR is a highly relevant field of research, particularly in context of a future in which ubiquitous MR becomes widespread. It is a logical next step, extending the concept of cross-device into immersive MR, and providing an overview of vital display and interaction techniques.

Chapter 2

Thesis Overview

Current MR research is focused on intra-reality interactions and shared VEs, with CR originating from problems that occur in these intra-realty systems such as bystander inclusion and awareness. However, CR has many potential applications that involve collaboration and transition across realities.

A vacuum exists around CR applications where AR is one of the realities being interacted with or from, due to the relatively recent introduction of consumer see-through AR devices [8]. See-through AR was previously achieved via devices with narrow field of view, such as HoloLens 1, or custom setups where researchers attached depth cameras onto the front of VR HMDs [291]. More capable see-through AR devices, such as HoloLens 2 and Magic Leap, were released at the outset of this thesis, which opened up opportunities to research immersive AR as part of CR. In the CR subfield of TIs, systems mainly focused on user transitions, specifically methods to move users between different manifestations of reality. However, since an interaction involves a subject and an object [133], there is an opportunity to investigate object transitions between realities, complementary to the earlier focus on user transitions. There is also a lack of insight into the future contexts in which CR systems have to operate, and whether a future with ubiquitous MR is a desirable one. CR is lacking a theoretical framework that unifies its solution space, that covers the whole field instead of separate sub-fields [295, 198, 323], and that goes beyond classification [8].

This thesis explores the design patterns, visualisations, and transition techniques necessary for developing CR systems. In doing so, it establishes a foundation for future applications such as user study tools, and contributes towards ubiquitous MR.

2.1 Research Objectives

I systematically explore CR through three research objectives, supported by the different aspects of human-computer interaction such as visualisation (output), interaction (input) [232, 187], and context [230] (Figure 2.1). In the areas of visualisation, interaction, and context, I address challenges like missing AR research [8], virtual object transitions, lack of future context, and the need for a unifying CR framework.

2.1.1 Research Objective 1: Visualising Virtual Environments to Augmented Reality Users

Visualising a VE to an AR user is a challenge, because by nature a VE takes up the user's entire view, making it unsuitable to display in AR. Therefore, the objective is to design and evaluate VE filter techniques to make it accessible to the AR user, which forms the foundation for subsequent work on CR interaction. Additionally, CR has application scenarios with different user roles, such as for guiding or observing the VR user. Hence, I develop an application to demonstrate how this type of visualisation can be applied to a VR user study scenario, and assess its added value.

2.1.2 Research Objective 2: Enabling Object and User Transitions Across the Reality-Virtuality Continuum

Transitions are one of the main interaction challenges in CR, and are further distinguished between object and user transitions. First, I investigate object transitions across desktops, AR and VR as a novel challenge, developing techniques that enable object transition between VR-AR and Desktop–AR. These potential solutions are evaluated for how effectively they enable users to complete CR object transitions. Although VR-AR can be achieved through a single input device, such as hand inputs or motion controllers, Desktop–AR requires different considerations, as there is more separation between display spaces and input modalities. Second, I envision a vertical surface TI, as an interface that allows users to freely switch to any point on the RV continuum, even mid-task. I aim to empirically explore how users employ and benefit from such an interface, and how it facilitates them in completing tasks on its surface.

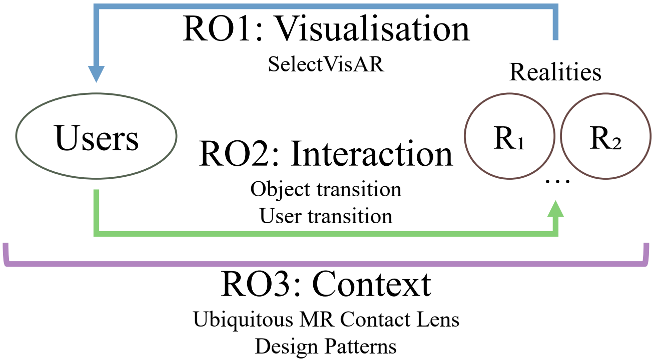


Figure 2.1: Research framework based on the different aspects of human-computer interaction such as visualisation (output), interaction (input) [232, 187], and context [230]. It describes a CR interaction between a user, or multiple users, and two or more realities. The interaction is described in terms of visualisation, addressed in the first research objective which investigates SelectVisAR systems, and interaction, addressed in the second research objective which centres around object and user transitions. Moreover, the framework takes into account the context of this interaction, addressed in the third research objective which speculates on ubiquitous MR contact lenses as possible future context, and addresses the challenge of creating a CR framework.

2.1.3 Research Objective 3: Context and Design of Future Cross-Reality

The third objective is to create a vision of the future context in which CR will be used, by speculating how advances in technology shape the future in the form of ubiquitous MR contact lenses. Furthermore, future CR systems lack a unifying framework to guide design, which I aim to address by analysing literature to identify the design patterns that make up CR systems.

2.2 Methodology & Contributions

Which chapter uses which methodology can be seen in Table 2.1.

	Literature	Concept	Quantitative	Qualitative	Demo	ISE
Chapter 3	◑	●				
Chapter 4	◑		●	●		
Chapter 5	◑			◑	●	
Chapter 6	◑		●	◑		
Chapter 7	◑	●			◑	
Chapter 8	◑		●	◑		
Chapter 9	◑		●	◑		
Chapter 10	◑			●		●
Chapter 11	●	●			◑	

Table 2.1: Overview of methodology, which shows the main contribution (●) and additional contribution (◑) of each chapter. Contributions include literature review, conceptual, quantitative data and statistics, qualitative data and thematic analysis, demonstration, and Immersive Speculative Enactment (ISE).

Literature review All chapters include a literature review to provide background specific to the topics of that chapter, discussing the relevant CR subfield and its problems and solutions. Moreover, chapter 11 has a more rigorous literature analysis of CR that forms the basis for a conceptual contribution of CR design patterns.

Conceptual Some of the chapters have a contribution that is solely conceptual, and is based on literature review and brainstorming with multiple researchers. The positional chapters, being chapters 3 and 7, provide a foundation for ideas that are further developed in the chapter that follows them (i.e. chapters 4 and 8.) Chapter 11 presents the concept of CR design patterns as its main contribution.

Quantitative data and statistics Quantitative data consists of questionnaires, like the System Usability Scale (SUS) [29] or NASA Task Load Index (NASA-TLX) [106], and logged metrics, like task completion times or accuracy. These quantitative metrics are further analysed according to best practices [222], for example, using parametric tests for normally distributed continuous data, and non-parametric tests for data that are discrete or non-normally distributed. All comparative studies use a within-group design, and counterbalancing with balanced Latin square to compensate for learning effects. The comparative studies in this thesis have a minimum sample size of 13 participants, which can be considered relatively small. Power analysis showed that for typical effect sizes a greater number of participants is required than realistically feasible (i.e. more than 46), thus in later chapters 24 is adopted as a 'rule of thumb' to provide a trade-off between statistical power and practicality.

Qualitative data and thematic analysis Qualitative data consists of recordings of conversations with participants during interviews or focus groups. Chapter 10 combines an Immersive Speculative Enactment (ISE) with a focus group, which allows qualitative data to be collected from a group of participants. For both interviews and focus groups, recordings were transcribed and then analysed. Extensiveness of the transcripts and analysis vary per chapter, such as using participant quotes in support of discussion (chapters 5 and 8), coding the transcript to derive themes from it in support of quantitative results (chapters 4, 6, and 9), or fully developing a thematic analysis over multiple iterations as a main contribution (chapter 10).

Demonstration When evaluating a toolkit such as in chapter 5, it is more important to show that it is useful rather than usable [156], i.e. a toolkit may have a very usable interface but not solve any problems. Hence, it is necessary to demonstrate the use of the toolkit to exemplify how it solves a problem. Similarly, chapters 7 and 11 contain a demonstration to support their contributions, albeit less extensively than chapter 5.

Immersive Speculative Enactment (ISE) ISE is a novel methodology introduced by our group [247], which draws on speculative design. It consists of envisioning a possible future that depicts a scenario which is not currently possible, via an immersive VE. Participants are then exposed to the VE, in which they are free to interact as they would in real life.

2.3 Thesis Structure

This thesis is divided into four parts, Part I to III address the three research objectives, Part IV summarises the contributions and presents conclusions and future perspectives.

Part I addresses research objective 1 (Visualising Virtual Environments to Augmented Reality Users) over three chapters. Chapter 3 is a positional work in which initial solutions are ideated, such as proximity or selection-based filtering of VEs. Next, chapter 4 concretises three techniques, static selection, proximity, and dollhouse. These techniques are evaluated against a baseline in two user studies (both $N = 13$), comparing user preference, event recognition, and qualitative feedback. Chapter 5 presents the Cross-Reality Study Tool (CReST), an application of the static visualisation techniques to user studies, to allow researchers to use AR to observe participants in a VR user study. CReST is demonstrated through replicated examples and a case study ($N = 17$).

Part II addresses research objective 2 (Enabling Object and User Transitions Across the Reality-Virtuality Continuum) over four chapters, in the following three contexts: object transitions between VR and AR, object transitions between Desktop and AR, and user transitions. Chapter 6 introduces the following three object transition techniques between VR and AR: *Virtual Magic Lens*, *Binary Transition*, and *Blended Space*. Furthermore, *Blended Space* consists of three variations with *Button Transition*, *Touch Transition*, or *Automatic Transition*. In two comparative user studies ($N = 20$ and $N = 16$) these transition techniques were evaluated on preference, usability, and efficiency. Next, chapter 7 presents initial ideas for a Desktop–AR prototyping framework, with the goal of providing developers with tools to facilitate the development of applications for this type of system. The chapter also describes how this framework would support transitions of objects between desktop screen space and AR, laying the foundation for the next chapter. Chapter 8 presents implementations of the following three Desktop–AR transition techniques: *hand-based*, *mouse-based*, and *modality switch*. In a comparative user study ($N = 24$) these techniques were evaluated in terms of user preference, usability and efficiency. In chapter 9, I investigated how CR can be *seamless*, where a transitional interface facilitates users to transition between four points on the RV continuum mid-task, supported by a mobile partition in the environment. An exploratory study ($N = 24$) revealed how participants used the freedom this interface gave them, and why they chose to transition at certain times.

Part III addresses research objective 3 (Context and Design of Future Cross-Reality) over two chapters, which includes the future context and a design framework for CR systems. Chapter 10 investigates the future context in which CR devices will be used through an ISE on the impact of near-future MR contact lenses on users' daily lives. ISEs and focus groups ($N = 16$, in four groups of four) allow envisioning how CR systems apply to this possible future, and which novel problems it would bring with it. Chapter 11 addresses the challenge of how to design CR systems, by providing designers with eleven design patterns that were described following a literature analysis. Each pattern description consists of 'intent', 'solution', and 'examples', and is supported with a diagram and figure of an archetypal example.

As the work in this thesis was conducted under close supervision and resulted from collaboration with other researchers, 'we' is used for the remainder of the text. Additionally, each chapter is preceded by a summary of how the chapter came to be, and what each author's contribution to it was[1], including the following information:

[1] https://www.elsevier.com/researcher/author/policies-and-guidelines/credit-author-statement

Each chapter is preceded by a box like this one, and contains the following information:

- Reference to the paper the chapter is based on. Shared first authorship is indicated with '*'.
- CRediT author statement, specifically the following roles: Conceptualisation, Methodology, Software, Investigation, Formal analysis, Visualisation, Writing - (Original Draft or Review & Editing), and Supervision.
- Brief background on the chapter.
- Link to video, if applicable.

Part I

Visualising Virtual Environments to Augmented Reality Users

Chapter 3

The Body in Cross-Reality: A Framework for Selective Augmented Reality Visualisation of Virtual Objects

[103] HAN, J.*, **COOLS, R.***, AND SIMEONE, A. L. The Body in Cross-Reality: A Framework for Selective Augmented Reality Visualisation of Virtual Objects. In *Proceedings of the International Workshop on Cross-Reality* (XR) Interaction co-located with 14th ACM International Conference on Interactive Surfaces and Spaces (ACM ISS 2020) (XR 2020). CEUR Workshop Proceedings, paper 6.

Jihae Han: Conceptualization, Writing **Robbe Cools:** Conceptualization, Writing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing.

This chapter is co-authored with Jihae Han, and was written at the start of the PhD. In it we present early ideas of how a VE and the VR user in it can be visualised to an AR user. It is the result of equal contributions in terms of conceptualising and writing.

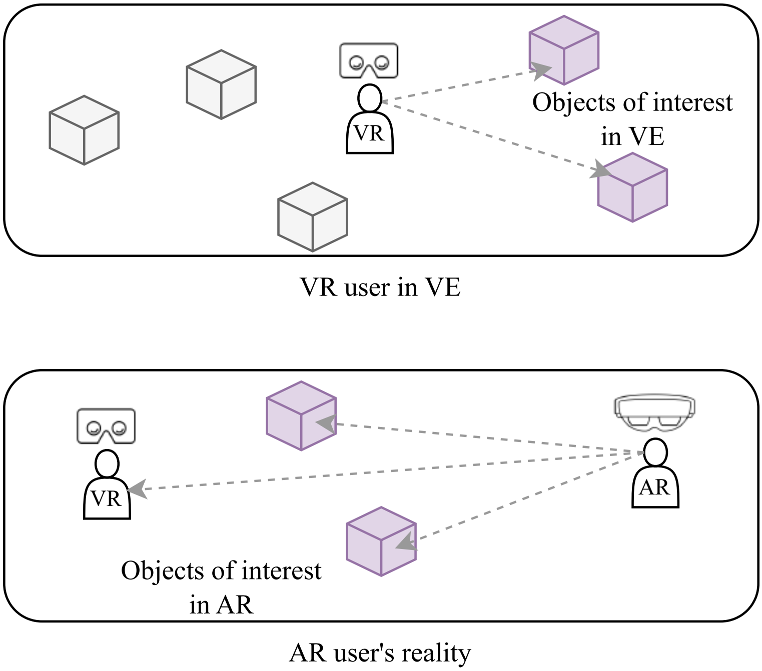


Figure 3.1: A selection of the elements of a Virtual Environment (VE) is based on the VR user's position in the VE. This selection is then shown to the AR user to give context to the VR user's physical actions.

abstract The body plays a communicative function in interaction. It expresses how we respond, experience and interact with the world through action, movement, and gestures. In this chapter, we investigate the impact of the body in Cross-Reality Interaction between users of different realities in the Reality-Virtuality continuum. We propose a Framework for Selective Augmented Reality Visualisation of Virtual Objects that enables an external Augmented Reality user to perceive an immersed Virtual Reality user against different levels of information. The augmented reality user may observe the real body of the user in the context of visualised objects from the virtual environment, selected according to three criteria: Proximity Threshold, Field of View, and Importance Ranking. We aim to investigate how much and what type of virtual objects need to be visualised in order to convey clear information on the activity and physical engagement of the immersed Virtual Reality user. Two use cases are presented to which this framework can be applied: vocational training on food hygiene and a virtual exhibition for architecture.

3.1 Introduction

3.1.1 Cross-Reality Interaction

Cross-Reality Interaction is an emerging field within Human-Computer Interaction that investigates how users of different realities can interact with each other – 'realities' referring to the real environment, the virtual, and anywhere in between. This spectrum of realities is identified in Paul Milgram's Reality-Virtuality continuum [173], which ranges from the reality we all experience to a completely controllable Virtual Environment (VE). For example, Augmented Reality (AR) is situated closer to our own reality compared to Virtual Reality (VR). Thus, a 'Cross-Reality Interaction' describes an interaction originating from a reality at one point of this continuum and affecting a reality at a different point.

For users of different realities to interact, they must first be aware of each other, which involves both an expression and perception of information. The question then becomes how much and what type of information should be conveyed in order to communicate clearly and effectively. In this work, we propose a framework that visualises the body of an immersed VR user in the context of select elements from the VR user's VE. Specifically, we aim to investigate how to design Cross-Reality interactions in scenarios where it is important for an external AR user to understand the physical actions and activity of an immersed VR user.

3.1.2 The Body in Cross-Reality

We experience, understand, and interact with the world through the body [145]. Especially with the rise of AR and VR, interactive technologies offer new possibilities for physical engagement. Users of immersive technologies can manipulate virtual objects, traverse through fictional landscapes, and interact with the VE through increasingly more complex and creative means. The actions of the immersed user are often expressed through interacting with specific virtual objects or the context of the activity, generating an importance of understanding the body in relation to the VE. Current Cross-Reality research includes sharing a field of view among multiple users across the Reality-Virtuality continuum [173, 36], tracking the positions of external users in a VE perceived by VR users [164], or synchronising the manipulation of select objects in the VE between different Cross-Reality users [89]. However, despite the extensive theoretical research in Cross-Reality Interaction, few focus on the impact of body language in interaction design.

3.2 Related work

Systems exist to make VR users aware of what is happening in their surroundings. NotifiVR [82], for instance, explores different notifications and interruptions. These interruptions can be physical, such a person or a pet, or digital, such as a text message or voice call. The VR Motion Tracker [248] presents a widget that enables VR users to track external persons. In RealityCheck [107] a system is presented that composites images from the real world into the VE. RealityCheck also enables external users to view the VE via a projection into the physical environment.

Projections can be used in different ways to visualise the VE to external users. ShareVR [99] presents a way to support interaction between VR and external users via a floor projection and a tracked display. The tracked display functions as a 'window into the VE'. The floor projection shows the spatial layout of the VE, projected by two projectors set up in the environment. An alternative to placing the projector in the environment is to make it head-mounted. A head-mounted projector has been explored for VR [292] and AR [108]. In both cases a small motor controlled projector was positioned on top of the user's head to project virtual content into the environment.

There are different ways of using an external screen to visualise the VE. One way is to mount one or more small screens onto the Head Mounted Display (HMD) [36, 100]. Tablets have also been used to enable collaboration between a VR user and non-immersed user. Grandi et al. [89] performed a study investigating collaboration of a VR and tablet AR user. They found that VR-AR asymmetric collaboration performed better than AR-AR collaboration but worse than VR-VR collaboration. TransceiVR [151] enables communication between an immersed VR user and an external tablet user.

Vishnu [38] is a system for a remote expert to assist a local agent in a maintenance procedure. The remote expert can perform the correct actions in VR, which are then shown to the local agent via AR. The local agent can then perform the correct actions as shown by the remote expert. We propose an approach that also uses a combination of AR and VR, in a different manner with the VR and AR user co-located. In our approach the VR user is the one under observation. The AR user is the one observing the VR user, and can simultaneously view the physical VR user and virtual elements from the VE they are in.

3.3 A Framework for Selective Visualisation

Our proposed framework uses AR to augment the external user's view of the VR user with elements from the VE. This enables the external AR user to view the VR user's gestures and body language in the context of the VE. This is important as both

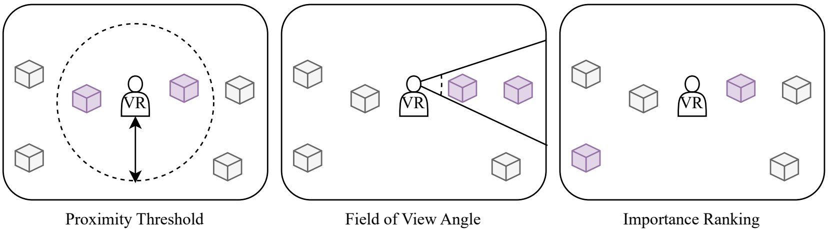


Figure 3.2: Criteria for selective visualisation: proximity threshold, field of view angle, importance ranking.

the stimuli (virtual object) and response (VR user's physical reaction) are involved in communicating an action-based interaction. To support this visualisation, two applications must be made, the main VR application and a companion AR application that can communicate with it.

We propose three criteria for selecting the scene elements that are visualised to the AR user: *Proximity Threshold, Field of View, Importance Ranking* (Figure 3.2). We expect different use cases for asymmetric interaction to require different selection criteria. This selection can also differ between users.

3.3.1 Proximity Threshold

We can selectively visualise objects nearby the VR user using a proximity threshold. We will implement this as the euclidean distance between the VR user's position in space and the position of the VE scene object. This way, only the relevant object within interaction distance may be selected for the augmentation.

3.3.2 Field of View

The Field of View refers to the area of the VE that the VR user is looking at, and how close the object is to this area. This is the angle between the VR user's forward vector and the vector between the VR user and the scene object on the yaw axis. This way, only the relevant objects in front of the VR user may be selected for augmentation. Eye tracking can potentially be used to more accurately identify which object has the VR user's attention.

3.3.3 Importance Ranking

Some objects in the VE are more relevant to display to external users than others. Importance ranking sets a predetermined selection of objects that are shown to the AR user. This selection can be binary, one group of objects to visualise and another group to hide in augmentation.

3.3.4 Technical Implementation

The software running on the VR system (eg. HTC Vive) and AR system (eg. Microsoft HoloLens) will be set up as networked applications. Two different versions of the application will be made, one for AR and one for VR. The VR application will be the host and contain the state of the VE. It will communicate this state as selected by the criteria described above with the AR client.

The AR application will be a client that contains minimal logic which displays the selection of objects as told by the VR host application. As users will be co-located, voice communication does not need to be networked. However, the AR user's position must be tracked in order to support a representation of this user. Having a representation of the AR user in the VE would prevent the AR user's voice from feeling disembodied and maximise immersion for the VR user.

3.4 Use cases

Current VR systems involve very physical types of interaction, such as walking and picking up objects. External observers may be interested in the relationship between the body and the interactive object in a VE, and body language can provide useful information about the nature of an interaction. We present two use cases in which we propose to apply our framework and evaluate how useful gestural information is to the perception of the AR user.

3.4.1 Vocational Training - Food Hygiene:

The employee (VR user) is immersed in the VE to become familiarised with the workplace, and the instructor (AR user) oversees the employee's performance in 'food hygiene' from how the VR user physically handles selectively visualised food objects.

- *Proximity Threshold*: The threshold is bound by the single station located by the VR user. Further food stations are not visualised.

- *Field of View*: The activity of handling food only occurs within arms-reach of the VR user, thus Field of View is a less relevant criteria for this use case.
- *Importance Ranking*: Visualising food and related tools is important to portray food handling, but other scene objects may be less relevant.

3.4.2 Virtual Exhibition - Architecture

The use case simulates an architectural review in which a client (VR user) explores and critiques an interactive 1:1 scale architectural project in the VE. The architect (AR user), who is already familiar with the project, perceives only the specific VE segments that the VR user is interacting with and referring to.

- *Proximity Threshold*: Less relevant as a 1:1 scale explorable architectural model is typically too large to effectively include within a visible threshold.
- *Field of View*: The VR user looks at different parts of the architectural proposal, and the AR user shares the VR user's line of sight when referring to specific architectural elements.
- *Importance Ranking*: Only relevant if the architect has included interactive architectural elements, such as furniture.

3.5 Evaluation

We seek to evaluate this approach in an experimental study. We will investigate changing the selection criteria (Proximity Threshold, Field of View, Importance Ranking) across different use cases, as well as conducting user experience studies focusing on how the AR user perceives the activity and bodily engagement of the VR user. As a baseline, we will also compare the selective visualisation framework to two other conditions: an external user that can see the VR user without AR, and a VR spectator which can see the entire VE without seeing the VR user.

3.6 Conclusion

In this chapter we presented a novel visualisation framework for selective visualisation of VE elements to an AR user. This visualisation allows the external AR spectator to view the physical movements of the VR user in context of the VE. All visual detail of the VR user's physical appearance may be preserved, including gestures and body language, while giving the context of the actions they are performing in the VE. As such, we aim to find the optimal amount of VE context to visualise to clearly and effectively convey the nature of an interaction between immersed users and their environment.

Chapter 4

SelectVisAR: Selective Visualisation of Virtual Environments in Augmented Reality

[48] **COOLS, R.***, HAN, J.*, AND SIMEONE, A. L. SelectVisAR: Selective Visualisation of Virtual Environments in Augmented Reality. In *DIS 2021 - Proceedings of the 2021 ACM Designing Interactive Systems Conference: Nowhere and Everywhere* (June 2021), Association for Computing Machinery, Inc, pp. 275–282.

Robbe Cools: Conceptualization, Methodology, Software, Investigation, Formal Analysis, Visualisation, Writing **Jihae Han:** Conceptualization, Methodology, Software, Investigation, Formal Analysis, Visualisation, Writing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing.

This chapter is co-authored with Jihae Han, and presents two studies that build upon the ideas developed in the previous chapter. It presents an equal contribution in research input between the two first authors in terms of conceptualising, implementing, conducting the user study, and writing the chapter. As supervisor Adalberto L. Simeone provided feedback and support for all aspects of this chapter.

https://youtu.be/IVzQCn9sKJo

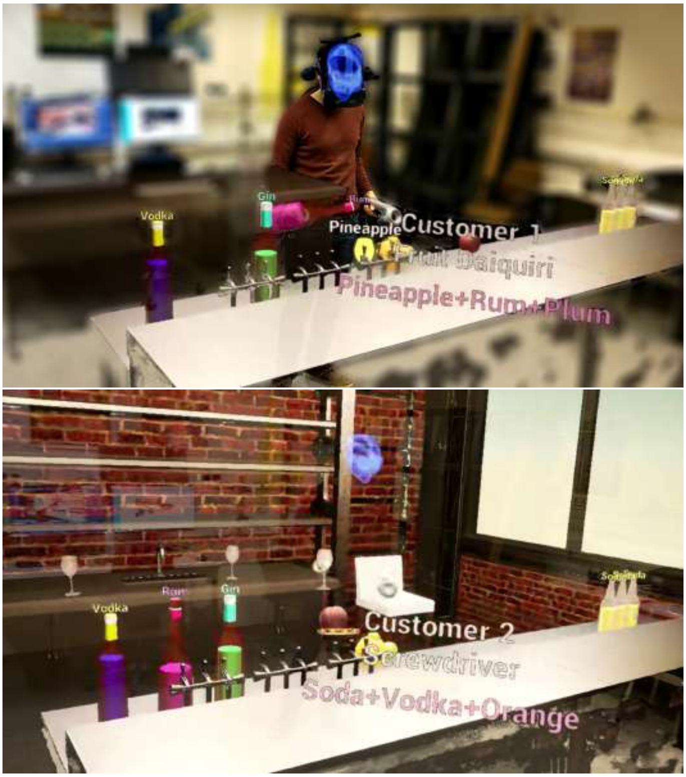


Figure 4.1: *Context* static selection technique (top) and *Everything* technique (bottom), AR User Perspective

abstract When establishing a visual connection between a virtual reality user and an augmented reality user, it is important to consider whether the augmented reality user faces a surplus of information. Augmented reality, compared to virtual reality, involves two – not one – planes of information: the physical and the virtual. We propose SelectVisAR, a selective visualisation system of virtual environments in augmented reality. Our system enables an augmented reality spectator to perceive a co-located virtual reality user in the context of four distinct visualisation conditions: *Interactive*, *Proximity*, *Everything*, and *Dollhouse*. We explore an additional two conditions, *Context* and *Spotlight*, in a follow-up study. Our design uses a human-centric approach to information filtering, selectively visualising only parts of the virtual environment related to the interactive possibilities of a virtual reality user. The research investigates how selective visualisations can be helpful or trivial for the augmented reality user when observing a virtual reality user.

4.1 Introduction

Virtual Reality (VR) and see-through Augmented Reality (AR) devices are becoming increasingly affordable. VR enables users to immerse themselves in a Virtual Environment (VE). A see-through AR device can overlay virtual content on top of the physical environment. VR and AR users can interact with each other through Collaborative VEs [313], and the interaction between a VR and AR user is considered a type of 'Cross-Reality Interaction'.

Cross-Reality (CR) is a field of research that looks at how users of different realities can interact with each other. These realities can be described through Milgram's Reality-Virtuality continuum [173, 251], which ranges from the real world to the virtual world and the spectrum of hybrid realities in between. This chapter focuses on scenarios that involve interactions between two different points in this continuum: AR and VR.

This type of scenario can be beneficial when the roles of the AR and VR user in the collaboration are asymmetrical. It is important to note the differences in how users perceive their VEs: VR users benefit from more immersion and AR users benefit from more nonverbal cues. In a CR context, nonverbal cues refer to the advantage AR users have over VR users when communicating with an external user — For instance, both VR and AR users can talk to an external user, but only AR users can see the physical body and gestures of an external user in real life. The VR user cannot see the external user, and at most can only perceive the external user's virtual avatar. As such, AR retains most nonverbal cues lost to VR users. In contrast, while VR users retain high immersion, AR users will be less immersed in the VE due to their vision of the physical environment.

For some users VR might be more desirable, such as for a training simulation where the user needs to have a sense of being at the place of the training. Other users can benefit from AR to enable them to see nonverbal cues of other co-located users. In this study, we aim to develop a CR scenario that exploits both the immersive benefits of VR and the nonverbal communication features of AR. We question how an AR user can spectate and interact with the VR user in their VE [103]. We propose a selective visualisation system that enables AR users to only see select virtual elements of the VE whilst VR users see the entire VE to maintain their immersion.

We investigated two design factors in terms of visualising a VE: level of information and scale. We designed a framework where we presented the VE to the AR user at a 1:5 dollhouse-scale and at 1:1 room-scale with three levels of information: no selection, a dynamic selection following the VR user, and a predetermined static selection. This study was then repeated with two improved dynamic and static selection techniques implementing feedback from the main study.

In our studies we found that participants felt they had a better overview of the VE at the small 1:5 dollhouse-scale; however, this had the drawback that it was more difficult to see smaller movements of the VR user. We found that our dynamic selection methods were preferred by fewer participants than the static selections. No significant differences were found in participant competence in recognising events in the VE.

4.2 Related Work

CR collaboration refers to users on different points on the Reality-Virtuality continuum [173, 251] working together. In order to support collaboration, users need to be aware of each other's "realities" and be able to interact with other users and their reality.

4.2.1 Tablet and screen-based VE Visualisation

Different technologies can be used to support CR visualisation and interaction. TransceiVR [151] enables a non-immersed tablet user to view the VE from the perspective of the immersed user by freezing the frame and making annotations that are communicated back into the VE. The real environment is disconnected from the VE, as the external user sees it from the perspective of the immersed user. FaceDisplay [100] mounted screens on the Head-Mounted Display (HMD), through which the external user can view the VE. This presents the VE from the perspective of the external user, however only when they are looking directly at the VR user's HMD.

Silhouette Games [147] presents an approach with a screen behind a one-way mirror. The screen displays a simulated reflection of the VE calculated based on the position of the non-immersed user. The non-immersed user can then view the VE and the physical reflection of the VR user simultaneously. Seeing both the VR user and their reflection caused some confusion in participants. Our AR-based approach does not rely on a reflection of the VR user, but visualises the VE directly around them. This does require the external user to wear an AR HMD which is more invasive than the approach presented in Silhouette Games.

4.2.2 Projection-based VE Visualisation

Wang et al. [292] mounted a projector on the HMD, which allows visualisation of the VE on the floor around the immersed user. The VR user had control over the content that was shown in the projection. In this chapter we investigate techniques that visualise the area of the VE around the VR user, changing the visualisation as the VR user moves.

ShareVR [99] combined a static floor projection, covering the entire space available to the VR user, with a handheld screen to enable interaction between an immersed user and an external user. ReverseCAVE [121] also used a projection-based approach, where the VE was not projected on the floor but on four translucent screens around the VR user that external users can then spectate from the outside. In our work we also investigated static visualisations covering the entire available space, however we used see-through AR instead of a projection.

4.2.3 Miniatures

Pham et al. [204] investigated the effect of the scale of AR visualisation on gestures, investigating models at 'in-air' scale, tabletop scale and room scale. They found that these different scales elicited different gestures from users. We will investigate the effect of the scale of the visualisation on an AR user spectating a VR user, inspired by Dollhouse VR [117] and World In Miniature [203]. We find further investigation on scale relevant as neither Dollhouse VR [117] nor its follow-up study [268] specifies or justifies the use of a specific scale when implementing the 'dollhouse', only detailing a relative size difference between visualisations. World In Miniature provides more but still relatively abstract detail regarding its implementation of scale, remarking that a World In Miniature may be 'hand-held' but not specifying a scalar value [203].

4.2.4 AR-based VE Visualisation

Grandi et al. [89] investigated collaboration between VR and tablet AR users. AR and VR users were co-located and collaborated on solving a docking task. The AR user saw the virtual objects with the same spatial orientation as the VR user. The shared virtual elements were limited to tabletop-size objects, whereas we investigated how to visualise the VR user in context of their VE to the AR user on room-scale.

ObserVAR [274] explores the use of see-through AR for a teacher to visualise their students, who are immersed in a VE using 3-DOF VR devices. Three visualisations were tested: *First Person View*, *World in Miniature* and *World Scale*. Participants found *World Scale* easier to use than *World in Miniature*, though scale was not the only factor because *World in Miniature* showed a separate miniature per VR user. The ObserVAR user study had multiple remote VR users, while our study had a single co-located VR user and focused on visualising them in context of the VE.

Interaction	VE Visualisation	Name
PR-VR	VR user's PoV (tablet)	*TransceiVR* [151]
PR-VR	VR user's PoV (tablet)	*FaceDisplay* [100]
PR-VR	top-down view of scaled 'dollhouse' VE (tablet)	*Dollhouse VR* [117]
PR-VR	own PoV (tablet) + top-down view (projection)	*ShareVR* [99]
PR-VR	VR user's PoV (projection)	*HMD Light* [292]
PR-VR	4 perspectives (CAVE)	*ReverseCAVE* [121]
AR-VR	AR/VR PoV	CR collab. [89]
AR-VR	'arrow' for VR user's gaze, AR/VR PoV, scaled VEs	*ObserVAR* [274]
AR-VR	selective 'filtering' of VE + scaling of VE	**SelectVisAR*
VR-VR	selective 'slices' of VE	*Slice of Light* [291]
VR	world-in-miniature of VE	miniature [203]
AR	scaling/sizing VEs	holograms [204]

**SelectVisAR = our own study, in context of related work*

Table 4.1: Table summarising related works.

4.2.5 Selection within VEs

Slice of Light [291] presents a method for a VR user to see and move between VEs of other VR users. The other users' VEs are visualised as slices around the user in that VE, the external user can then enter that VE be stepping towards it. The purpose of presenting the VEs as slices is so that multiple users in their VE can be shown at once. We investigated if filtering what is shown of the VE can improve AR users' understanding of the VR user's actions. To do this we tested both static and dynamic selections of virtual content. A summary positioning our work to the related work can be seen in Table 4.1. The table characterises the visualisation of the VE according to the Point of View (PoV) from one or a combination of VR, AR, and Physical Reality (PR) users.

4.3 Visualisation and Implementation

The selection methods we designed aim to visually emphasise the actions of the VR user to an AR spectator by selectively filtering relevant parts of a VE – specifically, the visual artefacts the VR user is interested in or interacting with. This VE is asymmetrically filtered only for the AR user; the VR user would see a fully visualised VE to maintain user immersion. We hypothesise that it is possible to remove a part of the VE without hurting task performance of how an AR user perceives the actions of a VR user.

4.3.1 Pilot Study

We conducted a pilot study with three HCI experts and a prototype $5\,\mathrm{m} \times 5\,\mathrm{m}$ virtual room as the VE. The HCI experts had varying degrees of experience with CR technologies: one expert, one with previous experience, and one without any experience. We conducted a *Think-Aloud Protocol*, with one researcher taking notes and the other assuming the role of the VR user, to prototype our framework and reduce the number of techniques being tested for the main study. We investigated six different visualisation techniques based around the interactive range and possibilities of the VR user, categorised as either static or dynamic visualisations:

Static visualisations select parts of the VE to visualise at all times of the simulation. **(1)** *Everything* visualises the entire VE to the AR user. This is the control condition in which AR users see the VE in the same way as the VR user. No changes are made to augment or filter information in the visualisation of the VE. **(2)** *Interactive* is a predetermined selection of interactive objects in the VE. This is inspired by literature that suggests only relevant information should be visualised to prevent overloading the user with irrelevant information [130, 90]. As 'relevant information' is an abstract term, we attempt to draw thresholds in information filtering using *Interactive*. Lastly, **(3)** *Dollhouse* visualises a smaller, scaled model of the VE. Directly based on Ibayashi et al.'s Dollhouse VR [117], this visualisation is grounded in previous literature [204, 203] that argue that scaled visualisations of VEs enable more efficient navigation of a VE. However, instead of a 2D top-down view of the VE as investigated in *Dollhouse VR* [117], we investigate a 3D scaled model of a VE in AR.

Dynamic visualisations select different parts of the VE to visualise, depending on where the VR user is located or what the VR user is doing. **(4)** *Head-Direction* only visualises the part of the VE that the user is facing towards. **(5)** *Proximity* visualises a radial area of VE nearby the VR user. The conditions *Head-Direction* and *Proximity* are inspired by *Slice of Light* [291], a visualisation which shows only part of the VE to the guest VR user and dynamically changes depending on the user's location or actions. We based these two conditions on common tracking methods for VR, head-direction tracking for *Head-Direction* and position-tracking for *Proximity*. Lastly, **(6)** *Dynamic-Interaction* visualises the virtual objects that the VR user is currently interacting with using the controllers. *Dynamic-Interaction* is a responsive implementation of the static condition, *Interactive*, which enables the users to filter information depending on their hand motions.

Based on a preference ranking and informal interviews, we decided to remove two visualisation techniques that participants liked the least for the main study: *Head-Direction* and *Dynamic-Interaction*. Regarding *Head-Direction*, participants found that frequent changes to a VR user's line of sight and thus the visualisation made the technique confusing. Regarding *Dynamic-Interaction*, participants found the visualisation difficult to understand as too little information was being shown. We also improved some visualisations, such as *Dollhouse* which participants complained that the visualisation was too small to see clearly. We thus increased the scale of the visualisation to 1:5 (1 m × 1 m) from its original 1:10 (0.5 m × 0.5 m) scale.

4.3.2 Selective Visualisation Framework

Using our pilot study to refine our design, we selected four visualisation techniques for the main implementation of the study:

- **Everything**: A VR-mimicking condition where the entire VE is visualised to the AR user at 1:1 scale. No modifications are made and the visualisation is symmetrical between the VR and AR user.
- **Proximity**: An 'arm's-reach' approach that dynamically visualises a 1 m radius of VE around the VR user, with an additional 0.5 m radius of decreasing opacity to fade-out the visible threshold of the technique. The parts of the VE visualised changes depending on the location of the VR user.
- **Interactive**: A static, predetermined visualisation of interactive movable objects in the VE. This is a selection of objects that the VR user can pick up and interact with using their controllers. As a static visualisation, all the interactive objects are visualised at all times.
- **Dollhouse**: A 1:5 scaled visualisation that provides a top-down overview of the VE, which hovers 1 m above floor-level. The walls and ceiling of the virtual model are removed to facilitate looking into its interior.

These techniques can be seen as diagrams in Figure 4.2 and from the AR user's perspective in Figure 4.3.

4.3.3 Implementation of System

We developed this selective visualisation system in Unreal Engine 4.25 as a networked application, which runs on two computers on the same LAN with one as the server and the other as a client. We used a HTC Vive Pro and a Microsoft HoloLens 2, both of which have their own coordinate systems: lighthouses managed by SteamVR and embedded camera-based tracking respectively. Using a custom calibration procedure the coordinate systems are aligned. This procedure consists of scanning two QR codes

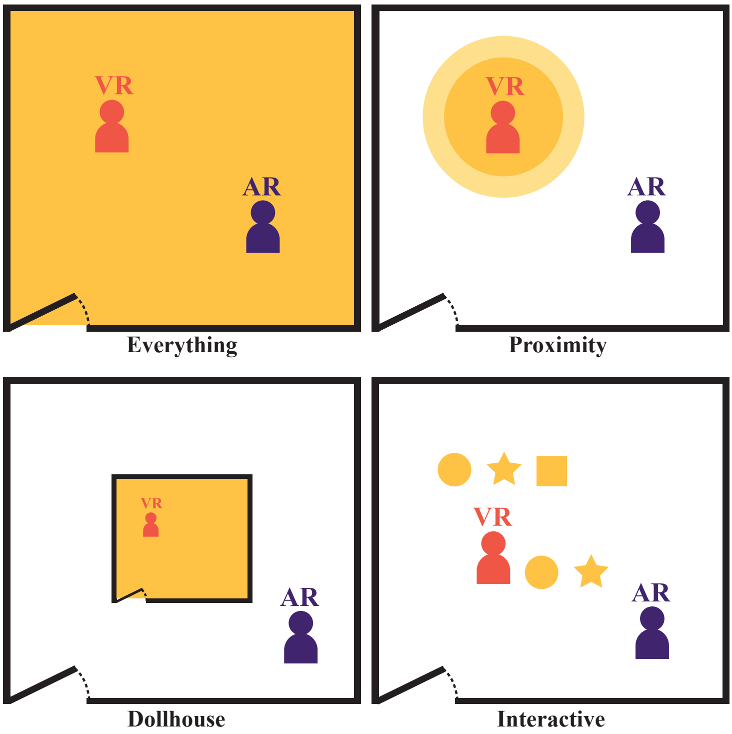


Figure 4.2: Selective Visualisation Techniques of VE: Entire VE is visible to the VR user in all conditions, while only coloured areas are visible to the AR user.

with the HoloLens, and placing the HTC Vive controllers on top of these codes. With two corresponding points in both coordinate systems known, the origin of the HoloLens coordinate system is transformed so that these points overlap with the corresponding points in the SteamVR coordinate system. In operation drifts between the coordinate systems can be observed up to a maximum of 5 cm.

4.3.4 The Virtual Environment

For the purposes of testing the visualisation system, we created a 'bartender simulation' as the VE. The participant assumes the role of the AR user within this simulation because we are investigating the AR user's perception of how a VR user interacts with a VE. The researcher assumes the role of VR user and conducts a 'performance' using a predetermined script for the AR user to observe. This performance consists of making cocktails using three different recipes, and the order of the recipes and the actions

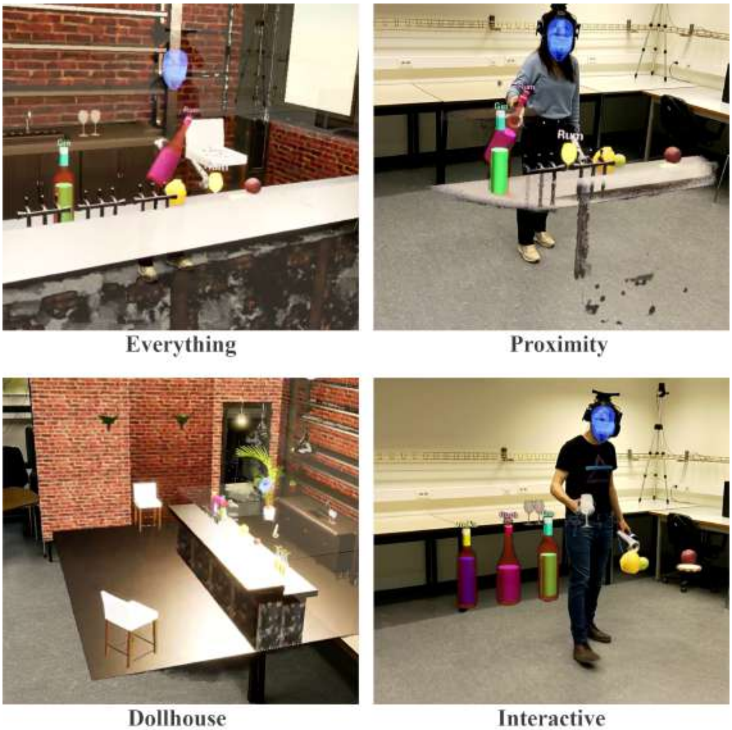


Figure 4.3: Selective Visualisation Techniques of VE (Mixed Reality Capture).

performed for making them differed between the different visualisation techniques. The VR user used three types of interactive objects to perform this script: fruits, bottles and glasses. All these objects can be picked up and moved with the VR user's motion controllers. The glasses can hold slices of fruit and liquids. The contents of the glass are indicated by floating text above it. A fruit is added on entering the collision box of the glass. Liquids are only added when the top of the bottle collides with the glass, to mimic a pouring motion. When the glass is held upside-down the contents are emptied. On the bar counter there is floating text indicating the current order and a simplified three-ingredient recipe. Below this text is a collision box that checks the glass contents on collision, and when the contents are correct empties the glass and advances to the next recipe. These are the different events that the AR user can perceive in the VE that are triggered by the VR user.

4.4 Main User Study

4.4.1 Procedure

We recruited thirteen participants for the main study, aged between 21 and 57 (M=30.62, SD=12.48; 6 male, 7 female). They had a low self-reported experience with VR and AR technologies (M=3.08, SD=1.19 on 7-point scale).

Participants were tasked with using the HoloLens 2 to observe a VR user that is performing a bartender simulation. The HoloLens was set to the highest brightness setting. *Holographic remoting*[1] was used to stream the image from the computer to the HoloLens. Participants were given an 'event recognition task', a list of events which they need to recognize as they happen. These events are triggered by the VR user's actions. The researcher used a HTC Vive Pro with the Vive wireless attachment to perform the role of the VR user, following a predetermined set of actions on each trial. During each trial the VR bartender made three drinks, consisting of combining three ingredients in a glass each. Participants performed four trails, one for each technique, during which they could move around the lab to adjust their viewpoint. The study lasted about 40 min.

Before taking part in the study, participants signed a consent form and filled in a demographics questionnaire, then we explained to them the event recognition task and instructed them on how to use the Microsoft HoloLens 2. Before starting the first trial participants were given some time to look around the bar environment and get to know the positions of all the objects. The techniques were presented in counterbalanced order, using a balanced Latin square. During each trial participants were required to pay attention to the VR user and the VE. After each trial participants filled in which events they saw happen, Slater-Usoh-Steed's (SUS) presence questionnaire [283], and Kennedy's Simulator Sickness Questionnaire [137]. After the last trial participants were asked to rank the techniques (1st, 2nd, 3rd and 4th), and were interviewed on their thoughts on the techniques and the experience in general.

The study took place during the COVID-19 pandemic. The keyboard, mouse and desk area used by the participants were disinfected before and after the study, as well as the HoloLens 2 for which a Cleanbox UV-C decontamination device[2] was used. Participants and researcher disinfected their hands before and after the study, wore face masks and maintained a distance of at least 1.5 m between them. There was at least 30 min between participants to avoid them meeting and allow time to disinfect and ventilate our lab. The study and COVID-measures were approved by the university's privacy and ethics board (PRET).

[1] https://docs.microsoft.com/en-us/windows/mixed-reality/develop/platform-capabilities-and-apis/holographic-remoting-player

[2] https://cleanboxtech.com/

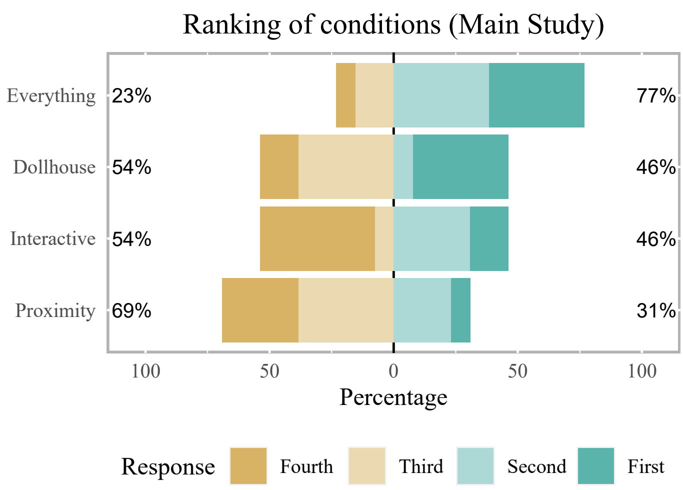


Figure 4.4: Participant rankings of techniques in the main study.

4.4.2 Results

Preference Ranking

Participant preference of the techniques can be seen in Figure 4.4. Following a pairwise Wilcoxon signed rank test, the *Everything* technique was ranked significantly higher than *Interactive* ($p<0.05$) and *Proximity* ($p<0.05$), with 77% ranking it at the first or second place. Both *Interactive* and *Dollhouse* were ranked first or second by 46% of participants. Only 31% of participants ranked *Proximity* in first or second place.

Event Recognition

We used a competence calculation (True Positive Rate - False Positive Rate) [302] to analyse how well the participants understood the VR user's actions during the event recognition task. 'Competence' is the probability of knowing a correct answer without guessing and not by chance, and in this context refers to the probability of an AR user correctly identifying the actions conducted by the VR user. A Kruskal-Wallis test showed no significant difference ($p=0.93$) in competence across the four visualisation conditions. However, we observed a marginal difference in the mean competence that favoured the filtered visualisations. The competence values range up to 1, representing a participant that only indicated the correct events. The mean competence ranges between 0.65-0.76, and from highest to lowest: *Interactive* ($M=0.76$, SD=0.32), *Proximity* (M=0.71, SD=0.31), *Everything* (M=0.67, SD=0.48), and *Dollhouse* (M=0.65, SD=0.42).

Interviews

We analysed the interview using a thematic analysis [28]. We categorised user responses in three themes: Firstly, the ability to focus on the simulation; secondly, the presence of the VR user; and thirdly, feedback on the visualisation methods. Participants found it harder to focus in the *Everything* condition, with five participants finding real objects distracting and two finding virtual objects distracting. In contrast, five participants found it easier to focus in the more visually filtered *Interactive* condition. This is higher than the number of people who stated that *Dollhouse* or *Proximity* helped focus, which was two. Regarding the presence of the VR user: five participants commented that rather than the physical appearance of the VR user, they found themselves focusing on the actions being conducted. Any mentions of the physical appearance of the VR user only arose from room-scale conditions, even if the VR user was co-located in all the conditions. Regarding feedback for the visualisation conditions: For *Interactive*, six participants found they missed the bar counter as a point of reference in the scene. For *Proximity*, two participants complained about having less control over the visibility of virtual artifacts, and two other participants about wanting to stay aware of the invisible part of the VE.

SSQ and SUS

A Kruskal-Wallis test showed no significant difference (p=0.91) between the SSQ's Total Score, with the following mean values: control (M=9.81, SD=8.26), *Dollhouse* (M=8.01, SD=9.96), *Everything* (M=9.94, SD=10.6), *Interactive* (M=6.64, SD=7.31) and *Proximity* (M=7.57, SD=7.54). The SUS questionnaire was analysed by counting the number of 6 and 7 answers. There was no significant difference, though the mean for *Everything* (M=1.15, SD=1.52) was higher than for the other conditions: *Dollhouse* (M=0.08, SD=0.28), *Interactive* (M=0.08, SD=0.28) and *Proximity* (M=0.15, SD=0.38).

4.5 Follow-up User Study

4.5.1 Changes to Selective Techniques

Issues with the selective techniques were found in the main study, these were addressed and the resulting improved techniques evaluated in a follow-up study.

In feedback given during the interviews, participants mentioned issues with our selective techniques: for *Interactive* they missed the bar counter as a point of reference and for *Proximity* it was confusing that the environment disappeared completely which removed all context in which to see the highlighted area around the VR user. We thus implemented two new selective techniques to address this feedback:

- **Spotlight**: An improved version of *Proximity* in which the VE in proximity of the VR user is rendered opaque, but modified to enable the AR user to see the rest of the VE as simple outlines. This allows users to see the highlighted area in context of the rest of the VE without it obstructing view of the physical environment.
- **Context**: A refinement of *Interactive* that responds to the participants' desire to see more of the VE. The furniture that supports the interactive objects can now be seen, i.e. the counter and sink.

These techniques can be seen as diagrams and from the AR user's perspective in Figure 4.5.

4.5.2 Procedure

The follow-up study followed the same procedure as the first study described in subsection 4.4.1, with the *Proximity* technique replaced by *Spotlight* and the *Interactive* technique replaced by *Context*.

For the follow-up study we recruited 13 participants, aged between 19 and 57 (M=29.77, SD=14.35; 6 male, 7 female). With low self-reported experience with VR and AR (M=2.92, SD=1.25 on a 7-point scale).

4.5.3 Results

Preference Ranking

Participant preferences can be seen in Figure 4.6. A pairwise Wilcoxon signed-rank test showed that *Spotlight* ranked significantly lower than the other techniques ($p<0.01$ for *Everything* and *Context*, and $p<0.05$ for *Dollhouse*) with 92% of participants ranking

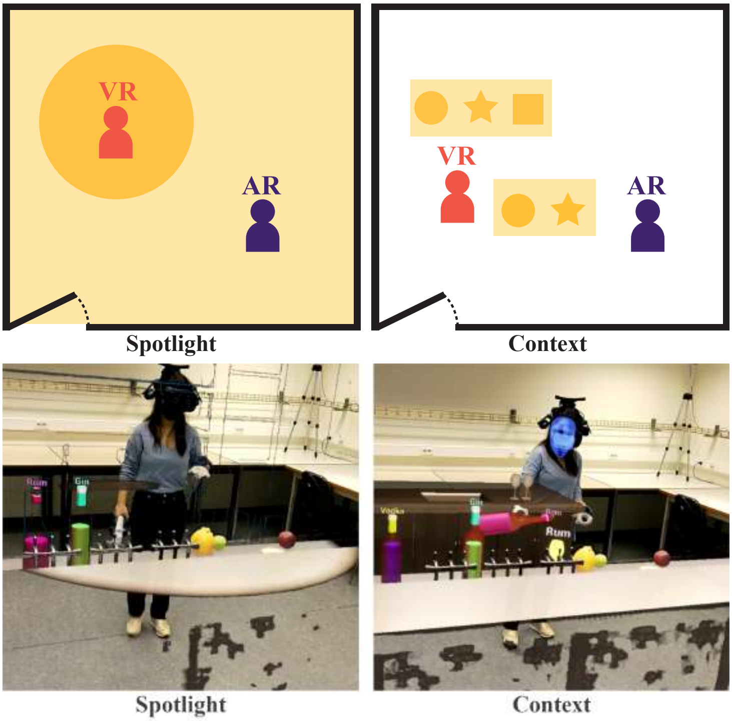


Figure 4.5: Improved Selective Visualisation *Spotlight* (left) and *Context* (right) Techniques diagrams (top) and as Mixed Reality Capture (bottom). The whole VE is visible to the VR user in all conditions, while only coloured areas are visible to the AR user (diagram).

it third or fourth. *Dollhouse* was ranked significantly lower ($p<0.05$) than *Everything* with 54% of participants ranking it third or fourth, and 77% ranking *Everything* first or second. 69% of participants ranked *Context* first or second.

Event Recognition

There were no significant differences in how well participants could recognise events (Kruskal-Wallis test, p=0.72). However competence values for selective conditions were marginally higher. The mean competence ranges 0.67-0.83, and from highest to lowest: *Context* (M=0.83, SD=0.25), *Everything* (M=0.76, SD=0.43), *Spotlight* (M=0.75, SD=0.31) and *Dollhouse* (M=0.67, SD=0.43).

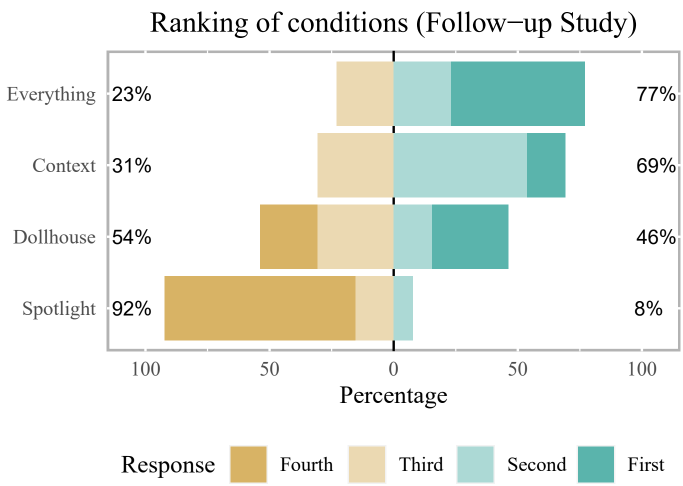


Figure 4.6: Participant rankings of techniques in the follow-up study.

Interviews

The same three themes were identified in the interviews as in the main study: ability to focus on the task, the presence of the VR user and feedback on the methods. Four participants found the VE distracting in *Everything*, while three participants were distracted by the VE represented as outlines in *Spotlight* finding that the out-of-focus environment was unclear. Three participants also found that *Spotlight* helped them focus more on the bartender. Four participants only saw the bartender as an avatar, while six others mentioned that they could see the physical person behind the avatar in the room-scale conditions. Three participants said they could not see the bartender well enough. Six participants found *Everything* and *Context* very similar, five participants even expressed difficulty in discerning these two techniques. Three participants expressed frustration with the limited vision in *Spotlight*. Nine participants found that *Dollhouse* gave them a good overview of the VE, five participants found it too small.

SSQ and SUS

A Kruskal-Wallis test showed no significant difference (p=0.96) between Total Score on Simulator Sickness, with the following mean values: control (M=4.72, SD=13.65), *Dollhouse* (M=3.21, SD=7.67), *Everything* (M=2.75, SD=5.17), *Context* (M=1.86, SD=3.47) and *Spotlight* (M=4.09, SD=10.19). The SUS questionnaire was analysed by counting the number of 6 and 7 answers. There was no significant difference (p=0.52), though the mean for *Everything* (M=1.46, SD=2.22) was higher than for the other conditions: *Dollhouse* (M=0.62, SD=1.45), *Context* (M=0.77, SD=1.69) and *Spotlight* (M=0.62, SD=1.19).

4.6 Discussion

We hypothesised that it would be possible to remove parts of the VE that are non-essential to the task being performed in it without altering an external user's perception of the task itself. Our results indicated that removing a large part of the VE indeed does not create a significant difference in how well an AR user can identify the events triggered by a VR user. However, our findings also reveal that competence does not necessarily correspond with user preference on the different visualisations, and we identified that participants preferred to see the supporting furniture of visible objects and did not prefer to lose control over the visualisation.

An initial static selection *Interactive* was also not preferred. Participants indicated that they could not see enough of the VE. We developed an improved iteration, *Context*, which showed relevant furniture in addition to the original objects of the VE. Participants expressed a more positive response for this visualisation, commenting that *Context* was similar to seeing the entire VE. Some of the participants even expressed being unable to tell the difference between this selection and seeing *Everything*. Between the main and follow-up study, the preference ranking of this static visualisation has increased by one rank.

On the effect of scale we found that *Dollhouse* provided a better overview of the VE, but also that many participants found it too small. We were able to use a 1:5 scale because our VE was only as large as the physical size of the room. Larger VEs need to be scaled down more, which can make the issue of them being too small worse, or not shown entirely which can make users lose their overview on the VE. In ObserVAR [274] participants found the *World Scale* condition easier to use than the *World in Miniature* condition which is supported by our results where participants preferred the *Everything* and static selection room-scale techniques over the *Dollhouse* technique. The results from ObserVAR indicate that their *World Scale* provided a better overview, which contradicts our results that indicated *Dollhouse* as the technique providing a better overview. This can be explained by the ObserVAR implementation of *World in Miniature* that visualises a separate miniature for each VR user, thus splitting up the information required by their user study participants.

Two types of selection were investigated, a predetermined static selection of objects and a dynamic selection that follows the VR user. Participants did not prefer the dynamic selection, citing lack of control over what they could see in the VE. A similar trend was cited in HMD light, in which external users looking into a VR user's VE wanted to have more control over the visualisation of the VE [292]. Further comparisons can be made with HMD Light regarding user preferences on the stability of the visualisation. Comparing a third person view with a first person view, most users in HMD Light chose third person view because as it was more stable and holistic than the 1st person view [292]. In our study, more holistic and static visualisations such as *Everything* and *Dollhouse* were preferred over more dynamic visualisations such as *Proximity* or *Spotlight*.

Participants in selective visualisations such as *Interactive* or *Context* could identify VE events more accurately than seeing *Everything*. Compared to other visualisations, the majority of the participants highlighted some distracting features in the *Everything* visualisation. In contrast, a number of participants cited *Interactive* in particular as useful for maintaining focus, despite the lower preference rating compared to *Everything*. However, it is important to note that statistically we have found no significant differences proving that more selective visualisations improve focus. We have only observed that the mean values for competence are marginally higher for the selective visualisations in this instance of a 'bartender' VE. Further investigation is necessary, perhaps with a range of different levels of information that incorporate tasks of greater complexity.

For researchers and developers in CR, we recommend different visualisations depending on the purpose and appearance of the VE. For a VE that requires the AR user to have an overview of the space, the *Dollhouse* condition has shown to be the most effective of those evaluated in this study. It is important to note that the scale of the *Dollhouse* depends on the size of the VE, as very large VEs are potentially limited by the physical space available even when scaled, and there exists a limit to how small a VE can be visualised before the AR user no longer understands what the VE represents. Additionally, the VE should be visualised as a static selection as opposed to a dynamic selection whenever possible, as static selections have shown to rank higher in terms of user preference. Lastly, it is possible to remove all non-essential information and preserve the recognition of events, but showing the immediate context matters for user preference.

Guidelines for SelectVisAR Systems

1. Use *Dollhouse* when the AR user requires an overview of the VE.
2. Use a static selection as opposed to a dynamic selection when possible.
3. Showing the immediate context improves user preference, however, it is possible to remove non-salient information and preserve the recognition of events.

4.7 Conclusion and Future Work

In two studies we investigated how a selective visualisation system of VEs can influence an AR user's perception of a co-located VR user. We looked at two variables: the level of visual information and the effect of scale. Regarding level of visual information, we observed that filtering specific selections of the VE did not significantly affect the competence of how well people could identify events in the VE. These selections were based on the interactive range and possibilities of the VR user. Regarding scale, users generally agreed that smaller visualisations provide a better overview of the VE, but had the chance of decoupling the user from the task at hand. In terms of user preference, our qualitative data showed that participants tended to prefer static visualisations over dynamic visualisations, disliking the lack of control they could exercise for visualising the VE.

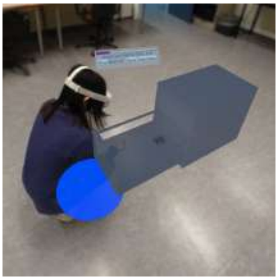
In future work we would like to improve these visualisations to apply to a greater variety of VE contexts. Techniques such as *Proximity* are generalisable, but techniques such as *Context* are very specific to the context of the VE as they use a predetermined selection of virtual objects. This selection of visualised objects can be made in different ways, instead of a predetermined selection future work can investigate the creation of an interface for the AR, or VR, user to make this selection themselves.

Moreover, we would like to apply these visualisations into an interactive implementation of this visualisation system, as currently the AR user only assumes a passive spectator role in the task. We could test our visualisations in a collaborative task that requires both the AR user and VR user to interact with elements from the VE.

Chapter 5

CReST: Design and Evaluation of the Cross-Reality Study Tool

- [53] **COOLS, R.**, ZHANG, X., AND SIMEONE, A. L. CReST: Design and Evaluation of the Cross-Reality Study Tool. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (New York, NY, USA, Dec. 2023), MUM '23, Association for Computing Machinery, pp. 409–419.
- **Robbe Cools:** Conceptualization, Methodology, Software, Investigation, Formal Analysis, Writing **Xuesong Zhang:** Software, Investigation, Writing - Review & Editing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing
- This paper is a collaboration with Xuesong Zhang, who provided the case study in section 5.6, in which we applied CReST to a project she was working on regarding graphic and interaction fidelity. Moreover, this paper uses previous work from myself [51], Stan Depuydt [247], and Jihae Han [104], as replicated examples in section 5.4.
- https://youtu.be/nvT5V2UdkS0



(a) Participant VR perspective (b) Researcher AR perspective

(c) Researcher Desktop UI

Figure 5.1: A trial in the case study from the participant's perspective (Figure 5.1a), the UI displayed on the desktop screen (Figure 5.1c), and from the researcher's Augmented Reality perspective (Figure 5.1b). The participant is opening the bottom drawer to find the number labelled '2', the blue circle on the floor indicates their starting position. The researcher UI shows a progress bar, indicates that the participant ID is 17, the study is in the drawer block, on the second trial of four, which is 'Hand Low Fidelity'. The buttons in the centre of the screen allow the researcher to make a note of when the participant performed a failed interaction, or to take a screenshot.

abstract In this work, we describe our experience developing and evaluating the Cross-Reality Study Tool (CReST), which allows researchers to conduct and observe Virtual Reality (VR) user studies from an Augmented Reality perspective. So far, most research on conducting VR user studies has centred around tools for asynchronous setup, data collection and analysis, or (immersive) replays. Conversely, CReST is centred around supporting the researcher synchronously. We replicated three VR studies as example cases, applied CReST to them, and conducted an interview with one author of each case. We then performed a case study, and recruited 17 participants to take part in a user study where the researchers used CReST to observe participant interaction with virtual drawer and closet artefacts. We make CReST available for other researchers, as a tool to enable direct observation of participants in VR, and perform rapid, qualitative evaluations.

5.1 Introduction

To facilitate synchronous participant observation we introduce the Cross-Reality Study Tool (CReST), which allows researchers to observe and control Virtual Reality (VR) user studies from an Augmented Reality (AR) perspective. With Cross-Reality (CR) we refer to a system where users are immersed at different points on the reality-virtuality continuum [173, 251] concurrently. We chose such a system because of the asymmetry in roles between actors involved in a VR user study. The participant needs to achieve immersion in the Virtual Environment (VE) for the study to provide valid results. Conversely, the researcher does not need to be immersed in the VE, and would benefit from situational awareness of both the real world and VE. Hence, CReST presents the researcher with an AR version of the VE, allowing them to see the participant's VE interaction while preserving real-world situational awareness. We aim to address the challenge of VR evaluation approaches [7] by facilitating synchronous qualitative evaluations.

Qualitative evaluation methods allow researchers to gain a deep understanding of a phenomenon, and formulate appropriate hypotheses that can be tested using quantitative methods. However, only few methods allow for qualitative data collection, with available tools focusing on quantitative data recording [169]. Tools that facilitate recording metrics [312] and in-VR questionnaires [234] have been investigated, however, observing participants remains a challenge due to the isolating nature of VR devices. Off the shelf VR equipment immerses users in a VE, while only providing bystanders with a limited mirrored view of what the VR user sees. While asynchronous immersive replay functionality allows the researcher to relive the study [114] in VR, synchronous observation of participants physical and virtual actions remains an unexplored challenge.

The contributions of this chapter are the following:

1. The iterative development of the Cross-Reality Study Tool (CReST), guided by three example cases and expert interviews. CReST was made available on GitHub[1].
2. Demonstration of CReST on a case study on graphic and interaction fidelity where three researchers used it to gather qualitative observational data from 17 participants.

5.2 Related Work

VR has been used for a variety of research applications, for example education [213], architecture, engineering, construction [321], and social psychology [317]. Hence, we

[1]https://github.com/AriaXR/CReST

first provide an overview of existing tools for supporting the different steps in setting up, running and analysing the data from a VR user study. Then we discuss CR, and motivate the CR methodology we employ in CReST.

5.2.1 User Study Tools

We discuss tools in support of the following five steps in running VR user studies: setup of interactive VE, procedure control, participant observation, data gathering, and data analysis. First, researchers have to set up an interactive VE. Second, the researcher conducts the study, and is in control of its progression. Third, they can observe the participant's interaction in the VE. Fourth, qualitative and/or quantitative data are gathered. And fifth, these data are analysed.

The *setup of the VE* has been facilitated by providing ready-to-use environments and UI elements [96], as well as procedural generation and furnishing of rooms [286]. Creating interactive VEs has been facilitated through trigger and toggle components [279], or visual editors to define user tasks [182]. However, available software, such as Unity or Unreal Engine, already adequately supports the creation of interactive VEs for researchers who are experienced with them [30], thus we chose not to focus on this aspect in CReST.

Progression through the *user study procedure* can be handled via UIs and researcher-initiated actions [159, 264], through a pre-defined order [96], or randomised [286]. However, within-group designs often use counterbalancing, for example, UXF [30] allows researchers to program a study procedure using the concepts of sessions, blocks, and trials, allowing the framework to automate the procedure. Similarly, VRSTK [312, 113] allows researchers to set up a list of stages, through which a controller component will progress the study. Tools can also include control panels [30, 312] with which the researcher can adjust settings while conducting the study. From our own experience in applying counterbalancing during user study procedures, we see value in supporting (partial) modelling and automation of the procedure and included similar functionality in CReST.

Support for *participant observation* is more limited, in addition to observing mirror views of the participant perspective, tools also allow the researcher to join the participant in the VE [312, 159, 264] or augmented environment [265] to observe while co-present. For VR studies, co-presence in the VE limits the researcher's view of the participant to their avatar, rather than their physical appearance. We see potential in more natural observation of co-located participants during the study itself, to allow the researcher to gain valuable insights quicker rather than during post-analysis. Insights that can be used during subsequent interviews with the participant, inquiring why they exhibited certain behaviour, or to allow for more rapid iteration on prototypical implementations.

Most tools we discuss here feature some form of *data logging* capability, such as replays [312, 159, 264, 114, 182], or scripts to more easily create log files [279, 30]. Tools have been developed to facilitate the recording of data in MR user studies. Since many tools already include data-logging capabilities, such as the Unreal engine replay system [2], we chose to not further investigate it for CReST.

Data analysis support ranges from R and Python script examples [312], to desktop dashboards and immersive analytics [114, 182, 35]. Both MRAT [182] and MIRIA [35] allow users to visualise collected data in-situ using AR. MRAT includes a visual editor in which users can easily select the data to collect, and is mainly event-based, while MIRIA centres around the visualisation of user movement by showing trajectories and heatmaps, with 3D models, pictures and videos giving context to the data. ReLive [114] combines both in-situ and ex-situ visualisation, allowing researchers to explore the data in the VE, but also analyse it through more conventional non-immersive visual analytics. IDIAR [287] is an AR and smartphone dashboard for researchers to monitor mobile intervention studies.

We chose to centre CReST around participant observation and procedure control, which both benefit the researcher while running a user study. The primary goal of CReST is to allow researchers to observe the participant in the VE, in order to gather useful insights related to their research question. To facilitate this, the second goal is to streamline supervised user study sessions, by allowing users to define a procedure consisting of sessions, blocks and trials so that advancing through the procedure can be automated, taking away this responsibility from the researcher. They can then focus on observing the participant using AR, for which we implemented a CR visualisation. Ledo et al. [156] identified demonstration, usage, performance, and heuristics as different evaluation types for HCI toolkits. From these we focused on demonstration as the most suitable for CReST, and used replicated examples and a case study.

5.2.2 Cross-Reality

Simeone et al. characterise Cross-Reality as ‘the transition between or concurrent usage of multiple systems on the reality-virtuality continuum’ [244]. Tools support CR prototyping through VR simulation [93]. CR systems include handheld displays or projectors [99], user transitions between VR and AR [210], content transitions between 2D and AR [239, 46]. In CReST one user is in VR and another in AR, which is categorised as a multi-user CR system [8]. This disparity in level of immersion has been explored before in context of a student-teacher scenario [274, 291, 311]. The student benefits from a high level of immersion, being present in the VE for an optimal learning experience. Conversely, the teacher has a supervising role and benefits from both real-world awareness, to guide non-immersed students or to help immersed

[2]https://docs.unrealengine.com/4.27/en-US/TestingAndOptimization/ReplaySystem/

students in avoiding collisions with real-world objects, and VE awareness, to provide students more information on the virtual content they are learning about. Additionally, teachers may wish to use a more comfortable interface when teaching multiple classes consecutively [311]. In the context of a user study, there is a similar disparity in roles, where the participant benefits from high immersion, while the researcher should maintain awareness of both the real world and VE.

Windows into the VE approaches have been explored [151, 45] as a technique where users have awareness of both real and virtual environments, by providing the user with a window of a limited size through which they can view the VE. This window approach, however, requires the user to aim to view different parts of the VE, adding extra complexity for the researcher. Thus, we explored a selective visualisation approach [48], where a selection of salient objects from the VE are shown in AR to allow the user to have awareness of both the real world and VE, without having to move around a window.

5.3 Developing the Cross-Reality Study Tool

CReST has two goals, first to allow researchers to observe participants in the VE through an AR visualisation, and second to streamline execution of the study procedure. By enabling the researcher to execute the procedure more easily, we aim to allow them to focus on participant observation rather than on which trial they have to set up next. We provide a flexible way of modelling the study procedure, which CReST can apply at run-time to allow the researcher to start and stop the trial through a button press.

We developed CReST in Unreal Engine 4.27, with support for HoloLens 2 and Quest Pro (following expert feedback, subsection 5.5.4). CReST is applied to a VR user study as an Unreal Engine Plugin, and runs as a second instance of the same project in AR, that is connected to the first instance running in VR (Figure 5.2). We will refer to the VR server as the 'participant instance', and the AR client as the 'researcher instance'. CReST functionality is divided in two parts, first we discuss participant observation, and then study control features.

5.3.1 Participant Observation

CReST's main feature is the ability for the researcher to observe the participant, and their interactions with virtual objects. The visualisation was developed with the goal of providing the researcher simultaneous awareness of both real and virtual environments.

We envision that increased situational awareness could help the researcher with guidance and observation of the participant. First, to ensure validity of results and safety, the

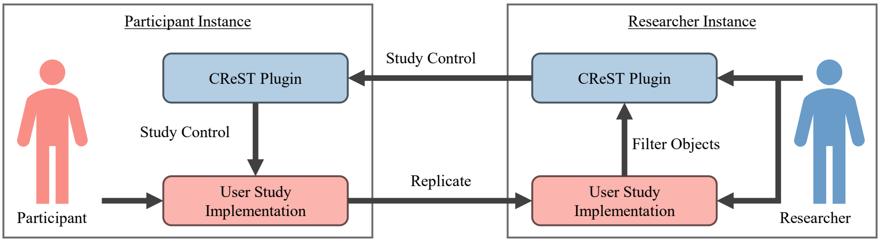


Figure 5.2: The CReST setup consists of two instances of the application, one for the participant and one for the researcher. The participants takes part in the user study. The researcher sees a replicated version of the participant's environment, which CReST filters for display in AR. The researcher can also interact with CReST through the UI, to control the progress of the study on the participant instance.

researcher should ensure that participants do not collide with real-world objects, and are not interrupted by unexpected events such as other people entering the room. Second, a real-world view allows the researcher to see participants' body language, and gain a more complete understanding of what is happening, such as how they interact with the Head-Mounted Display (HMD) and motion controllers. For these reasons, we chose not to immerse the researcher in the VE along with the participant and opted instead for a semi-immersive approach. Hence, we implemented a technique where the salient objects in the VE are filtered and shown in AR [48]. This approach allows the researcher to control their view on the participant and virtual elements, while preserving real-world situational awareness.

CReST supports setup of this selective visualisation by adding an extra button to the Unreal editor toolbar. This 'Tag' button allows the researcher to mark the selected salient objects for visualisation. Then, at run time the plugin will only visualise those objects to the researcher. The state of objects in the VE is replicated using Unreal's networking features [3]. Furthermore, when the participant moves without walking, such as by redirection [216] or teleportation [152], CReST ensures that the same locomotion is applied for the researcher to maintain alignment of virtual objects.

5.3.2 Study Control

To enable study control functionality, we modelled the user study procedure into sessions, blocks, trials, and steps similar to UXF [30]. The order of blocks and trials can be determined by filling in a matrix in which each row is the order for a certain

[3]https://docs.unrealengine.com/4.27/en-US/InteractiveExperiences/Networking/

participant (*participant ID* mod *#rows*). For example, the matrix can contain a balanced Latin square, or can omit trials to create more complex or between-group procedures. Researchers can distinguish between steps taking place in VR (VR step), for which CReST will apply the corresponding independent variable, and show a timer and taglist, and steps that take place in the real world (RW step), such as briefing the participant or having the participant fill in a questionnaire. An example model of a study session can be seen in Figure 5.4.

CReST is controlled by a distributed User Interface (UI), created with Unreal Engine UMG[4] and shown to the user floating in AR and on the desktop monitor (Figure 5.1c). Distributing the UI allows the researcher to take off the HMD, for example while the participant is filling in a questionnaire, and still use CReST. It can be interacted with in AR by pointing the motion controller and using the trigger button, or by pointing with the hand and making a tap gesture. We implemented the following shortcuts on the motion controller and keyboard: next step (thumbstick right, arrow right), previous step (thumbstick left, arrow left), and play/pause the timer (A button, space bar). CReST will initially prompt the user whether they want to start a new session, or resume the previous one (in case the application shut down unexpectedly). The UI consists of the following three parts: study progress, step content and settings.

Study Progress The UI displays the progress of the study by showing the researcher the following information: participant ID, block name, and step name. If the block consists of multiple trials, it will also show this (for example, 1/4) accompanied by the name of the independent variable. Progress can be controlled with 'next' and 'back' buttons, which advance the study to the next or previous step (if the researcher made an error and advanced to the next step too soon). CReST will log each step change, so that the study progression can be referred to later.

Step Content CReST adapts its functionality depending on if the next step in the sequence is a RW or VR step. The RW step's only functionality is a checklist of real-world actions, while the VR step has the following two functions: annotation and timing. For annotation of the trials, the researcher provides a list of pre-defined tags before starting the study, named for events that they expect could happen and wish to keep track of. CReST will then show a button for each tag, and create a log entry when the researcher presses this button. The VR step has a timer with a play/pause button that can be used to time the trial, and also tells the participant instance that it has started. The timer can also be configured to start automatically when the step is loaded.

[4] https://docs.unrealengine.com/4.27/en-US/InteractiveExperiences/UMG/

Settings Researchers can access options for calibrating the environment and setting the visibility of researcher and participant avatars via a settings panel. The 'calibrate' setting will display a gizmo with which the researcher can rotate and translate their environment to manually match it with the participant's environment, which is then saved and loaded so that the procedure does not need to be repeated. Both participant (P. Avatar) and researcher (R. Avatar) can be represented by an avatar, whose visibility can be adjusted. The researcher avatar resembled an abstract head and torso with a solid yellow colour.

5.4 Example cases

To guide the development of CReST, we identified three previously published studies to serve as Replicated Examples (REs; Figure 5.3). We interviewed an author from each study to understand how it was originally conducted. Successively, the tool was applied to a new trial of the study, which the (original) author was then invited to use and give feedback on. We selected the studies for different levels of environment and interaction complexity, ranging from objects with simple linear movement, to environments with more complex interactions and animated agents. REs were chosen from an initial selection of seven published user studies that were (1) implemented in Unreal Engine, (2) of which we could receive access to the source code, and (3) of which the researcher that conducted it originally was available to come to our lab to take part in the expert interviews. Especially due to this last requirement our final selection was limited to studies conducted at our own lab.

5.4.1 RE1 Distractor Interactivity

This study (Figure 5.3 RE1) investigated interactivity in distractors, which are virtual elements that appear in redirected walking systems to distract the user from the redirection that is taking place [51]. The authors developed three types of continuous redirecting distractors, with varying levels of interaction possibilities, called *Looking*, *Touching*, and *Interacting*. They compared them in a user study to a discrete reorientation technique, called *Stop and Reset*, in a task requiring participants to traverse a 30 m path. The study used a within-group design, counterbalancing the order of the four techniques using a balanced Latin square.

For CReST we selected the distractor, participant motion controllers, start location and destination marker as salient objects. The start location and destination marker were static objects, we applied replication to the location, rotation and animations of the distractor and motion controllers. In Figure 5.3 RE1, the destination marker is located outside of the walkable area, and gradually moves within reach as the redirection takes place.

5.4.2 RE2 Immersive Speculative Enactment (ISE)

This study (Figure 5.3 RE2), based on the concept of Immersive Speculative Enactments [247], studied a near-future scenario where an app can detect a pet's state, such as 'hungry' or 'thirsty', and compared two visualisation placements for this app: *smartwatch* and *dog collar*. The goal of the within-group study was to provide insights on the usability of the two designs. Participants were immersed in a VE representing a house, which was split in two parts: the interior house and its adjoining garden. Participants were tasked with completing household chores, and tending to the dog when the app indicated that it required their attention. The study was preceded by three tutorial stages where participants were explained the VR interactions and how to tend to the dog.

Because in the ISE nearly all objects could be picked up, or could be relevant to the user, we selected all VE objects in reach of the participant as salient. Only the house itself, its walls, ceiling and floor, and the ground plane and skybox outside, were not selected. For the interactive objects, the interaction was limited to picking them up, so replication of their location and rotation was sufficient. For the dog, watch and collar, we also replicated their animations.

5.4.3 RE3 Foldable Spaces

This research [104] introduces the *Foldable Spaces* locomotion technique, which folds the VE geometry to reveal new locations depending on the trajectory of the VR user. Three different folding techniques were developed: *Horizontal*, *Vertical*, and *Accordion*. In a within-group user study, the foldable techniques were compared against each other along with a similarly situated redirection technique, *Stop and Reset*.

For each foldable space we selected all objects as salient except for the room's floor, ceiling and walls. We replicated door rotations, so that the opening animation is replicated to the researcher instance. For the folding techniques we applied location and rotation replication to each segment of the environment. In Figure 5.3 RE3, adjacent rooms of in the VE are seen outside of the walkable area, and will come within reach of the participant as the locomotion techniques activate.

Figure 5.3: The replicated examples (RE1 [51], RE2 [247], and RE3 [104]) from the researcher and participants' perspectives.

5.4.4 Adapting Examples to CReST

In this section we reflect on the additional effort required to apply CReST to the examples, both in terms of inputting the study procedure and setup of the AR visualisation. Modelling the procedure was straightforward, especially for RE1 and RE3 which consisted of a simple counterbalanced design with four conditions. RE2 required slightly more effort to include three tutorial stages, for which CReST needed to be set up to load the correct maps.

The largest effort in setting up CReST was needed for network replication. For dynamic objects, this consists of checking the 'Replicates' and 'Replicate Movement' options the engine provides. However, objects with more complex behaviour require the implementation of remote procedure calls so that an effect with origin on one client can be invoked on the other. This is done in Unreal Engine by separating the outcome of an event into a custom event set to multicast to all clients. For example, the sphere lighting up and the orb on the controller switching colour in RE1, and triggering dog animations in RE2. RE3 also required modification because the environments were populated randomly at runtime, which caused the environments to be different on both instances of the application due to different random seeds. We encountered these issues because the REs were originally not designed with network replication in mind, and they should be mitigated when the application is developed with CReST from the beginning. Selecting which items needed to be included in the AR visualisation could be done with one button click, through the plugin we provide.

5.5 Expert Evaluation

For each example case we conducted an evaluation session with one of the authors, referred to as the 'experts' (N=3). The session consisted of the following three parts: first, we asked experts about the procedure of their study and general practices of conducting user studies. Second, they used CReST to run one trial of their study, staged for this purpose. Finally, we asked them to reflect on the added value of CReST and possible missing features to further guide development. The interviews were recorded (between 20-30 minutes), and transcribed, after which we conducted a thematic analysis [28]. We performed an initial coding on the transcripts, from which we identified the three preliminary themes, which we then refined further into the following three themes: 'verifying participant activity', 'monitoring participant progress', 'AR Participant Observation'. Third, we identified feature requests the experts had.

5.5.1 Verifying Participant Activity

During the interviews we established a need for researchers to verify that participants are performing the activities they intended for the study. Systems implemented for the purpose of the study may not be sufficiently tested and still contain bugs, expert 1 gave as reason that '*there's rarely enough time to check every possible situation that can occur.*' Moreover, unclear instructions may lead to participants misunderstanding the task. In studies with multiple independent variables, there is an additional possibility for the researcher to make an error in setting the independent variable. When the participant's unintended behaviour is caused by bugs, unclear instructions or error, the researcher benefits from early detection to avoid it having an effect on the data. Conversely, researchers should be careful not to bias participants when unintended behaviour follows from intended system behaviour. Tools available to the researcher to verify participant activity are limited to screen mirroring, where they can only see the VE from the participant's perspective, which makes it difficult to debug issues.

CReST provides researchers with means of verifying participant activity. The researcher's awareness of both real and virtual environments allows them to notice unintended behaviour early, in both environments. For example, if the researcher notices during the tutorial stage that the participant is not able to perform an action in the VE, they can see in the real world whether they are pressing the correct motion controller button, and respond by clarifying the instructions on the use of the motion controllers. Because CReST allows the researcher to stay focused on the participant, rather than on the computer monitor, it can help avoiding real-world collisions. CReST also helps by setting independent variables, and visualising the active independent variable, to '*streamline the study procedure to make it less likely that there are problems*' (expert 1).

5.5.2 Monitoring Participant Progress

In addition to verifying activities, monitoring participant progress can also be part of the procedure, where the researcher needs to perform specific actions once the participant has progressed to a certain point. Most common is ending the trial when the participant has finished the task, for which the end condition can be fuzzy, as was the case in RE2 and RE3. Expert 2 said, '*What I noticed with the study was that I didn't want to stop participants too early [...] I received useful feedback from people after they completed the task but while they were still in the simulation.*' Additionally, RE2 included three tutorial stages through which the researcher manually advanced the phase of the study, after asking the participant if they were finished with the stage.

Experts agreed that screen mirroring provided a limited view into the environment, only from the participant's perspective. This makes it difficult to see participants' progress and caused them to miss events. Asking the participant whether they are finished allows the researcher to verify that they are, but also causes a break in presence which is best avoided if they are still performing the study's task. Experts also proposed other ways to monitor study parameters, such as boolean values or counters in the CReST UI, would be useful to know when to progress the study. Expert 3 mentioned that the real-world steps could become less useful after repeatedly running the study, because they would become more familiar with the procedure.

5.5.3 Augmented Reality Participant Observation

Experts indicated that the AR monitoring in CReST would be especially useful for qualitative evaluation methods, to detect participant behaviour that would be difficult to capture with objective metrics. Expert 3 found that '*because you are in a spatial environment, you are able to read the situation, instead of looking at the screen*' and that it allowed them to identify behaviour such as the VR user looking outside out of the virtual windows.

The AR visualisation made experts focus on the participant, instead of going back and forth between the participant and the computer running the software. They found that only the VR view was limited, and preferred the AR view over a monitor mirror screen, especially for more spacious environments. Experts also found that the AR view would allow them to take notes while observing.

5.5.4 Expert Feedback Iteration

From the expert interviews we identified the following four features: VR user avatar, screenshots, UI toggle, and video see-through (VST) support, which we added to CReST.

Feature 1: having an avatar for the VR user We initially did not implement an avatar for the VR user, because this part of the implementation is done by the researcher themselves as the VR user could have different appearances depending on the study. However, experts requested an avatar because it would help with identifying that the coordinate systems were aligned by seeing the avatar align with the participant's HMD. We implemented the VR user avatar as a floating head at the HMD location, that is only spawned on the researcher instance. This choice was made to avoid influence of CReST on the study it is applied to, as the participant can see their own appearance in shadows and reflections.

Feature 2: saving screenshots Experts wanted a visual reference for the actions they tagged. Activating the tag feature saves two screenshots, in addition to logging the tag. Screenshots are made from the researcher's and participant's perspective.

Feature 3: ability to toggle the UI Experts requested the ability to minimise the UI, because its automatic follow behaviour can get in the way of the AR visualisation. In response to this feedback we changed the behaviour of the UI to follow the user's head direction at the bottom of their vision, as to not get in the way of the AR objects. Pressing the 'B' button on the Oculus touch controller transitions the AR UI to a minimised version that only shows the progress bar, trial name, timer and back/next buttons. The desktop interface is unaffected and remains maximised.

Feature 4: video see-through (VST) support Experts found limitations with the HoloLens 2, such as having a small field of view and poor visibility of darker objects. Following this feedback, we added support for Meta's video see-through HMDs, such as the Meta Quest Pro, for which we used the Oculus integration SDK [5]. A see-through layer with 10% opacity was overlaid on the virtual objects to make them translucent and avoid occluding the participant.

[5] https://developer.oculus.com/downloads/package/unreal-engine-4-integration/

Session	User Study																			
Block			Closet									Drawer unit								
Trial			Controller High Fidelity		Controller Low Fidelity		Hand High Fidelity		Hand Low Fidelity			Controller High Fidelity		Controller Low Fidelity		Hand High Fidelity		Hand Low Fidelity		
Step	Briefing	Show real closet	Interact with closet	Questionnaire	Interact with closet	Questionnaire	Interact with closet	Questionnaire	Interact with closet	Questionnaire	Show real drawer	Interact with drawer	Questionnaire	Interact with drawer	Questionnaire	Interact with drawer	Questionnaire	Interact with drawer	Questionnaire	Interview

Virtual Reality step Real World step

Figure 5.4: Procedure of the case study modelled in CReST.

5.6 Case Study: Graphic and Interaction Fidelity

In the previous section we gathered feedback from experts, however, we also wanted to demonstrate the utility of CReST by conducting a study with real participants. Thus, we used CReST for a study on graphic and interaction fidelity, where the main goal was to gather researcher feedback. The main author conducted 15 sessions, while two other experienced researchers from the same university conducted two more sessions. In this section we focus on the aspects of the case study relevant for the CReST demonstration.

We recruited 17 participants (7 female, 10 male) from the local university and personal contacts, aged between 18 and 59 ($M = 32.9$, $SD = 13.4$). Participants rated their experience with VR and video games on a 7-point scale, having low experience with VR ($M = 2.7$, $SD = 1.4$), and medium experience with video games ($M = 4.0$, $SD = 2.2$). Participants used a Meta Quest 2, while the researchers used a Meta Quest Pro. The study was approved by the university's ethics committee.

5.6.1 Study Procedure

An overview of the study procedure can be seen in Figure 5.4. Graphic fidelity had the following two levels: 'high' and 'low' to indicate the level of detail and effort required in the modelling process. Similarly, interaction fidelity had two levels 'controller' and 'hand', indicating which the participant used to interact with the object. The procedure consists of two blocks, one for each object, the closet (Figure 5.5) and the drawer unit (Figure 5.6), both of which could be opened. Participants' task was to search for three numbered labels, located in and around the virtual object, requiring them to open it to look inside. Before starting a block of trials, participants could interact with the real artefact (drawer or closet). Participants then wore the HMD and had to stand on a starting area marked with a blue circle on the floor in the VE. Then the researcher

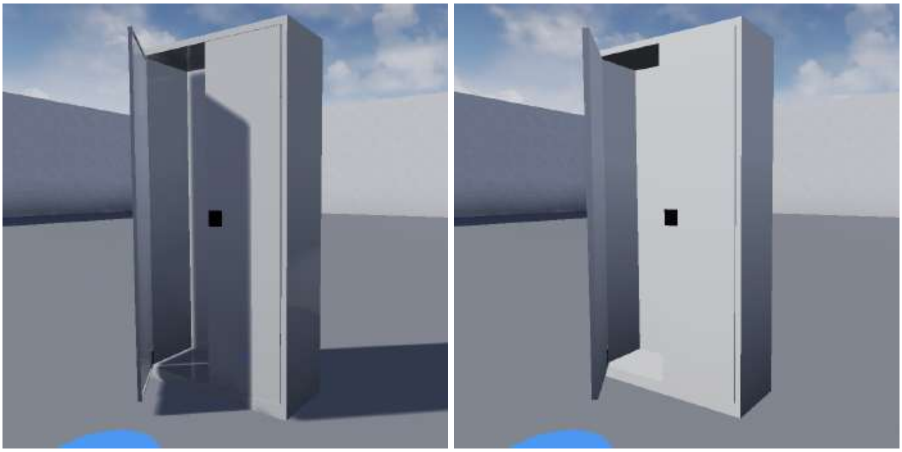


Figure 5.5: Closet artefact, High Fidelity (left), and Low Fidelity (right).

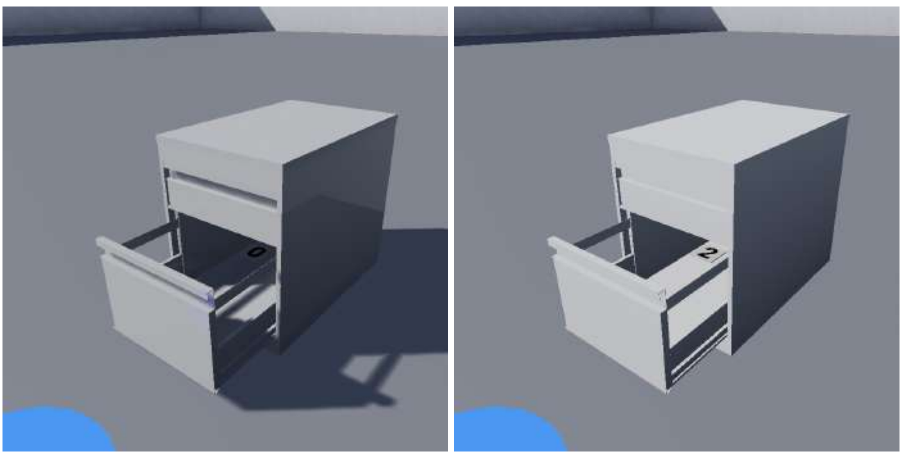


Figure 5.6: Drawer unit artefact, High Fidelity (left), and Low Fidelity (right)

instructed them to investigate the virtual closet or drawer from all sides, and search for the three numbered labels hidden around it.

We used CReST to collect task completion times, participants also filled in a custom fidelity questionnaire after each trial, and we conducted a semi-structured interview at the end of the study, during which the researcher made notes. In this chapter we only present feedback and results related to CReST, which are the observational notes made by the researchers, and the participants' response to the following interview questions:

'What did you think about the research using a second headset to observe you?' followed by 'Did you notice the researcher avatar in the environment?'

5.6.2 Integration with CReST

Conversely to the replicated examples in section 5.4, this study was implemented with CReST in mind, as it minimised the additional effort required to apply it retroactively. The procedure (Figure 5.4) was more elaborate than those of the replicated examples, however, because we opted to support multiple levels such as blocks, trials, and steps our model was flexible and could accurately describe it. CReST allowed us to counterbalance on the block level, alternating drawer and closet, and on the trial level, using a balanced Latin square for the order of the different combinations of graphic and interaction fidelity.

All closet and drawer objects, and the starting location, were tagged using the 'tag' plugin editor button, as to only visualise those objects to the researcher. We then implemented network replication on the closet door and drawer, to replicate their location when interacted with. As the numbered labels were positioned randomly, their random position was generated on the server, and then replicated to the client.

5.6.3 CReST Functionality

Because CReST automatically applied counterbalancing, researchers did not have to manually keep track of which trial was next, especially because the study had a total of eight trials. Moreover, CReST allowed the researcher to stay focused on the participant, rather than switching between participant and monitor. The two additional researchers that conducted study sessions did so without first having to discuss the procedure with the main author. They found it useful that CReST displayed the real-world steps because they were unfamiliar with the procedure, however they also recognised that the real-world steps would not be required after running multiple sessions because they would memorise the procedure.

By observing participants, we were able to identify different kinds of participant behaviour, such as unsupported, impossible and explorative behaviour. Participants tried performing interactions that were possible with the real artefact, but unsupported with the virtual one, such as rolling around the drawer unit, slamming shut the closet door or drawer, and bi-manual interaction. We also observed behaviour that was impossible with the real artefact, such as sticking their hands through the sides of the drawer unit to open the drawer, and walking into the closet without opening it. Additionally, we observed how participants turned towards more explorative behaviour after finishing the task, such as standing inside the closet. The AR visualisation allowed us to see the participant's real hands, as they tried out different hand poses to explore how well the hand tracking worked. Because we could observe these behaviours, we were able to enquire about them in the interview.

CReST allowed us to identify issues with the participant's VE. For example, when the Quest 2 went into stand-by mode in between trials, its internal coordinate system would change, and misalign the participant's VE. The researcher using CReST was able to observe that the participant's avatar was no longer in the correct location and make adjustments before starting the trial.

We asked participants about the researcher avatar, five participants did not notice it, while another five noticed but were neutral about its presence. The other seven participants noticed the avatar and found it helpful, feeling more free in their movements knowing where the researcher was and that they would not bump into them. They were also able to use the avatar to locate the researcher to hand over the controllers to them before taking off the headset.

5.7 Discussion

In this section we first discuss what we learned from applying CReST to three replicated examples, conducting interviews with a researcher involved for each, and a case study where we used CReST to gather qualitative observational data from 17 participants. Then, we discuss how CReST can be applied to different types of studies. Finally, we discuss limitations and future work.

5.7.1 Added Value of CReST

In terms of procedure control, our implementation closely resembles UXF [30], with CReST having an additional layer to its model dividing trials in multiple steps. Our main reason for doing this is to distinguish between VR and real-world steps, as to include instructions for actions outside of the VE. While this allowed researchers not familiar with the procedure to run the study, experts indicated that after multiple sessions this information might not be necessary anymore due to increased familiarity with the procedure. With CReST's study control functionality, we were able to have different researchers conduct the study consistently, after only a brief explanation of how to operate CReST. Additionally, the researchers had more freedom, as they could control the study from anywhere in the room rather than only when close to the desktop computer. This was especially useful in our case study, where it would have been difficult to reliably detect when the participant found the final number to stop the timer.

Observing participant behaviour in AR allowed the researcher to take an external perspective with an overview of what the participant was doing, as opposed to using the screen mirror where participants' hand interactions might be happening out of view. Interfaces where users take their own perspective have been shown to outperform mirrored perspectives on task success and completion time [151], and room-scale visualisations were preferred over miniatures [274, 48]. In our case study the room-scale AR visualisation allowed more detailed observation of participant movements, such as having to bow down to reach the bottom drawer, without elaborate full-body motion tracking setups to drive an avatar.

Because CReST provides the researcher with insights in participant behaviour synchronously, rather than during post-analysis, we found that insights gathered with CReST were of immediate use. Indeed, when the researcher noticed participants exhibited a certain behaviour in the VE, they could bring up these observations during the interview to enquire about the reason behind their actions. We envision rapid prototyping and iterative design processes to benefit from the immediate feedback provided by CReST.

Adding a researcher avatar benefited the study overall, and facilitated interaction between researcher and participant, though it can also influence study results [201]. In our study not all participants noticed the presence of the avatar, though those who noticed were positive about it. The researcher avatar facilitated handing over controllers, because the participant could easily locate the researcher in the VE after putting on the HMD. When receiving instructions from the researcher, such as where to stand to begin the study, participants felt it was more natural that the voice came from an avatar. We designed our avatar as partially diegetic [87], as a translucent yellow head and torso, to have it resemble another person but still feel separate from the study.

5.7.2 CReST for Different Study Types

We envision CReST to provide novel insights in studies where the participant is not completely detached from the real world, such as with haptic objects [304], Substitutional Reality [250], locomotion techniques [60], required interaction with the researcher [249], and the think-aloud protocol [320].

When the study involves *haptic objects* [304], the researcher is interested in both the behaviour of the physical prop as its virtual counterpart. For example, the researcher needs to ensure the active haptic [243] functions properly, while also keeping track of the behaviour of its virtual counterpart. CReST can provide a holistic view in which the researcher can see both real and virtual layers of their active haptic work concurrently. In *Substitutional Reality* [250] the shape of the virtual object and its haptic proxy may differ, CReST could provide researchers with novel insights because it allows them to see the participants interaction with the virtual object overlaid on top of their interaction with the real one. Moreover, CReST's AR room-scale visualisation of the VE can facilitate alignment of virtual props to their physical counterpart.

For research on *locomotion techniques* [60], the researcher aims to understand how the participant's real and virtual locomotion are related. With CReST they can understand how new areas of the VE come into the reach of the participant due to the locomotion technique, or how effective a redirection controller is in avoiding participant collisions with real-world objects.

Studies where the *researcher* is required to take an *active role* in the VE [249] could benefit from CReST, by allowing the researcher to perform their role without needing to be completely immersed in the VE. The researcher can thus have a direct interaction with the participant, while still being able to observe them in the real world. CReST could benefit from more intrusive observational methods, such as the *think-aloud protocol* [320]. In the case study, more expressive participants provided more useful observational notes. By using CReST, the researcher can see the context of the participant's remarks in a think-aloud study synchronously.

5.7.3 Limitations & Future Work

We only tested CReST with four studies (three replicated and one case study), which were all conducted within our own research group. While we selected the studies to vary in terms of procedure, interaction type, and virtual environments, this does limit the external validity of the evaluation. Future research should therefore study the applicability of CReST to other type of studies not considered here, and conducted by other labs that are working in different fields of VR research.

The time investment to incorporate CReST in the research represents another factor. We aimed to develop CReST as a plugin and made available editor tools, but incorporating it in a project requires a non-negligible effort. Applying CReST to the example cases was easier when proper software design practices were followed, as discussed in subsection 5.6.2. Moreover, the added complexity of CReST could introduce potential sources of error to existing applications.

In future versions of CReST we wish to better support studies with passive haptics by introducing granularity in the existing calibration method, to allow calibrating object locations individually. For example, if the research involves a table, CReST can allow calibration of the table as a single object, to match it to its physical counterpart without applying the calibration to the entire VE. This allows the researcher to use the AR visualisation to align real and virtual objects [11]. To make CReST a more complete tool we will add support for replay recording and post-analysis. This integrates with the synchronous observation, for example, real-time tagging of events could provide initial annotation of the recorded data.

5.8 Conclusion

In this chapter we introduced the Cross-Reality Study Tool (CReST), which allows researchers to control and observe VR user studies from an AR perspective. The AR perspective displays a selection of VE objects in AR, allowing the researcher to observe the participant in the real world and VE simultaneously. We applied CReST to three example cases, for which an original author of each case conducted one new session, and was interviewed. Based on these interviews we iterated on the design of CReST with four additional features. In a case study, we used CReST to gather observational data from 17 participants interacting with a virtual closet and drawer unit in low and high fidelity conditions. The study was conducted by three researchers, using CReST to ensure consistency between sessions. Researchers were able to make observations on what interaction possibilities participants expected, in which ways they tried to perform 'impossible' interactions, and how they further explored the artefacts after finishing the task. With CReST we allow streamlining of VR user study procedures, and enable researchers to adopt a more direct observational approach to user studies, to achieve feedback more rapidly and support qualitative evaluations.

Part II

Enabling Object and User Transitions Across the Reality-Virtuality Continuum

Chapter 6

Blending Spaces: Cross-Reality Interaction Techniques for Object Transitions Between Distinct Virtual and Augmented Realities

- [45] **COOLS, R.**, ESTEVES, A., AND SIMEONE, A. L. Blending Spaces: Cross-Reality Interaction Techniques for Object Transitions Between Distinct Virtual and Augmented Realities. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Oct. 2022), pp. 528–537.
- **Robbe Cools:** Conceptualization, Methodology, Software, Investigation, Formal Analysis, Visualisation, Writing **Augusto Esteves:** Writing - Review & Editing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing
- As the first author of this chapter I was responsible for conceptualising, implementing, conducting the study, and writing the chapter. This chapter was written in collaboration with Augusto Esteves, who provided support in writing and finalising the work.
- https://youtu.be/BqwdRGxvWkA

(a) Augmented Environment (AE) (b) Blended Space (c) Virtual Environment (VE)

Figure 6.1: (a) The Augmented Environment (AE), (c) the Virtual Environment (VE), and (b) The *Blended Space* which combines elements from VE and AE into one environment. Elements in the Blended Space remain rooted in their original environment, which is indicated using colour: blue for the VE and yellow for the AE. We introduce three object transition techniques based on blended space: *Manual Blended Space - Button Transition*, *Manual Blended Space - Touch Transition*, and *Auto Blended Space*.

abstract Cross-Reality (CR) involves interaction between different modalities and levels of immersion such as Virtual and Augmented Reality, as we explore in this chapter. Whereas previous work assumed similarity between their respective Virtual and Augmented Environment (VE and AE), we explore the case in which VE and AE are distinct. This gives rise to novel and critical problems, such as how to visualise and interact with the other environment. In this context we investigate the fundamental interaction of transitioning an object across environments, to which we contribute five interaction techniques. Two are inspired by literature: *Virtual Magic Lens* and *Binary Transition*; while the other three are entirely novel: *Auto Blended Space*, *Manual Blended Space - Button Transition* and *Manual Blended Space - Touch Transition*. In a study evaluating the first four techniques, we found that participants (N=20) performed a CR object manipulation and transition task significantly faster using our *Auto Blended Space* technique. We then modified *Manual Blended Space - Button Transition* into *Manual Blended Space - Touch Transition* in response to these results, and reassessed the four techniques in a more complex object manipulation task (N=16). We found that this type of task was better suited to manual transition methods rather than automatic methods. Taken together, our final contribution are five blended space design factors, and timely Cross-Reality transition design guidelines.

6.1 Introduction

In recent years, improvements in Virtual Reality (VR) technology have made it widely available to consumers. Not only VR but also see-through Augmented Reality (AR) devices have seen a similar increase in commercial availability. This increased availability gives rise to Cross-Reality (CR), which is described as 'the transition between or concurrent usage of multiple systems on the reality-virtuality continuum' [244]. The term CR thus captures crossing between Physical Reality, AR and VR; of which we focus on the latter two in this chapter.

Previous CR research centres mostly around external users accessing immersive environments [274, 89], or replicating an environment in a different reality. Systems involving VR and AR mainly focus on Virtual Environments (VEs) and Augmented Environments (AEs) replicating one another. For example, using VR to access a remote AE [150, 80] through video and point cloud visualisations in VR, or filtering a Virtual Environment (VE) to enable visualisation in AR [48]. Conversely, we focus on a VE and AE where the goal is not to replicate the other environment, but where environments are distinct instead of replicated.

To highlight the importance of distinct environments, we present the following four examples: museum curator, video game modelling, furniture shopping, and a simplified office task. The first museum curator example is also the scenario for our first user study. A museum curator wants to augment an exhibit in their physical environment with 3D models of artefacts and other information [309]. To do so, they can access a VE that shows the artefacts in a virtual recreation of their context of use and select those they wish to transition to augment their present-day remnants for the benefit of the visitors. Additionally, informative signs can also be transitioned from AE to VE to provide artefacts with extra information in-situ. The second example is that of a video game modeller, using AR for modelling to still have access to real-world reference materials [218]. The modeller then wants to to transition the models to the VE of the game to see how they look in their future context of use. The third example is an AR user wishing to shop for new furniture [118, 260]. They could first explore the catalogue in a VE, where the furniture is being showcased, and then transition a selection of their most-liked items into AR to see how they fit in their own space. This second and third example could also involve manipulation of objects within the same environment, in addition to transitioning objects. Because of this we introduce an additional fourth example, where users have to perform a combination of transitions and intra-reality manipulations, as the scenario for our second study. The fourth example involves a stamping task, where a user is presented with three stamps in the AE and three papers in the VE. Performing a stamping action, in either the VE or AE, then requires them to transition one object but manipulate the other without transitioning it.

Interaction between a distinct VE and AE introduces several additional challenges, as it is no longer possible to see and interact with all elements of the other environment. Users immersed in these distinct environments will need to express agency on objects that are currently not part of their native environment. As such, 'transitioning' virtual objects between realities represents one of the main CR tasks. Due to the previous examples being mainly AR-centric, as the core task such as 3D modelling or previewing furniture takes place in AR, we chose to approach this problem from the AR user's perspective.

To this end, we introduce the concept of a 'blended space' and compare it to traditional 'step into' and 'reach into' approaches. Blended space is a single, combined space in which objects from one dimension can be transitioned to the other. It allows users to see both source and destination environment simultaneously, as with a reach into approach, and in full, as with a step into approach. Based on the blended space concept we introduce two techniques, the *Auto Blended Space (ABS)* and the *Manual Blended Space (MBS)*. In *ABS*, objects transition automatically when they are manipulated in the blended space which prioritises speed over agency. In *MBS* users explicitly perform the transition, in the first study with the press of a button (*MBS-B*), and in the second study by touching with their hand (*MBS-T*), prioritising user control over speed. Due to the combined visibility of both environments, we expect that the blended space will allow users to perform CR manipulations faster.

We designed a user study (N=20) to evaluate our two blended space techniques, *ABS* and *MBS-B*, along with the following two techniques inspired by literature: *Virtual Magic Lens (VML)* and *Binary Transition (BT)*. Participants were required to match an object's location and rotation to a target in the other environment. We found that the *ABS* technique was significantly faster compared to the *VML* and *BT* techniques. Successively, we modified *MBS-B* based on feedback from the first study into *Manual Blended Space - Touch Transition (MBS-T)*. The original techniques, with *MBS-T* replacing *MBS-B*, were evaluated in a second study which centred on a task involving both transitions and intra-reality manipulations. In this second study (N=16), participants preferred *VML* and *MBS-T* over *BT* and *ABS*. Based on the findings we propose five blended space design factors, and a set of CR transition design guidelines.

Approaches for blending environments have been proposed for VEs [228], for CR prototyping in VR [93], and for visualising a VE in AR [48]. However, our contribution is novel in that (1) blended space supports environments with different levels of virtuality, such as VEs and AEs, (2) our implementation is focused on object rather than user transitions, and (3) we evaluate it as such and compare to other object transition methods.

6.2 Related Work

We first give an overview of CR [244] work that involves users in Virtual, Augmented and Physical Realities or systems that transition between them. Then we discuss works which are the inspiration for *VML* and *BT* techniques, as well as the works most closely related to our concept of blended space.

6.2.1 Cross-Reality Interaction

VR users are unaware of their physical surroundings, hence incorporating the physical environment into the VR experience [107] helps avoid conflicts and provides passive haptics [250, 79, 59]. Furthermore, external users interrupting immersed VR users can cause breaks in presence, which can be minimised through diegetic avatars [87] or asynchronous interaction [72]. In order for external users to further engage with Head-Mounted Display (HMD) users, they need to be provided with more visual information, for example, through screens or projections.

External users can use screens to interact with HMD users in different ways: handheld [151], mounted on an HMD [100] or placed in the environment [147]. ShareVR [99] combined multiple methods, a large TV that showed the VR user's perspective of the VE, a small motion-tracked display that the external user could move around to show the VE from different perspectives, and a floor projection. The projection was implemented through two projectors that covered the VR user's walkable area. Projections have also been used on transparent screens in front of the VR user [121], and mounted to the HMD [292] to allow the projection to follow the user. This area of research focuses on devices to allow interactions without instrumenting external users, whereas we focus on both users wearing an HMD where one is in VR and the other in AR.

The interaction between VR and AR users has not been researched extensively. Grandi et al. [89] investigated collaboration between a VR and tablet AR user. Participants performed a collaborative docking task with blocks that were visible to both AR and VR users. The visualisation was symmetric between participants, with asymmetric input modalities: the VR user used motion controls and the tablet AR user, a touch-screen interface.

Mini-Me [206] presents an adaptive avatar for collaboration between a remote VR and local AR user. In ObserVAR [274] see-through AR is used for a teacher to visualise a group of students in a VE. In SelectVisAR [48] see-through AR is used to visualise a VR user and their VE to an AR user, selectively visualising parts of this environment. These three examples focused on how to present a single environment through different visual and interaction modalities.

In our work the VE and AE are distinct hence users need a transition technique. We will further discuss two applicable techniques: lenses and user transitions. Through a lens, the user can view and interact with the other environment, whereas user transitions transport the user to the other environment entirely.

6.2.2 Lenses Between Virtual Environments

The original Magic Lens [18] presented the user with a transparent overlay on a 2D monitor, through which they could click with a cursor. This concept was later reprised in successive works. Spindler et al. [262, 261] presented a handheld Magic Lens that the user could move around to change their perspective. This version of the Magic Lens is used in combination with a tabletop projection and has multiple functions such as extension, filtering and manipulation.

ShareVR's [99] motion-tracked display applies the idea of the Magic Lens to VEs, allowing an external user to view and manipulate the VE from different perspectives. Garcia et al. [78] compared a motion-tracked tablet with which the external user could view the VE, and a virtual camera feed through which the VR user could view the real environment, with the purpose of creating virtual representations of physical obstacles. Clergeaud et al. [42] present 'doors' and 'windows' as methods for previewing and travelling between VE and AE. The door is fixed in the room, can be used to preview the VE and allows a VR user to travel between environments. The window is tangible and can be moved around by the user, allowing them to reach in to grab objects from the other environment, similar to our *VML* technique.

Previous implementations of the magic lens were tangible, whereas we designed a *Virtual Magic Lens (VML)*. Due to its virtual nature, our *VML* technique allows users to pass their hand through it from all sides, and reach into the other environment to manipulate and transition objects by moving them in and out of the lens.

6.2.3 Transitioning Between Virtual Environments

We discuss related work on user transitions with the goal of identifying techniques that can be extended to transition objects between VE and AE. VR transition methods move the user either to a different location within the same VE [176], or to a different environment [189].

Works by Bruder et al. and Steinicke et al. [31, 267] used portals to enable user transitions between environments. Oberdörfer et al. [189] explored three metaphors for transitioning between VEs: Simulated Blink, Turn Around and Virtual-head Mounted Display. Simulated Blink fades out the screen, changes the environment, and then fades back in. Turn Around requires the user to turn around, changing the environment behind their backs. Virtual-head Mounted Display requires the user to put on a virtual HMD. Husung and Langbehn [116] investigated six VE transition techniques: cut, fade, dissolve, Portal, Orb, and Transformation. They found that Portal and Orb techniques scored higher on presence, continuity and preference, without falling behind on usability. They describe the cut technique, which transitions the user between two frames, as the fastest. As such we will use cut as a baseline user transition in the form of our *Binary Transition (BT)*.

The MagicBook [20] is a system that allows users to view a 3D scene in AR, as part of a physical book. Users can then transition themselves into the scene and explore it in VR. George et al. [81] implemented transitions between the real environment and the VE using video see-through AR. They present two techniques: Sky Portal and Virtual Phone, which both provide a preview of the other environment. Sky Portal presents a top-down view into the other environment on the ceiling, pressing a button when gazing at the Sky Portal would transition the user between environments. The user was able to move around the Virtual Phone as a Lens into the other environment. Double-tapping a button would move the user between environments.

6.2.4 Blending Space

Previous work on blending spaces mainly centred around VEs, such as OVRLap [228] which presents a technique in which multiple locations in a VE are visualised simultaneously with lowered opacity. The technique allows users to visualise the other location to which they can transition and identify objects within it that they require before making the transition. Users can then transition objects along with them by holding the object while transitioning.

Smooth immersion by Valkov et al. [284] presents a gradual transition between the Physical Environment (PE) and VE, which is done by morphing from a replica of the PE to the VE. VRception [93] presents a system to prototype CR applications within VR. It allows for multiple users to develop and test a shared VE within which prototype PEs, AEs, and VEs can be created. VRception implements a slider that allows a user to transition between these environments.

'Transitioning' virtual objects between realities is a crucial task in CR, for which we propose the novel concept of blended space as a space that combines elements from both realities. We will evaluate two versions of the blended space: *ABS* and *MBS* along with two existing techniques applied to the CR context: *VML* and *BT*.

(a) BT (b) VML (c) MBS-B (d) MBS-T (e) ABS

Figure 6.2: To enable transition of objects between VE and AE we created the following five techniques: (a) *Binary Transition (BT)*, (b) *Virtual Magic Lens (VML)*, two version of *Manual Blended Space*: (c) *Manual Blended Space - Button transition (MBS-B)* and (d) *Manual Blended Space - Touch transition (MBS-T)*, and (e) *Auto Blended Space (ABS)*.

6.3 Cross-Reality Transition Techniques & Implementation

For our first user study, we designed four techniques to transition objects between VE and AE: *Binary Transition (BT)*, *Virtual Magic Lens (VML)*, *Manual Blended Space - Button Transition (MBS-B)* and *Auto Blended Space (ABS)*.

6.3.1 Binary Transition (BT)

We designed *BT* as a baseline 'step into' approach, where the user switches between being immersed in either the VE or AE. Inspired by Husung and Langbehn's [116] *Cut* technique, upon activation it changes environments in between two frames. Similar to previous work [228], we extended the user transition to objects held by the user as well. This approach has the drawback of users losing awareness of the environment after transitioning.

In order to initiate the transition we created a 'user widget' (Figure 6.2a), which follows the user's head movements to remain in sight. The transition is initiated by pressing an 'Augmented Reality' or 'Virtual Reality' button on this widget. The button for the reality in which the user is currently present is highlighted.

6.3.2 Virtual Magic Lens (VML)

VML is a baseline 'reach into' approach, and thus presents the user with a window into the VE measuring 30 cm × 30 cm (Figure 6.3, Figure 6.2b). The size of the window is

Figure 6.3: The *VML* technique allows the user to see both the VE and AE simultaneously.

calculated so that it takes up about half of the user's field of view when they stretch out their arm, allowing them to see both environments at once. With this approach users maintain awareness of both environments at the cost of limiting VE visibility to the size of the lens.

The *VML* follows the non-dominant hand, leaving the user's dominant hand for picking up objects. For instance, while the hand is inside the *VML*, it can only pick up objects in the VE and conversely, when it is outside the *VML*, only objects in the AE. Held objects can be transitioned between realities by moving them in and out of the *VML*, or vice versa.

6.3.3 Manual Blended Space - Button Transition (MBS-B)

We envision a blended space in which the user can access both environments in full, as with *BT*, but also simultaneously, as with *VML*; combining the benefits of both approaches. However, to combine both environments into one without overloading the user with visual information, the VE is filtered. As a filter criteria, all interactive objects and their supporting geometry were maintained [48]. This filter was chosen because it enables visualisation of the VE in AR. Though visualised simultaneously, objects in blended space remain rooted in their respective environments. In order to distinguish these we colour-coded the objects: AE in yellow and VE in blue (Figure 6.1b). To allow entering the blended space, we extended the user widget created for *BT* with a third 'Blended Reality' button (Figure 6.4, Figure 6.2c).

Figure 6.4: User in the blended space using the *MBS-B* technique, about to press the 'Virtual Reality' button to transition the chalice object to the VE. Left: the 'user widget' currently indicating that blended space has been entered. Right: chalice with 'object widget' hovering above it, the yellow colour-code indicates that the chalice is part of the AE.

To explicitly support changing the environment to which an object belongs, *MBS-B* requires users to press buttons. These buttons are implemented as an 'object widget', which hovers above each object in blended space. This object widget (Figure 6.4) contains two buttons, one for AR and one for VR, each transitioning the object to the corresponding reality upon being pressed. The button for the environment to which the object currently belongs is highlighted. We chose this approach because we deemed pressing a button (discrete action) faster than using a slider (continuous action) to change environments [93].

6.3.4 Auto Blended Space (ABS)

ABS is similar to *MBS-B* except that it operates under the assumption that all manipulations performed in the blended space are intended to also perform a transition. The object widget is thus omitted (Figure 6.2e), and transition is initiated automatically when an object is picked up in blended space. Thus, when a user wishes to manipulate an object from one location to another in blended space they will also transition it simultaneously. Users can then further manipulate the object without each subsequent pick up interaction triggering additional transitions, as *ABS* was designed to only allow each object to transition once after entering the blended space. Transitioning the object back to its original environment requires the user to leave the blended space, to AE or VE, and enter it again.

6.3.5 Implementation

We used Unreal Engine 4.26 for the implementation, together with the Windows Mixed Reality Toolkit[1] (MRTK) to provide hand interaction with objects and UI buttons. Objects could be manipulated using the MRTK near interaction mode, which allows picking up and releasing objects by opening and closing the hand.

As we chose to investigate the AR user's perspective, we developed for the Microsoft HoloLens 2, which uses optical see-through displays that allow users to view virtual content overlaid on top of the real world. To enable it to show the VE, it was set to maximum brightness in our artificially lit lab environment, and the VE rendered to its entire display area (43° by 29° field-of-view). This setup was sufficient to immerse participants, though to a lesser degree than with conventional VR devices. The benefit to this approach compared to using a VR HMD with video see-through capability is that it allows the participants to see the real world unaltered, rather than through a camera feed.

6.4 Transitions User Study

Our first study focuses on the efficiency of the techniques for performing transitions, for which we designed a user study where participants had to transition six objects. Results indicate that, on this task, participants were significantly faster when using *ABS* than with *BT* and *VML* techniques.

6.4.1 Environments & Task

Both the VE (Figure 6.1c) and AE (Figure 6.1a) were designed to fit within a 4 m × 4 m area. The VE represented a ruin in a forest in which different artefacts were located. The AE contained ten museum exhibit stands. Six stands were already exhibiting an artefact to give the AE the appearance of a museum. Three were empty and displayed a target in the shape of an artefact currently present in the VE. The remaining tenth stand contained three signs, each showing the name of an artefact. The corresponding

[1] https://github.com/microsoft/MixedReality-UXTools-Unreal

artefacts could be found in the VE, accompanied by a target in the shape of the sign to indicate where it should be placed.

Participants performed an object manipulation task that consisted of transitioning a total of six objects. They transitioned three artefacts from the VE to AE, and three signs from AE to VE to label artefacts in the VE. These six artefacts were randomly selected from a set of twelve, but located at six predetermined positions. Participants were required to match the position and orientation of each artefact and sign with that of a target (Figure 6.5).

6.4.2 Experimental Setup & Design

We recruited 20 participants for the user study (9 female, 11 male). They were aged between 21 and 58 ($\mu = 30.2$, $\sigma = 12.8$). Overall, they rated their experience with Mixed Reality as medium, on a 7-point scale ($\mu = 3.5$, $\sigma = 1.7$). 19 participants were right-handed and controlled the *VML* with their left hand, one was left-handed and controlled it with their right. We used a within-subjects design, conditions were counterbalanced using a balanced Latin square.

Participants first signed an informed consent form and filled in a demographics questionnaire. Then we explained to them how to use the HoloLens 2 and its hand tracking features. Before each trial, we reminded them of their task and explained how to use the active transition technique. They could then familiarise themselves with the technique by practising on a tutorial setup with a single artefact and a single sign, after which they moved on to the task. Four trials were conducted, one for each technique.

6.4.3 Metrics

We chose to measure completion time as a measure of transition efficiency, as opposed to placement accuracy. We expect an effect of technique on completion time due to each technique being made up of actions with different time requirements, such as button presses and hand movement trajectories. We measured task completion time with a timer that started when the participant picked up the first object, and stopped when they correctly placed the final object. Each object also had an individual timer that started when the object was picked up and stopped when it was placed correctly.

After each trial, participants filled out the NASA-TLX questionnaire [106] and were also able to write any additional comments they had. At the end of the study participants ranked the techniques from best to worst based on their experience. We then interviewed them about why they preferred the technique they ranked highest and asked them to

Figure 6.5: A participant placing one of the artefacts correctly during the transitions study. The artefact target is represented by a wireframe and changes colour from yellow to grey to indicate that the correct object was placed.

identify aspects of the techniques that they thought made completing the task particularly (in)efficient.

6.5 Transitions Results

We analysed the data collected during the transitions user study and found that *ABS* was significantly faster, preferred by participants, and resulted in a lower perceived workload. We then analysed the interviews to gather qualitative feedback on how the participants experienced the techniques. Unless otherwise stated, we used a Friedman test because the assumption of normality was not met, as resulting from a Shapiro-Wilk test.

6.5.1 Task completion times

Task completion times can be seen in Figure 6.6. The application of a Friedman test showed a significant difference between the techniques ($p < 0.01$, $\chi^2 = 13.1$, $f = 0.20$). Participants using the *ABS* ($\mu = 116$ s, $\sigma = 65$ s) performed the task the fastest, followed by *MBS-B* ($\mu = 156$ s, $\sigma = 76$ s), *VML* ($\mu = 171$ s, $\sigma = 125$ s) and *BT* ($\mu = 177$ s, $\sigma = 96$ s). A pairwise Wilcoxon test showed a significant difference between *ABS* and *VML* ($p < 0.05$, $f = 0.62$), *ABS* and *BT* ($p < 0.05$, $f = 0.65$), and a trend between *ABS* and *MBS-B* ($p = 0.08$, $f = 0.54$).

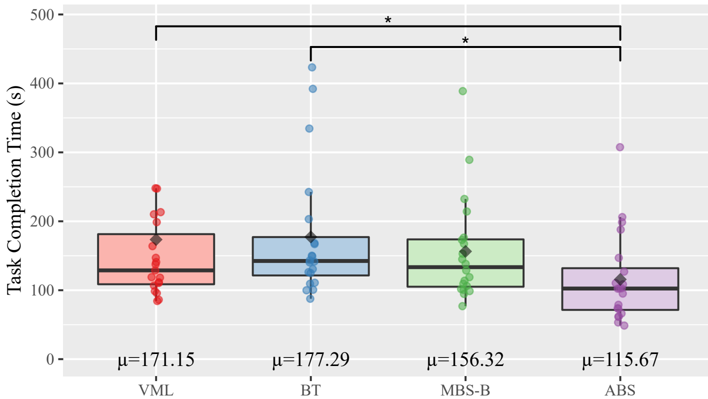


Figure 6.6: Task completion times for the transitions study.

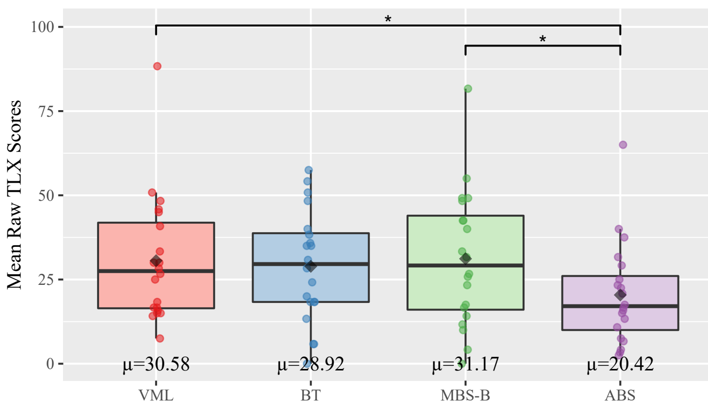


Figure 6.7: Raw NASA-TLX scores for the transitions study.

In addition to task completion time, we also recorded how quickly participants picked and placed individual objects by measuring the time between picking up the object and when it entered the target. The Friedman test showed a significant difference between conditions ($p < 0.001$, $\chi^2 = 22.0$, $f = 0.37$). The Wilcoxon signed rank test showed that *ABS* ($\mu = 10.31$ s, $\sigma = 12.21$ s) was significantly faster than *MBS-B* ($p < 0.001, f = 0.76$, $\mu = 19.62$ s, $\sigma = 21.43$ s) and *BT* ($p < 0.001, f = 0.83$, $\mu = 19.69$ s, $\sigma = 22.74$ s). With *VML* ($\mu = 15.08$ s, $\sigma = 24.80$ s) being between *ABS* and the other two techniques.

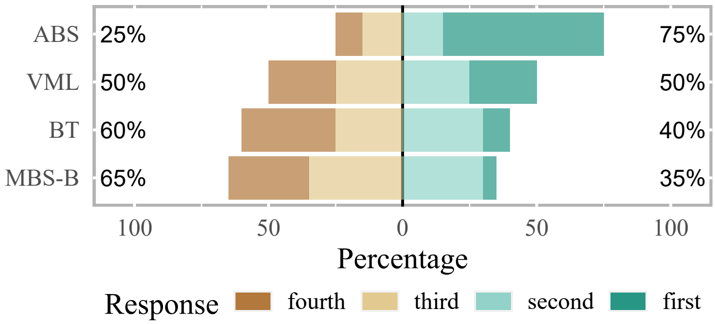


Figure 6.8: User preference for the transitions study.

6.5.2 NASA-TLX

Participants scored each of the dimensions of NASA-TLX [106]. These scores were averaged as 'raw TLX' scores between 0 and 100, where a lower score indicates a lower perceived workload (Figure 6.7). *ABS* scored the lowest ($\mu = 20.42$, $\sigma = 15.13$), followed by *BT* ($\mu = 28.92$, $\sigma = 16.59$), *VML* ($\mu = 30.58$, $\sigma = 18.79$) and *MBS-B* ($\mu = 31.17$, $\sigma = 20.16$). A Friedman test revealed a significant difference between the techniques ($p < 0.01$, $\chi^2 = 14.10$, $f = 0.23$). Following a pairwise Wilcoxon test, *ABS* scored significantly lower than *VML* ($p < 0.05$, $f = 0.67$), and *MBS-B* ($p < 0.05$, $f = 0.63$).

Further testing using pairwise Wilcoxon tests on the dimensions of NASA-TLX revealed that *VML* scored significantly higher than *ABS* on physical demand (*ABS*: $\mu = 18.5, \sigma = 15.0$; *VML*: $\mu = 31.8$, $\sigma = 22.7$; $p < 0.01$, $f = 0.72$) and effort (*ABS*: $\mu = 18.7, \sigma = 19.2$; *VML*: $\mu = 36.2$, $\sigma = 25.7$; $p < 0.05$, $f = 0.67$). *MBS-B* scored significantly higher, compared to *ABS*, on mental demand (*ABS*: $\mu = 20.5$, $\sigma = 20.0$; *MBS-B*: $\mu = 40.2$, $\sigma = 27.6$; $p < 0.05$, $f = 0.70$) and frustration (*ABS*: $\mu = 14.2$, $\sigma = 16.0$; *MBS-B*: $\mu = 29.8$, $\sigma = 23.5$; $p < 0.05$, $f = 0.67$).

6.5.3 Preference & Interviews

Participants' preference is shown in Figure 6.8, a Friedman test revealed significant differences between how users ranked the techniques ($p < 0.05$, $\chi^2 = 10.14$, $f = 0.17$). A pairwise Wilcoxon test showed that *ABS* was preferred over *VML* ($p < 0.01$, $f = 0.80$), *BT* ($p < 0.01$, $f = 0.85$) and *MBS-B* ($p < 0.01$, $f = 0.87$). *VML* was preferred over *MBS-B* ($p < 0.05$, $f = 0.63$).

We interviewed participants on why they preferred certain techniques and on what aspects of the techniques they found more efficient for completing the task. Transcripts of these interviews were then coded and analysed following a thematic analysis [28]. Multiple mentions of the same code by the same participant were only counted as one occurrence, all codes with at least three occurrences are reported in the following summary.

Participants expressed a preference for *ABS* and found it easy to use (7 participants), especially for the given task (3). They felt that, when using the blended space in general, they did not need to switch to the VE or AE (4) and that the blended space felt like only one world (10), as noted by P10: *'Because of the blended [space] you sometimes forgot the difference between VR and AR.'* Participants found that blended space gives them a better overview (4), as P5 noted *'The blended [space] had the advantage that you could see everything together.'*

In contrast to the blended space feeling as one world, participants using *BT* experienced the VE and AE as two clearly separated worlds (5), noting that effort was needed to remember what the other environment looked like (5), P1: *'The binary transition wasn't good because I had to remember how the AR environment looked like every time I switched back.'*

Participants had mixed reactions to the *VML* technique, finding it easy to switch with (3), but also inefficient (3) and clunky (3). P12 summarises this by saying *'Virtual Magic Lens I found handier because then you can see both dimensions at the same time, so that you can't mistake [what is in which environment]. On the other hand, it's also more difficult to see everything, because you only see the [VE] through a lens.'*

6.6 Mixed Manipulations User Study

The results of the previous transitions user study indicate that participants using *ABS* were significantly faster than with *BT* or *VML*. Though this shows blended space can be effective, its design aspects play a role in this effectiveness as *MBS-B* was preferred least. Hence we perform a second study, which investigates a different *MBS* transition method, and evaluates the previous techniques on a more complicated task.

In this study, we chose a task that combines object transitions with intra-reality manipulations. This task is representative of more complicated tasks because, for example, when using *ABS* intra-reality manipulations require participants to leave the blended space to manipulate objects without transition. In such a scenario, we expect *MBS-B* to be preferable over *ABS*, since it allows the user to perform the intra-reality manipulations within the blended space. However, since *MBS-B* was not received well by participants, we redesigned it for a more intuitive transition of objects.

6.6.1 Manual Blended Space - Touch Transition (MBS-T)

The button press transition in *MBS-B* was not well received by participants, who ranked it as their least favourite technique in the transitions study. As such, we chose to explore bimanual interaction as an alternative, inspired by *VML* where users could control the lens with their non-dominant hand. *Blended Space - Touch Transition (MBS-T)* adds a small virtual sphere to the user's offhand which they can use to 'touch' an object to swap it between realities, performing another swap on each subsequent touch (Figure 6.1g).

6.6.2 Environments

The environments of the study can be seen in Figure 6.9. The AE contained a table with three stamps on it, numbered between one and three. The VE represented a forest, in which there was a desk area with two tables on which papers were placed. The papers were labelled with letters 'A', 'B' and 'C'. The tables in the AE and those in the VE were positioned on opposite ends of the physical space.

6.6.3 Experimental Setup & Design

For this study we recruited 16 participants (8 female, 8 male) who were aged between 21 and 38 ($\mu = 26.75$, $\sigma = 4.57$), 8 of which also participated in the transitions study. Participants had medium self-reported experience with mixed reality ($\mu = 4.00$, $\sigma = 1.71$ on a 7-point scale). All participants indicated to be right-handed, and thus controlled the *VML* and *MBS-T* techniques with their left hand.

We used the same procedure (subsection 6.4.2) and metrics (subsection 6.4.3) as in the transitions study. However, participants were presented with a different AE and VE, and a different task to complete. Whereas the task in the previous study solely involved manipulating objects that also needed to be transitioned, participants now had to perform a task designed to involve both transitioning objects as well as manipulating objects within the same reality.

The task involved participants stamping three pieces of paper. In each trial, participants received three instructions indicating which paper ('A', 'B' or 'C') to stamp with which stamp ('1', '2' or '3') and in which environment (Augmented Environment or Virtual Environment). The instructions were presented as text that was visible in both AE and VE, and varied between the trials, always requiring two papers to be stamped in the AE and one in the VE.

Stamping in the AE required the participant to transition the paper from the VE to the AE and place it on the table, then take the correct stamp and touch the paper with it.

Figure 6.9: The VE (top) and AE (bottom) for the mixed manipulations study. Tables in both realities were located on opposite ends of the physical space.

Stamping in the VE required the participant to transition the correct stamp to the VE, and touch the corresponding paper in the VE with it. Each trial participants had to transition a total of three objects (two papers and a stamp) to the other environment, as well as perform two manipulations on objects in the AE without performing a transition (two stamps). This task was repeated four times, once with each technique.

We chose this task because it includes a transition of one object (paper), followed immediately by an intra-reality manipulation of another (stamp). As such it is representative of more complex tasks in which a mix of both types of interaction is performed, for example the video game modelling and furniture shopping use cases presented in the introduction.

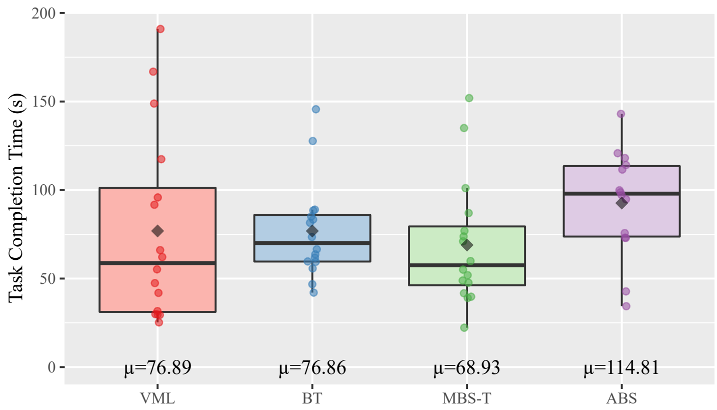


Figure 6.10: Task completion time for the mixed manipulations study.

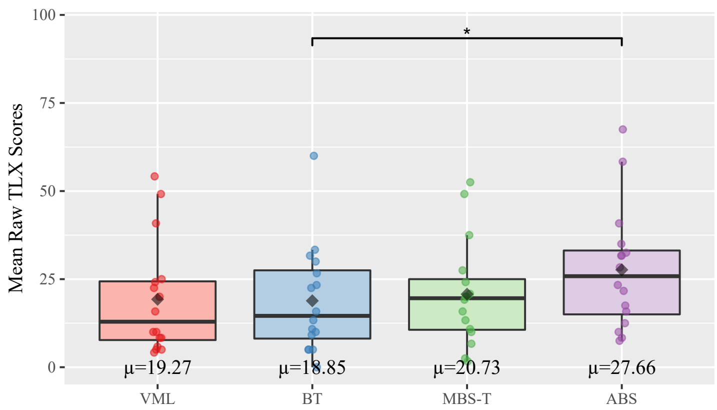


Figure 6.11: Raw NASA-TLX scores for the mixed manipulations study.

6.7 Mixed Manipulations Results

Data of the mixed manipulations study were analysed. *ABS* resulted in a higher reported workload, and was received poorly by participants who found it easy to make mistakes with. Participants preferred *MBS-T* and *VML* most for this task, which they experienced as intuitive.

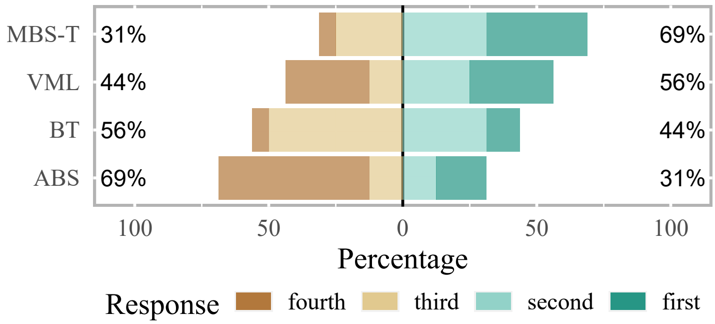


Figure 6.12: User preference for the mixed manipulations study.

6.7.1 Task completion times

Task completion times can be found in Figure 6.10. A Friedman test did revealed significant differences between conditions ($p < 0.05$, $\chi^2 = 9.3$, $f = 0.19$). Participants using *MBS-T* were the fastest ($\mu = 68.93$ s, $\sigma = 35.39$ s), followed by *BT* ($\mu = 76.86$ s, $\sigma = 27.57$ s), *VML* ($\mu = 76.89$ s, $\sigma = 53.42$ s) and *ABS* ($\mu = 114.81$ s, $\sigma = 72.61$ s). The pairwise Wilcoxon test did not reveal any significant differences between conditions, with the lowest p-value ($p = 0.079$, $f = 0.61$) between *ABS* and *MBS-T*.

6.7.2 NASA-TLX

Mean raw NASA-TLX scores were calculated (Figure 6.11): *BT* ($\mu = 18.85$, $\sigma = 15.26$), *VML* ($\mu = 19.27$, $\sigma = 16.09$), *MBS-T* ($\mu = 20.73$, $\sigma = 14.93$) and *ABS* ($\mu = 27.7$, $\sigma = 17.2$). A Friedman test revealed a significant difference between conditions ($p < 0.01$, $\chi^2 = 14.57$, $f = 0.30$). A pairwise Wilcoxon test revealed that *ABS* scored significantly higher than *BT* ($p < 0.05$, $f = 0.77$). Further testing using a pairwise Wilcoxon test revealed that *ABS* ($\mu = 38.8$, $\sigma = 27.4$) scored significantly higher on mental demand than *VML* ($\mu = 17.8$, $\sigma = 18.6$; $p < 0.01$, $f = 0.84$) and *BT* ($\mu = 21.9$, $\sigma = 22.4$; $p < 0.01$, $f = 0.86$).

6.7.3 Preference & Interviews

How participants ranked the techniques can be seen in Figure 6.12. A Friedman test did not reveal a significant difference between how participants ranked the techniques ($p = 0.14$, $\chi^2 = 5.48$, $f = 0.11$).

Participants found *ABS* to be easier to make mistakes with (9 participants), and annoying (3). They found *VML* to be intuitive (9), though some found it too small (4) and annoying to stretch out their off-hand to aim (4). *MBS-T* was experienced as intuitive (6), although participants had to recall that their off-hand was used for switching (3). Participants liked seeing both realities at once with the blended space techniques (5), and did not like that they were separated with the *BT* technique (3). P1 summarised *MBS-T* as follows: *'I could [be in] one reality, and then move things between both realities which felt very intuitive.'*

Though participants expressed that *ABS* was more difficult to learn and easier to make mistakes with, P10 was able to still use it very efficiently and summarised their approach as follows: *'I first switched everything to the space in which it had to be, when the two papers had to go to the other space I first moved those two papers, and then switch to the space in which you had to stamp.'* P8 also expressed *ABS* to have more potential once they got used to it: *'I think it is annoying but at the same time I think that if I get used to it then it would be much easier than the other ones.'*

6.8 Discussion

Our first study focused on object transitions and compared *Auto Blended Space (ABS)*, *Virtual Magic Lens (VML)*, *Binary Transition (BT)* and *Manual Blended Space - Button Transition (MBS-B)*. The second study involved a more complex mixed manipulation task and modified *MBS-B* by introducing the *Manual Blended Space - Touch Transition (MBS-T)* technique.

6.8.1 Design Guidelines on Transition Techniques

Based on the findings, we summarise in which scenario each technique is most appropriate:

Binary Transition (BT) is least suitable for transitioning objects because it requires the user to transition along with each object they wish to transition, causing them to lose track of the other environment each time. Therefore, *BT* is more suitable when the user's task requires them to perform manipulations in different environments, while transitioning objects between environments occurs infrequently.

Virtual Magic Lens (VML) is recommended for scenarios where the user's task requires them to remain present in their 'native' environment, while occasionally transitioning an object to or from a different environment. The main disadvantage of the technique is the limited view of the other environment, with the main advantage being that users found *VML* intuitive and easy to learn.

Blended Space (MBS-B, MBS-T, ABS) shows a task-dependency between its design variations. We recommend *ABS* for tasks consisting mainly of transitions and few intra-reality manipulations, such as the transitions task of the first study where *ABS* was significantly faster. *MBS* is more versatile and can be used for efficient transitions of objects in tasks with a mixture of transitions and manipulations. However, the transition method (*MBS-B* or *MBS-T*) was found to influence the efficiency of *MBS* and requires further investigation.

Guidelines for Object Transitions Between VR and AR

1. *BT* is most suited when user transition is the main requirement, and object transition only occurs infrequently.
2. *VML* is most suited when the task takes place in the 'native' environment, while object transitions are required occasionally.
3. *Blended Space* is most suited for frequent transitions, though it is task-dependent which variation (*MBS-B*, *MBS-T*, and *ABS*) should be used.

6.8.2 Future Design of Blended Spaces

We introduced the concept of blended spaces as an object transition technique and evaluated the following three implementations: *ABS*, *MBS-B*, and *MBS-T*. Participants perceived seeing both AE and VE at the same time as an advantage over seeing them separately. We identify the following five blended space design factors: transition method, reality distinction, environment filter, environment foundation, and scalability.

In this chapter we investigated the first factor, **transition method** (*ABS*, *MBS-B* or *MBS-T*), which was found to be task-dependant. For a task that only consisted of transitions, automatic was preferred, while a task with a mix of transitions and intra-reality manipulations benefited from a manual transition. Further exploration of automatic transition criteria will improve *ABS* for mixed manipulation scenarios. For example, making an object manipulated in blended space adopt the reality of others in its vicinity. With *MBS-B* and *MBS-T* we explored manual transition methods implemented using hand tracking, another approach is using motion controllers to allow initiating the transition with physical buttons or joysticks.

As a **reality distinction** method we implemented a yellow-blue colour coding to distinguish AE and VE. While effective, we recognise that changing the colour of objects could also interfere with tasks where colour is of importance (e.g. transition only the yellow flowers). As such, future blended spaces should consider other methods for distinguishing between realities. For example, showing an icon next to each object, or overlaying objects with visual patterns. Asymmetry could also be introduced, visualising AE objects normally but altering VE object appearance, such as with lowered opacity or monochrome rendering.

We used a static **environment filter** on the VE, which selects only the objects relevant for the task. With blended space we improve on related work [48], by also blending environments in addition to filtering. Furthermore, we envision this filter to be made dynamic in the following three ways: task, environment, and user driven. Firstly, when users are performing a task with predetermined stages, the task can select which objects are of interest in each given stage (e.g. the museum curator first fills a display case with small trinkets, and then moves on to finding an large artefact for display on a pedestal). Secondly, the environment can drive selection, for example based on proximity of objects only the supporting geometry that is in use can be selected (e.g. only show the tables with objects on them, not the empty ones). Finally, users can have agency over the selection, for example when the VR user has finished working with an artefact they can select it for visualisation in the blended space, allowing the AR user to transition it and continue the work.

In our study the **foundation environment** for the blended space was the AE, which was combined with select objects from the VE. Alternatively to basing the blended space in reality, as with the AE, it could also be based in virtuality. Our approach of basing it in the AE is appropriate when real world objects are involved, such as placing virtual objects on physical tables. Additionally, the foundation environment can be asymmetric between users, such as AE for the AR user and VE for the VR user.

As a final factor, we wish to explore **scalability** of the blended space techniques to more crowded environments. As possible solutions to objects overlapping in blended space, we envision the use of opacity to visualise overlapping objects [228], or a technique that transposes the VE so that overlap with the AE is minimal in the blended space by either moving the VE as a whole or by transposing single objects.

We envision blended space to enable other interactions, in addition to object transition. For example, environments can be asynchronous [72], with the user being in control of the passage of time in the reality they are blending with. Or as another example, linking objects together while they remain in different realities, such as visualising an object in the VE onto a real physical screen.

6.8.3 Limitations

The task in this chapter was limited to transitioning small, handheld objects. While this is the case, we expect blended spaces to generalise to larger objects as well. For example, the *MBS* techniques allow the user to transition an object without having to hold it; while *VML* requires objects to be moved in or out of a much smaller window.

In this chapter, we chose to evaluate the efficiency of the transition techniques, from the perspective of the AR user. In future work, we will implement our transition techniques for the VR user using video see-through (VST) AR similar to the work of George et al. [81]. This enables us to further investigate the foundation environment as described in the previous section.

Moreover, implementation of the techniques for VR users would allow for an evaluation based on a collaborative task between VR and AR users. We envision that the blended space will benefit this collaboration as users can reference elements of both environments simultaneously and then transition objects as required. We also envision visualisation of the other user's state to arise as a novel requirement in this scenario, as users would need to know whether the other is in the AE, VE or blended space. Hence, future studies will investigate the different design factors of blended space on collaborative tasks between VR and AR users.

6.9 Conclusion

In this chapter we proposed novel interaction techniques for transitioning objects between distinct VEs and AEs. Two techniques were inspired by literature: *VML* and *BT*. Three other techniques were based on the novel 'blended space' concept: *ABS*, *MBS-B* and *MBS-T*. These techniques were evaluated in two user studies, the first study focused on transitioning objects, while the second study presented the user with a task requiring a mixture of transitions and intra-reality manipulations.

We found the *BT* technique to be the least appropriate for transitioning objects, because it forces the user to transition themselves as well, which causes them to lose track of the other environment. *VML* is most suitable when the task requires transitions sporadically, as participants found it intuitive and easy to learn. However, it only provides them with a limited view of the VE which makes it harder to navigate complex VEs. Blended spaces provide users with an efficient method for performing CR manipulations, as demonstrated by *ABS* in the transitions study. However, we also found that the efficiency of the technique is task dependent.

Based on these findings we discussed five blended space design factors and formulated a set of design guidelines for Cross-Reality transitions between distinct Augmented and Virtual Environments.

Chapter 7

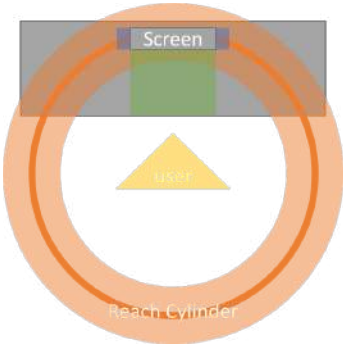
Towards a Desktop–AR Prototyping Framework: Prototyping Cross-Reality Between Desktops and Augmented Reality

[46] **COOLS, R.***, GOTTSACKER, M.*, SIMEONE, A., BRUDER, G., WELCH, G., AND FEINER, S. Towards a Desktop–AR Prototyping Framework: Prototyping Cross-Reality Between Desktops and Augmented Reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Oct.2022), pp. 175–182.

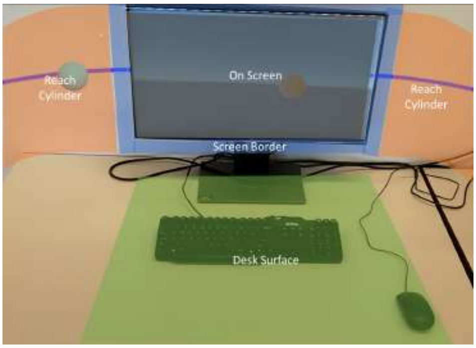
Robbe Cools: Conceptualization, Writing **Matt Gottsacker:** Conceptualization, Writing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing **Gerd Bruder:** Writing - Review & Editing **Greg Welch:** Writing - Review & Editing **Steven Feiner:** Writing - Review & Editing

This chapter is a collaboration with Matt Gottsacker from the University of Central Florida, who shares the first co-authorship (*). During writing we received feedback from Matt's supervisors Gerd Bruder and Greg Welch from the University of Central Florida, as well as Steven Feiner from Columbia University where Matt was involved in an internship at the time of writing.

https://youtu.be/amBV7uX0r9c



(a) Top-down view.



(b) View from user perspective.

Figure 7.1: Subdivision of the display space: (a) top-down view of user and their computing space, (b) computing space from the perspective of user marked by yellow triangle in (a). *On Screen* is the 2D space displayed on the screen. *Screen Border* is an area around the screen where AR content can be anchored to the screen. *Reach Cylinder* is an area all around the user that is within reach from their seated position. *Desk Surface* is the tabletop space in front of the user.

Abstract Augmented reality (AR) head-mounted displays (HMDs) allow users to view and interact with virtual objects anchored in the 3D space around them. These devices extend users' digital interaction space compared to traditional desktop computing environments by both allowing users to interact with a larger virtual display and by affording new interactions (e.g., intuitive 3D manipulations) with virtual content. Yet, 2D desktop displays still have advantages over AR HMDs for common computing tasks and will continue to be used well into the future. Because of their not entirely overlapping set of affordances, AR HMDs and 2D desktops may be useful in a hybrid configuration; that is, users may benefit from being able to work on computing tasks in either environment (or simultaneously in both environments) while transitioning virtual content between them. In support of such computing environments, we propose a prototyping framework for bidirectional Cross-Reality interactions between a desktop and an AR HMD. We further implemented a proof-of-concept seamless Desktop–AR display space, and describe two concrete use cases for our framework.

7.1 Introduction

Augmented reality (AR) head-mounted displays (HMDs) can add virtual content to users' physical surroundings. This property gives AR HMDs great potential in numerous computing contexts, many of which are yet unexplored. In this chapter, we focus on using AR HMDs in stationary single-user personal computing environments. In this domain, researchers have explored using them for extending users' 2D screens [70, 127, 33] or replacing them altogether [69, 61] to give the user a larger digital workspace. These devices also show promise for computing tasks that require users to manipulate and study objects in 3D, such as in 3D modelling and design tasks [218, 6, 57, 63] or when annotating 3D objects for educational purposes [319].

Some researchers have even posited that AR HMDs will eventually replace desktop computing environments entirely. However, for many tasks, desktop computer displays currently have advantages over AR HMDs. The displays themselves are often better: they provide higher resolutions and better colour representations. Further, AR HMDs are less ergonomic than desktop displays because they add weight to the user's head, which causes fatigue over time, and restrict the user's field of view. Desktop screens are also less expensive and more widely available than AR glasses. While some of these differences may become marginal as more investment is made to solve the grand challenges facing AR displays [19, 122], it is our position that the semi-fixed nature of desktop displays will continue to be useful for computing tasks long into the future. Rather than imagining a world in which one kind of display replaces another, it is perhaps more interesting to imagine ways in which different kinds of displays can interoperate so that people may use different displays (or even both displays simultaneously) to benefit from their complementary capabilities.

Such interactions are characterised as Cross-Reality (CR) because they involve the 'transition between or concurrent usage of multiple systems on the reality-virtuality continuum [244].' The reality-virtuality (RV) continuum [173] describes systems that immerse users in virtual content to varying degrees along an axis from fully physical to fully immersive virtual reality (VR). To contextualise our CR area of interest in the RV continuum, desktop displays are fixed in the physical world, and AR displays are at a more virtual point along the continuum because they add virtual content to the physical space. We define the Desktop–AR CR space as encompassing systems that include both a desktop display and an AR HMD and that support interactions and transitions among the distinct display spaces provided by each device. The Desktop–AR space is not to be confused with Desktop VR (or Fish Tank VR [298]), where a virtual environment is displayed on a stereoscopic desktop display and coupled to the user's head position. Nor is it to be confused with Desktop AR (sans en dash), where a user can interact with AR content displayed using a tabletop 3D monitor (e.g., Schmandt's stereoscopic workstation [229]). In this chapter, we focus on a single user using both a desktop display and an AR HMD simultaneously. Because the aim of this chapter is to produce a tool that is immediately useful, we concentrate on optical see-through AR HMDs because presently they allow the user to view the desktop display in its actual resolution (compared to video see-through HMDs that can only view the physical environment at the same resolution as virtual objects).

Currently, prototyping Desktop–AR applications is difficult because it requires custom implementation of the system on AR and desktop applications separately. In this chapter, we propose a unified prototyping framework that combines desktop displays and AR HMDs in their current forms to support Desktop–AR developers in prototyping these kinds of systems for different use cases. We discuss characteristics of the interactions afforded by this combination and present novel CR interaction and transition techniques that will be supported by our framework. Our proof-of-concept implementation allows users to select and manipulate content on and across a 2D desktop display and the 3D space extended from the desktop displayed through an AR HMD, using either the mouse or their hands as input devices for the displays.

Our goal is to discuss and refine the Desktop–AR prototyping framework with members of the CR community to gather feedback, and further guide development. In future work, the framework will be made available to the community as an open-source Unity plugin.

7.2 Related Work

In this section, we present related research on CR interactions and transitions. We then review related work on user interfaces that incorporate ideas applicable to the Desktop–AR design space.

CR systems can include either 1) multiple users at different points on the RV continuum or 2) a single user using multiple systems at different points on the RV continuum, either alternating or concurrently. Previous research on CR with users at different points on the continuum includes interaction between HMD and non-HMD users as well as interaction between HMD users with different levels of virtuality, such as VR and AR. The HMD-to-non-HMD interaction involves HMD-user situational awareness [87, 226], as well as allowing non-HMD users to interact with virtual content [100, 99, 292]. Research on HMD users with different levels of virtuality is centred around making content of one reality accessible to the other [89, 48].

Prototyping CR has been investigated in VRception [93], where a system was presented to create prototypes of CR applications across the entire RV continuum [173] spanning physical reality, AR, augmented virtuality, and VR. While VRception allows for simulating the Desktop–AR scenario, our prototyping framework supports prototyping on actual hardware and makes it possible for developers to incorporate peripheral devices such as a mouse.

In our prototyping framework, a single user uses both AR and desktop interfaces, and is thus interacting with content at different points on the continuum. Connecting interfaces with different interaction affordances introduces interaction asymmetry. There is much research on the general problem of reconciling these cross-device differences and providing meaningful and useful interactions across both interfaces. Indeed, a recent review by Brudy et al. sorted through 510 papers on the subject [32]. In this work, we focus specifically on bidirectional interactions between desktop and head-worn AR devices, which is characterised in the Cross-Device Taxonomy [32] as single-user, synchronous, spatially or logically distributed, semi-fixed, and personal.

In 1991, Feiner and Shamash [70] noted the asymmetric benefits of using 2D and 3D displays together to visualise and interact with virtual content and explored 'hybrid user interfaces' that allowed users to move 2D windows between each display completely or partially (i.e., part of a window could be visible on a desktop display while the remainder was simultaneously visible on the AR HMD). This hybrid design effectively treated the displays as complementary.

Benko et al. [12] built 'cross-dimensional gestural interaction techniques' to transition virtual content between a 2D display and a 3D AR HMD. Their system used an AR HMD, a tracked glove, and a projected 2D display, and it allowed users to pull and push virtual objects between the 2D display and the HMD's 3D space. More recently, Roo and Hachet [224] presented the OneReality system, an instrumented Mixed Reality (MR) environment that allowed users to transition virtual content among projected tabletop displays, handheld displays, and AR and VR HMDs. In XD-AR, Speicher et al. [259] presented a framework for cross-device interaction and transitions between handheld, projected, and head-worn AR displays. BISHARE [323] presented a design space for single-user interaction between head-worn and phone-based AR, They explored using the phone for spatial interactions and hand-tracked interactions with the phone display, similar to how our framework describes CR interaction between desktop and AR. Zhu et al. emphasised the different interaction strengths for each device (e.g., mobile phones are useful for efficient and precise input as well as high-resolution and full-colour 2D content, while AR HMDs are good for displaying 3D imagery in the user's spatial environment) as well as differentiating between phone-centric and HMD-centric interactions [323].

Related to the desktop input side of the Desktop–AR CR space, Zhou et al. [322] created a 'depth-adaptive cursor' that integrated a conventional 2D desktop mouse into 3D space viewed through a VR HMD. Additionally, Kim and Vogel [142] extended 2D mouse interaction into projected AR through a cursor that moves along the 3D geometry of virtual objects in the projected AR space.

Applying these CR concepts, Wang et al. [297] implemented a data visualisation system for physicists that allows seamless usage of both screen and AR space. Their application can show 3D visualisations both on the screen and in AR. Similar to our work, they also allowed the mouse to be used in 3D space, though with a button to switch between spaces rather than moving the mouse over the screen edge.

7.3 Proposed Desktop–AR Prototyping Framework

In this section, we propose a general-purpose CR framework that would allow developers to easily implement Desktop–AR CR interaction techniques. The framework would provide a seamless display space between desktop and AR, with support for mouse-based and hand-tracking input, as well as basic CR interactions and transitions. Developers can use this framework to prototype applications that make concurrent or alternating use of Desktop and AR.

To help define CR interactions in our framework, we divide the Desktop–AR display space, similar to Zhu et al. [323], into *On Screen (OS)* and *Spatial*. The spatial display space is further subdivided into *Screen Border (SB)*, *Reach Cylinder (RC)*, and *Desk Surface* (see Figure 7.1). The *Reach Cylinder (RC)* is an approximation of where the user can reach in 3D space. To aid in this approximation, we can use 'joint-centered kinespheres' [161] that model users' nearby reachable space. We discuss interactions and transitions across different subdivisions of the space, but with limited emphasis on the *Screen Border* space because we envision it being mainly used to add extra user interface elements to *On Screen* applications.

We first present our choice of input devices for the Desktop–AR framework: mouse and hand tracking (the mouse being the traditional input on the desktop screen and hand tracking being the traditional spatial input). Then we identify the ways in which these can be used to interact 'across realities,' by which we mean using the input device outside of its traditional display area or performing an interaction in one display area with effects in the other. Finally, we use these CR interactions to perform CR transitions, and describe methods by which content can be moved between display spaces. Our goal is to implement these proposed interactions and transitions into the prototyping framework, allowing developers to use them when prototyping Desktop–AR applications.

7.3.1 Input Modalities

Our framework considers two input modalities: mouse and hand tracking. First, the mouse serves as a traditional 2D input device, moving a cursor with two degrees of freedom (DOF). We chose to include this input device because it is the most common for desktop computing. Second, hand tracking allows for directly interacting with content in 3D space, and is enabled by the AR HMD. Current MR devices offer both hand tracking and motion controllers as options for spatial input, yet we chose not to include motion controllers in our framework. The main reason for focusing on hand tracking instead of motion controllers is the ease of switching between modalities, as hand tracking avoids the need to put down the motion controller when switching to the mouse. To maintain a seamless input modality in a different way, the desktop could be equipped with a touchscreen, so the user could use their hand to interact with content displayed on the AR HMD and on the desktop. In this chapter, we focus on the more common personal computing case of a user using a keyboard and mouse primarily for their tasks.

Both input modalities, mouse and hand tracking, have strengths and weaknesses. The mouse is more precise, but typically limited to 2D movements, though extensions to 3D environments have been investigated [322]. Hand tracking is less precise, but allows for direct and intuitive 3D interaction. Because of these differences in precision and dimensionality, mouse and hand tracking are complementary input modalities.

7.3.2 Desktop–AR CR Interactions

We consider the following cases of CR interaction in our framework:

- using an input device outside of its traditional display space
- using an input device in one display space with effects in the other
- using an input device with effects in both display spaces
- bimanual interaction in the same or different display spaces

Table 7.1 and Table 7.2 summarise Desktop–AR CR interactions.

Using an input device outside its traditional display space. we consider the following interaction scenarios: using the mouse in 3D space, or using hand interactions on the screen. As a basic mouse interaction, extra user interface (UI) elements added to the *screen border* can be easily accessed. For more distant interactions, we envision the mouse to remain a 2D input device, thus limiting its reach to the surface of a cylinder in 3D space rather than adding a third axis of movement. This approach allows the mouse to move off the screen onto the *Screen Border* and along the *Reach Cylinder* surface. However, we expect areas farther along the *Reach Cylinder* to be more difficult to access with the mouse in a traditional desktop setup. For example, the area opposite the desk and behind the user would be difficult to access as the mouse needs to remain on the desk. Users may position content nearby but outside of the HMD field of view (FOV), or outside of their human FOV once AR HMD FOV is wide enough, so it is necessary to help users maintain awareness of out-of-view objects [94, 92, 91]. As a possible solution to this limited reach and view, extra functionality could allow the *Reach Cylinder* to be rotated, moving along all the content attached to it.

Hand interaction could be enabled on the screen via raycasting, as the hand cannot pass the physical screen for direct interaction with the objects. Alternatively, the *On Screen* objects could be augmented with handles that stick out of the screen to allow hand manipulation. Hand tracking *On Screen* would provide less precision than mouse input, but it may be useful if users are working primarily with 3D object manipulations.

Users may also perform an **interaction in one display space with effects in the other**. For example, users could use the mouse for fine manipulation of an object displayed *On Screen* that is reflected in other display spaces such as on the *Desk Surface* to allow multiple simultaneous perspectives. As another example, users may use their tracked hands to perform coarse manipulations (e.g, 90° rotations) of an object on the *Desk Surface* that are reflected *On Screen*.

In the Desktop–AR CR space, it is also possible to **use an input device with effects in both display spaces**. This CR interaction builds on the previous one and adds that the user may manipulate an object that is duplicated in two different display spaces. This case uses the same interaction techniques as the previous one but offers additional views on the virtual object being manipulated.

The Desktop–AR CR space also affords **bimanual interaction in the same or different display spaces**. That is, users may use a tracked hand and mouse input simultaneously to interact with virtual objects. When the user's mouse and tracked hand are in the same display space (e.g., both *On Screen*), the mouse can be used for fine-grained manipulations and the hand can be used for coarser direct manipulations. When the user's mouse and tracked hand are in different display spaces, the virtual object could be mirrored in each display space to give the user different perspectives on and interaction affordances for the object. Alternatively, in the distributed input scenario, the user could use the separate input display spaces to cause a CR transition.

7.3.3 Desktop–AR CR Transitions

We use *CR transition* to refer to transitioning content between 2D and 3D space in either direction. We envision the main method for CR transitions to be based on spatial positioning; that is, content positioned in the space behind the physical display is rendered in 2D and transitioned to 3D as it moves out of this area at the sides or front of the physical display. The Desktop–AR display space is much larger than the limited area of the physical display, so moving objects may require covering a greater virtual distance than users may prefer. Thus, this CR task benefits from novel transition techniques.

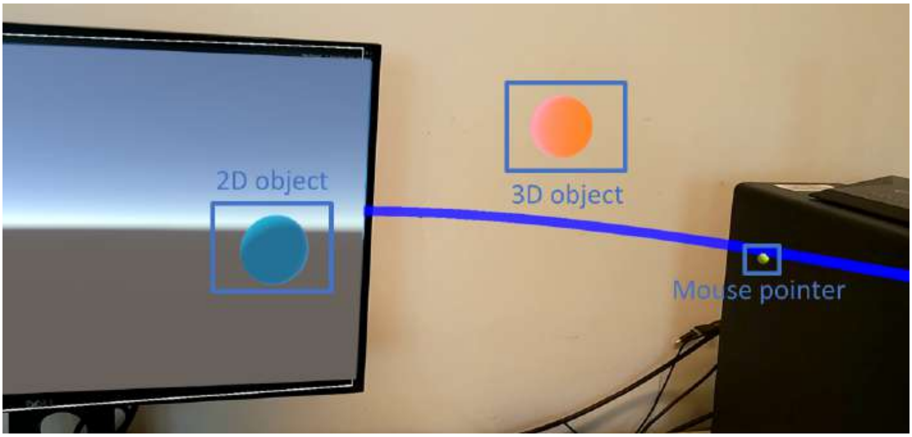


Figure 7.2: MR capture of proof-of-concept system.

We propose the following two novel techniques to more efficiently transition objects: bimanual and batched. First, the bimanual technique uses the mouse and hand tracking at the same time. The user first makes a gesture with the hand not using the mouse to indicate a position in 3D space. Then, objects on the screen that are selected with the mouse transition to this position. In contrast, the user may make a gesture in 3D space to mark objects that are then transitioned to the mouse's on-screen position when clicked. Second, the batched technique allows the user to select multiple objects, either *On Screen* or in 3D space, and then switch modality only once, after which the selected objects can be transitioned sequentially, either by clicking in screen space or making a gesture in 3D space.

Another design aspect of the transition is that objects that transition can either be moved or duplicated. In traditional systems, changes made to an object would not be reflected in a copy. In our framework, however, one may wish to manipulate an object on the screen, but visualise the changes in AR. In this case, a transitioned object needs to remain synchronised with its *On Screen* counterpart. When multiple such distributed visualisations exist, it could become difficult to know which ones belong together; thus, some indication of their connection is required. For example, a line could visually connect corresponding *On Screen* and spatial content.

7.4 Proof-of-Concept Desktop–AR Workspace

As a first step towards validating our prototyping framework, we created a proof-of-concept Desktop–AR workspace in Unity Engine 2021.2.16f. The system consists of two application instances running simultaneously, one on a desktop computer with a physical flat-panel display and the other on a Microsoft HoloLens 2. Both are instances of the same project and are networked to synchronise the virtual environment (VE) state, but differ in the way they visualise the VE. A virtual orthogonal camera is positioned in front of the physical display, and records the VE with the same viewport size. The resulting image is rendered onto the desktop display's application. The HoloLens 2 application renders the same VE in world space, blocking the part already covered by the physical display.

The system supports two interaction types: mouse and hand interaction. The mouse is implemented as a cursor in 3D space, moving on a plane that is flat on the display but curved towards the user along a cylindrical surface in 3D space (Figure 7.2). This allows the 2-DOF mouse to work in the 3D *Reach Cylinder*. Curving the plane avoids steep viewing angles onto the workspace as the cursor moves farther to the side, similar to how some large computer monitors are curved. Objects near the cursor can be moved by holding the left mouse button. As a second interaction technique, we implemented MRTK[1] hand tracking, which allows manipulation of content away from the mouse curve freely into 3D space. The MRTK 'far interaction' technique can be used to manipulate objects on the screen or at a distance.

As shown in the demo video, the proof-of-concept system supports transitioning content between 2D and 3D space based on its position. Content positioned behind the display in 3D space is rendered on the 2D display. This 2D content transitions to 3D when moved outside of the screen space, and vice versa. While this transition method is intuitive, it might not be the most efficient as it requires large mouse movements to cover the distance between spaces. Additionally, it does not support more complex interactions such as copying objects into 3D space rather than moving them. It also would be useful to further explore the rendered position of virtual content. For example, the AR HMD could render all content in front of or behind the desktop display, with only content positioned in the plane of the desktop display being rendered on the desktop display. This would allow traditional 2D content to be rendered with the full resolution and stability of the desktop screen, while 3D content and closer or farther 3D windows would be rendered on the AR HMD. We plan to investigate the *asynchronous* Desktop–AR CR space as well (e.g., user switching between a VST AR HMD and a desktop display: when the user puts on the HMD, the content transitions from 2D to 3D, and vice versa when they take off the HMD). As future work, we will refine the proof of concept with the functionality described in section 7.3.

[1] https://docs.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/

7.5 Use Cases

We explore the computing context of a user working in three dimensions, such as when creating a virtual world or doing computer-assisted design (CAD). These tasks involve creating, shaping, and manipulating 3D virtual objects. Desktop displays are useful for detailed viewing of 3D objects and for fine-grained edits and manipulations. AR HMDs are also useful in this case because they allow users to view 3D objects in actual 3D space. This feature allows users to more naturally view and manipulate the objects from multiple perspectives by directly moving the object in front of them or moving their head around it. This can improve users' spatial ability and help them more reliably perceive and mentally represent the objects in three dimensions [163]. For these reasons, an important CR interaction for the virtual world builder or CAD worker is easily transitioning objects to and from the desktop and AR HMD. Using the AR HMDs hand tracking, they should be able to grab objects from the surface of the desktop display and bring them onto the *Desk Surface* space in front of the monitor where they can manipulate them with their hands, similar to Benko et al. [12]. At the same time, the objects could be duplicated on the original desktop display so that the user may make finer adjustments with the mouse.

The virtual object the user is designing may consist of numerous nested objects. Thus, they might want to work on a specific piece of it while maintaining awareness of how that part relates to the whole. The user could work on the part *On Screen* and in the *Desk Surface* space while the whole virtual object could be displayed through the AR HMD in the *Reach Cylinder*. The whole object could have an augmentation highlighting the part being worked on and the current perspective from which the user is viewing the part. The user could do detailed work on the part *On Screen* and in the *Desk Surface* and perform more macro-level manipulations (e.g., scaling the object to fit in with surrounding objects) on the whole model in AR space.

7.6 Conclusion

In this chapter, we propose a prototyping framework for combined mouse and hand inputs across a Desktop–AR display space. To describe this framework, we discussed the input modalities involved and outlined CR interactions and CR transitions that the framework would support. We then presented a proof-of-concept implementation of such a Desktop–AR display space, and explored two potential use cases for it. In the next chapter (chapter 8) we further investigate object transitions in the proposed Desktop–AR system.

Table 7.1: Desktop–AR CR Interactions part 1. This table lists examples illustrating why a developer or prototyper may want to implement certain CR interactions made possible by supporting mouse and tracked-hand input modalities both separately (cases 1–6), and simultaneously for bimanual interactions (cases 7–15, shown in the subsequent table). The *Input Modality* columns describe the space where each input modality is interacting; OS stands for *On Screen*, RC for *Reach Cylinder*, DS for *Desk Surface*. Interaction scenarios we consider unlikely are marked as such.

#	Input Modality		Display space of primary object		
	Mouse	Hand	On-Screen [OS]	Reach Cylinder [RC]	Desk Surface [DS]
1	OS	-	Traditional 2D interaction; fine manipulations of 3D objects.	User is primarily working in 2D screen but wants to interact with a 2D window that was extended into RC space (e.g., [70, 69]).	User is primarily working in 2D screen but wants to interact with a 3D object on the DS using the mouse, as in [142].
2	RC	-	User is interacting with secondary object in RC that affects the primary object (e.g., a window with view controls is placed in RC)	User performs fine-grained manipulations on 3D object.	(Similar to 2[OS]) User is interacting with secondary object in RC that affects primary object on DS.
3	DS	-	User is interacting with secondary object in OS space that affects the primary object in DS space.	User is interacting with secondary object in RC space that affects the primary object in DS space.	User is performing fine-grained manipulations of object in DS space.
4	-	OS	User is performing direct 3D manipulation of 3D object displayed in high resolution.	User is directly manipulating something on screen that affects an object in RC space.	User is directly manipulating something displayed on screen that affects an object in DS space.
5	-	RC	User is rotating a large object in RC space while a part of that object is shown and updated live in OS space.	User is directly manipulating an object in RC space.	User is rotating a large object in RC space while a part of that object is shown and updated live in DS space.
6	-	DS	User is directly manipulating an object in DS space that is mirrored in high resolution in OS space.	User is directly manipulating an object in DS space that is part of a larger object displayed in RC space. Manipulations that affect the DS object affect the whole RC object.	User is directly manipulating an object in DS space.

Table 7.2: Desktop–AR CR Interactions part 2. This table lists examples illustrating why a developer or prototyper may want to implement certain CR interactions made possible by supporting mouse and tracked-hand input modalities both separately (cases 1–6, shown in previous table), and simultaneously for bimanual interactions (cases 7–15). The *Input Modality* columns describe the space where each input modality is interacting; OS stands for *On Screen*, RC for *Reach Cylinder*, DS for *Desk Surface*. Interaction scenarios we consider unlikely are marked as such.

Case	Input Modality		Display space of primary object		
	Mouse	Hand	On-Screen [OS]	Reach Cylinder [RC]	Desk Surface [DS]
7	OS	OS	User uses the mouse to finely rotate an object while using their hand to scale the object.	(unlikely)	(unlikely)
8	OS	RC	Quickly move an object from OS space to RC space.	Quickly move an object from RC space to OS space.	(unlikely)
9	OS	DS	User performs fine adjustments in OS space while performing coarse manipulations (e.g., 90° rotations) in DS space.	(unlikely)	User directly manipulates object in DS space while adjusting fine parameters in OS space.
10	RC	OS	User makes a gesture to quickly move object from OS space into RC space.	User clicks object in RC space to quickly move it into OS space.	(unlikely)
11	RC	DS	(unlikely)	User performs fine translations of an object with the mouse in RC space while scaling or rotating the object using a virtual control in DS space.	Users makes a gesture in DS space to quickly move object into RC space.
12	RC	RC	(unlikely)	User performs fine manipulations of object with the mouse and coarse manipulations with their hand.	(unlikely)
13	DS	OS	User performs fine adjustments in DS space while performing coarse adjustments in OS space.	(unlikely)	Users performs fine manipulations in DS space while performing coarse adjustments in OS space.
14	DS	RC	(unlikely)	User selects an object in RC space and makes a gesture to quickly move it to DS space.	User clicks on objects in DS space to quickly move it to the position the user is pointing in RC space.
15	DS	DS	Perform fine and coarse manipulations on object and see them displayed in high resolution in OS space.	(unlikely)	Perform fine and coarse/direct manipulations on object in DS space.

Chapter 8

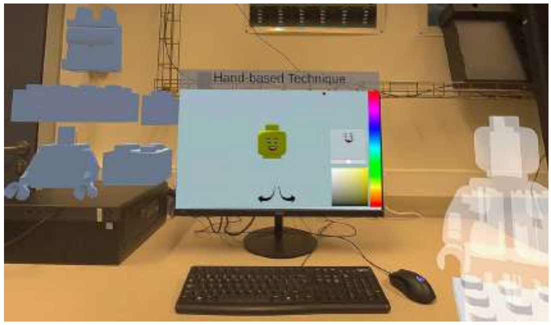
Comparison of Cross-Reality Transition Techniques Between 3D and 2D Display Spaces in Desktop–AR Systems

[49] **COOLS, R.***, MAEREVOET, I.*, GOTTSACKER, M., AND SIMEONE, A. L. Comparison of Cross-Reality Transition Techniques Between 3D and 2D Display Spaces in Desktop–AR Systems. *IEEE Transactions on Visualization and Computer Graphics* (2025).

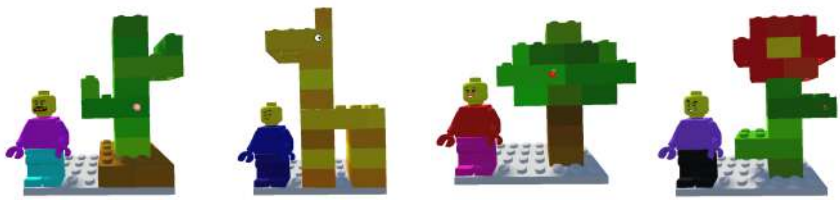
Robbe Cools: Conceptualization, Methodology, Formal Analysis, Visualisation, Writing **Inne Maerevoet:** Conceptualization, Methodology, Software, Investigation, Formal Analysis, Visualisation **Matt Gottsacker:** Conceptualization, Methodology, Writing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing

This chapter is based on the master thesis of Inne Maerevoet, and came about following the collaboration with Matt Gottsacker on chapter 7. Inne was supervised by Matt and I in regular meetings. The chapter itself was written mostly by me, based on Inne's study and with contributions from Matt.

https://youtu.be/_PdrcbaqqNg



(a) The Desktop–AR Setup.



(b) Builds used in the study.

Figure 8.1: A user completing the building task (Figure 8.1a), where they had to transition bricks from a palette on the left side, to the *2D* display space where they could perform detailed manipulations such as selecting a colour and placing stickers. The bricks then had to be transitioned again to the *3D* display space where they had to be placed into a build. Participants had to make four different builds (Figure 8.1b).

abstract The aim of this study is to develop an understanding of how virtual objects can be transitioned between 3D Augmented Reality and 2D standard monitor display spaces. The increased availability of Augmented Reality devices, in combination with the prevalence of conventional desktop setups with mouse and keyboard input, gives rise to future hybrid setups in which users may need to transition virtual objects between display spaces. We developed three virtual object transition techniques: Mouse-based, Hand-based, and a Modality Switch (where users can only use the input methods in their respective display spaces). The three techniques were evaluated in a user study (N=24) alongside a fourth condition in which participants could freely switch between Hand- and Mouse-based techniques. Participants were tasked with transitioning virtual bricks from 3D space onto the screen, then using the mouse to make fine adjustments, such as choosing the colour of the brick and placing decorations, to then transition them back into 3D space to build with. The Modality Switch technique was not preferred due to higher mental demand. Participants preferred the mouse-based technique, which allowed them to transition the virtual bricks faster.

8.1 Introduction

Recent advances in the availability of Mixed Reality (MR) devices give rise to scenarios in which they are used in conjunction with conventional desktop computing environments. We refer to this scenario as Cross-Reality (CR), a system that transitions between or has concurrent usage of multiple points on Milgram's Reality-Virtuality continuum [173]. More specifically, we refer to a Desktop–AR system as one where head-mounted display (HMD) Augmented Reality (AR) and a conventional desktop setup with mouse, keyboard and monitor, are used simultaneously [70].

Compared to conventional 2D displays, stereoscopic displays such as AR HMDs have been shown to address common challenges with visualisation in 3D space [305], as well as provide benefits for remote collaboration [316]. Nevertheless, 2D displays are still advantageous for precise manipulation tasks. Therefore, it is attractive to create Desktop–AR systems that provide both kinds of benefits to users. Indeed, there are several examples of Desktop–AR applications that have proven beneficial, such as visualising vascular models [111], analysing user study data [114], autopsies via CT scan data [211], and analysis for the treatment of Parkinson's disease [132]. However, more research is required on how these hybrid and CR interfaces can be best leveraged [242, 115], especially with respect to transitioning content between each display space [68].

In this chapter, we used consumer HMD technology (Meta Quest 3) which supports high-resolution display of virtual content, but its ability to render the real world through its video see-through (VST) technology has lower resolution and can cause issues with reading small text or details on desktop screens. Note, however, that this is a technical limitation that can be solved by using higher resolution VST or by using an optical see-through (OST) AR HMD. Additionally, the purpose of our study was to focus on users' perceptions of the interaction techniques themselves, unaffected by technical display limitations. For these reasons, we used the desktop setup for its input devices but simulated the 2D display space as a high-resolution virtual object in the HMD rather than rendering it on the desktop monitor. This kind of system has the advantage that the user benefits from the precise and familiar inputs of the 2D desktop setup, as well as intuitive 3D representations through AR [46].

Desktop–AR systems involve both *2D* and *3D* display spaces, each with their own customary input modalities. Mouse input is typically used in 2D space, and hand-based input (e.g., via controllers or hand tracking) is commonly used for 3D displays. In this work, our aim is to design for a system that preserves the advantages of each input modality in its respective (customary) display space. In other words, we do not wish to extend one modality to replace another. Instead, we focus on designing interaction techniques to facilitate the transition of virtual objects between 2D and 3D display spaces.

These transitions have directionality, either being from *3D* to *2D* or from *2D* to *3D*, both of which need to be supported by the technique. Based on the conventional modalities offered in both display spaces, mouse and hand inputs, we designed the following three Desktop–AR virtual object transition techniques: *Mouse-based*, *Hand-based*, *Modality Switch*. *Mouse-based* extends the mouse's interaction space past the edges of the screen, to allow users to drag objects between display spaces. *Hand-based* allows the user to directly grab objects from the *3D* display space and drag them to *2D*, and grab objects from *2D* to drag them out into the *3D* display space. Conversely, *Modality Switch* is designed to have users switch modality rather than extending modalities across display spaces, and works by having users select an object (with pinch or click) to then switch modalities to transition it (again, with pinch or click). These were evaluated alongside a fourth condition, *Modality Choice*, which allows users to choose between *Mouse-based* or *Hand-based* techniques at runtime, allowing them to use a different technique depending on the directionality of the transition. We evaluated the transition techniques in a task-based context involving conventional interaction with the mouse in *2D* space, and with hand gestures in *3D* space, in addition to using these modalities for the transition.

We conducted a user study (N=24) in which participants used the transition techniques to complete a brick building task. They first had to select a brick from a palette in the *3D* display space. Second, they transitioned it to *2D*, where fine-grained manipulations were performed, such as selecting a colour and placing stickers on the bricks. Finally, they transitioned the brick back to the *3D* display space, in order to place it in a build (consisting of a giraffe, flower, cactus, or apple tree, together with a Minifigure). Each build consisted of 16 pieces that needed to be transitioned and placed. We found participants performed the task significantly faster with *Mouse-based* compared to *Modality Switch*. We also found the *Mouse-based* and *Modality Choice* to result in better usability than *Hand-based* and *Modality Switch*. Additionally, we found that participants in the *Modality Choice* technique mostly used *Mouse-based* for transitions to *2D* and *Hand-based* for transition to the *3D* display space. From the results we derive two guidelines for designers of *Desktop–AR* systems: 1) *Mouse-based* transitions are most effective when objects remain in the plane the mouse can reach, and 2) Users prefer transition techniques using the modality of the target display space, in order to keep using it after the transition, and to minimise switching modalities.

8.2 Related Work

In this section, we present related work on CR interactions and transitions as well as user interfaces and interaction techniques relevant to our proposed CR transition techniques.

CR interaction systems involve interactions between users and/or objects at different points on the Reality-Virtuality (RV) continuum [173], mediated by MR devices. A variety of CR interaction methods have been explored, including techniques to facilitate external users to interrupt Virtual Reality (VR) users [87, 193], productive and meaningful collaboration and presentation across the RV continuum [99, 100, 85], and enable an MR user to interact with multiple integrated interfaces that incorporate varying stages of the RV continuum [75, 70, 294]. Research on CR transitions focuses specifically on methods for connecting phases of the RV continuum while providing a seamless and intuitive user experience. Researchers have created techniques for transitioning users [20, 44, 71, 84, 86, 116, 210, 233, 266, 284] and objects [45, 296, 70], as well as substituting real and virtual objects [250]. A subset of CR systems are hybrid interfaces, which consist of multiple display and interaction modalities. For example, researchers have investigated combinations of handheld displays with shared surfaces [144], AR [153, 241, 14] or VR [17].

In this chapter, we focus on Desktop–AR[1] CR transitions that involve a single user working with virtual objects in both a 2D standard desktop computing environment and a 3D spatial computing environment as seen through an AR HMD. In Brudy et al.'s Cross-Device Taxonomy [32], this scenario is characterised as single-user, synchronous, spatially or logically distributed, semi-fixed, and personal. Besançon et al. [15] classify our scenario as an object manipulation task with hybrid input. Such Desktop–AR systems have a variety of applications, including supporting immersive analytics practices [297, 238, 237, 157].

Researchers have explored how 2D and 3D displays can be combined for complementary purposes. Feiner & Shamash [70] first presented a system for 'hybrid user interfaces' that enabled a user to move 2D windows between 2D desktop and 3D AR displays to support interactions leveraging the benefits of each kind of display. Their system also provided support for partial CR transitions in which a window could be displayed partly in 2D and partly in 3D. Benko et al. [12] introduced a system that supported using a tracked glove to pull and push objects between a projected 2D display and a 3D AR HMD. Roo and Hachet's OneReality system [224] enabled users to transition virtual content between handheld displays, projected table-based displays, AR and VR HMDs. Speicher et al. [259] presented a CR interaction and transition framework for unifying the input and output spaces of handheld, projected, and AR HMDs. Zhu et al. [323] presented a design space focused on handheld devices and AR HMDs. Cools et al. [46] presented a Desktop–AR prototyping framework and prototype to enable CR transitions between a 2D desktop and AR HMD. Each of these examples from the literature has emphasised the different interaction strengths of each kind of device: 2D displays are best for precise input, and 3D AR HMDs are best for spatial interaction with 3D content.

[1] Desktop–AR is not to be confused with Desktop AR as in a fixed stereoscopic AR workstation [229] or with Desktop VR, e.g. "Fish Tank" VR [298]

To design CR transitions between the 2D and 3D spaces in this study, we considered how common forms of input for each space (mouse and tracked hand) might be used or extended. Tracked controllers are common input devices for AR HMDs as well, but we excluded them from this study to avoid users having to switch between multiple input devices.

Fully extending 2D mouse input into 3D space requires adding depth as a third degree of freedom to the cursor. Zhou et al. [322] presented a solution for this problem by raycasting from the user's viewpoint in the direction of the cursor to place the cursor at the depth of the first object encountered in 3D space. Kim et al. [140] created a technique to enable 2D mouse input to move along the 3D geometry of projected AR environments. For interacting with 3D objects with 2D mouse input, Plasson et al. [209] separated the mouse movements along horizontal and vertical planes. A challenge of using 2D input in 3D space is selecting objects located behind the object that is closest to the user. In this case, however, it is possible to use selection refinement techniques to disambiguate the user's intention, even using a single hand [180]. In our *Desktop–AR* system we employ a raycasting method similar to Zhou et al. [322] to detect which object the user intends to click on, however, we do not variably modify the cursor's depth position. Rather, the 3D space is a flat extension of the 2D display space. Alternatively, one could use a curved screen extension of the 2D display [46] modeled on the users' reachable space [161].

Related to the design of hand interactions for CR systems, Benko et al.'s system [12] supported hand-based pushing and pulling gestures to transition objects between 2D surfaces and 3D spaces. Additionally, for a CR system with a 2D projected display and an AR HMD, Fischer et al. [73] found that hand-based interactions were perceived as more immersive. In a recent elicitation study, Wang et al. [296] found that users preferred hand-based grab, tap, and drag interactions to transition virtual content between monitors and AR. In this chapter, we investigate how hand- and mouse-based interaction techniques affect users' performance and perceptions of usability.

8.3 Desktop–AR Design Rationale

To develop our Desktop–AR setup and the techniques that enable users to move objects between its display spaces, we had to make a number of design decisions. These were based on the following three user requirements that we identified as essential: combined display space, native interaction techniques, and transitions.

As a first requirement the user needed to be able to see both 2D and 3D spaces as a **combined display space**, requiring a sufficiently large field-of-view so that when focusing on one display space they can see the other display space in their peripheral vision. Initially we opted to use the HoloLens 2, which as an optical see-through AR device offers the user the ability desktop monitor at its full resolution [46]. However, testing revealed its horizontal field-of-view (43 degrees) to be too small for our purposes. Thus we opted to use Meta Quest 3, whose VST functionality was sufficient to capture the environment and allow the user to use the desktop input devices, but did not capture the monitor with sufficient quality. Ultimately we expected field-of-view to have more impact on the user experience and results, thus chose to use Quest 3 for its wider field-of-view and to simulate the 2D display space in the HMD.

As a second requirement the user needed to have **native interaction techniques** to interact with objects in these two display spaces, where we identified moving the object as the main interaction to be supported. Moving the object in 2D space is done simply by clicking it with the mouse and dragging it while holding down the mouse button. In terms of 3D interaction there was a choice to be made between hands and motion controllers. Pilot testing quickly revealed that motion controllers required the user to first put down the controller before they could switch to using the mouse, which caused inconvenience, thus we opted to use hand interaction for moving objects in 3D space, which is also recommended by Wang et al. [296]. Hand gestures provide intuitive manipulation of objects in the *3D* display space, here the user makes a grabbing gesture with their hand while touching the object. The object is then grabbed and is only released when the hand gesture is released.

The third requirement is that, given the combined display space and a native interaction technique in each part of it, users need **transitions** between 2D and 3D spaces to leverage displaying and interacting with content in both of them. In this case it is not clear what such a transition technique looks like, and how it should be designed for efficiency and usability. Thus we developed three transition techniques based on the native interaction techniques, which we present in the next section and evaluate in a comparative study.

8.4 Design & Implementation of Transition Techniques

This section discusses the creation of three transition techniques to move a virtual objects between *3D* and *2D* display spaces, which are *Hand-based*, *Mouse-based* and *Modality Switch*. Each technique is designed to support bi-directional transitions between *2D* and *3D* display spaces. The design rationale behind our first two techniques, *hand-based* and *Mouse-based*, is to extend the respective native interaction techniques to allow them to cross the borders between display spaces (thus not to serve as a replacement for the native technique in the other display space.)

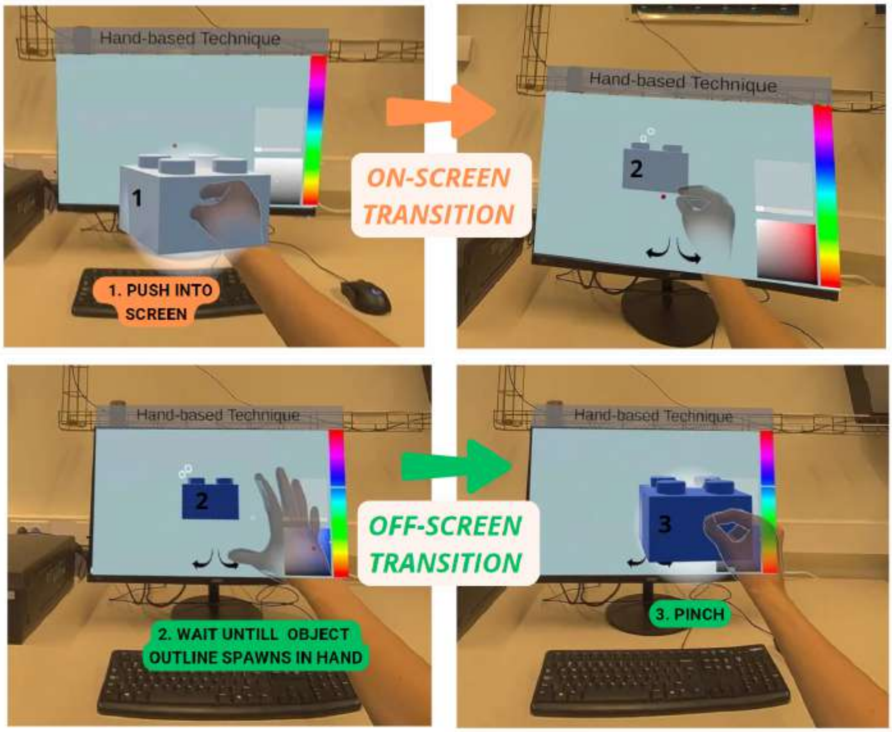


Figure 8.2: The *Hand-based* technique allows the user to transition between display spaces based on grab and drag gestures.

The first technique, *hand-based* (Figure 8.2), extends upon the hand native interaction technique so that it can be used for transitions. We designed the technique to extend the grab and drag interactions to enable transition, as also Wang et al. [296] found participants to favour in their elicitation study. Implementing these grab and drag interactions, however, needs to take into consideration that the user cannot move their hand into the physical display. When the user drags the object closer than 5 cm to the front of the screen it transitions to the *2D* display, accompanied by a small additional translation to completely move into it. When snapping into the screen the object will automatically be centred. The transition back into *3D* space is based on a distance grab, where the user points their hand at the on-screen object and pinches to transition it into their hand. We chose to prioritise direct interaction as it was found to be more accurate as compared to ray interaction [240].

The second technique, *Mouse-based* (Figure 8.3), extends upon the mouse native interaction technique. To enable transitions between display spaces, we extended the

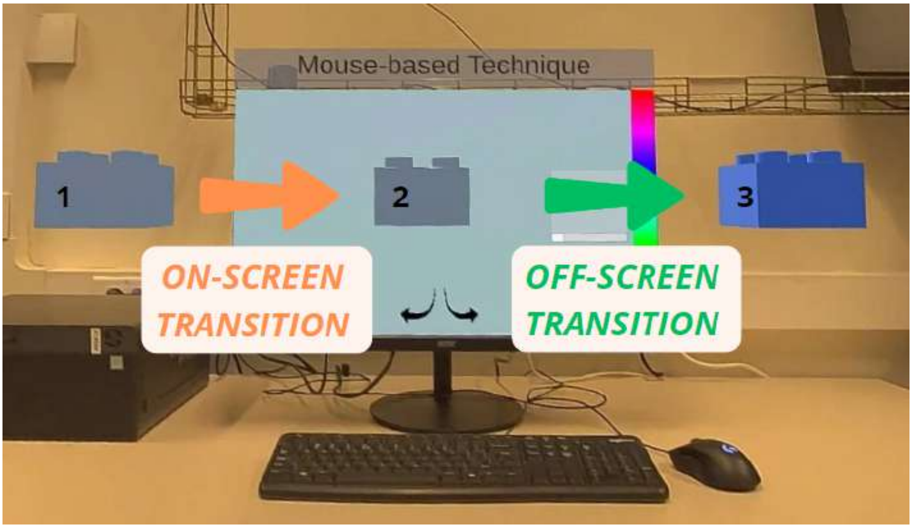


Figure 8.3: The *Mouse-based* technique allows the user to extend the mouse past the edges of the screen to drag objects to and from the *3D* display space.

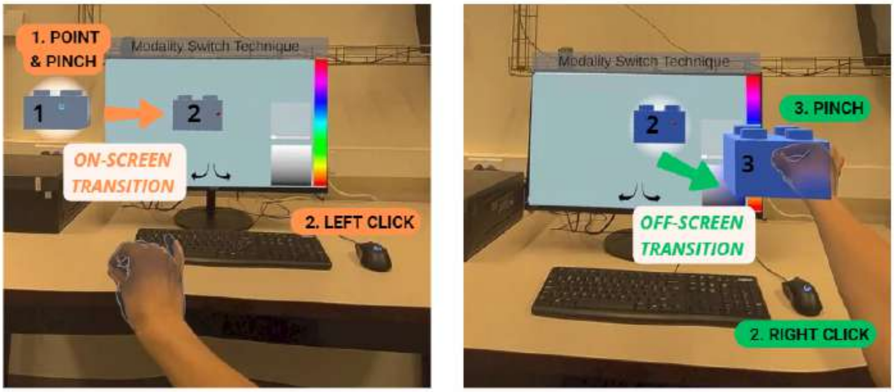


Figure 8.4: The *Modality Switch* technique avoid the need for extending modalities across display spaces by having users switch between modalities during the transition.

reach of the mouse to a plane coming out from the screen into the *3D* display space. To move an object located on this extended plane in *3D* space, the user moves their mouse to it, holds down the left mouse button, and then drags it into the *2D* display space, releasing the left mouse button when the object is at the desired location. Similarly, objects already in the *2D* display space can be dragged out. The mouse moved with the same speed in both display spaces.

The third technique, *Modality Switch* (Figure 8.4), is designed as an alternative to the previous two, where native interaction technique are kept isolated to their respective display space and the user has to switch between these modalities for each transition. It separates the selection of the object and selection of its desired destination, each to be performed in one of the native interaction techniques with a switch between them in between. For example, an object in the *3D* display space is transitioned by pointing at it with the hand, making a pinch gesture to select – switching modalities – and completing the transition by clicking at the desired position in the *2D* display space. Similarly, objects are transitioned from *2D* to *3D* by right clicking them with the mouse – switching modalities – and making a pinch gesture with the hand at the desired location in the *3D* display space.

8.5 User Study

We conducted a user study in which the three transition techniques: *hand-based*, *mouse-based*, and *modality switch*, were evaluated together with a fourth condition *modality choice* in which participants could choose between *hand-based* and *mouse-based*. The study followed a within-group design, in which each participant used all four techniques in counterbalanced order. We recruited 24 participants, of which 9 identified as female and 15 as male. Participants were aged between 21 and 53 ($M = 24.6, SD = 6.24$), and rated their experience with AR as 3.0 on a 7-point scale on average. Our setup consisted of a desktop computer with 27" Acer xv272u monitor, logitech g403 mouse, and Meta Quest 3 HMD connected through a Meta Link Cable. The study was approved by the university's ethics review board.

8.5.1 Procedure

The study involved the participant in a stationary single user setup, sitting in front of a desk with a desktop computer on it (Figure 8.1a). The participant then received an explanation about the usage and possible side effects of wearing the HMD, filled in a demographics questionnaire and gave their informed consent to participate in the study. They then completed the task four times, once with each transition technique in counterbalanced order following a balanced Latin square. At the end of each trial they filled in NASA-TLX [106] and SUS [29] questionnaires. The study ended with participants ranking the techniques based on their preference, after which they were interviewed.

8.5.2 Task

The design of the task had two main requirements: perform virtual object transitions in both directions, and use the display spaces to their advantage. First, we wanted to ensure the transition techniques were evaluated in both directions, so the task had to involve both transitions from 2D to 3D and 3D to 2D. This is especially important since techniques such as *Hand-based* and *Modality Switch* employ different interactions depending on the direction of the transition. Second, we wanted to evaluate the transition techniques in a scenario where it made sense to transition the objects. We allowed the user to perform more fine-grained manipulations in the 2D display space using mouse input. The 3D space was used as an extension of the screen to provide a palette of objects, and to perform the placing of objects which benefits from a 3D representation.

Thus, we chose to have participants complete a building task inspired by LEGO (Figure 8.1), where they were given 16 bricks with which they had to create a structure (giraffe, flower, cactus, or apple tree, as seen in Figure 8.1b) and character. These four structures consisted of the same number of objects, and were designed to be of similar difficulty. We had participants build different structures to avoid learning effects caused by them building the same structure each time. The task took about five minutes to complete. Participants were presented with neutrally coloured bricks floating to the left of the screen, as well as with a printed example on paper of what the finished build should look like. They then had to transition the bricks from 3D to 2D space, where fine-grained manipulations were made such as changing the block colour and attaching decorations (Figure 8.1a). The participants' goal was to make the bricks match those in the example build they were given. Once a brick was completed, they transitioned it into 3D and placed it in the build, where the bricks snapped into placeholders when released.

8.5.3 Metrics

At the start of the session participants filled in a questionnaire which enquired on their age, gender, and AR experience. Then, at the end of each trial, participants filled in the SUS [29] and NASA-TLX [106] questionnaires. At the end of the session participants ranked the four conditions of the study based on which they preferred for completing the task, and were are asked in an interview to describe their experience with each of the transition techniques. We logged the participants' task completion time, from when they grabbed the first brick, until they placed the last brick in the build. For *Modality Choice* we also logged which technique participants chose, for which direction of the transition (2D to 3D or vice versa).

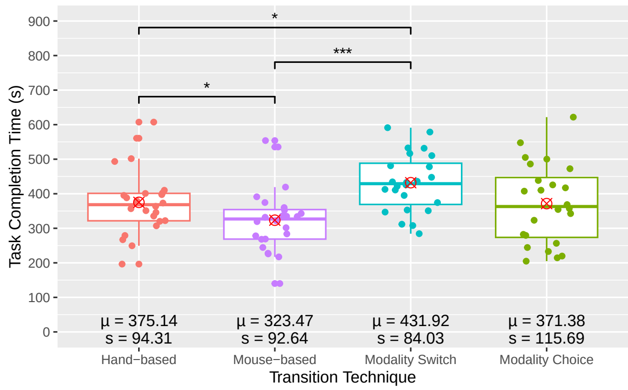


Figure 8.5: Task completion time in seconds, lower is better. Significant differences are indicated with * ($p < 0.05$) and *** ($p < 0.001$).

8.6 Results

We analysed our data using null hypothesis significance testing, following best practices [222] such as using non-parametric tests when normality assumptions are violated, and Holm's sequential Bonferroni procedure to correct for inflated type I error rates when doing multiple comparisons. We tested our only continuous metric, task completion time, for normality using the Shapiro-Wilk test, which did not reject the assumption of normality (p=0.26). Thus task completion time was analysed using ANOVA followed by a pairwise paired-samples t-test. For the discrete data, usability, workload and preference, we used the appropriate non-parametric tests: Friedman test followed by pairwise Wilcoxon signed-rank test.

Task Completion Time Results for task completion time can be seen in Figure 8.5. There was a significant effect of transition technique on task completion time ($F(3, 92) = 4.98, p = 0.003$). Post-hoc tests revealed that participants using the *Mouse-based* technique were significantly faster than those using *Modality Switch* ($p < 0.001$) and *Hand-based* ($p = 0.044$), and that *Hand-based* was significantly faster than *Modality Switch* ($p = 0.049$).

Choice of Modality For the *Modality Choice* condition we logged which technique, *Hand* or *Mouse-based*, participants chose to perform the transition. We found that for all transitions from *3D* to *2D* display space, 76.5% were performed with the mouse (and thus 23.5% with the hand), conversely for transitions from *2D* to *3D* display space

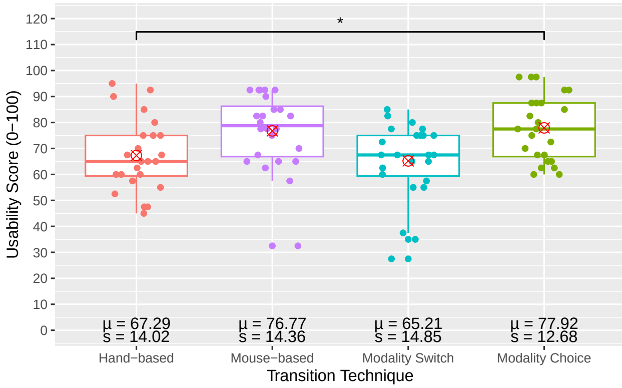


Figure 8.6: Usability score (0-100) resulting from the SUS questionnaire, higher is better. Significant differences are indicated with * ($p < 0.05$).

we found only 5.6% to be performed with the mouse (and thus 94.4% of them with the hand modality).

Usability Results for usability can be seen in Figure 8.6. There was a significant effect of transition technique ($\chi^2(3, N = 96) = 13.4, p = 0.004$). Post-hoc tests revealed that *Modality Choice* was significantly more usable than *Hand-based* ($p = 0.035$). The subscales, on a 7-point scale, revealed participants wanting to use the *Modality Choice* ($M = 4.29, SD = 0.55$) technique significantly more frequently than *Hand-based* ($M = 3.54, SD = 0.98; p = 0.018$), *Modality Switch* ($M = 3.46, SD = 0.83; p = 0.020$), and *Mouse-based* ($M = 3.58, SD = 0.93; p = 0.022$).

Workload Results for workload can be seen in Figure 8.7. There was a significant effect of transition technique on workload ($\chi^2(3, N = 96) = 16.6, p < 0.001$). Post-hoc analysis revealed that *Mouse-based* resulted in lower workload than *Modality Switch* ($p = 0.029$) and *Hand-based* ($p = 0.003$). Analysing the sub-scales we found a significant difference in terms of mental demand ($\chi^2(3, N = 96) = 18.9, p < 0.001$), where *Modality Switch* ($M = 39.2, SD = 21.1$) had a higher mental demand than *Mouse-based* ($M = 20.6, SD = 19.7; p = 0.0013$) and *Modality Choice* ($M = 22.7, SD = 14.7; p = 0.011$).

Preference Results for preference can be seen in Figure 8.8. The analysis of preferences found significant differences ($\chi^2(3, N = 96) = 15.2, p = 0.0016$), where

Modality Choice ranked significantly higher than the *Modality Switch* ($p = 0.033$) and *Hand-based* ($p < 0.001$) techniques. Additionally, we transcribed the interviews where participants explained why they preferred certain techniques in support of the discussion in section 8.7.

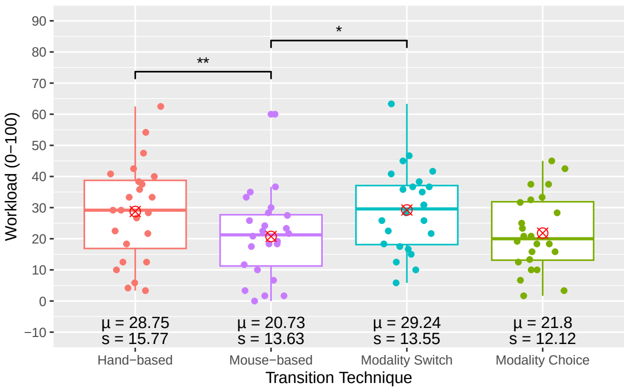


Figure 8.7: Workload as Raw TLX score, lower is better. Significant differences are indicated with * ($p < 0.05$) and ** ($p < 0.01$).

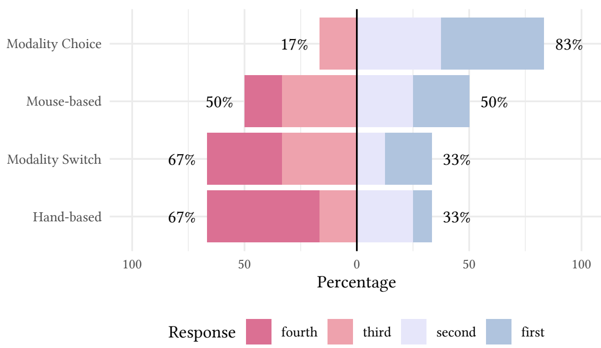


Figure 8.8: Techniques ranked based on preference. Modality Choice ranked significantly higher than Modality Switch ($p = 0.033$) and Hand-based ($p < 0.001$) techniques.

8.7 Discussion

The *Hand-based* transition resulted in lower usability and was generally disliked by participants. Participants did find the technique intuitive, as participant 13 noted *'And then the [Hand-based technique] was also nice to use. Because it was also very intuitive.'* They mainly complained about inconsistency when the object snapped to the screen or the hand, as well as it being more physically demanding to reach out with their hand. Thus, future work can explore hand gestures that allow users to remotely transition objects to the *2D* display space, such as throwing [296].

Participants found the *Mouse-based* transition technique to be familiar, which allowed them to use it most effectively as evidenced by faster task completion times and higher usability scores than *Hand-based* and *Modality Switch* techniques. We identify legacy bias and task specificity as two reasons. First, participants indicating the *Mouse-based* technique as familiar could point to legacy bias [16] having an effect on task completion time. Thus incorporating legacy devices such as mice into *Desktop–AR* setups allows its users to leverage their existing knowledge of these devices to perform virtual object transitions. Second, we note that a major limitation of the mouse transition is that it can only grab objects where the mouse is able to move, on a plane extending the monitor. This led to annoyances with participants having to reach very far to grab the object after transitioning it to *3D*, as indicated by participant 15 *'With the [Mouse-based] technique, the fact that you always had to reach past the screen to pick up objects makes it difficult for you to sit in your chair.'* Thus for transitions where the objects are already within reach of the mouse (as was the case in our study) or where they need to be transitioned to a location on the plane, the *Mouse-based* transition was more efficient.

Modality Switch resulted in the longest task completion time out of the four conditions, as well as the lowest usability and highest mental demand. Participants found the technique to be inefficient, indicating that it was difficult to use its selection methods. As participant 12 noted *'With the [Modality Switch] technique it was indeed more difficult, I thought it was difficult to select. Less intuitive. Because you're just like, if you want to grab something, just go there. And here you really had to point and pinch.'* However, they also recognised its potential to be more efficient, and expected that with training their performance would increase. As noted by participant 21 *'Once I become very familiar with it, I would definitely rank [Modality Switch] higher.'*

The results for the *Modality Choice* condition provide more insights into how participants experienced *Mouse-* and *Hand-based* techniques. We found that participants divided the task in two parts, mostly using the mouse to transition objects from *3D* to *2D*, and the hand to transition them from *2D* to *3D*. As noted by participant 4 *'With [Modality Choice] I had the best of both worlds. I could very easily insert new blocks into the screen with the mouse and remove them with a hand pinch.'* This way they do not need to switch modality after transitioning the object, and can continue with the task right away. For example, after transitioning the object to *2D* they would already have their hand on the mouse to interact with the *2D* UI. Similarly, when transitioning the object to *3D*, they are already holding the object with their hand to continue manipulating it into the correct location. Participant 22 found that *'[dragging] your block out of the screen with the mouse felt like an extra action. After editing with the mouse, immediately switching back to manipulating by hand felt easier.'* While this freedom to choose between techniques resulted in better subjective metrics, such as higher usability and lower workload, it did not cause participants to complete the task significantly faster than *Mouse-based*, which resulted in the lowest completion times.

8.7.1 Design Guidelines

We summarise our findings into the following two design guideline for seamless Desktop–AR transitions:

Transition into reach of the target interaction modality. Designers have to be considerate of the reach of transition techniques and modalities to ensure that the transition ends with the object easily reachable by the target modality. As evidenced by participants disliking the mouse for the 2D to 3D transition because it leaves the object at the edge of their reach, requiring high physical effort to grab it. Concretely for our system we noticed that the plane extending from the screen, upon which the mouse moved in the *Mouse-based* technique, was too far for users to easily reach. They noted that this was an issue for both *Hand-based*, when initially grabbing the objects, and for *Mouse-based*, when grabbing the objects after they are transitioned out of the screen. Thus, for any type of screen extension we recommend angling or curving the plane to bring objects closer to the user [46].

Minimise forced modality switches. Transition techniques should minimise how often users have to switch between modalities, as evidenced by *Modality Switch* having a higher mental demand and being indicated as more difficult to use by participants. It is worth noting that our user study could be completed with two modality switches per brick, independent of the technique used for the transition, which merely varies at which point the user has to switch. However, having users switch at a specific moment as part of the transition between display spaces was seen as not intuitive and difficult to use.

Guidelines for Future Desktop–AR Systems

1. Ensure objects are within reach of the target interaction modality after a transition.
2. Minimise forced modality switches.

8.7.2 Limitations & Future Work

As a first limitation, we only considered the most common input modalities (mouse and hand) for the design of our transition techniques, as such other modalities and input devices could also provide users with usable and efficient means of transitioning virtual objects. However, as *Desktop–AR* becomes more commonplace as a computing paradigm, future work could evaluate using a single specialised input device (e.g., a mouse that is tracked with six degrees of freedom [299]) for both stable 2D interactions and spatial 3D interactions. As a second limitation, to ensure sufficiently wide field-of-view in the *3D* display space we used video see-through technology, which forced us to render the *2D* display space in the HMD. However, advancements in optical see-through AR [122], such as wider field-of-view, are expected to overcome this limitation.

In future work we will explore Desktop–AR transition techniques in different contexts, as the context in which the transition takes place likely has an effect on which techniques are most effective. For example, when investigating Desktop–AR transitions in isolation (i.e., without performing specific tasks before and after the transition) Wang et al. [296] found that users prefer techniques based on hand gestures. However, in our study we found that for a task which involves both hand and mouse input, mouse-based transitions become more important to minimise forced modality switches.

8.8 Conclusion

We developed a Desktop–AR system consisting of *2D* and *3D* display spaces, which supported the following three virtual object transition techniques: *Hand-based*, *Mouse-based*, and *Modality Switch*. In a user study (N=24) where we compared the three techniques with each other and a fourth *Modality Choice* condition, we found that *Mouse-based* was significantly faster than *Modality Switch*, and that both *Mouse-based* and *Modality Choice* resulted in better usability than the other two techniques. For *Modality Choice* participants mainly used the mouse for transition to *2D* and hand for transitions to *3D*. From these results we derive two design guidelines.

Chapter 9

How a Vertical Surface Supports Cross-Reality Transitional Interface Tasks at Different Virtuality Levels

[44] **COOLS, R.**, BALCI, O., NGUYEN, B. V. D., VANDE MOERE, A., AND SIMEONE, A. L. How a Vertical Surface Supports Cross-Reality Transitional Interface Tasks at Different Virtuality Levels. In *Proceedings of the 2024 International Conference on Advanced Visual Interfaces* (New York, NY, USA, June 2024), AVI '24, Association for Computing Machinery, pp. 1–5.

Robbe Cools: Conceptualization, Methodology, Software, Investigation, Formal Analysis, Visualisation, Writing **Ozan Balcı:** Software, Visualisation, Writing - Review & Editing **Binh Vinh Duc (Alex) Nguyen:** Software, Writing - Review & Editing **Andrew Vande Moere:** Resources, Writing - Review & Editing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing

This chapter presents the unabridged version of the corresponding paper, which is a collaboration with Ozan Balcı, Alex Nguyen, and Andrew Vande Moere from the RxD research group at KU Leuven. They provided the robotic partition used in the study, and the technical support to set it up, as well as providing feedback on the text.

https://youtu.be/nirXpxbSHAw

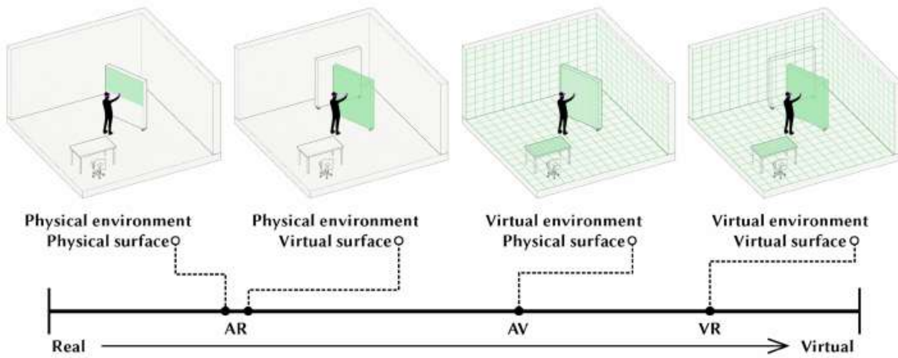


Figure 9.1: The four levels of virtuality supported by our Transitional Interface (TI). *Physical environment-physical surface* projects Augmented Reality (AR) content onto the surface via an HMD. By switching the surface mode to virtual, the partition is replaced by a virtual twin which no longer provides haptic feedback when touched, i.e. *Physical environment-virtual surface*. *Virtual environment-physical surface* is commonly referred to as Augmented Virtuality (AV), and provides visual stimuli virtually while still providing haptic feedback. *Virtual environment-virtual surface* is commonly referred to as Virtual Reality (VR), providing only virtual stimuli.

abstract Cross-Reality Transitional Interfaces, typically manifested as tabletop surfaces, handheld artefacts or head-mounted display based, are able to support tasks that require interactions at different levels of virtuality. However, there is little knowledge on how such Cross-Reality capabilities can be scaled on vertical and life-sized surfaces, by offering haptic feedback, transitioning the surface between physical and virtual environments, and moving the surface at room-scale. The aim of our study is therefore to generate an empirical understanding of which combinations of mixed reality and a user-controlled movable robotic surface (resulting in four virtuality levels) are better suited to execute five different tasks, which differed in terms of haptic feedback and spatial elements. During our user study (N=24), physical environment-physical surface was used most due to its familiarity and haptic feedback, while the virtual environment mode was experienced as more relaxing and fun, and virtual surface mode as more convenient and safe.

9.1 Introduction

Recent consumer Mixed Reality (MR) head-mounted displays (HMDs), such as Meta Quest Pro[1] and Vive XR Elite[2], put an increased emphasis on video see-through (VST) technology. VST allows the user to transition seamlessly across the reality-virtuality continuum [173, 251], which enables support for Cross-Reality (CR) scenarios, described as 'the transition between or concurrent usage of multiple systems on the reality-virtuality continuum' [244].

We developed a life-size vertical surface Transitional Interface (TI), which is a type of CR interface [8] that allows its user to transition on the reality-virtuality (RV) continuum [173]. TIs have previously been explored for small handheld [20] and tabletop sized [224] systems, while in this work we developed a life-size interactive vertical surface TI that can be repositioned at room-scale. Combining Augmented Reality (AR) with large-scale displays improves the user experience [123] and facilitates exploration of datasets [217]. As a vertical surface we used a robotic partition [183], which we chose because of its real-world architectural function when the user is not instrumented. By moving the partition to different locations, it can function as a room divider, block unwanted sights, or in our case move the interactive surface. Hence, we see potential in combining a mobile robotic surface with CR, as it provides a physical surface to which virtual content can be anchored, as well as provides haptic feedback for users interacting with the content. We then define environment and surface as two independent dimensions, which can each by physical or virtual. By combining these two dimensions our TI discretises the RV continuum into four levels of virtuality (Figure 9.1), between which the user can transition. We aim to generate an empirical understanding of how users employ the TI and its four virtuality levels to complete five different tasks, which we designed to benefit from the different types of haptic feedback the surface can provide (continuous, discrete, or no haptic feedback).

This work makes the following two contributions: (1) the development of a life-size interactive vertical surface TI, and (2) the results of a study investigating how participants used the TI to complete five tasks. In our study participants completed the following five tasks: Puzzle, Drawing, Maths, Map, and 3D manipulation. We found that they preferred to use physical environment, because it allowed them to maintain real-world awareness and because its design fit better with the tasks. Participants felt the virtual surface was more convenient and safer, however, they still preferred to use the physical surface, especially for tasks where they could receive continuous haptic feedback. Taken together, these findings guide the development of future room-scale TIs.

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

[2] https://www.vive.com/us/product/vive-xr-elite/overview/

9.2 Related Work

In this section we provide a general overview of how robots have been used together with Mixed-Reality (MR) applications, then discuss Cross-Reality (CR), and elaborate on Transitional Interfaces (TIs) as the type of CR interface that we implemented in this chapter.

9.2.1 Mixed Reality & Robots

Robots have been used to support MR applications, such as to communicate with bystanders, or to provide haptic feedback. For example, robotic room dividers create a separate area for VR users [197] to avoid conflict with nearby persons, and require less effort than manually moving dividers. In Blended Agents [231], a virtual agent moved a physical robot ball, which was controlled by a robot so that the agent could have real-world effects. Participants found that the agent's physical manipulations made it more present, and were more enjoyable and memorable.

The most common use of robots for MR, however, is for haptic feedback. Our robotic partition provides a vertical surface, on to which AR content is projected, and touched to receive haptic feedback. Haptic displays add a physical counterpart to a virtual object [250, 39], so that users can touch and feel it. This is achieved by aligning the virtual and physical objects in space, either manually if the object is static, or via motion trackers if it is dynamic. Haptic displays take many forms, such as being ground-based, holdable, or wearable [293]. For example, ZoomWalls provides encounter-type haptic feedback [318] with autonomous wall-shaped robots. Although they enhanced immersion, users were not always comfortable with not being able to see them. Hence, we designed our robotic surface to always have a visualisation when it is in use. RoomShift [271] allows robots to move different kinds of furniture, such as tables and chairs, to match a virtual layout. CoVR [26] is a ceiling-mounted robot that provides strong kinesthetic feedback, allowing users to lean against it and even be transported. CoboDeck [179] uses a robot platform which follows the users, in order to provide quick encountered type haptic feedback with the robot arm mounted on the platform.

9.2.2 Cross-Reality

Related work on CR [8] can be divided into two categories: multiple users on different points on the continuum, or a single user interacting across the continuum. Multi-user CR systems often focus on bystander inclusion and awareness, with goals such as collision avoidance, interruption, or collaboration. Providing VR users with awareness of nearby persons helps with avoiding collisions and allowing the nearby person to interrupt the VR user, however, the interruption should be designed to avoid breaking the VR user's immersion [167, 87]. Providing bystanders with the means of engaging with VR [99, 100] or AR users [292, 108] allows the bystander to join the experience uninstrumented, facilitating users that cannot or do not want to wear an HMD in joining the collaboration. This asymmetry in interface also supports an asymmetry in collaborative roles, such as a teacher using a more comfortable desktop interface, and students using an immersive HMD [311]. Asymmetry in collaborative roles can be supported by immersing multiple users in different realities concurrently, such as one user being in VR and another in AR [48, 274], which allows the AR user to monitor a VR user in a VE while maintaining real-world awareness.

A single VR user that is made aware of the physical environment, can avoid obstacles and utilise physical objects [250, 290]. While introducing physical objects into the VR experience was found useful, for example to be able to sit in a chair, it was also found problematic because it became difficult to distinguish real from virtual objects [107]. In addition to integrating the physical environment into VR, a user can also interact across multiple virtual and augmented environments [45], for which blending was found most efficient, but a lens interface easier to learn. Conversely to interacting across realities, or integrating one reality into the other, a transitional interface allows the user to move between different points on the reality-virtuality continuum, thus changing their level of virtuality.

9.2.3 Transitional Interfaces

The MagicBook [20] is a physical book artefact, that is used as such and can be read in the real world. However, it is also an interface that allows the user to transition between different points on the reality-virtuality continuum. The user transitions into AR, which augments the pages of the book with 3D models, and then immerses themselves in the virtual scene in VR. Up to six levels of virtuality are identified [224], of which level 1 closely corresponds to physical environment-physical surface except that we use an HMD to provide the visual stimuli rather than projection. Level 3 describes physical environment-virtual surface, as it decouples the virtual from the physical surface. Level 4 is similar to virtual environment-physical surface, except that we do not allow teleportation. Finally, level 5 most closely resembles virtual environment-virtual surface, where everything is virtual.

For transitions between physical and virtual environments with portals, handheld portals were preferred over ceiling-mounted [81], and users benefit from being able to peek into the other reality before fully transitioning. Transitional interfaces are also beneficial when one user has a supervising role, as AR gives them an overview of multiple VEs in the form of a hub, between which they then choose a VE to step into in support of its local user. A foot-based interface to initiate the transition is found to be natural, but also to cause fatigue [291], while simple animations for the transition itself are preferred [210]. Hence, we opted to use a conventional GUI with a cut transition animation. A transitional interface can also be developed across multiple devices, such as desktop computer, handheld AR, and immersive VR. In dyadic collaboration, one user often kept using the same device while the other switched, either because taking on and off the HMD was too inconvenient, or because they wanted to maintain a real-world overview [233].

9.3 Vertical Surface Interactions Across Realities

In this section we give an overview of our approach in designing the TI and the interactions it supports. First, we discuss the dimensions of the TI, which are the environment virtuality, surface virtuality, and surface repositioning. Then, we outline the characteristics of possible interactions with the TI, and how they might be influenced by the dimensions.

Capabilities of MR HMDs and the Robotic Vertical Surface are complementary. MR allows the user to visually alter their environment to any point on the RV continuum. However, what they lack is haptic feedback, which is often limited to controller vibration, and does not allow users to touch any virtual objects. HMDs provide input methods such as controllers or hand tracking, allowing users to interact with virtual objects, which enables them to conduct various tasks involving virtual objects. Our mobile robotic surface allows the user to move a vertical surface around the room, to bring it closer to them or to move it out of their way. As such, we can 'remove' the surface by parking it at the edge of the room, and 'enable' it again by moving it back towards the user. Moreover, we can give the user control over the position of the surface in the room, enabling them to reposition it to suit their task. When the surface is close to the user, augmentations shown on it provide haptic feedback when touched.

9.3.1 Dimensions of the Transitional Interface

Our TI has three dimensions, enabled by the mobile robotic surface and MR HMD: environment virtuality, surface virtuality, and surface repositioning. We first discuss the design variations for both environment and surface, which we limited to two virtuality levels, being physical and virtual.

Environment Virtuality has the following two levels: physical and virtual environment. The physical environment was an almost empty room in which a table and chair were placed, this gave both the user and mobile robotic surface sufficient space to

Figure 9.2: The user interface for transitioning between virtuality levels.

move around, and gave the physical environment a simple appearance. For the virtual environment we designed it to be distinct from the physical environment, as a replica of it would not provide any incentive to transition. The virtual environment resembled a forest, as exposure to virtual nature has positive effects [160], and we wanted it to provide possible benefits over the physical environment.

(RQ1) How and why do users switch between the virtual and the physical environment?

Surface Virtuality has the following two levels: physical and virtual surface, which changes whether the surface is present physically, or only as a virtual replica. We chose to use a virtual replica with the same dimensions and appearance as the physical surface. The user transitioned the surface or the environment between physical and virtual modes via two toggle buttons (Figure 9.2). To allow users to intuitively understand the modes, we used conventional 'AR' and 'VR' labels for the environment mode, and 'real' and 'virtual' labels for the surface mode.

Surface Repositioning is enabled by a mobile robot, which can move the surface around the room. As an additional dimension to our transitional interface, the user can reposition the surface to any desired locations. The technique of repositioning the surface is analogous between virtual and physical surface modes: the user points down at the floor with their motion controller to position a marker at the desired location and then confirms the selection, after which the surface will initiate the repositioning. The movement to a new location differs between modes, as the physical surface is required to physically reposition itself, which took between 1 and 30 seconds depending on how far it needed to move. However, we designed the virtual surface to take advantage of not having this limitation, and allowed it to move between positions instantly. Following this design decision, we expect the virtual surface to be easier to reposition. Additionally, the user could optionally rotate the surface prior to initiating to move.

(RQ2) How does surface virtuality affect surface repositioning?

9.3.2 Characteristics of Possible Interactions

We designed the tasks (Figure 9.4) in such a way that users could complete them in any configuration, as to not bias them to prefer one configuration over the other. Thus, components that are essential to the task are always visible, appear the same, and have the same functionality independent of the configuration of the TI, such as interfaces, objects or information that are required for successful completion. However, for vertical surface-based tasks, we identify the following two interaction characteristics that impact how the TI supports the task: haptic feedback and environment connection.

Haptic Feedback informed the design of our tasks, for which we identified the following three levels: continuous, discrete, and no haptic feedback. With continuous haptic feedback we refer to an interaction where the user drags their finger across the surface, for example drawing, dragging an item across the surface, or moving a slider. Discrete haptic feedback refers to when the user only receives haptic feedback at certain instants, such as selection an object or pressing a button. We also take into account tasks without haptic feedback, where all manipulations happen in front of the surface without the user touching it. We expect users to benefit more from the physical surface as the task integrates more haptic feedback.

(RQ3) For tasks with different levels of haptic feedback, how and why do users switch between virtual and physical surface?

Environment Connection is a second aspect that informed task design, as elements connected with the task placed at different locations in the room enable the user to benefit from repositioning the surface. We designed the spatial elements to be available at all virtuality levels. The spatial elements had varying levels of virtuality: physical object, virtual anchored to a physical object, or virtual floating mid-air. The design

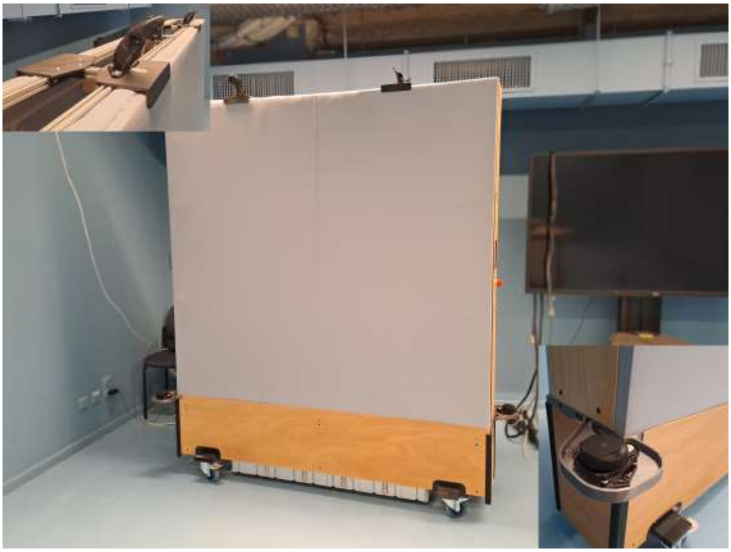


Figure 9.3: The mobile robotic surface. Quest Pro motion controllers were added to 3D printed mounts on the top of the surface (top left insert). LiDAR scanners are located at the bottom of the surface on both sides (bottom right insert).

rationale behind the spatial elements was that participants had an incentive to reposition the surface, either by rotating it and getting surface and relevant information into one view, or by repositioning it to move it closer to the spatial element required for the task.

9.3.3 Implementation

The CR transitional interface consists of two interconnected parts, a modular robot platform implemented within the ROS (Robotic Operating System)[3] framework, which provides the physical part of the interface, and a Meta Quest Pro running a Unity[4] application that provides the virtual stimuli, and controls the robot platform.

The robotic surface was developed primarily for use in architectural studies of human-building interaction [185], and consisted of an aluminium frame, closed off with fabric and wood panels and supported by four caster wheels (Figure 9.3). The bottom of the frame was integrated with a customisable, industrial-level mobile service robot, including two KELO Robile drive wheels[5], a battery and a controller module. To avoid

[3] `https://www.ros.org/`
[4] `https://unity.com/`
[5] KELO Robile: kelo-robotics.com

collision, the robotic surface perceives its surroundings with two LiDAR scanners. To update its position in the virtual environment, two 3D-printed Quest pro controller mounts were attached to the top of the robot, in which we slotted one of the controllers. Users held one controller in their dominant hand, and the other was attached to the top of the robot, as shown in Figure 9.3.

The semi-autonomous movement of the robotic surface relied on a custom software developed within the ROS framework, allowing it to localise and navigate within a known map of the study room. At the same time, the LiDAR scanners allowed it to perceive unpredictable obstacles, such as the user, to stop its movements instantly or navigate around them. During the study, a customised Python script subscribed to a TCP connection, on which it received the navigation goals as geometrical coordinates. The ROS software then navigated the robotic surface to these given coordinates, while ensuring safety for the user at all time.

9.4 User Study

To gain an understanding of how users would use the TI, and transition across the four virtuality levels, we conducted a user study. We recruited 24 participants (12 female, 12 male), aged between 21 and 60 ($M = 29.4$, $SD = 9.8$). They were recruited through a mailing list at the local university, and had a mean MR experience of 3.1 ($SD = 1.6$) which they rated on a 7-point scale (1—never used before, 7—daily use). Participants spent 19.9 min ($SD = 6.3$ min) on average to complete the task, the total duration of the study including the questionnaires and interview varied between 30 and 40 minutes. One participant was left-handed, and used the left controller.

9.4.1 Interactive Surface Tasks

Task	Haptic feedback	Spatial element
Drawing	continuous	virtual/physical gamepad
Map	none	map legend anchored to physical table
Puzzle	discrete	mid-air floating text
3D manip.	none	mid-air floating objects
Maths	continuous	none

Table 9.1: Table of tasks and their type of haptic feedback received from the mobile surface and spatial element, in order of physicality of the spatial element. Ranging from a physical reference object, to virtual text anchored to a physical table, floating text, and virtual 3D objects.

(a) Puzzle

(b) Drawing

(c) Maths

(d) Map

(e) 3D manipulation

Figure 9.4: The five tasks, with the surface positioned in proximity to their spatial element. The environment and surface modes were set to virtual.

Participants performed the following five tasks: *Puzzle, Drawing, Maths, Map, and 3D manipulation* (Figure 9.4, Table 9.1), which we implemented based on the characteristics discussed in subsection 9.3.2.

Puzzle: Participants solved two puzzles, each consisting of two sets of three labelled buttons. They referenced a sheet to find which numbers corresponded to the label, and then pressed the two buttons that together summed to the indicated amount. The button press provides a moment of discrete haptic feedback.

Drawing: Participants sketched a gamepad on a virtual whiteboard, while using a gamepad that was lying on the table as a reference. The reference was presented in the physical environment with a physical gamepad, and in the virtual environment with a virtual model. The whiteboard consisted of a white plane on which they could draw by touching it with their finger, providing continuous haptic feedback, and a set of seven colours and an eraser between which they could choose.

Maths: Participants used the same whiteboard as in the previous task to write down the answer to five calculations.

Map: Participants grabbed and moved three markers to landmarks on a map. To find the landmark they first referred to a legend located on the physical table to find the number corresponding to the landmark, and then find this number on the map. Markers were moved in front of the surface, and did not cause haptic feedback.

3D manipulation: Participants had to match the position and orientation of five platonic solids from the environment, to those of the targets located in front of the surface. The solids could be grabbed and moved with the controller trigger button. Solids were placed in front of the surface, providing no haptic feedback.

9.4.2 Procedure

To avoid a bias for the first level of virtuality the participant experienced, we alternated it so that for each of the four levels of virtuality, six participants experienced it initially when donning the HMD. Then participants went through a training phase, in which they were instructed to change environment and surface modes, and to reposition the surface. When participants were finished exploring the TI, they were reminded that they could change the modes through the entire study, and continued by completing the five tasks. To provide participants with an incentive to change virtuality level, in between tasks the virtuality level was randomised (each level was selected exactly once per session). In addition, the surface also relocated to a different location from a set of four predetermined locations (Figure 9.5). The participants then received a reminder to change virtuality level, and reposition the surface, before they started the next task.

9.4.3 Metrics

During the study the Unity application automatically logged entries such as mode changes and task start or completion events. An entry in the log contained a timestamp, locations of the user and surface, as well as the name of the event that triggered. We also logged the user and surface position every 300 ms. After the study participants filled in a custom questionnaire in which they ranked the four virtuality levels by preference, per task. The initial ordering of the levels in the custom questionnaire was set to be randomised, to avoid bias of the initial order. After which we conducted a semi-structured interview, with leading questions ‘How did you experience the difference between the VR and AR settings?’, ‘How did you experience the difference between the physical and virtual surface?’, and ‘Was it helpful that you could move the surface?’.

9.5 Results

In this section we present results from the data logged on the HMD, questionnaires, and interviews.

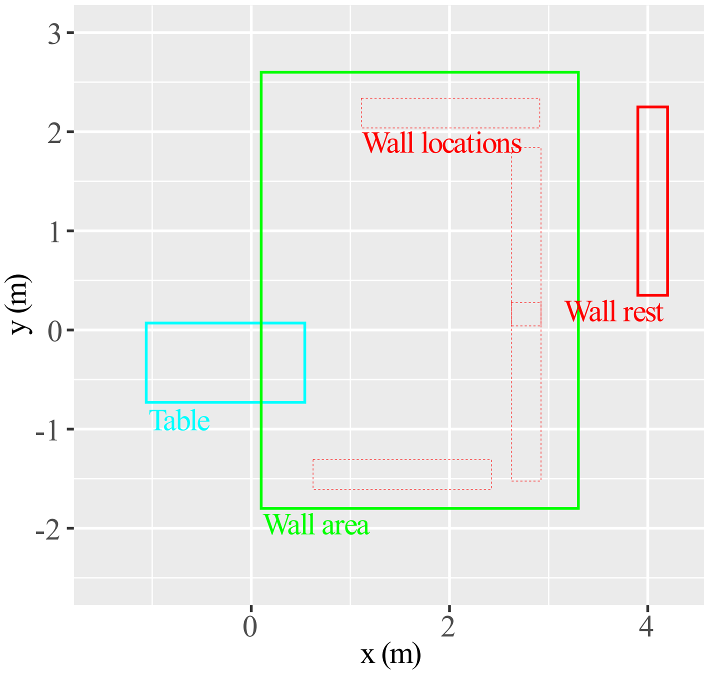


Figure 9.5: Layout of the room.

9.5.1 Quantitative Data: Time, Surface Movement, and Questionnaire

We first identify which metrics to derive from our log files for analysis, based the following three hypotheses.

(H1) participants will have a preference for an environment, in which they will spend more time and which will rank higher.

(H2) virtual surface will be easier to reposition, and will be repositioned more often and used more for tasks without haptic feedback.

(H3) participants will prefer the physical surface for tasks with haptic feedback, and will rank it higher and use it more for the drawing and maths tasks (continuous haptic feedback), and puzzle (discrete haptic feedback).

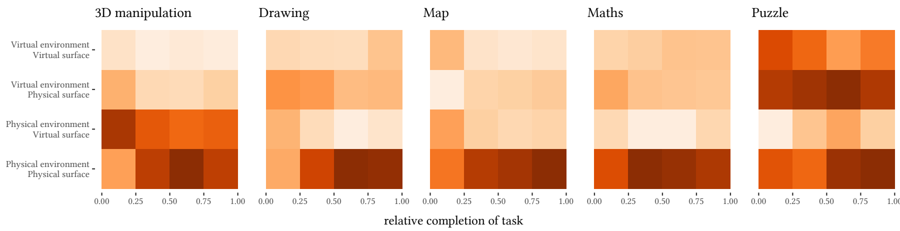


Figure 9.6: Overview of virtuality level transitioned to by participants during the course of the five tasks. Darker indicates more participants were at that level, for the indicated part of the task.

Following these hypotheses we analysed our logged data for dependent variables time (per virtuality level and per task-surface combination), and number of times the surface was repositioned by the participant (per surface mode and per task-surface combination). Then the results of the preference questionnaire were analysed (per environment mode and per task-surface combination). Following a Shapiro-Wilk test normality was rejected for all data, thus we used a non-parametric analysis. Friedman test was used, and is reported in the text and on graphs with two independent variables. For data with more than two levels, we performed a pairwise Wilcoxon signed rank text which is reported in the graphs, alongside means and standard deviations.

To verify the data and gain initial insight into how the TI was used over the course of each task, we created a heatmap that shows how participants were distributed across virtuality levels for four quadrants of task completion (normalised: 0 is task start and 1 is completion; Figure 9.6). Participants used physical environment-physical surface most, which can also be observed in the heatmaps as the bottom row showing a darker shade of red. We can also observe that participants chose to use the virtual environment more for the puzzle task, though some shifted towards physical environment in the last quadrant. For 3D manipulation, most participants opted to start the task in physical environment-virtual surface mode, though also switching towards the physical surface.

To gain insight into how participants used the TI, we analysed the *time* in two ways: per virtuality level, and per task-surface combination. Because completion times between participants varied (between 9 and 31 minutes; $M = 20, SD = 6.3$), we normalised the times between 0 and 1. We found significant differences between virtuality levels ($\chi^2 = 15.05; p = 0.002$), with participant switching to and using physical environment-physical surface, longer than the other three (Figure 9.7). There were also differences between which surface mode was used in combination with which task ($\chi^2 = 32.5; p < 0.001$). Participants used virtual least for the maths task, which was used significantly less than the two most-used combinations: physical surface for drawing and physical surface for the puzzle task (Figure 9.8).

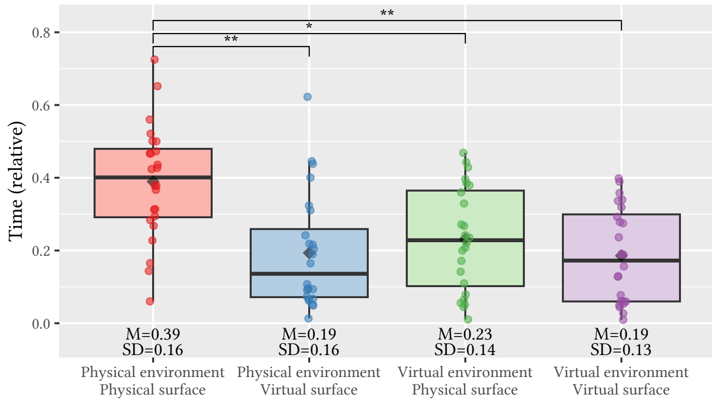


Figure 9.7: Relative time participants spent at each virtuality level.

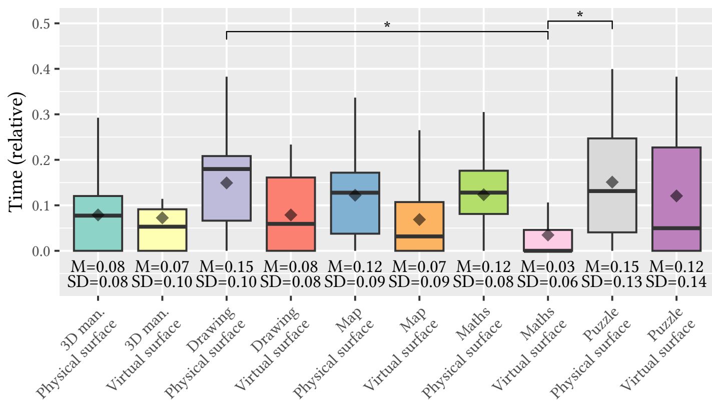


Figure 9.8: Relative time participants spent in each task-surface mode combination.

To compare how often the surface was moved, we calculated the *number of surface moves per minute*, which was not significantly different between virtual and physical surface ($\chi^2 = 0$; $p = 1$; Figure 9.9). However, the virtual surface has a greater standard deviation, and more participants that moved it over once per minute. There was no significant difference between how often the surface was moved in the different tasks ($\chi^2 = 2.3$; $p = 0.51$; Figure 9.10).

We analysed ranks participants assigned to the virtuality levels as numbers ranked between 1 (ranked first) and 4 (ranked last). To isolate environment as a single parameter we averaged the rank of the two virtuality levels that contained that environement, for all

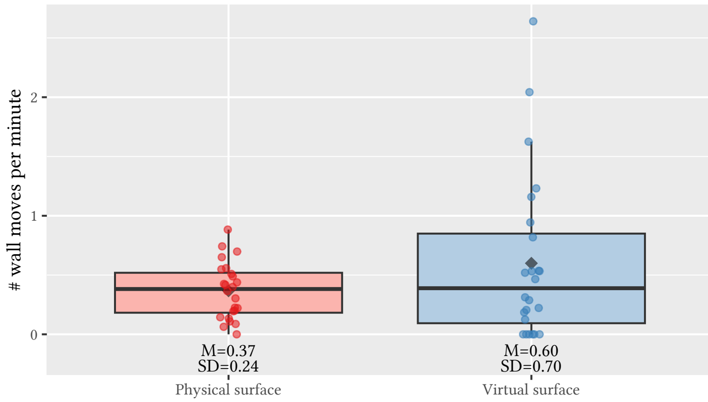


Figure 9.9: How often participants moved the surface per surface mode.

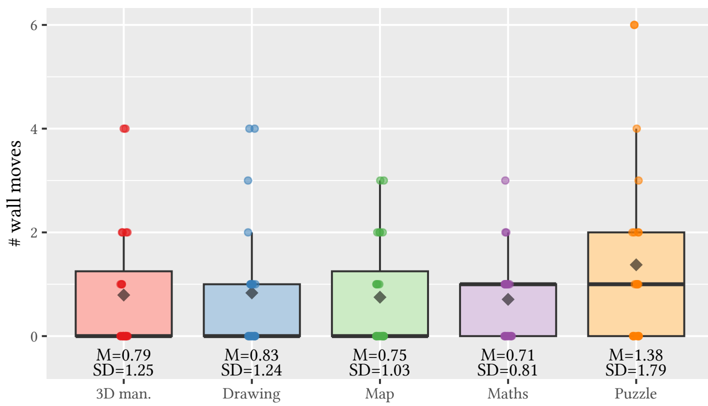


Figure 9.10: How often participants moved the surface per task.

tasks, which showed participants preferred physical environment ($\chi^2 = 6.0$; $p = 0.014$; Figure 9.11). Comparing which surface mode they preferred for which task, there is a significant difference ($\chi^2 = 24.2$; $p = 0.004$; Figure 9.12), which the post-hoc test reveals to be for the 3D manipulation task where virtual surface is preferred over physical surface.

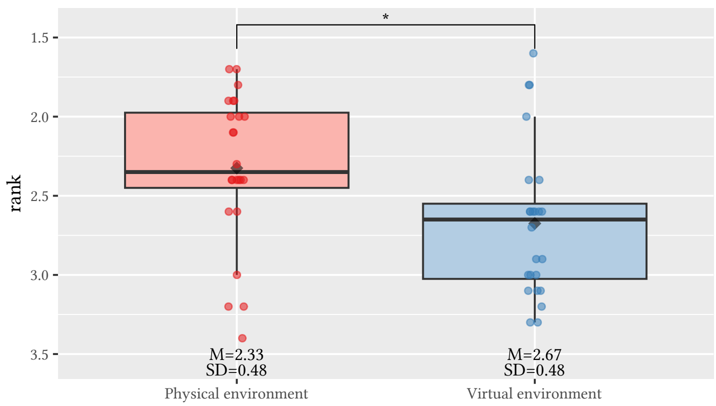


Figure 9.11: How participants ranked the environment modes.

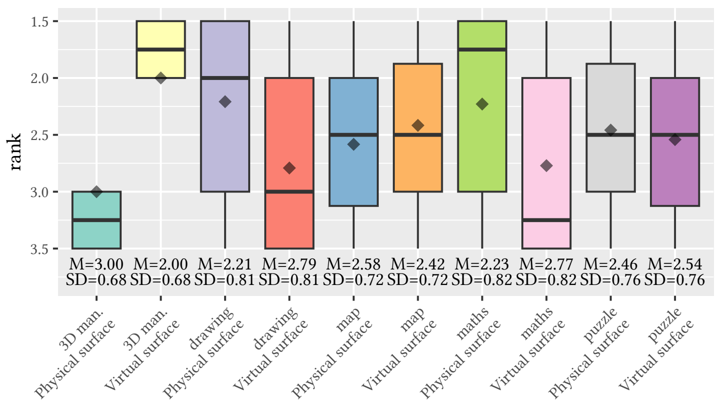


Figure 9.12: How participants ranked the surface mode for each task.

9.5.2 Qualitative Data: Interview

We transcribed the interviews, and performed an inductive thematic analysis [28] in two iterations. Coding was done by the main author, after which the other authors reviewed the themes. This resulted in 43 codes, which were grouped into six themes. We report on all codes mentioned by more than one participant (34 codes).

Environment design. Whether participants' transitioned to the physical and virtual environment was mainly based on personal preference (9 participants), such as participant 5 who preferred VR '*it was a personal preference that I use VR because it's* *more relaxing like I was in the nature*', and participant 24 who preferred the physical environment '*I like the design of the forest, but I prefer the augmented reality*'. Some participants found that there was little added value to the forest environment (3), and even that it was distracting (3), and artificial (3). They also found the physical environment to be simpler, compared to the VE (2), indicating the design of the environment to be more important that whether it was physical or virtual (2).

Consequences of real-world awareness. Participants found they could find the spatial components of the tasks more easily in the physical environment due to better spatial awareness (4), which is explained by participants finding the physical environment more familiar (5). As indicated by participant 23 '*VR was like I was in a whole new world, and in AR I felt like I knew what was happening and what was around me*'. Additionally, participants also indicated to benefit from being able to refer to physical objects (3), for example to draw the gamepad in the drawing task, and found that it provided a clearer view of the physical surface (2).

Qualities of VR. Contrary to the more practical benefits of physical environment, participants found more qualitative benefits in VR. They found VR more relaxing (4), and related it with being pretty (3), fun (3), and enjoyable (3). VR felt more like a game to them (2). Participants also mentioned that VR was preferred when the focus was on the surface, and they did not need to reference the environment (3). Participant 22 used these qualities as a basis for their choice between the environments, stating '*for example for the maths questions I thought, ah calm nature, so that's what I based myself on for the differences [between the environments]*'. Participants expressed that the choice for both surface and environment modes could be task dependant.

Task dependency. Participants found that the haptic feedback from the physical surface helped them with the drawing and writing tasks (14), and that the virtual surface was more suited for 3D manipulation (8), the map task (5), and the puzzles (3). As participant 3 summarises: '*So I think if I want to draw or write down something, I prefer to use the real surface. But if I did some other jobs, I think virtual surface is better for me*'. In general participants found that the physical environment was more suitable to perform the tasks they were given (3). Two participants mentioned the physical environment being more appropriate for the map task (2). Overall, comments on task dependency were mostly related to the surface, while participants found there was little influence of the environment on the task (3). As indicated by participant 20 '*Just to perform tasks like these, augmented reality seems better to me*'.

Moving the surface. Participants mostly moved the surface closer to the task area to facilitate task completion (10), as participant 17 said '*Because I had to for the various tasks it told me to like, go to a particular point in the room. So to be in proximity, it was nice to be able to move the surface closer to where I had to be*'. Less frequently, they also moved the surface to uncover parts of the environment it was blocking (3), such as participant 10 '*in the beginning my clues were behind the surface, so I had to* *move the surface again to look at them.*' Some participants also did not find moving the surface useful to facilitate the tasks (3), though they also mentioned it would be more useful when the working area were larger (4). In general, participants found the physical surface less convenient (4), and virtual surface more convenient to move (11) due to it repositioning faster, which participant 9 summarises as '*there the advantage of virtual surface was that I did not have to wait for the real surface to move*'.

Safety. Participants remarked that they felt safer around the virtual rather than the physical surface (3), and that the physical environment also helped them feel safer around the robot (3), participant 17 '*I feel more safe and more comfortable with the [physical environment] mode because in the virtual [environment] I cannot actually see the the wall and I would be worried about hitting the wall if I walk.*' Additionally, participants were sometimes confused by whether the surface was physical or virtual (4), which could lead to safety issues. Participant 20 '*with the AR I felt, Yes, I felt that I was more aware of the environment and I found it more reassuring, otherwise I was afraid I was going to run into things or something*'.

9.6 Discussion

We implemented a transitional vertical surface that allowed participants to reposition it in the room, and change between four levels of virtuality. To gain an understanding of how users would use the TI, we then conducted a user study (N=24) in which participants used the surface to complete five tasks.

9.6.1 Physical Environment Preferred for User Study Tasks (RQ1)

We found that participants tended to prefer the physical environment and use it more (H1), for which we identify real-world awareness and environment design as the main factors for this trend. Participants saw real-world awareness as an advantage, as it provides a more familiar environment which users typically prefer [278]. They also experienced better spatial awareness, making it easier to navigate, which could be related to the physical environment being simpler and more familiar. Moreover, it solves typical CR problems such as interruptions [87] and collision avoidance [315], which made participants feel safer. For example, participants reported feeling safer around the robot in the physical environment mode, more than when the robot moved when they could not physically see it [318]. Hence, real-world awareness contributes to physical environment being better suited, and thus experienced as more functional, for the specific activities participants were tasked with in our study.

With environment design we refer to the visual appearance and layout of the environment, being a room with a table, chair and large empty space for the physical environment, and resembling a forest for the virtual environment. Conversely to real-world awareness, this is not inherent to the virtuality mode but rather a choice that is made by the designer of the system. Albeit, with the virtual environment being easier to design for as it does not require changing physical locations. Participants indicated that the design of the physical environment was simpler, and fit better with the tasks, making it more functional. Conversely, performing the tasks in a forest seemed artificial, and out of place. Participants still reported the positive aspects we expected the virtual environment to have, such as being more relaxing, fun, game-like, and enjoyable. However, these did not provide sufficient added value to induce switching to the virtual environment. Hence, in our study we did not succeed in designing the virtual environment in such a way that participants would prefer it over the physical environment for the tasks they were given.

We designed our tasks for participants to complete them in any configuration of the TI, however, we envision that this is not always possible, giving rise to the following three categories of activities: fully transitional, semi transitional, or non-transitional. A fully transitional activity, such as the tasks in our study, is one that can be completed independent of how the environment is designed. A semi transitional activity, is one that can be completed in any environment, but for which some environment offer benefits over others. For example, to learn about forest animals in an environment resembling a forest. A non-transitional activity is one that cannot be completed in any other environment than the one it was originally designed for. For example, when the environment offers functionality tied to the activity, as when it revolves around navigating the environment. Moreover, for transitional activities, users' individual environment preference can change depending on the circumstances, such as transitioning to the virtual environment to achieve isolation as the physical environment becomes more crowded with colleagues [225].

We found that participants preferred the physical environment for its more familiar design and the benefit of real-world awareness, and when given the choice between physical and virtual environments, they transitioned to the physical environment more. However, we also recognise that this finding is specific to the activities in our study, and discussed how the nature of the activity (fully transitional, semi transitional, or non-transitional) could impact these findings.

9.6.2 Virtual Surface is More Convenient (RQ2), Physical Surface is Used More (RQ3)

Participants reported that the virtual surface was easier to reposition, however, we found no evidence of them moving it more often than the physical surface because of this

(H2). We did find that the number of physical surface moves participants performed per minute to have a lower upper limit and smaller standard deviation than those of the virtual surface. This indicates that it is possible to move the virtual surface more extensively than the physical surface, even though most participants refrained from doing so. Likely, the marker that is shown prior to confirming the new surface position was sufficient for participants to confirm the new position was as intended, not requiring any further movements to correct errors. Additionally, participants reported feeling safer around the virtual surface than around the physical surface.

Participants benefited from the physical surface mode for tasks that allowed for continuous haptic feedback, as evidenced by them using the physical surface more for the drawing task, and the virtual surface less for maths (H3). For the 3D manipulation task without haptic feedback, participants significantly preferred the virtual surface over the physical surface, however, this did not cause them to use it more. Thus participants switched to the physical surface for its haptic feedback, but were unlikely to switch back to virtual once in physical surface mode. This led to participants using the physical surface more across all tasks, especially in combination with the physical environment mode. They preferred physical surface for its more realistic visual appearance compared to the virtual twin as participants noted in the interview that the physical surface gave them a clearer view. As noted by participant 13 '*For example, to draw the gamepad I had to quickly turn my head ... and then it was better if I saw [the surface] correctly*'.

While physical surface mode was used more by participants because they preferred its appearance and haptics, this trend could be affected by the scope of our study, with different patterns arising for prolonged use of the system. We expect that for short activities users are unlikely to switch to physical surface mode, for example, when making a small modification to existing content on the surface. For more extensive collaborative activities, users could be more likely to switch to physical surface mode.

Guidelines for Vertical Surface Transitional Interfaces

1. Users are inclined to transition to the physical environment mode of the Transitional Interface, and should be given the option to do so.
2. The virtual surface is easier to reposition.
3. Users benefit from the physical surface mode for tasks that allowed for continuous haptic feedback.

9.6.3 Limitations & Future Work

In this study we only looked at tasks that benefited from vertical surface haptic feedback, as this is what our TI is limited to. This constrained the set of tasks to those that involve vertical surfaces, from which we chose a subset to include in the study. In our study we only looked at a single user, whereas TIs are often used by multiple users concurrently [291]. Thus a task involving two or more participants would provide us with different insights on how the TI performs for collaboration.

While switching between the physical and virtual environment was instant, switching between physical and virtual surface took some time for the robot to reposition. An implementation such as Mortezapoor et al. [179] could have decreased this time. They programmed the robot to follow the user while not providing haptic feedback, so that it would be close when it was required.

Our virtuality level UI consisted of two toggle buttons, which the participants used to toggle the environment and surface modes separately. While effective, this approach had downsides, for example, two presses were required to transition between physical environment-physical surface and virtual environment-virtual surface. In future work we wish to iterate on the UI by comparing it with different designs, such as sliders [93]. Additionally, because users expressed some confusion, we wish to develop visualisations that help users distinguish between the virtuality levels.

9.7 Conclusion

In this chapter we developed a transitional interface, using an HMD and mobile robotic surface, that allowed for virtuality switching between the following four levels: physical environment-physical surface, physical environment-virtual surface, virtual environment-physical surface, and virtual environment-virtual surface. With this TI participants performed five interactive surface tasks that differed in terms of haptic feedback and environment connection: puzzle, drawing, maths, map, and 3D manipulation. From the study we gathered quantitative data on how the participants used the TI, and qualitative data on the participant's experience with it and its different virtuality levels.

We found the physical environment-physical surface was used most by participants, gravitating towards the more familiar physical environment and haptic feedback from the surface for the tasks in the study. When participants used the other virtuality levels, it was due to the relaxing and more fun appearance of the forest in the virtual environment mode, and due to the increased safety and convenience of the virtual surface mode, especially for tasks without haptic feedback.

Part III

Context and Design of Future Cross-Reality

Chapter 10

The Impact of Near-Future Mixed Reality Contact Lenses on Users' Lives via an Immersive Speculative Enactment and Focus Groups

[52] **COOLS, R.***, VENEMA, R.*, ESTEVES, A., AND SIMEONE, A. L. The Impact of Near-Future Mixed Reality Contact Lenses on Users' Lives via an Immersive Speculative Enactment and Focus Groups. In *Proceedings of the 2024 ACM Symposium on Spatial User Interaction* (New York, NY, USA, Oct. 2024), SUI '24, Association for Computing Machinery, pp. 1–13.

Robbe Cools: Conceptualization, Methodology, Formal Analysis, Writing **Renée Venema:** Methodology, Software, Investigation, Formal Analysis, Writing **Augusto Esteves:** Supervision, Writing - Review & Editing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing

This paper is based on the master thesis of Renée Venema, who shared first authorship. The implementation and initial analysis was done by Renée, which I followed up with another iteration of thematic analysis and writing of the chapter based on the thesis she wrote.

Figure 10.1: Participants took part in the ISE and focus group in groups of four.

(a) Virtual recipe button (b) Smoothie recipe options (c) Smoothie recipes (d) Fruit baskets with a virtual arrow indicating which fruit to grab

Figure 10.2: The different steps of the recipe task, which is one of the four activities participants could engage in as part of the ISE.

abstract In this chapter we investigate the impact of near-future Mixed Reality (MR) contact lenses on users' everyday lives via an Immersive Speculative Enactment (ISE) and focus groups. If or when MR technology advances to the same level of ubiquitousness of current smartphones, this is likely to have a large impact on people's everyday lives. To gain qualitative insight on this impact, we created an ISE in which participants could experience a simulated MR lens prototype together in groups of four, thereby expanding the ISE method to multiple participants for the first time. This was followed by a focus group, in which the impact of the MR lenses was discussed. Participants raised concerns about the future of social interactions and expressing agency over the device, while also recognising how it could have practical applications. Based on these findings we formulate three guidelines for future MR contact lenses.

10.1 Introduction

Arguably the next computing paradigm, in this chapter we investigate the impact of near-future Mixed Reality (MR) contact lenses on users' everyday lives through Immersive Speculative Enactments (ISEs) [247] and focus groups. ISEs are a form of speculative design where participants are immersed in a fictional scenario in an open-ended virtual environment (VE), that would be difficult to recreate in real-life (in this case, due to the proposed MR lens not being feasible with current technology). The enactment thus serves as a means to study users' behaviours when confronted with the speculative probe, ideally providing insights for designers or practitioners to build on when the technology does becomes feasible. We use MR as an umbrella term [258], encompassing multiple points on the reality-virtuality continuum [174] such as Virtual and Augmented reality technologies. A future ideal MR technology could be realised in a variety of ways, such as small form-factor MR glasses, MR contact lenses, brain-computer interfaces, etc. We opted for contact lenses as the most suitable method for the speculative enactment, as small form-factor MR glasses do not provide a sufficient advancement from the current state of the art, while brain-computer interfaces capable of supporting MR interactions will require many theoretical and technological advancements beyond what is feasible in the foreseeable future. To implement the ISE, we created a multi-user VE where participants are presented with a prototype lens for which they could enable and disable its functionality without seeing the device on each other's avatars.

Currently, research on lenses focuses on how to engineer these with embedded chips to display a virtual layer on top of the real world [9, 119, 177]. The development of these lenses is driven by health concerns, for example, to help the visually impaired see better and live a more independent life [9]. But the concept has also broad commercial applications, and several companies are looking to promote lenses to augment and improve people's everyday life [119, 177]. If these lenses do become available to everyday users, it is still unclear the impact they will have on their lives and to which near-future challenges this will give rise. For example, ubiquitous everyday MR could lead to dark patterns such as harmful virtual-physical perceptual manipulations [277] or memory manipulations [22]. The contact lens form-factor presents additional opportunities and challenges, as it greatly facilitates integration with everyday life and is nearly imperceptible to others. However, other potential ramifications of the use of ubiquitous MR lenses remain unclear, which is why we chose to investigate the following two research questions:

1. How do potential users envision a future with ubiquitous MR contact lenses?
2. How desirable do potential users find a future with ubiquitous MR contact lenses?

As the MR lenses we envision currently do not exist, there is no empirical research on their impact on user's lives. To create a broad base for possible future research, the study we present encompasses multiple different aspects of MR lenses such as a pervasive heads-up display and spatially placed 3D content, to gather feedback on possible benefits and issues. In order to create our ISE, we made the following assumptions about the possible future it depicts [227]:

1. The lenses will be *context aware*, this way the lenses will be able to react to various objects and situations;
2. The users can interact with the lenses in *3D space* with hand movements, such as pressing a button mid-air;
3. The lenses are meant for *everyday use*, meaning users will wear them at all times.

We conducted our user study with four groups of four participants (N=16), which consisted of two phases. First, they collectively experienced the ISE, each exploring the different MR activities while in the same room. Second, we conducted a focus group, guided by how participants experienced the ISE and a list of possible applications for the device, in which they discussed their experience with the speculative prototype, how it would impact their lives, and whether they found the imagined future desirable.

Analysis of the focus groups revealed a consensus that the capabilities of the lenses should be limited to functional aspects, for example as a driving assistant. Participants also envisioned various disadvantages, such as isolation (i.e., fewer social interactions), being more susceptible to misinformation and advertisements, somehow conceding control of their lives to these lenses, losing touch with reality, privacy issues, and safety concerns – all of which we discuss in detail later in the chapter. We found that, in general, participants did not find the envisioned future desirable as thus we present guidelines for future MR lens development. In sum, our two contributions are: (1) an extension of the ISE method with multiple users and focus groups; and (2) speculative insights on near-future MR contact lenses.

10.2 Related Work

This section discusses relevant papers concerning both ISEs and MR devices. First, several speculative methods are described, concluding with a comparison between these methods and ISEs. Second, the idea of MR contact lenses is further developed, defining the terminology of Augmented Reality (AR), Virtual Reality (VR), MR and smart glasses, explaining the current state of the art of MR devices. Third we discuss social acceptability of public MR, and last the ethical and privacy concerns that could arise.

10.2.1 Immersive Speculative Enactments

Speculative Design is a speculative method that reflects on possible futures and alternative presents to facilitate public discussion [10, 227]. By using prototypes, the constraints of commercial design are mitigated, and critical discussion on technology and society can be raised without technological or ethical limitations [54]. The purpose of *Speculative Design* is to step out of reality and attempt to investigate some imagined possibilities, similar to thought experiments. This imaginative, open aspect is what differentiates *Speculative Design* from mainstream, commercial design, which has a stronger focus on problem solving in the present [10].

User Enactments take place in an environment created by the designers. Therefore, they use all types of materials and props. One enactment [190] required researchers to recreate a meal planner application, which would be part of the kitchen. In these physical environments, participants will be asked to follow a loosely scripted scenario, while also having freedom to explore and improvise within the setting. Odom et al. claim that the constructed scene allows users to experience the imagined situation in a more realistic way. As the environment is not the main focus, the scene does not have to be worked out in detail. The most important aspect of a *User Enactment* is creating a shared reference to mediate dialogue between participants and designers [190].

Speculative Enactments on the other hand imitate the scenario in real-life environments and public spaces. Elsden et al. [65] claim that the lab setting of *User Enactments* (although disguised) risks breaking the suspension of disbelief. An aspect of *Speculative Enactments* that distinguishes it from other speculative methods is the emphasis on real-life consequences. The setting of public spaces allows consequences of a social and emotional sort to influence participants' behaviour, such as meeting new people or feeling awkward. The other characters in the speculation could either be hired actors or other participants. These consequences are expected to be conducive to more real experiences and genuine interactions and reflections of the imagined prototype and future [65].

Immersive Design Fiction (IDF) complements *Design Fiction* [25, 128], by creating a VR storyworld instead of solely using the participants' imagination to envision the views conveyed by the designers. As a result interactions in VR are possible with the prototype and context, allowing researchers to speculate about those interactions and the surrounding context rather than only investigating the interface of the prototype. The VR scenario contains both pre-scripted and interactive elements. However, they cannot alter the outcome of the story in VR. So *IDF* is more open than a linear story, but more constrained than *Speculative Enactments*. It also has less emphasis on real-life consequences, as the whole experiment takes place in VR, compared to real-life *Speculative Enactments* [166].

Immersive Speculative Enactments (ISEs) are a combination of the previously mentioned methods. They are described as an extension of *Speculative Enactments* into VR [247]. Conventional speculative enactments are limited to what is practically feasible with current technology. What sets ISEs apart from IDFs is that an ISE is set in a sandbox scenario that attempts to provide users with natural interaction possibilities. Rather than exposing users to a specific narrative, the ISE places users in an open-ended scenario and prompts them to act as they would, were the situation real. Further, VR supports the suspension of disbelief needed for users to envision the proposed prototype [65, 247], which can help if the possible future envisions a significant departure from our present.

10.2.2 Towards Mixed Reality Contact Lenses

The MR device that we study in this chapter is the result of speculation. It does not exist in the way it is portrayed here. However, building such a device is driving current efforts. On the one hand, glasses and contact lenses are becoming more capable of rendering MR content [9, 119, 177]. On the other hand, recent standalone MR devices are adopting smaller form factors, such as the Meta Quest 3, Pico Neo 4, Lynx R-1, or the Apple Vision Pro [5].

In 2013 Google released the Google Glass [83], which was not successful and was terminated only a few years later [162]. This failure was due to privacy concerns and the lack of acceptance from the general public. Williamson [129] attributes this on the lack of a mental model of the bystanders. They did not understand how the device worked, and thus they were not ready to accept the Google Glass and the people who wore them around them. Snap Inc, the company of Snapchat, distributes their own type of smart glasses, Spectacles, which evolved into an AR device in the latest version, Snap Spectacles 3 [254]. Lastly, XREAL Air Glasses look like regular sunglasses, but act as an AR headset. This means that outsiders have no way to differentiate between someone wearing sunglasses and someone wearing the AR glasses. The XREAL glasses are meant for recreational use, for example to watch movies or play games. They do not support interaction with the environment, however [188].

MR contact lenses represent another research direction. Prototypes by Mojo Vision [177] and Innovega [119] are implementing lenses with passive overlays, not meant for extensive interactions. However, at the time of writing these devices are not available to the public. Other potential applications of MR lenses are geared towards helping people with bad eyesight through intensifying specific lights [9]. The MR device envisioned in this chapter goes beyond all of these products. Essentially, it will act as traditional contact lenses and thus be invisible to others.

10.2.3 Social Acceptability & Practical Concerns of Public Mixed Reality

The social acceptability of head-mounted displays and public VR has been extensively researched [3, 98, 306, 129, 178, 220, 235]. However, it is not clear how well these findings extend to contact lenses. Social acceptability is a complex emotion, dependent on several factors, such as social conventions, context, individual preferences, and culture [76]. It takes into account both the user's and the observer's social acceptance [178]. Thus, whether the user felt comfortable, awkward or relaxed, which leaves them with a positive or negative impression of the acceptability. And whether the spectator watching the user, in different settings such as public and private, has a positive or negative impression of the acceptability. Taking into account both views on the experience produces the overall social acceptability of certain actions [178], which also varies depending on the situation [235]. The problems with social acceptability of current Head-Mounted Displays (HMDs) include unsubtle input modalities [3], and cross-reality challenges [8], such as real-world awareness, interruptions, and bystander inclusion [308].

It has been found that subtle and less noticeable input gestures result in increased acceptability, in order to avoid capturing unwanted attention [3] and infringing on nearby person's personal space [168, 308]. We expect the proposed MR lenses to be more subtle than an HMD, because they are nearly invisible to bystanders, and thus we could expect their use to be accepted more easily. However, it is also possible that there exists a point where the technology becomes too subtle, and it is no longer possible to discern between who is and who is not using it. This could then again have a negative effect on social acceptability. It is possible that mid-air gestural interactions become more familiar and normalised as they become more prevalent [168].

Another element hampering social acceptability are cross-reality challenges, such as real-world awareness, interruptions, and bystander inclusion [308]. Real-world awareness is especially a problem for VR users, who are visually disconnected from the real world, which can cause them to unknowingly impede on other person's private space or make them unable to react to real-world events [315]. Because of this it is difficult for a bystander to interrupt the VR user, as they do not know whether the VR user is aware of their presence [87]. Moreover, any type of HMD-based MR causes bystanders to be excluded from the interaction, unless additional devices display relevant virtual content to them [48, 99]. Future MR interfaces would therefore benefit from seamless transitions of users and content between realities [81, 45, 44], especially in the case of MR contact lenses where physically removing the device to transition back to and from reality is cumbersome and uncomfortable.

10.2.4 Privacy & Ethical Concerns

The introduction of MR devices into homes and public spaces leads to privacy and ethical concerns [101], for both users of the device and bystanders [223]. Input modalities supported by, e.g., microphones and cameras, can lead to private information being recorded from users. While output modalities supported by, e.g., speakers and displays, can provide users with unethical virtual stimuli looking to influence or cause them harm.

MR devices require always-on sensing to function, thereby constantly capturing privacy-sensitive information about the user and their environment, which malicious applications can capture and abuse [223]. Data that can be gathered from these sensors includes video and audio, but also derivatives, such as user height, movement and emotional state [1]. Therefore, existing permissions systems should be adjusted for use on future MR devices, to provide fine-grained permissions to only allow an application to access the data it requires [102, 1]. Privacy concerns extend towards bystanders, who are also at risk if they are in the vicinity of the MR user [192]. Similarly, virtual objects in shared environments require a form of access control to identify which MR users can manipulate them [155, 285]. Thus, it is important to form standards on privacy for MR devices [1, 165, 2].

In addition to privacy concerns, there are ethical concerns related to how virtual content can deceive users [155]. For example, in puppetry attacks the user is redirected into potentially dangerous areas, which could cause them to fall or break physical objects. Similarly, mismatching attacks exploit mismatch between virtual and physical environments, such as when the user expects a virtual chair to have a physical counterpart, not having one may cause them to try and sit in it and fall [277]. User avatars can be impersonated, where one user with malicious intent impersonates another user's avatar [1, 252]. Conversely, user's own appearance can be augmented to different degrees, which raises questions on to which degree it is ethical to augment oneself, and whether a user has the right to augment others [23]. Moreover, MR could be used to alter a user's memories in three ways: at encoding, at pre-retrieval, or at retrieval. For example, MR can alter the memory as it is being made, can remove reminders of the memory to cause forgetting, or display modified replays of the memory to change it after the fact [22].

10.3 The Immersive Speculative Enactment

We implemented the ISE using Unity version 2021.3 [282], with Photon Engine [205] for networking and Microsoft Rocketbox for user avatars [171]. The project ran as a standalone application deployed on Meta Quest devices, and was set to automatically connect to the Photon server at startup.

This ISE was developed based on previous ISE guidelines [247]. All objects in the VE had realistic physics and supported natural interactions, i.e., users were able to pick up all objects present in the VE that they could lift with one hand in a real-life situation. Fragile objects broke if they fell on the floor and water flowed from the faucet when it was turned on or from a glass when it was poured. Minimal environmental sounds were also added to the ISE. In addition, it was expected that participants would make sound while talking to each other, thus creating a believable experience. We also chose to allow users to user their hands to naturally interact with items instead of controllers. Hand tracking and interactions were provided by the XR Interaction Toolkit package [281] and required users to pinch their thumb and index finger to grab objects. The goal was to create a mundane, believable, and everyday experience.

The VE is developed to present a living room and kitchen of a small apartment as realistically as possible. The room has an area of $4\,\text{m} \times 4\,\text{m}$, the size of which corresponds with the size of the physical room the user study took place in. This allows for natural walking without any need for locomotion techniques and avoids collision with physical objects or walls. A sofa, a cabinet, some plants, and a little kitchen are present in the apartment. A training environment was developed with the purpose of getting participants accustomed to VR and its interaction techniques beforehand. It is important to note that the VE exists in two layers of virtuality, one representing the real world and one representing the MR layer and objects only visible through the lenses. For simplicity, in the rest of this section the first layer of virtuality (representing the real objects in VR) will be indicated as *real*. The second layer, coming from the MR lenses, will be referred to as *virtual*. Both these *real* and *virtual* objects are thus part of the VE of the ISE.

Every participant will have a small permanent overlay on the top left of their field of view, which states the current date and time. This overlay is passive and can not be interacted with, but gives constant information to the user. We chose to include a pervasive overlay to convey the always-on character of the MR lenses. The ISE consisted of four non-mandatory tasks that could be completed, and were designed to represent the following five aspects of MR in different capacities: diminishing reality, immersive advertisements, the lens being always-on, functionally assisting a real world task, and virtual decoration.

1. Preview a sofa purchase in-situ We displayed an advertisement for a new sofa floating above the existing sofa in the living room, which can be closed by pressing the *X*-button (Figure 10.3a). However, pressing the *browse*-button will lead to a new panel, presenting four different sofas (Figure 10.3b), all of which can be chosen to be previewed in the apartment. This virtual sofa then replaces the real sofa (Figure 10.3c)

(a) Sofa scene with an advertisement

(b) Virtual sofa shop

(c) Previewing a virtual sofa

Figure 10.3: The different steps while previewing a potential new sofa in-situ.

and can be grabbed and placed anywhere in the room. To show the previewed sofa was not real, it is slightly transparent and a pop up announces to be mindful of the real sofa since it is still present in the apartment, but not visible when trying new sofas. We made this choice because it would be difficult to imagine the new sofa in the apartment, if the real one is visible through the previewed one. This task is meant to show the practical advantages of the lenses, but also allows for some critical thinking on advertisements, and the disparity of virtual and real objects.

2. Decorate a room In another corner of the room the lenses will show a button to add virtual decorations in the room (Figure 10.4a), when clicked on, a panel with different choices of decorations appears (Figure 10.4b). An object can be selected by pressing it with the index finger, and will then appear on the cabinet (Figure 10.4c). This object can then be grabbed, placed wherever the user wants it to be, and it can be scaled by pinching the object with both hands and consequently moving the hands closer together or further apart. Decorations can be deleted as well by pressing the *'Remove items'*-button at the bottom of the decorations panel. This highlights all removable virtual decorations in red (Figure 10.4d). When a decoration element is selected with the pointer finger in this mode, it will be deleted. This task lets participants reflect critically on the idea of personalised environments, the advantages of always being able to change your surroundings, and the effects of privacy leading each user's creation to remain individual and not visible by others.

3. Follow a recipe In the apartment, a kitchen is present with a faucet, a blender and different types of fruit. The lenses display a button above the blender (Figure 10.2a), which, when pressed, shows five smoothies that can be made (Figure 10.2b). Pressing one of the options results in a panel showing the steps to make the smoothie, highlighting the current step, based on the ingredients already in the blender (Figure 10.2c). On top of that, an arrow points to the ingredient needed to complete the recipe (Figure 10.2d), and warns when the last of an ingredient is used, prompting users to decide whether or

(a) Virtual decoration button (b) Virtual decorations panel (c) Placing virtual decorations (d) Removing decorations

Figure 10.4: The different steps of the decoration task.

not they would like to add the fruit to their shopping list. Participants will experience how context aware the MR lenses can possibly be and the advantage of the assistance they could provide in their daily life, as well as the disadvantages of always being watched and followed by the lenses. To create an immersive experience, participants poured water from the tap into a glass and subsequently into the blender. Following ISE guidelines [247], we implemented glasses to break when dropped. This, as well as the blender activating, were accompanied by sound effects to maintain the expectations of being in a realistic environment.

4. Turn the MR lenses on and off For the last task participants had to turn the lenses on and off, via a wrist-based button. Turning the lenses off disables the virtual layer, and all the objects associated with it. This way participants can make a clear distinction between real and virtual objects, as most of them have the same graphical quality in the ISE, making it difficult to distinguish between real and virtual if the lenses are constantly turned on.

10.4 User Study

The aim of the user study is to determine the general public's attitude towards the use of MR lenses in everyday life. To this end, we ran a focus group where multiple participants get to experience a glimpse of a life where those lenses are possible, made possible via an ISE. Successively, they discussed and speculated about the possible future with this device in a focus group. This group discussion was recorded and analysed with the goal of gauging people's attitudes and feedback on the potential use of MR lenses in the real world, and if or when they will become possible. The study lasted between 100 and 120 minutes, and was approved by the university's ethics review board.

Table 10.1: Participant demographics (age, MR experience, gender composition, and group description) across focus groups.

		min	max	mean	sd	M/F
Everyone	age	19	57	30.6	14.5	8/8
	experience	1	5	2.31	0.946	
Group 1	age	53	57	54.8	1.71	2/2
	experience	2	3	2.25	0.5	
Group 2	age	23	25	23.5	1	1/3
	experience	1	3	1.75	0.957	
Group 3	age	19	22	20.2	1.5	4/0
	experience	2	5	3	1.41	
Group 4	age	22	25	23.8	1.26	1/3
	experience	2	3	2.25	0.5	

10.4.1 Participants

We recruited four groups of four participants (N=16), of which eight identified as female and eight as male (Table 10.1). They were all European and aged between 19 and 57 years ($M = 30.6, SD = 14.5$). On a 7-point scale, where a score of 1 was equal to *no experience, never used it before* and 7 represented *everyday use*, participants reported low previous experience with VR and AR ($M = 2.31, SD = 0.95$). For the experiment four HMDs were needed at the same time, for which we used a variety of Meta Quest devices due to varying availability in our lab. The first focus group was conducted with a Meta Quest Pro, two Meta Quest 2, and a first generation Meta Quest, while for the latter three focus groups we used a Meta Quest Pro, and three Meta Quest 2.

Participants indicated heavy phone-use, having it within reach always (5) or often (10). Only one participant indicated they did not use their phone that often. It was mainly used for contact and communication, including text-messaging apps, such as Whatsapp and Messenger, or more general uses, such as *'for mails'*. Internet search was mostly used for news and translation, while practical applications included payment and authenticator apps. From the 16 participants, only one stated to have used AR before once, and eight of them replied to have at least had one experience with VR using a HMD. Most of these eight VR users added that their experience was limited, as only two had used a VR HMD more than twice before and none of them used it more than five times. However, six participants mentioned to have used a smartphone in combination with specific VR headset or cardboard accessory to experience 3-DOF VR, and four participants told that they had been in a theme park in a roller coaster while wearing a HMD to experience a different environment. Five participants brought up they had used AR on their smartphone, for example with Pokemon Go, Snapchat filters, or Google Maps AR.

10.4.2 Procedure

Prior to taking part in the study, participants were provided with a short explanation of what the lenses are, the concept of ISEs, and a list of 30 possible purposes of MR lenses. Firstly, we felt it was important that all participants were up to date with terminology such as VR, AR, MR and smart glasses. We then proceeded to explain them the concept of MR contact lenses, and what they are capable of in the possible future we envisioned (as outlined in section 10.1). Secondly, the next step was to clarify the concept of ISEs, based on ideas and examples from Simeone et al. [247]. The immersive aspect of the VR experience was emphasised, along with importance of behaving in the virtual world as participants would in the real world and to try and see the virtual room as if it was a real room. It was pointed out that everything should behave naturally and that the participants were allowed to freely explore the apartment. Emphasis was put on the idea of two virtual layers within the ISE, one representing reality and one representing virtuality through the lenses. Thirdly, to give participants the right mindset and to get them thinking about the MR lenses a list of possible purposes for the lenses was provided to them. The list consisted of several images from other papers or websites encountered by the authors while researching the topic of MR contact lenses and ubiquitous MR in preparation for this work.[1] We went over the list of applications with the group and explained every image individually, making sure they were presented in a neutral way. The list contained examples from the following application types: navigation & direction, personalising environments, buying products & interactive advertising, enhancing social media, educational, instruction task overlays, permanent overlays, work-related, architecture & construction, medical, and home automation. Participants were free to discuss other applications, which form they would take, and ultimately whether their effect would be a positive or negative one.

The participants were told to ask practical questions right away, but leave ethical concerns or problems they envisioned for later. After the introduction was completed, participants were immersed in a virtual tutorial environment in which they could learn the basic VR interactions needed for the ISE. Then participants where immersed in the ISE together, and received a maximum of 15 minutes to walk around and explore (Figure 10.1). When the ISE was finished, participants filled in a demographics and a Slater-Usoh-Steed presence questionnaire [283], after which the focus group started.

[1] The full list can be found as supplementary material to the paper that corresponds to this chapter.

10.4.3 Focus Group

This research primarily focuses on qualitative data obtained from the focus group and observations made during the ISE. The researcher served as a facilitator for the focus group, asking questions, activating disengaged participants, and facilitating discussion among the participants. The questions asked during the focus group were classified into three categories: introductory questions, key questions, and closing questions. The key questions included 'Was there anything surprising or unexpected about the MR lenses?', 'What do you think of the possibilities of MR lenses?', and 'Can you see yourself or others ever using the MR lenses and what for?' The full list of questions in the following:

Introductory Questions

1. Please introduce yourself. Tell us your name and what your previous experience with VR and AR is.
2. How often do you use your phone and what for?
    - How often is your phone within arms reach?
3. Are you familiar with AR, VR and smart glasses?
    - Could you explain the difference?
4. Were you immersed in the environment?
    - What could have immersed you more?
    - Was there something missing?

Key Questions

1. What was the hardest part about using MR lenses?
2. Was there anything surprising or unexpected about the MR lenses?
3. What do you think of the possibilities of MR lenses?
    - What aspects did you or did you not like?
    - Which concerns did you have?
4. Can you see yourself or others ever using the MR lenses and what for?
    - What might keep people from using MR lenses?
5. Do you think MR lenses will exist in the future?
    - What aspects will or will not exist?
    - What could be done to improve the idea of MR lenses?
    - What features would you add?

Closing Questions

1. Does anyone have anything else they would like to add or want to discuss?

2. Summarise what has been said. Does everyone agree with this summary?

10.5 Results

The conversations from the focus groups were recorded, resulting in four recordings between 40 and 55 minutes (46 min on average). The recordings were then transcribed using a speech-to-text tool, and any errors in the transcript manually corrected, resulting in four transcripts between 7 401 and 13 339 words (10 308 words on average). We then followed an approach based on an inductive thematic analysis [28] to analyse the data. One author conducted an initial analysis, familiarising themselves with the data, developing codes, and themes. This author was a last-year Computer Science student specialised in Human-Computer Interaction and VR. Two iterations of theme development resulted in 63 codes and the following eight themes: *Encouraging (anti-)social behaviour*, *Maintaining agency*, *losing connection to reality*, *security of sensor data*, *safety*, *social acceptance*, *functional and recreational usage*, and *future existence*. The themes were then further developed by a second author, who is a fourth-year PhD researcher. Another round of thematic analysis based on the initial coding (familiarising, coding, and refining themes) added 12 additional codes, and resulted in the following four final themes:

- *Privacy, Security, and Perceptual Agency*
- *Social Acceptability of Increased Virtuality*
- *Excessive use Taking Away From Real Life*
- *Future Existence for Practical Application*

We noticed that participants were primed by the list of applications. They often referred back to immersive advertisements, remote meetings, public AR social media, public place (and specifically bar) decoration, assembly instructions, and driving assists; while other examples arose organically from the discussion. Additionally, participants rated their presence with a mean of 4.4 ($SD = 1.8$) on a 7-point likert scale, and an average 'count of 6 and 7 answers' [283] score of 1.5 ($SD = 1.3$) out of the six questions.

10.5.1 Privacy, Security, and Perceptual Agency

As participants realised the amounts of data that would need to be gathered to enable the lenses' functionality, they raised privacy and security issues. They believed it to be necessary that it was clearly communicated what kind of data would be used and how, similar to current smartphone apps. They found it dangerous for a limited number of companies to have access to all user data, as this would give them too much power. It was difficult for them to trust companies that they would not misuse their data, and that they would communicate clearly on the purposes of gathering the data. Several participants also indicated to be afraid of getting hacked as well, where all their information could be misused or made public by a malicious actor. Regarding privacy and security, all participants agreed that there should be regulations established by a government authority, as an extension of current regulation to MR contact lenses. In addition to privacy and security, participants also brought up the novel issue of perceptual agency.

With perceptual agency, we refer to agency of a person over what they are seeing and how they are seen by others, which participants identified as a challenge in addition to privacy and security. On the one hand, participants wanted to maintain agency over what the lens shows them, as indicated by participant 9 '*Besides, then your lens decides for you which information you see. That's really the scariest thing ever, in my opinion, that you don't have the ability to decide that yourself.*' All participants found immersive advertisements to be an extreme drawback, and agreed this should never happen. They recognised the power of advertisements on people's behaviour and one participant compared it to current social media, stating it should not be allowed to influence people. Participants also saw issues with fake information being spread and being subject to the information feed provided by the lenses. Participants imagined a type of lens that would simplify the user's life, by providing information on specific objects or situations without having to specifically request it. To them, there is an important difference between actively looking up information or it being provided it without asking. Participants also wanted to configure the information shown on the lens to their preference, to avoid getting overwhelmed by all the information that is displayed. All groups mentioned the dangers of the virtual objects resembling real ones too much, as people would not be able to distinguish between the two, which could lead to safety issues. Participant 1 said: '*Because if it becomes less evident over time to distinguish between what is real and what is virtual, then I think that poses a risk for people.*'

On the other hand, participants indicated they wanted agency over how they were seen by others, and that it was not ethical to change another person's appearance without their permission. One group initially suggested they could make everybody look like monkeys, but then quickly realised the seriousness of this issue, as people with malicious intent could make real people look less human, thus lowering the bar to harm them.

The issue of perceptual agency requires solutions not found in current devices such as smartphones, due to the always-on nature of the contact lens and the additional complexity of taking it out of the eyes. Participants stressed the importance of some type of indication that shows whether or not the lenses are actively showing an overlay. This could be an outline around the whole field of view or some way to highlight every virtual object separately. Additionally, as indicated by participant 4, '*You should always be able to turn it off, so that you have control over it.*' Three of the groups doubted they would trust the lenses to be really turned off, and indicated they would rather wear them only when they used them. Participant 12: '*You can turn it off, but you still have the lenses on. You can never be completely sure that it's off.*' These issues highlight the differences between the MR lenses compared to MR glasses, which participants indicated would solve this problem, as they could take them off more easily. For participant 10 it was important to '*be able to return to reality practically immediately.*'

10.5.2 Social Acceptability of Increased Virtuality

Participants brought up issues with social acceptability related to input methods, that are already identified for current HMDs. Participants could use their hands to touch and grab virtual objects, resembling real-life actions, resulting in extensive hand movements. They indicated that during the study they were *'wondering what they looked like'*, while interacting with the virtual objects, and stated that they *'probably looked pretty stupid'*. While speculating on a future including MR lenses, participants started wondering on the social acceptability of interacting with invisible objects. They found that initially it would be very strange, but that user acceptance would grow if the lens would become more prevalent. As summarised by participant 12: '*I think initially it will be very socially unacceptable, but eventually it will be fine.*' However, the lenses would also benefit from an additional interaction technique, other than having to explicitly point to virtual objects [307], for example when a user's *'hands are full'* and they want to quickly interact with the lenses. Suggested alternatives included voice recognition, a small object, like a ring, close to the hands that could control the lenses, or gloves to manipulate the virtual layer. In general, participants suggested that more subtle interaction techniques would help with social acceptability.

On the one hand, we noticed a trend with participants being more negative about the lenses as applications introduced more virtuality, leading to reduced social acceptability. For example, participants were afraid the lenses would replace human interaction, or people would be so immersed in the virtual world, they would not be approachable. The lenses were compared to noise cancelling headphones and how participants found it hard to approach someone who is wearing them, and the resulting interactions feeling awkward, as it was assumed that person would be rather left alone. They would imagine in this future that users of the lenses would seem inapproachable as well, immersed in their own world and not open for conversation with a real human. On top of this, it is not visible to see whether or not someone is mentally present in the real world or immersed in a virtual world.

On the other hand, participants also identified positive effects on social interactions. For example, the application of holograms facilitating remote meetings. They did not think that these interactions could replace normal human interaction, stating that it would replace phone calls to stay in contact with family or friends in another country more frequently. Participants could see the benefits of displaying people's interests publicly, with the aim of connecting like-minded people through the lenses, when meeting someone in public. However, this could cause more antisocial behaviour, as people would start filtering who they want to speak to, before they have even met them. As a result, this did not seem to be a solution for the problem of diminishing human interactions, and one group even agreed that it should be prohibited to add this feature to the lenses to show this information publicly.

When referring back to the social media example, three of the four groups found that the lenses could further amplify the stimuli received from it. They referred to the idea that notifications from social media applications would pop into view and would be constantly visible, in contrast to a smartphone, where they are only visible when looking at or using the phone. Despite it would be fair to assume that the invasiveness of such notifications would be customisable, this idea was something they did not look forward to. As participant 11 said '*people might become more antisocial because they have enough stimulation from those lenses rather than from other people.*' However, participants recognised that the problem was not only with the MR lenses, but also with the idea of social media on itself, stating that current social media do not at all help people be more social. Participant 13: '*I think that social media, in general, makes people more antisocial because then they are more focused on that.*' The lenses would provide a more extensive view on social media, where it could be used in public to instantly befriend people you meet. Although in this manner, the lenses provided more tools to be social and meet new people, participants thought this would backfire, and make people more antisocial.

Every group discussed how they were scared that people would start living in their own virtual world, created by the lenses. Some stated they would engage more with features and content related to the lenses, and thus be less socially present. As indicated by participant 15 '*[I think] that they might lead to isolation if everyone is more immersed in their own little world.*' One group indicated that 'sharing' the virtual world of the lenses would become a necessary feature. This would allow users to invite friends in their virtual world, where they would be able to see the same virtual content. This positive twist on the individualistic bubbles, could, according to them, create more social experiences. To encourage this social behaviour, this should be able to happen spontaneously, as it causes too much friction to find someone's VE, it will create a barrier for people, resulting in reduced social behaviour.

10.5.3 Excessive use Taking Away From Real Life

Participants could imagine the possible benefits of permanently enabling certain augmentations, especially by enhancing their environment, stating they would enjoy their commute to work more if it could be surrounded by more nature through the lenses. Similarly, it was suggested that the lenses could be used to go on a virtual trip, for example for those who were not able to get there. However, in the same discussion it was suggested that this virtual nature would never be able to replace the experience of the real outdoors. This caused a discussion on the merits of such applications. It was stated that socially disadvantaged people could replace their real environments with nicer virtual environments. There were two counterarguments, firstly a financial issue, where those people would not have the means to purchase the devices, and secondly that this simply tries to conceal the problem, rather than solve it.

A danger that was recognised in replacing the reality by a virtual world, was in the idea of escapism. Participants feared people getting stuck in the virtual environment, as it could be more appealing to them than reality. Parallels were mentioned with gamers escaping in game worlds and how this virtual world would immerse them even more, and the elderly stuck in a nursing home, who could experience the outside using the lenses. The question was raised whether or not it was their own task to handle it responsibly or that it should be monitored from higher up. At the end of the discussion most participants agreed that it would not be a desirable functionality of the lenses to transport people to a whole new virtual world.

Similarly, participants opposed the idea of decorating the physical environment. An example that was often referred back to, was that people in a bar could choose how the bar would be decorated, by choosing a style or putting up some objects themselves. This topic encountered a lot of opposition as well, stating it would lead to more individualism and takes away the charm of life. Participant 15 found that '*If everyone experiences something different, you also lose a bit of the human aspect of shared experiences.*'

Finally, they thought the lenses could make life too easy, which could result in people becoming too reliant on the lenses and experiencing difficulties when living without them. There was some discussion about whether excessive use of the lenses should be regulated. Whereas some participants thought it was adequate to prohibit certain applications from being deployed, to protect people from becoming addicted or too dependent, others believed that it was up to each individual person to take on this responsibility, citing government overreach.

10.5.4 Future Existence for Practical Application

Almost all participants agreed they would not use the lenses for recreational purposes, and limit their use to more practical applications, such as work-related to communicate easier, for displaying instructions to assemble things easier and as driving assistant by showing the route and possible dangers. One of them was even very enthusiastic about using them during a city trip to show information about historical buildings, as they would not have to carry around a city guide anymore. There was clear gap between the opinions of this first group and the other three, where five out of the eleven participants from groups 2-4 indicated they would not want to use it at all. The reason being that they saw no benefits in the lenses, and thus the disadvantages, such as being more isolated, more easily influenced, and more susceptible to data misuse, outweighed the advantages for them. Participant 8: '*I wouldn't even do it. You know what I mean, if it exists, I wouldn't be inclined to buy it.*'

Some of them were opposed to the idea of ever using it, but they indicated that eventually, when the lenses would be integrated into society, they would have no choice but to use them as well. They assumed society would indirectly obligate the use of the lenses, as life would become more difficult without them similar to what happens with smartphones right now. They preferred a minimalist use of the lenses, but feared that eventually, due to peer pressure, they may start using the lenses for recreational purposes as well. One of the participants opposing the lenses, said they would prefer to occasionally use an HMD or smart glasses for the practical benefits instead. Five other participants stated they would want to use it, but in a very limited way. They all referred back to the example of using it as a driving assistant, showing the route to take and alerting the user for unexpected situations, other benefits were extremely limited, and none of them stated to want to use it for recreational purposes.

To conclude the key questions of the focus group, participants were asked whether or not they thought these lenses would exist in the future. All participants answered this question positively, stating they were a *'100% sure'* and that it would be *'inevitable that it would be used'*. Some participants could only envision a future with a limited form of the lenses, with a small overlay and less information than the idea they were presented with. Others were convinced it would start with this simpler form, but that it would quickly enough evolve to a device with more and more functionality, eventually resulting in a device that would resemble the overarching idea of the lenses they speculated about. Few participants stated they would expect some push back from people, especially concerning their privacy and safety.

10.6 Discussion

Our focus groups revealed potential issues with future MR contact lenses with regard to perceptual agency and social interaction (RQ1). Participants found mostly practical applications for the MR lenses desirable, while finding situations in which the lenses replaced reality to be less desirable (RQ2). Based on these findings we highlight key takeaways, as well as limitations to our approach.

10.6.1 Perceptual Agency as new Issue

Participants recognised already identified issues concerning privacy of users surrounded by MR devices that collect information about them through their sensors [223], and agreed that third parties such as governments, companies, or even individuals such as hackers, should not be able to access this information. Moreover, the ubiquitous nature of MR contact lenses gives rise to new issues with perceptual agency, which participants brought up concerning how their view of the world was augmented, as well as how they were augmented by others [191].

The contact lens form-factor led to additional issues with agency, not identified for smart glasses or current HMDs, mainly caused by the lenses being more difficult to remove and harder to see when worn by someone else. Conversely to glasses, lenses cannot be taken out as easily. Participants indicated the need for a reliable way of turning the lenses off without taking them out of their eyes, out of fear that their view would be augmented without them knowing. As it is hard to see if another person is wearing the lenses, and whether they are using them at that time, participants identified that an external visual indication of the lenses state would be required. It is difficult to extend existing HMD solutions, such as attaching screens [100] or LEDs [88], because of the contact lens form-factor. Hence we identified the issue of **state visualisation**, to visualise the current state in which the lenses are to their user and external persons, as being more prevalent for the lens form-factor. This could be achieved on the lens itself, e.g., by affecting the wearer's eye colour, or via the other users' own MR devices (e.g., via an augmented overlay).

10.6.2 Social Interaction and Acceptability

Our results show issues with social acceptability that were already identified [308], such as making large hand gestures or lowered awareness of the real world. Additionally, MR contact lenses possibly pose new challenges related to social interaction that can hamper acceptability. For example, participants noted that the MR lenses provide a more isolated experience, and that it is not possible to physically share the display as is possible with smartphones. Hence, for the MR lens to become part of social interactions, such as smartphones now [112], intuitive sharing of content with other nearby users (or even bystanders without MR lenses) should be explicitly supported [100]. Furthermore, as the lenses provide the user with the ability to immerse themselves in a completely virtual world, it is not apparent to bystanders whether a user of the lenses has real-world situational awareness or can be interrupted. Participants found that users of the lenses required an external indication to communicate to which degree of virtuality they are immersed and whether they can be interrupted or not. As we recommend above, this can be achieved on the lenses themselves or via other users' own MR devices.

A vision of the lenses supporting social interaction rather than hampering it was met with increased acceptability. Participants found it should augment social interaction rather than replace it, for example by supporting remote communication, or by publicly displaying information about a person as a conversation starter [221]. We expect that **sharing virtual content** with co-located users will be a pivotal issue with MR lenses, as, conversely to smartglasses or current HMDs, the lenses cannot easily be passed on to another person to show them virtual content or experiences.

10.6.3 Towards a Desirable Future

Participants raised concern about excessive use of the MR lenses taking away from real life, expecting that the lenses would be used to escape into virtual worlds and that everyone's experiences would become individual. While similar concerns were raised around the turn of the millennium [24], the advent of the internet and ubiquitous smartphones has not replaced real life but rather complemented it. For example, people will publicly use their smartphone while waiting but disengage from it as friends arrive [112], online friendships are valuable [64], and online classes can replace most but not all of a student's basic learning needs [310]. Thus, the lenses would provide convenient access to many applications for communication, instruction and navigation, which would support people with tasks in their daily lives.

We believe future users of the contact lenses should be **supported in using them responsibly**, by providing them with a high degree of personalisation, as suggested by participants. Similarly to smartphones, the technology is not inherently good or bad, but can encourage undesirable behaviour. For example, the increase in screen time caused by the current prevalence of technology is known to cause more sedentary behaviour, which is unhealthy [186]. Thus users can be encouraged to reduce screen time [324, 214]. Current smartphone solutions, such as reducing notifications have been shown to be effective [196]. However, most solutions for reducing smartphone screen time rely on physically removing the smartphone from the user's vicinity [196], which may not be feasible for the lenses. For the MR contact lenses, applications should be developed to help users keep track of how they are dividing their time between the virtual and the real world, as well as configure the virtual content to their personal preference.

Guidelines for Future MR Contact Lenses

1. Reliably communicate the state of the MR contact lenses to the wearer and others around them.
2. Facilitate social interactions, rather than replace them, through intuitive sharing of virtual content.
3. Support future users in using the MR contact lenses responsibly, through time monitoring and personalisation.

10.6.4 Limitations

Since we did not want participants to focus on the social concerns around sensing, fashion, visibility, etc., we chose a (arguably impossible) MR display concept that would invite participants to only consider how it is ultimately used, excluding concerns about how others perceive the wearers of said devices. As such, the ISE does not take into account the hardware challenges, such as large enough field of view, responsiveness of the interface, comfort over long duration that are required for adoption. Additionally, during the study participants used current VR devices, which could have influenced their perception on how the MR lenses would work. However, we see this as a positive since familiarity with the state of the art helped participants in envisioning the future MR lenses. As we chose to focus on prospective end users, and how they expect the device to influence their daily lives, we did not incorporate experts into our focus groups, which could have given us a different perspective.

Participants indicated that the avatars caused the experience to be somewhat unrealistic, as noted by participant 2 '*In my case, I found that the avatars most compromised the realistic experience.*' We identified three factors that contributed towards the avatars causing a break in immersion. First, there were no facial animations as this was not supported by all hardware. Second, when the participant crouched to reach a lower area of the VE, their avatar would sink through the floor rather than display the proper crouching pose. And third, when participants' hands were no longer tracked their hand location would revert to a default position.

In terms of photorealism we were limited by the capabilities of the Meta Quest hardware, as we deemed managing desktop computers connected to HMDs for the four participants simultaneously as too impractical. As such, all groups stated the ISE could benefit from a more realistic environment, stating that the space was too sterile, the textures too flat, the images not the highest quality or that they saw the pixels of the room and its objects. However, these problems were mentioned only briefly in each group.

10.7 Conclusion

The objective of this chapter was to speculate on a possible future in which MR contact lenses are as ubiquitous as smartphones are today. We conducted a user study using an ISE, a speculative method to reflect about possible futures in VR. Multiple participants were immersed in the same ISE to provide a social aspect in researching the possible benefits or disadvantages of the MR lenses, and would allow a focus group to be done afterwards. The main purpose of the focus group was to retrieve opinions on the concept of MR lenses. There seemed to be a consensus that the functionalities of the lenses should be limited to practical aspects, as no participant mentioned recreational advantages or indicated that they themselves would want to use it. Most participants indicated they did see the added value of the lenses, for assisting them in daily tasks. However, there seemed to be a general feeling that when participants had to choose between the lenses with all their aspects, practical and recreational, or no lenses at all, a significant part of the participants preferred the lenses to not be developed at all. This is due to the many disadvantages participants envision for this future, the most important of which are isolation, reduced social interactions, being more susceptible to misinformation and advertisements, feeling out of control of the lenses, losing touch with reality, privacy issues and safety concerns.

Based on the results, we discussed the issues of perceptual agency, with 'state visualisation' being a perceptual agency issue especially relevant for the contact lens form-factor; we discussed how the lenses impact on social interaction influences their acceptability, with intuitive content sharing being one of the main issues; and we discussed how in a desirable future the lenses would complement rather than replace real life, by supporting users in using them responsibly. Based on this discussion we then formulated three guidelines for future MR contact lenses.

Chapter 11

From Display to Interaction: Design Patterns for Cross-Reality Systems

[47] **COOLS, R.**, HAN, J., ESTEVES, A., AND SIMEONE, A. L. From Display to Interaction: Design Patterns for Cross-Reality Systems. *IEEE Transactions on Visualization and Computer Graphics* (2025).

Robbe Cools: Conceptualization, Methodology, Visualisation, Writing **Jihae Han:** Conceptualization, Writing - Review & Editing **Augusto Esteves:** Conceptualization, Writing - Review & Editing **Adalberto L. Simeone:** Supervision, Writing - Review & Editing

This chapter was developed over the latter part of the PhD, and went through multiple iterations of possible ideas for theoretical contributions to the field of CR, before settling on design patterns.

https://youtu.be/FHstaOlQh_U

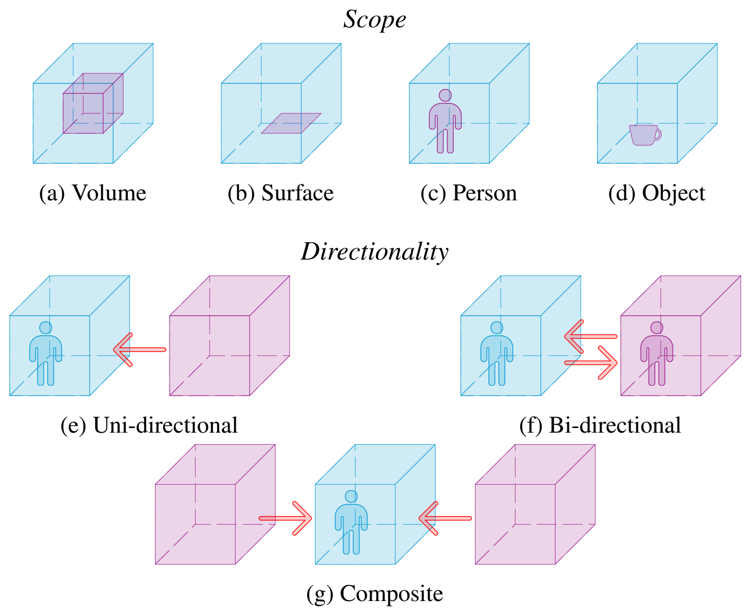


Figure 11.1: Diagrams for the first category of our CR design patterns: **fundamental** patterns, consisting of scope (volume, surface, person, or object) and directionality (uni-directional, bi-directional, and composite). Other categories include origin, display, and interaction design patterns.

abstract Cross-reality is an emerging research area concerned with systems operating across different points on the reality-virtuality continuum. These systems are often complex, involving multiple realities and users, and thus there is a need for an overarching design framework, which, despite growing interest has yet to be developed. This chapter addresses this need by presenting eleven design patterns for cross-reality applications across the following four categories: fundamental, origin, display, and interaction patterns. To identify these design patterns we analysed a corpus of 60 papers, with the goal of identifying recurring solutions. These patterns were then described in form of intent, the solution, and application examples, accompanied by a diagram and archetypal example. This chapter provides designers with a comprehensive set of patterns that they can use and draw inspiration from when creating cross-reality systems.

11.1 Introduction

Cross-Reality (CR) is an emerging field of research concerned with systems operating across different points on the reality-virtuality (RV) continuum [173, 251, 244]. It envisions a future where the boundaries between the virtual and the physical become increasingly seamless, until platforms are no longer device-dependent and transitions between the different points of the reality-virtuality continuum are indistinct from experiencing a single, unified reality. Hence 'reality' no longer refers solely to our physical world, but also to distinct experiences that virtually stimulate the senses. Current technological trends have begun to reflect this CR future, where head-mounted displays facilitate both Virtual Reality (VR) and Augmented Reality (AR) applications. Importantly, CR addresses the research challenge of technological isolation, where users in physical, augmented, or virtual realities cannot communicate with each other, creating safety hazards such as unaware VR users colliding with non-immersed bystanders [315] or social tensions where the act of wearing a headset in public elicits discomfort in spectators because they do not know what the user is seeing or doing [100]. Moreover, CR also explores how accessing multiple realities provides benefits in asymmetric collaboration, e.g., between AR and VR users [48] or how these CR transitions should unfold to support task performance and seamless interaction [44, 81].

Despite the growing body of CR literature and applications, there has yet to be a design framework that encapsulates the knowledge from tackling these CR challenges, from bystander inclusion to speculating new CR futures. A design framework is essential in providing a structured approach to identify the problem and ideate potential solutions, leading to tailored and effective CR applications. In the context of CR, our research objective is to aid designers of CR systems by providing a list of possible solutions that they can draw inspiration from and cater to the specific problems present in the CR systems being developed. Therefore, we propose eleven 'Cross-Reality Design Patterns' as a first step towards a language for designing and prototyping CR applications that: facilitates communication about CR system designs amongst team members; and helps develop CR systems and identify their trade-offs.

In this work we analysed a sample of 60 CR papers towards eleven CR design patterns that describe how realities in CR systems are connected. Each design pattern presents an intent and description of its solution, as well as application examples drawn from current research. We divided these design patterns into four categories according to the overarching challenges they tackle: (1) Fundamental (*scope* and *directionality*), (2) Origin (*live copy*, *replay*, and *notification*), (3) Display (*miniature*, *windows*, *transformed environments*, *combined visual modalities*), and (4) Interaction (*inclusion of physical objects* and *different interaction modalities*).

11.2 Related Work

In this section we introduce Mixed Reality (MR) and CR, successively we highlight the use of frameworks and design patterns in Human-Computer Interaction (HCI) and MR research.

11.2.1 Mixed Reality & Cross-Reality

We use MR as an umbrella term encompassing all points on the RV continuum. However, there is no consensus on what MR entails [257], but the umbrella term characterisation is most commonly adopted in research and aligns with its original meaning [173, 251]. Other frameworks have described the dimensions of MR, such as number of environments, number of users, level of immersion, level of virtuality, degree of interaction, extend of world knowledge, and coherence [257, 251]. Conversely our framework describes pattern in the connection between multiple realities in CR systems.

CR, as conceptualised by Simeone et al., is 'the transition between or concurrent usage of multiple systems on the RV continuum' [244]. CR enables users to collaborate with others in different realities [233], which is beneficial if the roles of users differ [311]. Users can transition [210], which is referred to a using a Transitional Interface (TI), allowing them to complete tasks in which they engage with multiple realities. Moreover, CR allows users to engage simultaneously with elements belonging to different realities [46].

As the domain of CR involves multiple users and realities, a wide variety of terms are used with little common language between descriptions of systems [8]. Hence, in this section, we explicitly introduce all the terms used in describing the patterns in this chapter. Our goal is to use terms that are specific and descriptive, yet recognisable. The conceptual framing begins with the levels of virtuality described by Milgram et al. [173]. To describe the rightmost parts of the RV continuum, the established terms *Virtual Environment (VE)* and *Virtual Reality (VR)* refer to a reality that is entirely virtual and presented to the user through artificial stimulation of sight and hearing. Conversely, *Physical Environment (PE)* and *Physical Reality (PR)* refer to the unmediated 'real world'. *Augmented Reality (AR)* describes when the environment is primarily physical, but contains some virtual elements. In contrast, *Augmented Virtuality (AV)* refers to the situation in which the environment is primarily virtual, but augmented with physical elements. *Mixed Reality (MR)* is used as an umbrella term that encapsulates all realities that are not entirely physical, as described by previous researchers [8, 257].

In this context, *immersed user* refers to those that are present in a mediated reality, and thereby may lose awareness of the PR as well as see content that cannot be seen by others. As such, users outside of the immersive reality relevant for the system are referred to as *external users*. Furthermore, users' physical locations are important for CR applications, where we refer to remote users if they are not at the same physical location, and co-located users when they are.

11.2.2 HCI Theory & Design Patterns

Our work on CR design patterns is influenced by HCI theory research such as the four stages of interaction [187] and activity theory [133]. The four stages of interaction between a person and computer are intention, selection, execution, and evaluation [187]. Similarly, activity theory conceptualised HCI as a subject-object interaction [133]. Complementary to these high-level theories, design space exploration systematically analyses design variations based on parameters. It has been effective for sub-fields of CR such as single-user applications [295], and within types of CR systems such as interactions between smartphone and AR [323]. However, to analyse the field of CR as a whole we opted to use design patterns.

Design patterns were first described by Alexander et al. as: *'Each pattern describes a problem that occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice.'* [4] Design patterns were adapted to software engineering. [77] (visitor, composite, singleton, etc.), where they facilitate development, maintenance and reusability of code. These patterns have been used for a wide variety of software development, including VR [212].

Tidwell describes patterns for interaction design that lay the foundation for HCI design patterns [276] across desktop, mobile and web applications. Furthermore, Kruschitz et al. [148] identify the anatomy of HCI design patterns in terms of form, organising principles, and relationships. To describe our patterns, we took inspiration from both the Alexandrian ('context', 'problem' and 'solution') [4] and Tidwell ('use when', 'why', and 'how') [276] forms. Our form includes both a figure of an archetypal example [276], as well as a diagram depicting the pattern [4].

Other researchers have described design patterns for MR applications [66, 34, 215], which focus on a single reality. Examples of these design patterns include point of interest, voice commands [66], world consistency, reduction of physical effort [34], signifiers and actual affordances, design for movement [215], etc. Moreover, Piumsomboon et al. [208] described two patterns for user disengagement from a VE, which are 'disengagement strategies and policies' and 'strategic escalation'. In some cases these patterns may be larger and could be implemented using our CR patterns, such as asymmetric multi-player [66]. Design patterns originate from literature review [66, 34] or through analysis of application development processes [215]. Distinct from MR design patterns, however, we focus on not one but multiple interconnected realities for CR patterns.

11.3 Methodology

We derived our design patterns from a sample set of 60 CR papers, for which we followed a integrative review process. We chose to base ourselves on an integrative review, as it is most suited to synthesise research to enable new frameworks on a certain topic [255], for which a systematic review was deemed too rigid. Furthermore, without widespread MR adoption CR systems are still mainly prevalent in literature, thus it was not feasible to derive the design patterns from analysis of real-world systems [215]. We followed best practices for the integrative review [256] and collected a corpus using database searches (for keywords such as ‘Cross-Reality’, ‘transitional interface’ and ‘asymmetric virtual reality’), conference programmes (IEEE VR and ISMAR, as well as ACM CHI, DIS, and SUI), and references in selected papers, resulting in a total of 519 MR-related papers[1]. After familiarisation with the corpus, we made a representative selection of 60 papers by 1) ensuring that the three types of CR systems (transitional, substitutional, and multi-user) from the systematic CR review by Auda et al. [8] were represented in our sample; 2) including papers that deal with the problem statements present in the larger corpus, which are remote expert [150], bystander awareness [87], bystander inclusion [99], collaboration [89], and physical object interaction [250]; 3) ensuring inclusion of scenarios with both remote [150] and co-located [89] users; and 4) including CR involving all points on the RV continuum: physical reality, augmented reality, augmented virtuality, and virtual reality.

Before analysing the sample set, we specified a scope for the patterns, which we did based on the concept of pattern scale [4]. For example, a pattern such as asymmetric multi-player [66] is closely related to CR but defined at a larger scale, which could be achieved by using a CR pattern such as mirror [134]. Thus, we limited the scale at which we looked for these patterns to what is most relevant for CR: patterns in the connection between realities.

Each pattern was named and discussed in regards to the following three topics: the user’s intent, the solution, and examples of application. To aid the reader in understanding the patterns, we included an archetypal example of each pattern in the origin, display, and interaction categories (Figure 11.3, Figure 11.5, and Figure 11.7) as well as accompanied each pattern with a schematic representation (Figure 11.1, Figure 11.2, Figure 11.4, and Figure 11.6). Table 11.1 provides a guide to reading the textual description of each pattern, as organised under the section titles ‘Intent’, ‘Solution’, and ‘Examples’.

[1] A full list of papers can be found in supplementary material to the paper that corresponds to the chapter.

CR systems often involve spatial, geometric relationships that should be represented visually [4], for which we developed a visualisation alongside the design patterns (Figure 11.1, Figure 11.2, Figure 11.4, and Figure 11.6). This differs from conventional software design patterns [77], where class diagrams depict relationships between classes. For each visualisation, we abstractly depict the realities as rectangular cuboids, as often, MR experiences are constrained by cuboid-shaped rooms or safe areas. The cuboids are populated with iconographic representations of (or part of) the reality's contents, such as users and objects. We use colour codes such as cyan for physical (■) and violet for virtual (■) to communicate the parts that make up a reality. We build upon this to show the connection between realities for each design pattern. The user depiction is central to display from which reality the user typically employs the pattern, and whether multiple users are involved. Dotted lines indicate a transmission of information between realities. Furthermore, we use a timeline to indicate if a reality is a recording of the past, an bell icon to signify the action of receiving a notification, icons for display and input modalities, such as mice, motion controllers, and monitors. Following our methodology we identified a non-exhaustive list of eleven CR design patterns.

Table 11.1: Reading the textual description of each CR Design Pattern

Intent	*To do ... across realities.*
Solution	*Pattern is ... and consists of ... components.*
Examples	*Pattern is used to ..., as shown in archetypal example in Figure.*

11.4 Fundamental Patterns

While developing the CR design patterns, we identified two fundamental patterns (Figure 11.1): **Scope**, and **Directionality**. As these patterns are more fundamental in nature, they permeate through all other categories, and have no separate archetypal examples.

11.4.1 Scope

The user is presented with part of the other reality that is of interest, as categorised into *volume*, *surface*, *person*, and *object*. A user's access to another reality may be tailored to their needs by representing only part of the reality, i.e. the scope. This list is not exhaustive but represents how CR systems typically scope the data transfer between realities. Different scopes may be combined, and other ways of scoping parts of realities may emerge in future work.

VOLUME

Intent. To access a part of a reality.

Solution. *Volume* (Figure 11.1a) refers to a 3D cutout, such as cubes, cylinders [48] or wedges [291]. Volumes can be centred around an object, such as a bookshelf [41], above a table [80], or around a workbench [120]. Furthermore, volumes can be used to communicate barriers to immersed users [315].

Examples. *Volumes* are used to capture part of a physical environment to enable remote collaboration [150], here the users have to ensure that the elements relevant for the collaboration are within the volume (Figure 11.3a and Figure 11.3b).

SURFACE

Intent. To interact with a surface across realities.

Solution. *Surfaces* (Figure 11.1b) are often of interest for collaboration, such as whiteboards or tables [95], where the surface is being synced between realities. In this scenario the surface can appear for users in different locations, for example when they are remote. Moreover, surfaces of different shapes can be remapped [58]. In terms of transitional interfaces, when the user is performing a task on a surface, the surface can persist even as the environment around the user transitions between realities [44].

Examples. *Surfaces* are used to anchor virtual content such as blocks for the docking tasks [58, 89] in Figure 11.5e and Figure 11.7b, or as a medium where virtual content is projected onto to enable external users to participate in games [99, 124] (Figure 11.5d and Figure 11.5c).

PERSON

Intent. To engage with another across realities.

Solution. Including nearby *persons* (Figure 11.1c) helps to avoid physical collisions between co-located people and facilitates communication [288]. Transferring a person between realities can be done based on a recording of their physical appearance [288]. However, it is often better to represent them with avatars that fit into the reality to avoid breaks in presence for other users [87, 82, 149], and the remote users themselves can also make use of avatar representations [150]. Additionally, persons do not need to be represented at real-time, e.g., when the immersed user should not be disturbed, the other person may be recorded for playback at a later moment [72].

Examples. Including co-located *persons* is used to facilitate multiplayer games [124] (Figure 11.5c). However, other applications such as remote collaboration may require recording the person [150] (Figure 11.3a) or displaying them via an avatar (Figure 11.3b). Similarly, when a non-immersed user wishes to interrupt a VR user, they need an avatar or other virtual representation so the latter is aware of them [87] (Figure 11.3c).

OBJECT

Intent. To interact with an object across realities.

Solution. *Objects* (Figure 11.1d) are usually central to interaction, such as the blocks used in a docking task (Figure 11.5e and Figure 11.7b) or physical props that are incorporated into a VR experience (Figure 11.7a).

Examples. In multi-user CR, sharing salient *objects* is used to facilitate collaboration, hence researchers have investigated filtering specific objects from a VE to enable AR users to see them [48, 89] (Figure 11.7b).

11.4.2 Directionality

As a second fundamental pattern, we identified the directionality of connections between realities as one of *uni-directional*, *bi-directional*, or *composite*.

UNI-DIRECTIONAL

Intent. To access one reality (B) from another (A).

Solution. The user is presented with a system in which there is an information flow from reality B to reality A in order to make it available to the user (Figure 11.1e). Information flows only one way, thus reality B cannot access reality A.

Examples. *Uni-directional* is used in CR systems that centre around the various ways in which bystanders can be included in VR experiences, such as participation in games through external displays [100, 99] or projections [99, 124] (Figure 11.5d), and allowing spectatorship through AR [48] or monitors [67]. This helps VR users avoid colliding with bystanders, yet maintains their privacy as bystanders cannot see the VR user's environment.

BI-DIRECTIONAL

Intent. To access another's reality (B) while it accesses ours (A).

Solution. The users are presented with a system where information flows both ways (Figure 11.1f). Part of reality *A* is captured and transmitted to reality *B* and vice versa. Additionally, in a transitional interface the bi-directionality could take the form of portals that show the reality being transitioned to [81].

Examples. *Bi-directional* remote presence is used to allow both users to access each other's environments, for a remote expert to observe a novice, but also for the novice to observe the remote expert as they are performing a demonstration [150] (Figure 11.3a). *Bi-directional* transfer is achieved by combining systems for uni-direction transfer, such as including a real-world view into a VR application and simultaneously projecting the VR application into the real world [107].

COMPOSITE

Intent. To access two or more realities simultaneously.

Solution. The user can enter an additional reality made specifically for the purpose of accessing the other realities (Figure 11.1g). For example, if a user needs to access reality *A* and reality *B*, then reality *C* is created which combines both *A* and *B*.

Examples. The *composite* reality is used to scale up a transitional interface to allow the user to choose between multiple different realities to enter [291]. Another example is blended space [45] which blends virtual and augmented realities with the purpose of facilitating object transitions. The concept of *composite* realities can also be expanded towards systems in which the user can activate a type of uni- or bi-directional transfer, for example when a portal between reality *A* and reality *B* is not active they are in reality *A*, but when the portal gets activated they cease to be in the unaltered version of reality *A*, but enter a slightly different reality (reality *C*) which is a *composite* of *A* and *B*. In the magicbook [20] there are three realities to the transitional interface, the real-world physical book, a VE, and also a third *composite* which shows the VE as an augmented miniature on top of the book.

11.5 Origin Patterns

Origin patterns analyse where the content transferred between realities originates from, in which we found the following three patterns: **live copy**, **replay**, **notification**. For each pattern we created diagrams (Figure 11.2) and collected an archetypal example (Figure 11.3).

11.5.1 Live Copy

Intent. To access a remote physical artefact or environment.

Solution. The user accesses a *live copy* (Figure 11.2a), i.e. an original reality or part of it, that is being copied and updated at real-time. This *live copy* is synchronised with its physical counterpart, which allows users to access the physical artefact or environment at a distance. Because users can only directly interact with the physical reality, not its virtual copy, the copy provides different affordances and needs to support each interaction the users wants perform.

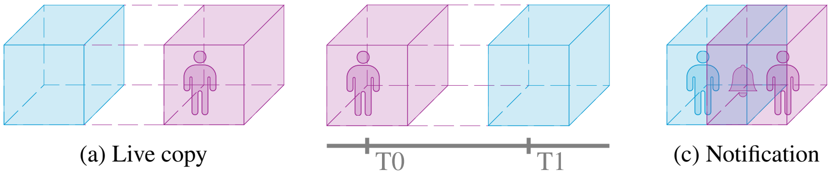
A *live copy* consists of the following components: source reality, capture, transmission, visualisation, and optionally, feedback. The source reality is usually remote, but could also be co-located [319]. The pattern has also been applied with a virtual reality as the source, where a 3D reconstruction is made when the geometry is not accessible [275]. Capture can vary in terms of scope (subsection 11.4.1), for example to only focus on a person [288]. 3D capture is often used [80, 120, 150]. Alternatively, a static 3D replica can be used in which changing elements, such as persons, are dynamically represented [206]. Or remote users can view live 360 videos [207], in which they no longer have agency over their location but only the viewing direction. Transmission allows the copy to update real-time, however, researchers also investigated a static copy of a reality, for example to achieve diminished reality, or to allow users to edit and play back the environment. It can then be visualised to the local user, who could optionally be given a way of providing feedback to the source reality, such as with annotations.

Examples. Live copy is used for remote expert guidance, Figure 11.3a shows how a remote student can be taught how to play the guitar by capturing the environment and remotely displaying it to a tutor, who can then give feedback and teach the student how to play. Thus the *live copy* is a suitable solution for remote presence [150], for example in surgical telemonitoring [80]. Alternatively, local users may also benefit from a co-located *live copy* to facilitate annotation without getting in each other's way [319].

11.5.2 Replay

Intent. To revisit a past version of a reality.

Solution. The user accesses a copy that is recorded and *replayed* [150] (Figure 11.2b). This typically involves controls similar to video playback, that allow users to play a recording, pause, rewind, change speed, or loop. As such users of the *replay* have all



(a) Live copy

(b) Replay

(c) Notification

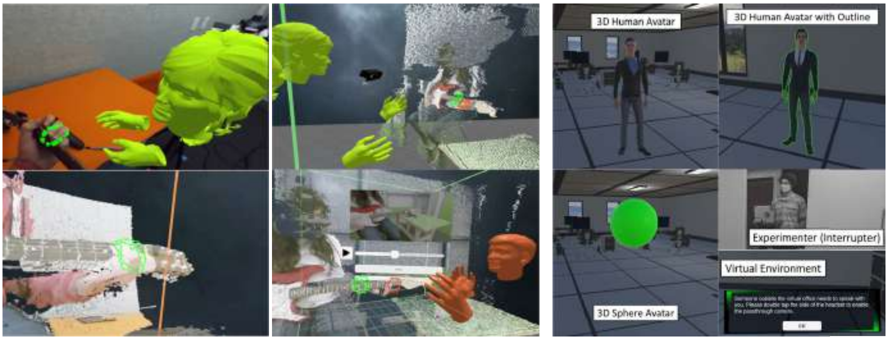
Figure 11.2: Our three **origin** patterns describing where the content transferred between realities originates from. The diagrams show a user in the reality that is the target of the pattern being: a *live copy*, *replay*, or a user in PR as the origin of a *notification* received by a user in VR.

the controls they need to go through and search for relevant events in past recordings. Additionally, when the user does not have agency over their viewpoint (such as in the mirror pattern), replays can be useful to rewind to when areas of interest were in view [151].

Replay consists of the following components: source reality, recording method, storage, display. The user is often already located in the source reality, but recordings can also be made of remote realities [150]. Recording can happen in different forms, such as regular video [151], 360 video, or point clouds [150]. This recording is then stored for playback at a later time. The user can then *replay* the recording as it was captured, but it is also possible to modify the data before it gets visualised.

More advanced *replay* systems may include the ability to synthesise events before they are displayed, such as including a visualisation of the trajectory that objects followed in addition to replaying their movement [41]. Or even to build causality graphs, which allow users to *replay* events out of sequence without violating causality, i.e. if an event requires another event to happen first the system will know and play that first [72].

Examples. Figure 11.3b shows a *replay* in the same context as the example from subsection 11.5.1, where student and tutor are reviewing the student's performance in learning to play the guitar through a *replay* [150]. This allows them to gain a better understanding of the past performance together, and allows the student to see their own performance from another perspective. Moreover, replays are used when collaborating on video game assets, to catch up on progress of other users in the real world after having entered a VR focus mode [72]. In the context of user studies, replays can be used to gain further understanding of participant behaviour.



(a) Live Copy (b) Replay (c) Notification

Figure 11.3: Archetypal examples of CR systems representing our **origin** design patterns. Image credits: Kumaravel et al. [150] and Gottsacker et al. [87] (©2021 IEEE).

11.5.3 Notification

Intent. To be made aware of events happening in another reality.

Solution. The user receives a *notification* (Figure 11.2c) of events that are happening in another reality, which contains information based on which they can decide on action.

Notification consists of an event in a source reality, transmission, and display in target reality. The event can be of multiple types, such as other persons or smartphone alerts. It can also originate from different realities: the immersed user is not always the one that has to be notified, for example external users can be notified of the immersed user's upcoming movements by projecting an arrow on the floor in front of them [292]. The *notification* is then transmitted, and displayed to the user. How and with which information content the event is displayed is the main design decision for this pattern.

When a bystander enters a VR user's area they can be notified through an avatar representation, abstract non-diegetic representation, text notification [87], widgets [149], footsteps, audio, or haptic vibrations [82]. When text notifications are used, it is important to consider where they are placed and when they are delivered [37].

Examples. Figure 11.3c shows an immersed user being notified of an interruption by an external user. Representing the external user with a (partially) diegetic avatar improved the user experience as compared to non-diegetic or text notifications. Content of the *notification* can be a bystander entering the area [87, 149, 82], or smartphone-like apps such as text messages and email [37].

11.6 Display Patterns

Display patterns describe how content from other realities is visualised, via **miniature**, **windows**, **transformed environments**, and **combined visual modalities**. For each pattern we created a diagram (Figure 11.4) and collected an archetypal example (Figure 11.5).

11.6.1 Miniature

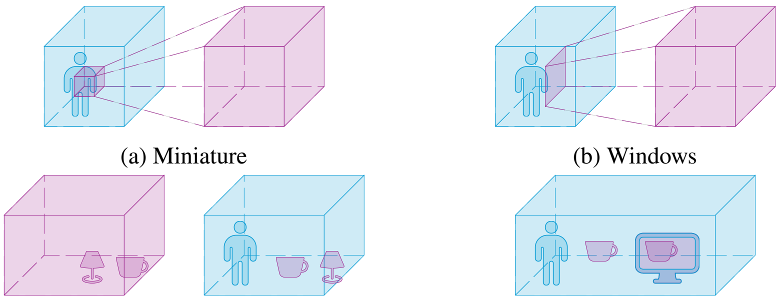
Intent. To attain an overview of another reality.

Solution. The user is presented with a *miniature* (Figure 11.4a), which is a scaled-down version of the other reality. Danyluk et al. [56] present a design space for miniatures, which consists of the following dimensions: size, scale, scope, abstraction, geometry, reference frame, links, multiples, and virtuality. In CR, miniatures most often take the form of a VE that is presented to an AR [20, 48] or external user [158], or conversely, a physical environment of which the live virtual copy is being displayed as a *miniature* [150].

Miniature consists of a source reality, scale factor, and display. The source reality needs to be virtual, or captured (such as with live copy), in order to scale it down. Additionally, the pattern may apply to different scopes, for example, by scaling down the user perspective, they may see the world larger rather than smaller [207], or the user avatar itself may be miniaturised to fit within the limited field of view for other users [206]. The scale factor determines the size of the *miniature*, moreover, the scale factor can also be greater than one, effectively scaling up the environment to scale down the user perspective. Furthermore, the *miniature* can be displayed through HMDs [48] or in combination with the mirror pattern [158].

Examples. Figure 11.5a shows a *miniature* used to give an AR user an overview of a VE. The VR user is performing a task in which they mix drinks in a virtual bar, and the AR user observes these actions in the *miniature* so that 1) they are not immersed in the environment and maintain real-world awareness, and 2) they can capture all VR user actions without having to look back and forth [48].

Generally *miniatures* are used for an exocentric view of VEs, such as when the user wishes to read a book and see a virtual scene that supports the story simultaneously [20]. Similarly, when learning about large-scale objects, i.e. a volcano, it is useful to look at smaller scales [224]. In multi-user scenarios, users can be assigned perspectives that best fit their roles, such as in RoleVR [158] where the immersed user was assigned a spatial role to support their sense of presence, and the user with the exocentric



(a) Miniature (b) Windows

(c) Transformed environments (d) Combined visual modalities

Figure 11.4: Our four **display** patterns describing how content from other realities is visualised. These include a depiction of the user and the pattern, which consists of *miniature*, *windows*, *transformed environments*, and *combined visual modalities*.

perspective with a temporal role to support their overarching view of the immersed user.

11.6.2 Windows

Two-dimensional windows are a familiar paradigm in desktop computing, which we extend into CR as a connection between two environments (Figure 11.4b), much like how real windows connect different rooms or indoor and outdoor spaces. Windows are extensively used in a range of variations, namely: *mirror*, *lens* and *portal*.

MIRROR

Intent. To passively spectate another's activities in their reality.

Solution. The user is presented with a *mirror* view where they see the environment from the viewpoint of the immersed user, hence 'mirroring' their viewpoint without providing agency over it.

The *mirror* consists of a source reality, the *mirror*'s location in it, transmission, and display. The source reality can be either physical or virtual. The *mirror* is located somewhere in this source reality from which it shows a view into it. For example, it can vary the location of the viewpoint, for example, around the immersed user's head [100], a fixed location encompassing the area of interest [311], the ceiling [81], and the floor [99]. The application of the pattern can also have limited degrees of freedom, such as allowing the user to freely rotate the view while only mirroring the location [303]. Additional information can be transmitted alongside the *mirror*, such as the immersed user's heart rate [134]. This view is then transmitted for display to the user, which can be implemented on desktop monitors [275, 134, 311], but can also make use of HMDs [81, 274], handheld devices, projectors [99, 292], or custom setups [100].

However, important drawbacks are that users cannot take on their own perspective in the environment, making it difficult to refer back to out of view areas [151], and that generally, mirrored perspectives can be shaky and hard to follow [303] if implemented naively.

Examples. Figure 11.5d is an example of a mirrored view that shows the environment from the perspective of the VR user on a television screen. Additionally, it shows how a floor projection can be used to *mirror* the VE from a top-down perspective. Mirrors facilitate observers in applications where user roles are closely coupled, such as game streaming [67], student observation [274], or games where users' focus is on the same area of the VE [134].

LENS

Intent. To explore another reality via a non-immersive interface.

Solution.

The user is provided with a *lens*, which we characterise as a device or object showing a view into another environment, and where the user has agency over the viewpoint.

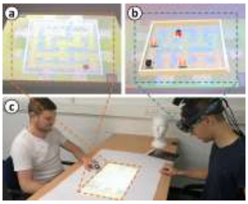
The *lens* consists of a source reality, its location in it, transmission, display, and user input. For display, lenses were first implemented through projection systems, but can also make use of commercial devices such as smartphones or tablets [89, 195], or purpose-built devices such as motion-tracked handheld displays [99].

The user is given a feedback mechanism that allows them to control its view into the reality. Control can be achieved by replicating the movement of the *lens* in the user's reality to the other reality [99, 89], or locomotion methods such as flying [195]. Other conventional input methods often paired with 2D displays to enable locomotion in VEs can also be used in CR applications, such as joysticks, mouse and keyboard, touchscreen, etc. Lenses may also be combined with other patterns, such as transforming parts of the environment to meet the needs of the user, for example by creating separate interactions for whiteboards and sticky notes to support meetings [135]. They can show simplified representations like maps [219], or miniatures [233].

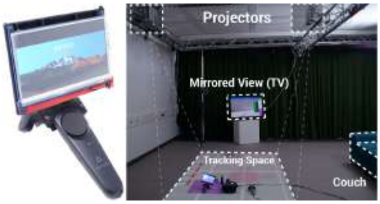
Examples. Figure 11.5d shows an example of a *lens* in the form of a handheld display, which allows the external user to view into and interact with the VE so that they can play games together with the VR user. Lenses are most useful in application where

(a) Miniature

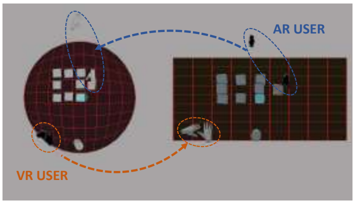
(b) Portal



(c) Combined Visual Modalities



(d) Lens and Mirror



(e) Transformed Environments

Figure 11.5: Archetypal examples of CR systems representing several of our **display** design patterns. Image credits: Cools et al. [48, 45], Gugenheimer et al. [99], De Bauw et al. [58], and Jansen et al. [124].

the external user actively engages with the immersed user's environment, such as collaboration [233], where the *lens* user is assigned a role in support of the immersed user [195]. Remote meetings [135], where the *lens* user can join and interact with a virtual meeting room where immersed users are present. In XRDirector, lenses are used to move virtual cameras to capture and direct a virtual movie scene, in which the actors are immersed [181]. Though mirrors are commonly used for remote spectatorship of VEs [67], lenses can also be used to allow the spectator to move around the environment and give them more presence within it [275].

PORTAL

Intent. To transition oneself, or objects, between realities.

Solution. The user is presented with a *portal* which, conversely to lenses and mirrors, allows for transitions through it.

Portal consists of a source reality, display, and transition. The source reality can either be the user destination, or contain objects they wish to access. Portals are usually displayed virtually with an HMD [45, 81, 107, 210, 291], though exceptions exist, such as implementing the *portal* as a physical box [200]. Given this visual connection between user and source reality, they can perform a transition by spatially moving through the *portal*. For example, portals allow users to walk through them to transition themselves into other realities [291, 210], reach into them to interact with the other reality, or transition objects by moving them in and out of the *portal* [45]. Objects can transition between physical and virtual in a controlled setting, where the object is modelled beforehand [200], but switching between manifestations ad-hoc is more difficult, especially if the user wishes to reflect the manipulations to the object in the other manifestation, and requires printing and scanning of the object [300].

Examples. Figure 11.5b shows a *portal* that allows transitioning an object between virtual and augmented environments, where the user can reach into the *portal* to grab the chalice [45]. Portals are used to smooth user transitions between environments, where the user first brings up the *portal* to then fully transition into the other environment, either with a button press [81] or by walking through it [210]. Multiple portals can also be used, so that the user can choose to which environment they intend to transition [291]. Another purpose of portals is to interact with the real world while immersed in VR, such as sitting in physical chairs [107] or taking a drink.

11.6.3 Transformed Environments

Intent. To engage with others in physical environments with disparate layouts.

Solution. The user is presented with a transformed version of the environment (Figure 11.4c), in which the elements within it are rearranged to better fit their local configuration. This approach draws from research on redirected walking, where similar remapping techniques are used between physical and VEs [141], and between VR users in different physical environments [263].

Transformed environment requires a source reality with objects in it, rules for a transform on these objects, and a target reality in which the objects are displayed. The transform between environments can either be discrete or continuous. Discrete environment transformations are achieved through shifting user positions [95], possibly after a step in which the environment is scanned and positions are detected [139]. However, to achieve a continuous transformation, a mapping function is required for the whole environment [43] or part of it [58].

AR spaces have similar problems, where elements can be rearranged to fit better with each physical location in a multi-user setting, to avoid the issue that elements that fit well for one user clip into physical objects for another [139]. Moreover, transformations of objects can be more conceptual in nature, such as showing them in a map rather than in a 3D environment [219]. When the transform is applied to avatars, it is beneficial to redirect their gaze as well [206].

Examples. Figure 11.5e shows an example of how a transformation facilitates completion of a building task for remote users working on different shape tables. By warping space on top of the tabletop surface an AR and VR user with a differently shaped tabletop can still make use of its entire surface [58]. Other applications include collaborative activities such as dancing [263], or playing a game [95].

11.6.4 Combined Visual Modalities

Intent. To access various visual representations as one.

Solution. The user is provided with a combination of visual modalities (Figure 11.4d), so that they can engage with elements in both realities simultaneously. AR is combined with another modality in the real world, often to extend familiar paradigms with 3D content. As such, *combined visual modalities* implements one modality that is visible to external users, and another that is only visible to the immersed user [124].

Combined visual modalities consists of two realities with different visual modality, each with its own characteristics, and a connection between them. Characteristics of realities are complementary, such as whether they are personal [217], their dimensionality (2D or 3D), size, and display quality. The connection between realities consists of transitions, or of each using *different interaction modalities* that extend inputs across realities.

Combined display spaces can be implemented using different physical display modalities, such as desktop monitors [46], touch surfaces [218], large displays [217], projection [124], or phones [323].

AR can be used to extend menus and display spaces over the edge of a display [218]. It can also be embedded into physical displays, show links between points on the display, or hinge and curve out of it to add additional content [217]. Moreover, because the AR content is rendered through an HMD it can be personalised per user [217]. Developers can assign content to the visual modalities so that certain visual information is always displayed in certain modalities [124]. Alternatively, the application can support transition of objects between display spaces dynamically [323, 46].

Examples. Figure 11.5c depicts a multi-user scenario with an external and immersed user, where the immersed user perceives the content through a combination of visual modalities, being AR HMD and projection. This allows for only the projection to be shared with the external user, giving rise to asymmetric games [124]. The pattern is used in instances where users benefit from both familiar and 3D visual modalities, such as information visualisation [217], or computer aided design [218].

11.7 Interaction Patterns

Interaction patterns describe how content from other realities can be interacted with, such as by **inclusion of physical objects** and **different interaction modalities**. We created a diagram (Figure 11.6) and identified an archetypal example (Figure 11.7) for each pattern.

11.7.1 Inclusion of Physical Objects

Intent. To interact with, or avoid, a physical object while in an immersive reality.

Solution. The user is presented with virtual objects at the same location as the physical objects (Figure 11.6a).

Inclusion of physical objects consists of a source reality with the physical object, a target reality in which the object is represented, and a connection between the objects. Representation in the target reality differs between closely representing the physical object and changing its appearance. The application needs to know object location, either preset if it is a static object, or via motion tracking if it is dynamic [40].

The virtual objects can differ in appearance, shape [250], and size from their physical counterpart. Virtual objects can be static if their physical counterpart does not move, or motion tracked if it does [250, 315]. Moreover, actuated physical objects allow users to feel effects of their interaction with the virtual object, or allow virtual agents to express agency on them [231]. The pattern can be extended to other kinds of physical feedback across realities, such as persons. Furthermore, AR applications may centre around a physical object that is augmented [224].

Examples. Figure 11.7a shows a physical environment that is substituted by a VE that matches the dimensions of the objects in it. This allows the user to touch, or avoid, the objects in the environments, while still presenting them with a VE that fits with the theme of the application [250]. Moreover, markers in the physical environment can be represented as virtual barriers to demarcate safe areas for the VR user [315].

11.7.2 Different Interaction Modalities

Intent. To interact with affordances tailored to the role one adopts in a specific reality, e.g., spectator or guiding roles.

Solution. Users in different roles are provided with *different interaction modalities* [158, 311, 233] (Figure 11.6b), e.g. one user intuitively explores graphs in VR while another provides input on a desktop computer.

(a) Inclusion of Physical Objects (b) Different Interaction Modalities

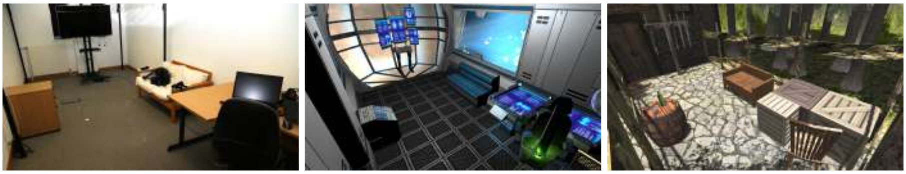
Figure 11.6: Our two **interaction** patterns describing how users can interact with content from other realities: *inclusion of physical objects* and *different interaction modalities*.

Different interaction modalities consists of a source reality and its input modality, then there is a transfer which adjusts it to the target reality, in the target reality the user then has affordances. Source input devices may include phones, tablets, and desktop inputs, which then need to be designed for to provide affordances in the target reality. These affordances are designed to mimic those of the immersed user, or to differ and fit with the role of the external user.

Implementation of the pattern consists of designing the reality to receive input from two devices rather than one, for example the tablet input method can be designed to mimic what the immersed user with motion controllers can achieve [89]. However, providing different modalities does not necessitate different input devices, for example, both users equipped with motion controllers may still be designed to have different affordances in the VE [99, 303].

Typically, an immersive reality is interacted with through motion tracking of either hands or controllers, so that its user can engage with the content in the reality. The *different interaction modalities* pattern comes into play when a second modality is introduced, which allows a second user to engage with the immersive reality [89]. This difference in modality may stem from one user being remote, and so having different affordances than local users [80]. Generally, the designer needs to make a decision of how to design the input modalities and which affordances they have, for example whether external tablet users utilise the tablet as a lens [89] or with a locomotion method such as flying [195]. For single user CR applications, different interaction modalities are used together with combined visual modalities to present the user with a different interaction modality per visual modality, for example to combine AR and hand tracking with smartphone [323], or desktop [46] devices.

Examples. Figure 11.7b shows VR HMD and tablet AR users being provided with the same affordances on a virtual object, allowing them to collaborate on a docking task. Enabling the tablet user to match the affordances of the immersed user requires careful design of their input modality and interaction techniques supported by the application [89]. Other examples of *different interaction modalities* to fit user roles includes providing teachers with a desktop interface because it provides more comfort



(a) Inclusion of Physical Objects

(b) Different Interaction Modalities

Figure 11.7: Archetypal examples of CR systems depicting our **interaction** design patterns. Image credits: Simeone et al. [250], Grandi et al. [89] (©2019 IEEE).

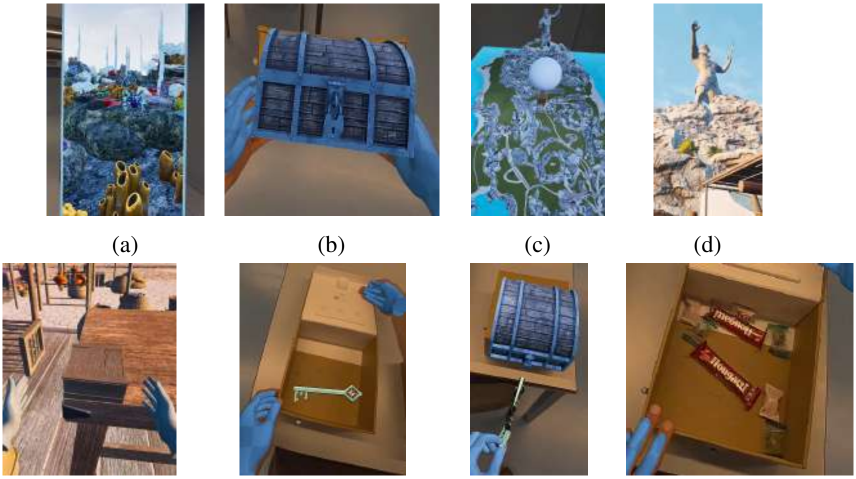
than an HMD when teaching for a longer duration [311], allowing a remote expert to use a pen to make annotations in the context of surgeries [80], asymmetric games [100, 99], or single users who can switch between modalities to interact with both 2D and 3D versions of objects [323, 46].

11.8 Example Applications of CR Design Patterns

To demonstrate the usage of our CR design patterns, we present two conceptual examples of how design patterns support CR system design. Additionally, the first example is also implemented as a demonstrator, while the second was only developed as a concept.

11.8.1 Example 1: Story-based Transitional Interface

In the first example we tell a story in which the user, as the protagonist, is searching for a hidden treasure (Figure 11.8 and Figure 11.9). We developed the story so they get to visit the following two locations: an underwater environment where they find the treasure chest, and an island on which they can find the key to the chest. This example was implemented, in Unreal Engine 5.4, to serve as a novel CR demonstrator (Figure 11.8). We used the *composite* pattern in combination with *miniature* and *portal* to allow the user to travel from the physical environment to the two VEs. Further,



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 11.8: Demonstration of how CR patterns support a user's journey. In this novel demonstrator, the user is searching for a hidden treasure across two environments. The experience exemplifies the use of, among others, *portals* (a) and *miniatures* (c).

inclusion of physical objects was implemented for the two objects the user needs to retrieve so they can actually physically pick them up. They begin their journey in an augmented environment that is a *composite* of the following two VEs: an underwater world accessible through a *portal*; and an ancient island accessible through a *miniature*. After entering the *portal* (Figure 11.8a) they find a locked treasure chest, which they pick up as a *physical object* (Figure 11.8b). After bringing the chest back to the physical environment, they go look for its key on the island, which they enter by interacting with the *miniature* (Figure 11.8c). The miniature interaction consists of pinching a simplified model of a hot air balloon. On the island (Figure 11.8d) they find a *physical* (Figure 11.8e) box which contains the key (Figure 11.8f). To return to the augmented environment, users are presented with a simplified model hot air balloon which they pinch to initiate the return transition. They can then open up the treasure chest (Figure 11.8g) and get to the treasure (Figure 11.8h).

11.8.2 Example 2: Seamless Bystander Inclusion

In the second example (Figure 11.10) we address the common problem of bystander inclusion into VR experiences, specifically with the goal of facilitating communication between bystander and VR user. Here the goals of our proposed system are to (1) allow the immersed user to see and talk to a co-located external user; and (2) to provide

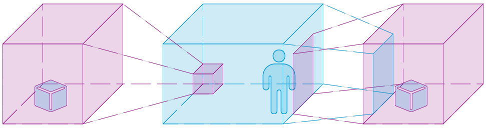


Figure 11.9: Diagram leveraging our patterns to depict the narrative-driven transitional interface (*Example 1*). In this diagram we used object colours to depict that the box is a physical object (cyan fill colour), that has a virtual representation in the VE (purple outline).

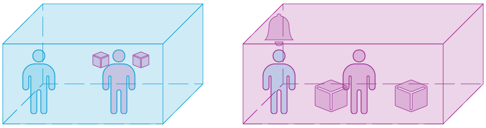


Figure 11.10: Diagram leveraging our patterns to depict seamless bystander inclusion (*Example 2*). We used the distinction between outline and fill colour of the user icons to communicate user's states, where the fill indicates user presence, and outline indicates the reality in which they are visible, as VE (purple) or PE (cyan).

the external user with a glimpse into what the immersed user is doing without being too intrusive. Thus, bystander inclusion should be designed as a *bi-directional* CR system [8, 107], to provide (1) the immersed user with awareness of the bystander, and (2) the bystander with awareness of the immersed user's VE. First, for the immersed user we employ the *person* and *notification* patterns, to visualise the external user in the VE and play a sound when they enter the immersed user's vicinity, similar to other researchers [87]. Second, to provide VE awareness for the bystander we used *object* and *miniature* patterns to show a selection of objects around the VR user in a non-intrusive way. Practically this solution can be implemented either by giving the external user an AR device, or with a custom HMD for the VR user [100].

11.9 Discussion

We presented eleven design patterns, which we discussed in terms of *Intent*, *Solution*, and *Examples* of their application. The pattern collection serves as inspiration, example, and an overview of design alternatives, not as a definite guide for CR system design. Designers using our patterns, should do so for inspiration to draw from when designing systems, taking into account their own unique context. Moreover, we provide the reader with examples which they can use as references when designing their own system. Ultimately, the chapter is an overview of the CR solution space, from which designers can compare alternatives, and adapt patterns to the right solution.

First, we discuss how the fundamental patterns permeate throughout the other three categories, and highlight how they can be used together. Then, we discuss limitations and future work.

11.9.1 Fundamental Pattern Opportunities

Directionality describes an abstract connection between realities, as visualised by the red arrows in Figure 11.1, which can be concretised through one of the other patterns. For *bi-directional* there is the opportunity to use a different pattern in each direction, depending on user roles, as demonstrated for bystander inclusion in subsection 11.8.2. In subsection 11.8.1 we use *composite* to allow user transition from the augmented reality to two VEs. For the underwater VE we used a *portal*, which works the same in both directions, i.e. the user can see the other reality through it and walk there. Conversely, the island *miniature* is only applied in one direction, i.e. the user has no miniature of the augmented environment to interact with to return.

Similarly, *scope* can be combined with other patterns. In this instance, however, some of the patterns already include a notion of scope, while others can be combined to alter their scope. For example, the *notification* pattern mostly originates from persons or smartphones, and *inclusion of physical objects* is as the name suggests concerned with objects. Looking at the patterns through the different scopes opens up new opportunities. For example, *live copy* can be applied to surfaces, persons, or objects, to only capture what is relevant for the application. *Transformed environments* can similarly be applied to surfaces, persons, or objects.

11.9.2 Limitations & Future Work

We identify limitations due to the sample set as source, and its limited size. First, our eleven patterns are not an exhaustive list of all patterns that occur within the sample, let alone all existing CR literature. Hence, as future work there are more patterns yet to be identified. Second, the decision to base our patterns on related work does not consider future real-world systems and their requirements. Hence, our patterns serve as a basis to expand upon for specific use cases and future requirements. When CR systems become more widely adopted, researchers should expand on our work with patterns that occur in development of CR applications in the field.

11.10 Conclusion

We presented eleven CR design patterns divided into the following four categories: fundamental, origin, display, and interaction. Each pattern presented an intent and solution statement, application examples, as well as a diagram and archetypal example. This sequence of patterns is both a summary, and an index of patterns that can be combined, layered, merged, and extrapolated into new contexts. In two examples we demonstrated how patterns are applied to design novel CR systems. We hope that this chapter serves as a base for future CR researchers to make a language of their own by choosing the patterns most relevant or inspiring to them, or even contribute new patterns as CR systems continue to evolve in the future.

Part IV

Conclusions

Chapter 12

Conclusions

This chapter summarises all research contributions, reflects on the findings, discusses limitations, and identifies opportunities for future work.

12.1 Summary of Contributions

The contribution is divided into three research objectives that are addressed in Part I, II, and III of this thesis. These objectives are visualising VEs to AR users (research objective 1 in Part I), user and object transitions across the RV continuum (research objective 2 in Part II), and speculating on the future usage context and creating a design framework for CR systems (research objective 3 in Part III).

12.1.1 Research Objective 1: Visualising Virtual Environments to Augmented Reality Users

Chapter 3 introduced the concept for visualising VEs for AR users, referred to as SelectVisAR, presenting three solutions to solve the problem: visualising the area surrounding the VR user based on a *proximity* threshold, displaying the VE in front of the VR user based on a *field of view angle*, and visualising specific objects around the VR user based on an *importance ranking*. The proposed SelectVisAR solutions from chapter 3 were implemented and evaluated through a pilot study in chapter 4, leading to the elimination of *field of view angle*, due to rapid changes in selected objects after head movements. Instead, we opted for a *dollhouse* technique in which the VE was presented to the AR user at 1:5 scale, to investigate the effect of scale. We then implemented the *proximity* technique, while *importance ranking* took the form of a static selection of *interactive* objects. Subsequently, the *proximity*, *interactive*, and *dollhouse* techniques were compared in a user study ($N = 13$) with a baseline showing the VE without filtering (*everything*). The findings from the study were then used to enhance the *interactive* and *proximity* techniques. *Spotlight* was a version of *proximity* that highlighted the area around the VR user, while still showing the surrounding area as a wireframe. *Context* was a version of *interactive* that expanded the selection of objects with more context, such as supporting furniture. The two new techniques were then compared to *dollhouse* and *everything* in a second study ($N = 13$). Based on the findings from the two studies in this chapter, we formulated the following three design guidelines. First, use Dollhouse when the AR user requires an overview of the VE. Second, use a static selection as opposed to a dynamic selection when possible. Third, showing the immediate context improves user preference, however, it is possible to remove nonsalient information and preserve the recognition of events.

Chapter 5 presented the 'Cross-Reality Study Tool' (CReST), an application of SelectVisAR to visualise a participant and their VE in AR to a colocated researcher. We identified this as a suitable application based on our own experience with VR user studies, where it is typically difficult for the researcher to fully comprehend the participant's interactions in the VE [51, 247, 45]. The *context* visualisation technique from chapter 4 was expanded to allow its use in lab studies, such as including a researcher interface to control the study progression. CReST allows the researcher to join the participants' VE semi-immersively, while being in control of advancing the study to subsequent steps at the appropriate times. We evaluated CReST on the following two aspects: First, whether it can be applied to a variety of studies with different requirements, by replicating three previously published studies as examples [51, 104, 247]. Second, whether the tool can be used with real participants, by conducting a case study ($N = 17$). In the case study, CReST enabled us to make observations on what interaction possibilities participants expected and how they further explored artefacts after finishing the task. With CReST, researchers adopt a qualitative observational approach to VR user studies to gain rapid feedback on prototypes.

Summary of research objective 1

1. Three guidelines for SelectVisAR systems.
   (a) Use *Dollhouse* when the AR user requires an overview of the VE.
   (b) Use a static selection as opposed to a dynamic selection when possible.
   (c) Showing the immediate context improves user preference, however, it is possible to remove non-salient information and preserve the recognition of events.
2. Demonstration of the Cross-Reality Study Tool (CReST), which uses static selection techniques to allow researchers in AR to observe participants in VR user studies.

12.1.2 Research Objective 2: Enabling Object and User Transitions Across the Reality-Virtuality Continuum

Chapter 6 adapted 'step into' and 'reach into' transition approaches from literature [81, 198], and conceptualised blended space as a third approach to the novel problem of objects transition. These approaches resulted in the following three interaction techniques for object transitions between VR and AR: *Binary Transition*, *Virtual Magic Lens*, and *Blended Space*. *Binary Transition* was based on a user transition between two realities, allowing the user to step into the other reality, which we expanded upon by allowing users to transition together with held objects, hence allowing for object transitions. *Virtual Magic Lens* presented the user with a portal into the VE that they could aim to see different parts of it, and reach into to retrieve virtual objects. *Blended Space* was a novel technique based on the *static selection* from Part I, which blends the augmented environment with a selection of VE objects. *Blended Space* facilitates the user in seeing both the target object and destination in the reality they wish to transition it into. The technique had three variations that differed in how the user achieves the transition: by pressing a *Button (MBS-B)*, via a *Touch (MBS-T)* with their offhand, or *Automatic (ABS)* when the object is moved around in blended space. We compared task completion time, workload, and user preference in two studies with different tasks, one solely focused on transitions ($N = 20$), and the other on combining transitions with manipulations ($N = 16$). We found that *Blended Space* was the fastest and most preferred, while *Virtual Magic Lens* was perceived as the most intuitive. In summary, we derived the following three design guidelines. First, *Binary Transition* is most suited when user transition is the main requirement, and object transition only occurs infrequently. Second, *Virtual Magic Lens* is most suitable when the task takes place in the 'native' environment, while object transitions are required occasionally. Third, *Blended Space* is most suitable for frequent transitions, though the choice of which variation (MBS-B, MBS-T, or ABS) to use depends on the task.

In chapter 7 we proposed a Desktop–AR prototyping framework that extends a desktop monitor (2D display space) with AR (3D display space), and described mouse and hand input combinations for object interaction and transition across this hybrid display space. Transitions to the 2D display space were based on proximity to it, following that hand inputs or extending the mouse beyond the screen could be used to drag objects into it. Additionally, bimanual and batched transition techniques offered a more efficient method of performing object transition. We presented a prototype Desktop–AR system with mouse and hand inputs to transition objects between 2D and 3D display spaces.

Chapter 8 uses the framework from chapter 7 to conceptualise and implement the following three Desktop–AR transition techniques: *Mouse-based*, *Hand-based*, and *Modality Switch*. *Mouse-based* was similar to the prototype from chapter 7, where the mouse extended beyond the screen and could grab and drag objects between spaces. *Hand-based* required the object to snap into the screen when sufficiently close, and allowing the user to remotely grab it, as the user was unable to physically move their hand into the screen. *Modality Switch* allowed users to select an object, switch modalities, and select a target for the object to transition to in the other display space. These techniques were evaluated ($N = 24$) alongside a *User Choice* condition where users could freely switch between *Mouse-* and *Hand-based* modalities during the task. We found that *Mouse-based* allowed participants to complete the task faster and with a lower workload. However, participants tended to favour using the mouse for 3D to 2D transitions and the hand for 2D to 3D transitions, citing that this felt the most intuitive and made the task less physically demanding. Based on these findings, we formulated the following two design guidelines. First, ensure that the objects are within reach of the target interaction modality after a transition. Second, minimise forced modality switches.

In contrast to the TI in chapter 6 where users were required to engage with all realities to complete the task, necessitating transition, we created a lifesize vertical surface TI in chapter 9. This TI allowed users to freely transition between four distinct virtuality levels during the task, providing an empirical understanding of how it enhances user experience. The lifesize vertical surface TI centred around a vertical surface, for which we integrated it with a robotic partition. The combination of HMD and robotic partition allowed the TI to manifest the following four levels of virtuality, in order of increased virtuality: physical environment-physical surface (AR with haptic feedback), physical environment-virtual surface (AR without haptic feedback), virtual environment-physical surface (AV) and virtual environment-virtual surface (VR). In an exploratory user study ($N = 24$), participants performed five surface-based tasks, during which we recorded their interactions with the TI and conducted interviews to gather insights into their experience. We discovered that participants preferred the physical environment to enhance their awareness of surroundings, while also recognising that the isolation offered by the VE could be advantageous in real-world situations. Additionally, participants were more likely to switch to the physical surface mode, particularly when the task they were performing allowed for direct interaction with it.

Summary of research objective 2

1. Three guidelines for transition techniques between VR and AR.
   (a) *BT* is most suited when user transition is the main requirement, and object transition only occurs infrequently.
   (b) *VML* is most suited when the task takes place in the 'native' environment, while object transitions are required occasionally.
   (c) *Blended Space* is most suited for frequent transitions, though it is task-dependent which variation (*MBS-B*, *MBS-T*, and *ABS*) should be used.
2. Two guidelines for transition techniques between desktop and AR.
   (a) Ensure that objects are within reach of the target interaction modality after a transition.
   (b) Minimise forced modality switches.
3. Two guidelines for vertical surface transitional interfaces
   (a) Users are inclined to transition to the physical environment mode of the transitional interface, and should be given the option to do so.
   (b) The virtual surface is easier to reposition.
   (c) Users benefit from the physical surface mode for tasks that allowed for continuous haptic feedback.

12.1.3 Research Objective 3: Context and Design of Future Cross-Reality

In chapter 10 we speculated on near-future ubiquitous MR contact lenses via an ISE, and presented results of focus groups ($N = 16$, in four groups of four) where participants reflected on the impact of the MR lens on their daily lives. Thematic analysis of the focus groups resulted in the following four themes: 'privacy, security, and perceptual agency', 'social acceptability of increased virtuality', 'excessive use taking away from real life', and 'future existence for practical application'. From these themes, we formulated the following three design guidelines for future MR contact lenses. First, communicate reliably the state of the MR contact lenses to the wearer and others around them. Second, facilitate social interactions, rather than replace them, through intuitive sharing of virtual content. Third, support future users in using the MR contact lenses responsibly, through time monitoring and personalisation. These findings correlate with CR challenges such as facilitating communication and sharing between immersed and external users, and allowing users to personalise in which reality they wish to consume content.

Chapter 11 presented an integrative review on a sample of 60 articles to identify recurring solutions as design patterns [4]. Eleven design patterns were identified in the following four categories: fundamental, origin, display, and interaction. Each pattern included a description of the user's intent, the solution, and examples. Additionally, we visually supported the patterns with an archetypal example from literature, and a diagram for which we developed our own visualisation, which is summarised in Figure 12.1.

Summary of research objective 3

1. Three guidelines for near-future MR contact lenses.
   (a) Reliably communicate the state of the MR contact lenses to the wearer and others around them.
   (b) Facilitate social interactions, rather than replace them, through intuitive sharing of virtual content.
   (c) Support future users in using the MR contact lenses responsibly, through time monitoring and personalisation.
2. Eleven design patterns for CR systems (Figure 12.1).

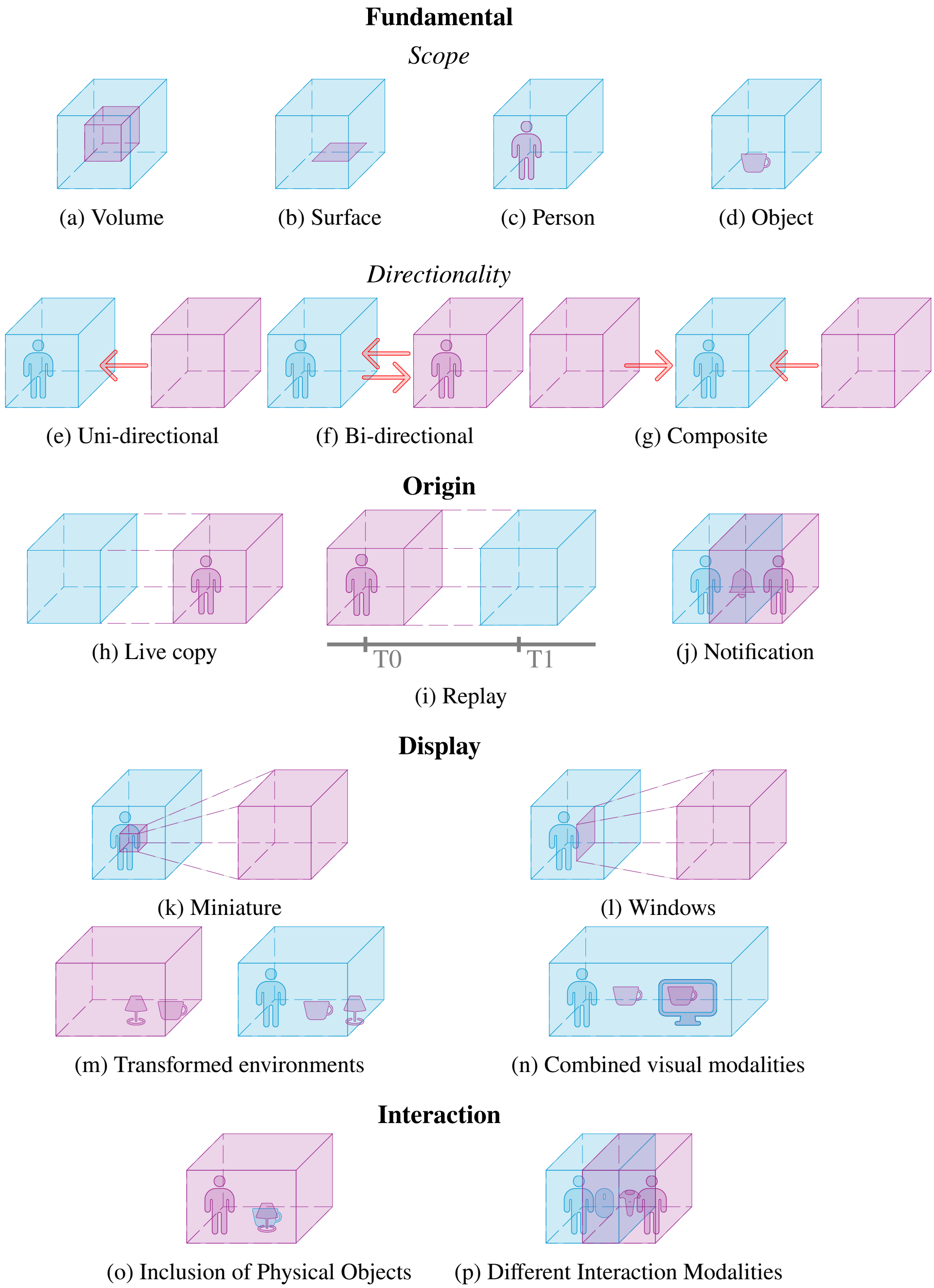


Figure 12.1: Overview of all the design pattern from chapter 11. The diagrams show how virtual (■) and physical (■) realities can be connected.

12.2 Reflection

Chapter 11 presents a CR design framework in the form of design patterns (Figure 12.1), providing a solution space for the CR subfields of section 1.3. CR design patterns were developed after chapters 3 through 10 were completed, prompting us to take a look back at these chapters through the framework properties of directionality, scope and origin. Furthermore, we reflect on generalisability of the findings, CR for everyday MR, and when an MR system should be CR.

12.2.1 Directionality, Scope, and Content Origin as Aspects of Cross-Reality

We identify directionality, scope, and content origin (Figure 12.1) as previously unconsidered aspects of CR, by comparing the research framework from section 2.1 (context, visualisation, and interaction) with the four categories of design patterns from chapter 11 (fundamental, origin, display, interaction). We reflect on directionality, scope, and origin in the CR systems of this thesis: SelectVisAR from chapter 4, CReST from chapter 5, blending spaces from chapter 6, Desktop–AR from chapters 7 and 8, and the vertical surface TI from chapter 9.

We identify the **directionality** of connections between realities in this thesis' systems, which can be uni-directional, bi-directional, and composite. The systems from chapters 4 and 8 are uni-directional, in both cases information only flows to AR. The AR observer can see part of the VE, but cannot transition anything from their own environment into it (chapter 4). Similarly, the Desktop–AR system is used from an AR perspective, and transitioning objects to the 2D display space simulated in the HMD does not show them in PR (chapter 8). This contrasts the Desktop–AR system from chapter 7, which is bi-directional because its 2D display space uses the physical display, enabling object transitions between AR space and 2D space visible in PR. Chapter 6 presents augmented and virtual environments that are connected bi-directionally, as users and objects from one can transition to the other. However, the blended space technique introduces a third composite reality, which combines the virtual and augmented environments mentioned above. Blended space clearly demarcates which of its parts originates from which reality. Conversely, chapter 9 presents four composite realities that centre around a surface with virtual content, where the virtual content does not 'belong' to any of them but is presented equally in each.

Objects are the dominant **scope** pattern in chapters 4, 6, 7, and 8, focusing on displaying or transitioning virtual objects between realities. Additionally, a volume is used for the context technique in chapter 4, and a surface for the TI in chapter 9. Hence, we consider how our systems scale to different scopes, and objects with different shapes and sizes.

To scale SelectVisAR for larger environments and objects, additional considerations are necessary to prevent visual overload for the AR user. This could involve selecting only objects that fit within the AR user's PE or introducing granularity to display only parts of objects that are too large to fit. This issue becomes apparent in Figure 5.3, where replicated example 2 shows an environment with a lot of furniture and a rug that largely covers the PE, and replicated example 3 an environment that greatly extends the PE and extends virtual objects into the physical walls.

The **origin** of content in our systems is virtual, only incorporating real-world objects to support the virtual content (chapters 9 and 11). Conversely, the origin patterns of Figure 12.1, live copy, replay, and notification centre around capturing the real world and presenting it remotely, replayed, or in notifications. In order to apply SelectVisAR to PR, where we selectively visualise parts of a PE to a VR user, 3D capture and segmentation algorithms are required [273] to isolate individual objects that meet the selection criteria. Then, the 3D capture has to be filtered, allowing an immersed VR user to observe an external user's actions in PR. Regarding object transition, creating a 3D scan or virtual replica effectively transitions an object from physical to virtual, but requires consideration on how the representations are coupled. Physical and virtual objects can either exist as one co-located object, or become two instances that are manipulated separately. Moreover, enabling transitioned objects to be manipulated involves 3D scanning and printing, which are subject to accumulative errors and model degradation [300].

12.2.2 Generalising Findings

Achieving external validity through replication in multiple case studies [154] was not within the scope of the thesis. However, we discuss how our findings can be generalised across different types of CR systems and applied to collaborative contexts.

Techniques for statically filtering salient objects from a VE (Part I) generalise well to TIs for both object (chapter 6) and user transitions (chapter 9). Blended space is built upon merging salient parts of two environments, which allows users to enter a composite reality where they can perform object transitions more efficiently. Similarly, the vertical surface TI is built on filtering the surface from the environment and allowing it to be displayed with different physicalities and in different environments. Generally, we expect saliency filtering to extend well to other CR systems if they include spatial environments, i.e. an environment that would take up the user's entire view but of which only a part is of interest. However, saliency filter techniques do not generalise to systems with a constrained display space like in chapters 7 and 8, as these already constraint the users view into the environment without need for filtering.

The main difference between VR-AR and Desktop–AR object transitions is visibility of the other space, which is always visible in the hybrid Desktop–AR interface, while without introduction of a separate technique VR and AR spaces are separated. Chapter 6 introduces techniques to visualise the VR space from AR in the form of a portal or a blended space. The AR-VR results do not translate directly into Desktop–AR, as in Desktop–AR the 2D display space is static and fixed on the desk, which negates the AR-VR portal's drawback of having to aim it around. Furthermore, the findings from chapter 8 are not directly applicable to other systems in this thesis, as it is the only system with multiple modalities.

Users' tendencies to transition to the PE (chapter 9) do not generalise and are contradicted in chapter 4. In the scenario of an AR user observing a VR user (chapter 4), the AR user tended to prefer more context in the most virtual condition where they saw the entire VE. However, for the TI that allowed users to control the virtuality level, they tended to favour the PE for enhanced real-world awareness. When another person is in VR, observers want a higher level of immersion to feel more connected with them, while when users are alone in a VE, they prefer to transition to the PE, where they are more likely to encounter others, suggesting the impact of social presence [146]. Conversely, the participants in chapter 9 noted that the VE is beneficial for intentional isolation from others in the PE.

Chapter 9 highlights that the physical surface is better for haptic interaction and easier to reposition, though these advantages are specific to this type of system. In line with our findings on the benefits of the physical surface interaction, inclusion of physical objects has been found by other researchers to lead to positive effects on user experience [250].

The findings on future MR contact lenses (chapter 10) are high-level system properties, such as the need to visualise the user's state of immersion. We anticipated a similar need for state visualisation if the systems from chapter 6 and chapter 9 were to be evaluated in a collaborative setting. Collaboration requires users to be aware of each other's activities, which adds additional requirements to the CR system, such as view sharing and avatar representations [233]. Furthermore, collaborative settings allow for opportunities, such as the ability to transition objects between individuals [50].

12.2.3 From Cross-Reality to Everyday Mixed Reality

Our CR research introduces techniques that address pivotal challenges for everyday MR, such as agency, personalisation, and content sharing (chapter 10). Similar to opening windows side by side on a conventional personal computer, blended space allows users to combine multiple MR spaces into one, supporting multitasking and object transition between spaces. Object transitions between MR and physical displays (chapter 8) are central to enable intuitive content sharing (chapter 10). Which environments users transition to is influenced by preference, therefore, users of ubiquitous MR systems should be given agency over the manifestation of reality at which they use the MR application. Similar to forced modality switches (chapter 8), forced reality switches are undesirable in everyday MR. Furthermore, even when using the same application on the same device, changing contexts can alter user requirements. When colleagues enter a shared office space, this may prompt MR users to either engage with their colleagues, or seek isolation from them to better focus on their work, thereby necessitating a transition on the RV continuum (chapter 9).

CR systems are by nature more complex than intra-reality systems, which is best exemplified in chapter 5, which presented three VR applications as replicated examples and used CReST to add support for an AR observer. Even disregarding the development of CReST itself, applying it to the VR applications required effort to ensure replication between application instances and to enable VE control for the researcher. CR systems consist of multiple viewpoints, devices, or users that need to be accommodated, making them more complex to design and develop, thereby necessitating toolkit support. Intra-reality systems are currently well supported with toolkits and frameworks (i.e. Unity XR Interaction Toolkit [280] and Meta's all-in-one SDK [170]), but CR support is limited. For example, the Meta all-in-one SDK only supports opening a 'passthrough window' into the real world as CR interaction. Other researchers created frameworks for prototyping CR systems, which focused on simulating the CR scenario, yet required developers to implement their own CR interactions and visualisations [93]. The conceptual findings from chapter 11 form a foundation for a 'CR interaction toolkit', which could serve as standard CR implementations for developers to use and adapt.

12.2.4 When Should a Mixed Reality System be Cross-Reality

We recognise that the CR systems presented in this thesis are not a replacement for conventional MR, but rather solutions to a specific subset of MR problems that are relevant depending on the context in which the system is deployed. To more precisely address the question 'when should an MR system be CR?', we divided MR systems into the following four levels to indicate to which degree a system incorporates CR: 0) no CR; 1) generic CR; 2) optional CR; and 3) core CR.

0. No CR: system is isolated to a single point on the RV continuum.
1. Generic CR: generic CR interaction, not specific to the application.
2. Optional CR: application gives the user the freedom to use it at different points on the RV continuum.
3. Core CR: application requires users at different points on the RV continuum.

Bystander inclusion and real-world collisions are challenges encountered in MR systems that were designed to operate solely within a single reality, and disregard the user's presence in the physical environment (level 0). Hence, all MR systems should provide generic support for CR (level 1), in the form of guardian systems that inform users of real-world safe areas while they are immersed. Moreover, recent devices such as the Apple Vision Pro implemented further generic CR support, through physical dials that allow users to change the extent to which they are immersed in a VE, and external facing screens that show the user's eyes to nearby persons.

We characterise level two as MR applications in which CR interactions are supported but optional, to allow users to freely choose in which environment they want to experience the application (chapter 9). This is similar to a migratory interface [32] in cross-device research, where the interface adjusts to different form factors. Games like Elite Dangerous [74], Microsoft Flight Simulator [172], and Phasmophobia [143] enable users to switch between desktop and VR modes, while the Meta home environment on Quest HMDs allows users to switch between VE and passthrough modes. However, research on TIs evaluates them with tasks that force users to switch to ensure they experience all environments [233], which is not representative of the aforementioned application scenarios. Hence, more research is required on TIs that offer users a free choice between virtuality levels, as we pursued in chapter 9.

Level three consists of applications that are designed inherently to be CR. At this level, not all MR applications should be CR, rather designers should consider whether CR is the right solution to the problem they are solving. For example, when two users require different perspectives on the task at hand [198], or when the task itself is distributed between multiple realities that require transition [233], which is reflected in the application domains from subsection 1.4.1.

Our key takeaway is to design for reality-independent applications first, rather than intentionally designing for PR, VR, and AR. This approach allows CR to emerge during the design process or be determined by the end user if the application does not require a specific point on the RV continuum.

12.3 Limitations

For the sample sizes in chapters with a quantitative approach, we deemed it not feasible to collect the high number of participants indicated by a power analysis, but rather opted for a mixed-method approach, where quantitative data are complemented by qualitative data from interviews. To illustrate what we mean by 'high number of participants indicated by a power analysis', we choose completion time in chapter 8 as an example. It is one of the few instances in which parametric tests were used, facilitating power analysis that is not extensively documented for non-parametric tests [222]. We found an effect size of 0.25, which is considered large for ANOVA tests [222]. Power analysis shows that a sample size of $N = 46$ is required for a significance level of 0.05. For the non-parametric tests in chapter 6 we found effect sizes ranging between 0.06 and 0.37, which is considered medium for non-parametric tests. Therefore, it is safe to assume that to meet this criteria for statistical power, more than 46 participants would be needed for each study, as the required sample sizes for non-parametric tests are typically larger [222]. We deemed it not feasible to collect this many participants for each of the mixed-method studies, rather we aimed to complement the quantitative data with qualitative data in the form of participant interviews. For qualitative data, we used the saturation criteria, which we could reach in the sample sizes that were used (between 13 and 24 participants), considering that the interviews were semi-structured to narrowly define the topic.

Chapter 10 uses qualitative methods in the form of focus groups, where we are uncertain whether saturation was reached with only four groups, as the data from group 1 differed from the other three. Ideally, more focus groups should have been conducted to verify whether the results of group 1 were repeated and if other topics would arise. Furthermore, this thesis uses convenience sampling, which limits participant diversity. During the analysis of chapter 9, we found that the collected data were of insufficient quality to draw in-depth conclusions about participants' use of the TI, despite a sample of 24 participants. The interviews conducted at the end of the study were not sufficient because the participants only superficially discussed the choices they made during the study.

The presence of the researcher during the studies could have introduced more bias than in typical VR studies, as participants frequently adopted AR perspectives, allowing them to see the researcher. In chapters 4 and 5, the researcher played an active role in the study, assuming the role of a second user in a two-user scenario. To ensure consistency between participants and trials, the researcher carried out their activities following a set of pre-programmed text instructions over the course of the study (chapter 4). However, chapters 6, 8, and 9 present single-user scenarios, where any influence of the researcher on participants is undesired. Especially for studies where users switched between PE and VE, the presence of another person could have influenced them to prefer the PE, in which they could see the researcher.

Another limitation of this work is that ideation of potential solutions was done by the chapter's authors, which constrained the variety of systems that were developed. Structured approaches were followed, such as identifying design dimensions (like static/dynamic, scale, step-in, reach-in, etc.) Informed design decisions were made based on extensive review of related work, leading to separate conceptual contributions in chapter 3 and chapter 7. However, by involving more experts than just the authors, we could have expanded the potential solution space for each research objective before settling on the solution we implemented.

12.4 Future Research Directions

In this section, we highlight potential future research, mainly stemming from research directions that were uncovered and considered over the course of the PhD, but ultimately not pursued.

12.4.1 Cross-Reality Social Presence

Presence [253] and social presence [105, 146] are established concepts for VR experiences, which measure to what extent users of MR systems feel present in the environment or with others. Multi-user CR systems allow for asymmetry in how users experience these feelings of presence, both in the environment and with each other. For example, in our SelectVisAR study (chapter 4) the VR user experienced presence in the VE and social presence with the AR user which they could see through an avatar. The AR user, on the other hand, did not experience any presence in the VE as they remained in the real world. We expect them to have experienced a high degree of social presence with the VR user which they were co-located with. However, despite experiencing reality differently, both users remained part of the same experience and ended the experiment with a common understanding of the VR user's interactions within the VE. Future work should define what it means to share presence with someone in a different reality and standardise a method to measure it, similar to the presence and social presence questionnaires [253, 105] that currently exist for VR.

12.4.2 Cross-Reality State Visualisation

In chapter 10, we found that near-future MR contact lenses have to reliably communicate their state to the wearer and others around them. Similar issues already exist with current generation VST HMDs, where it remains unclear to which degree users are aware of the real world. HMD users could be fully aware of the real world (PR), partially aware with some of their vision blocked by virtual objects (AR), or entirely immersed, unaware of their surroundings (VR), while external persons remain unaware of the HMD user's current state. Future research should explore methods to convey this state to external persons, i.e. a slider that represents the user's point on the RV continuum, colour codes, or icons. We envision that this could be achieved by adding additional hardware to the HMD, such as LEDs [88] or screens [36], or with AR if external users are equipped with it. AR could offer the opportunity to communicate more information than just the user's state, for example, in another work we investigated how we can present AR users with abstracted versions of objects, thereby providing them with an indication of other user's activities while minimising distraction from being shown too much information at once [285].

12.4.3 Cross-Reality Agents

The CR research in this thesis is concerned with users, environments, and objects, leaving open an opportunity to research virtual agents across realities. Virtual agents are autonomous, often humanoid, entities with which the user can interact. They can help the user perform various tasks, such as playing the role of an antagonist in game-like experiences. Future work into CR agents could investigate the following two aspects: how users interact with agents across realities, and how the agent itself interacts across realities [231]. For example, the agent could be the interface with which users interact across realities by allowing users to give directions to an agent that is in a different reality.

12.5 Closing Remarks

In this thesis we explored cross-reality through the following three research objectives: 'Visualising Virtual Environments to Augmented Reality Users', 'Enabling Object and User Transitions Across the Reality-Virtuality Continuum', and 'Context and Design of Future Cross-Reality'. The first research objective was to display virtual environments to augmented reality users. It resulted in three design guidelines, and a tool for AR spectatorship of VR user studies. The second research objective resulted in three AR-VR object transition guidelines, and two Desktop–AR object transition design guidelines. Furthermore, it resulted in three guidelines for user transitions with a vertical surface transitional interface. In the third objective, we speculated on near-future MR contact lenses in an immersive speculative enactment, resulting in three guidelines for future systems. Furthermore, we developed a framework with eleven CR design patterns.

These guidelines and patterns form a foundation for both practical cross-reality applications, as well as future ubiquitous mixed reality systems. With big tech companies pursuing development of small form-factor mixed reality devices, mixed reality becoming an ubiquitous and everyday technology seems almost inevitable. We should be considerate of the way in which we want to embrace this exciting new technology in our lives, in which cross-reality is critical to ensure that this novel technology brings users together rather than isolates them from one another.

Bibliography

[1] Abraham, M., Saeghe, P., Mcgill, M., and Khamis, M. Implications of XR on Privacy, Security and Behaviour: Insights from Experts. In *Nordic Human-Computer Interaction Conference* (New York, NY, USA, Oct. 2022), NordiCHI '22, Association for Computing Machinery, pp. 1–12.

[2] Adams, D., Bah, A., Barwulor, C., Musabay, N., Pitkin, K., and Redmiles, E. M. Ethics emerging: the story of privacy and security perceptions in virtual reality. In *Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security* (USA, Aug. 2018), SOUPS '18, USENIX Association, pp. 443–458.

[3] Alallah, F., Neshati, A., Sakamoto, Y., Hasan, K., Lank, E., Bunt, A., and Irani, P. Performer vs. observer: whose comfort level should we consider when examining the social acceptability of input modalities for head-worn display? In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, Nov. 2018), VRST '18, Association for Computing Machinery, pp. 1–9.

[4] Alexander, C., Ishikawa, S., and Silverstein, M. *A Pattern Language: Towns, Buildings, Construction*. OUP USA, 1977. Google-Books-ID: hwAHmktpk5IC.

[5] Apple. Wwdc 2023 - june 5 | apple. https://www.youtube.com/watch?v=GYkq9Rgoj8E, Jun 2023. Accessed: 06-06-2023.

[6] Arora, R., Habib Kazi, R., Grossman, T., Fitzmaurice, G., and Singh, K. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2018), CHI '18, Association for Computing Machinery, pp. 1–15.

[7] Ashtari, N., Bunt, A., McGrenere, J., Nebeling, M., and Chilana, P. K. Creating Augmented and Virtual Reality Applications: Current Practices,

Challenges, and Opportunities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2020), CHI '20, Association for Computing Machinery, pp. 1–13.

[8] AUDA, J., GRUENEFELD, U., FALTAOUS, S., MAYER, S., AND SCHNEEGASS, S. A Scoping Survey on Cross-Reality Systems. *ACM Computing Surveys* (Sept. 2023). Just Accepted.

[9] AZALEA VISION. Smart vision solutions. https://azaleavision.com/, 2023. Accessed: 09-01-2023.

[10] BARENDREGT, L., AND VAAGE, N. S. Speculative Design as Thought Experiment. *She Ji: The Journal of Design, Economics, and Innovation 7*, 3 (Sept. 2021), 374–402.

[11] BEEVER, L., AND JOHN, N. W. LevelEd SR: A Substitutional Reality Level Design Workflow. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Mar. 2022), pp. 130–138. ISSN: 2642-5254.

[12] BENKO, H., ISHAK, E., AND FEINER, S. Cross-dimensional gestural interaction techniques for hybrid immersive environments. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* (Mar. 2005), pp. 209–216. ISSN: 2375-5334.

[13] BENKO, H., WILSON, A. D., AND ZANNIER, F. Dyadic projected spatial augmented reality. In *Proceedings of the 27th annual ACM symposium on User interface software and technology* (Honolulu Hawaii USA, Oct. 2014), ACM, pp. 645–655.

[14] BESANÇON, L., ISSARTEL, P., AMMI, M., AND ISENBERG, T. Hybrid Tactile/Tangible Interaction for 3D Data Exploration. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (Jan. 2017), 881–890. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[15] BESANÇON, L., YNNERMAN, A., KEEFE, D. F., YU, L., AND ISENBERG, T. The State of the Art of Spatial Interfaces for 3D Visualization. *Computer Graphics Forum 40*, 1 (2021), 293–326. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14189.

[16] BEŞEVLI, C., BURUK, O. T., ERKAYA, M., AND ÖZCAN, O. Investigating the Effects of Legacy Bias: User Elicited Gestures from the End Users Perspective. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems* (New York, NY, USA, May 2018), DIS '18 Companion, Association for Computing Machinery, pp. 277–281.

[17] BIENER, V., SCHNEIDER, D., GESSLEIN, T., OTTE, A., KUTH, B., KRISTENSSON, P. O., OFEK, E., PAHUD, M., AND GRUBERT, J. Breaking the Screen: Interaction

Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers. *IEEE Transactions on Visualization and Computer Graphics 26*, 12 (Dec. 2020), 3490–3502. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[18] BIER, E. A., STONE, M. C., PIER, K., BUXTON, W., AND DEROSE, T. D. Toolglass and Magic Lenses: The See-Through Interface. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques - SIGGRAPH '93* (New York, New York, USA, 1993), ACM Press.

[19] BILLINGHURST, M. Grand Challenges for Augmented Reality. *Frontiers in Virtual Reality 2* (2021).

[20] BILLINGHURST, M., KATO, H., AND POUPYREV, I. The MagicBook - Moving seamlessly between reality and virtuality. *IEEE Computer Graphics and Applications 21*, 3 (May 2001), 6–8.

[21] BILLINGHURST, M., WEGHORST, S., AND FURNESS, T. Shared space: An augmented reality approach for computer supported collaborative work. *Virtual Reality 3*, 1 (1998), 25–36. Publisher: Springer London.

[22] BONNAIL, E., TSENG, W.-J., MCGILL, M., LECOLINET, E., HURON, S., AND GUGENHEIMER, J. Memory Manipulations in Extended Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2023), CHI '23, Association for Computing Machinery, pp. 1–20.

[23] BONNER, J., MATHIS, F., O'HAGAN, J., AND MCGILL, M. When Filters Escape the Smartphone: Exploring Acceptance and Concerns Regarding Augmented Expression of Social Identity for Everyday AR. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, Oct. 2023), VRST '23, Association for Computing Machinery, pp. 1–14.

[24] BORGMAN, A. Holding on the Reality: The Nature of Information at the Turn of the Millennium. *Journal of social work education 35*, 3 (1999), 487–487. Place: Washington Publisher: Taylor & Francis Ltd.

[25] BOSCH, T. Sci-Fi Writer Bruce Sterling Explains the Intriguing New Concept of Design Fiction. *Slate* (Mar. 2012).

[26] BOUZBIB, E., BAILLY, G., HALIYO, S., AND FREY, P. CoVR: A Large-Scale Force-Feedback Robotic Interface for Non-Deterministic Scenarios in VR. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2020), UIST '20, Association for Computing Machinery, pp. 209–222.

[27] BOYER, S. A virtual failure: Evaluating the success of Nintendo's Virtual Boy. *The Velvet Light Trap*, 64 (2009), 23–33. Publisher: University of Texas Press.

[28] BRAUN, V., AND CLARKE, V. Using thematic analysis in psychology. *Qualitative Research in Psychology 3*, 2 (Jan. 2006), 77–101.

[29] BROOKE, J. Sus: a "quick and dirty'usability. *Usability evaluation in industry 189*, 3 (1996), 189–194. Publisher: Taylor & Francis.

[30] BROOKES, J., WARBURTON, M., ALGHADIER, M., MON-WILLIAMS, M., AND MUSHTAQ, F. Studying human behavior with virtual reality: The Unity Experiment Framework. *Behavior Research Methods 52*, 2 (Apr. 2020), 455–463.

[31] BRUDER, G., STEINICKE, F., AND HINRICHS, K. H. Arch-explore: A natural user interface for immersive architectural walkthroughs. In *3DUI - IEEE Symposium on 3D User Interfaces 2009 - Proceedings* (2009). ISSN: 08872171.

[32] BRUDY, F., HOLZ, C., RÄDLE, R., WU, C. J., HOUBEN, S., KLOKMOSE, C. N., AND MARQUARDT, N. Cross-device taxonomy: Survey, opportunities and challenges of interactions spanning across multiple devices. In *Conference on Human Factors in Computing Systems - Proceedings* (May 2019), Association for Computing Machinery, p. 28.

[33] BUTZ, A., HOLLERER, T., FEINER, S., MACINTYRE, B., AND BESHERS, C. Enveloping users and computers in a collaborative 3D augmented reality. In *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)* (Oct. 1999), pp. 35–44.

[34] BÖRSTING, I., KARABULUT, C., FISCHER, B., AND GRUHN, V. Design Patterns for Mobile Augmented Reality User Interfaces—An Incremental Review. *Information 13*, 4 (Apr. 2022), 159. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[35] BÜSCHEL, W., LEHMANN, A., AND DACHSELT, R. MIRIA: A Mixed Reality Toolkit for the In-Situ Visualization and Analysis of Spatio-Temporal Interaction Data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2021), CHI '21, Association for Computing Machinery, pp. 1–15.

[36] CHAN, L., AND MINAMIZAWA, K. FrontFace: Facilitating communication between HMD users and outsiders using front-facing-screen HMDs. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2017* (New York, NY, USA, Sept. 2017), Association for Computing Machinery, Inc, pp. 1–5.

[37] CHEN, K.-W., CHANG, Y.-J., AND CHAN, L. Predicting Opportune Moments to Deliver Notifications in Virtual Reality. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), ACM, pp. 1–18.

[38] CHENECHAL, M. L., DUVAL, T., GOURANTON, V., ROYAN, J., AND ARNALDI, B. Vishnu: Virtual immersive support for HelpiNg users an interaction paradigm for collaborative remote guiding in mixed reality. In *2016 IEEE 3rd VR International Workshop on Collaborative Virtual Environments, 3DCVE 2016* (Sept. 2016), Institute of Electrical and Electronics Engineers Inc., pp. 9–12.

[39] CHENG, L. P., CHANG, L., MARWECKI, S., AND BAUDISCH, P. ITurk: Turning passive haptics into active haptics by making users reconfigure props in virtual reality. *Conference on Human Factors in Computing Systems - Proceedings 2018-April* (2018), 1–10.

[40] CHENG, L.-P., OFEK, E., HOLZ, C., AND WILSON, A. D. VRoamer: Generating On-The-Fly VR Experiences While Walking inside Large, Unknown Real-World Building Environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Mar. 2019), pp. 359–366. ISSN: 2642-5254.

[41] CHO, H., KOMAR, M. L., AND LINDLBAUER, D. RealityReplay: Detecting and Replaying Temporal Changes In Situ Using Mixed Reality. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 7*, 3 (Sept. 2023), 90:1–90:25.

[42] CLERGEAUD, D., ROO, J. S., HACHET, M., AND GUITTON, P. Towards seamless interaction between physical and virtual locations for asymmetric collaboration. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, Nov. 2017), VRST '17, Association for Computing Machinery, pp. 1–4.

[43] CONGDON, B. J., WANG, T., AND STEED, A. Merging Environments for Shared Spaces in Mixed Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, 2018), vol. 18, ACM.

[44] COOLS, R., BALCI, O., NGUYEN, B. V. D., VANDE MOERE, A., AND SIMEONE, A. L. How a Vertical Surface Supports Cross-Reality Transitional Interface Tasks at Different Virtuality Levels. In *Proceedings of the 2024 International Conference on Advanced Visual Interfaces* (New York, NY, USA, June 2024), AVI '24, Association for Computing Machinery, pp. 1–5.

[45] COOLS, R., ESTEVES, A., AND SIMEONE, A. L. Blending Spaces: Cross-Reality Interaction Techniques for Object Transitions Between Distinct Virtual and Augmented Realities. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Oct. 2022), pp. 528–537. ISSN: 1554-7868.

[46] Cools, R., Gottsacker, M., Simeone, A., Bruder, G., Welch, G., and Feiner, S. Towards a Desktop-AR Prototyping Framework: Prototyping Cross-Reality Between Desktops and Augmented Reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Oct. 2022), pp. 175–182. ISSN: 2771-1110.

[47] Cools, R., Han, J., Esteves, A., and Simeone, A. L. From Display to Interaction: Design Patterns for Cross-Reality Systems. *IEEE Transactions on Visualization and Computer Graphics* (2025).

[48] Cools, R., Han, J., and Simeone, A. L. SelectVisAR: Selective Visualisation of Virtual Environments in Augmented Reality. In *DIS 2021 - Proceedings of the 2021 ACM Designing Interactive Systems Conference: Nowhere and Everywhere* (June 2021), Association for Computing Machinery, Inc, pp. 275–282.

[49] Cools, R., Maerevoet, I., Gottsacker, M., and Simeone, A. L. Comparison of Cross-Reality Transition Techniques Between 3D and 2D Display Spaces in Desktop–AR Systems. *IEEE Transactions on Visualization and Computer Graphics* (2025).

[50] Cools, R., and Simeone, A. Mobile Displays for Cross-Reality Interactions between Virtual and Physical Realities. In *Proceedings of the 20th International Conference on Mobile and Ubiquitous Multimedia* (New York, NY, USA, Dec. 2021), MUM '21, Association for Computing Machinery, pp. 217–219.

[51] Cools, R., and Simeone, A. L. Investigating the Effect of Distractor Interactivity for Redirected Walking in Virtual Reality. In *Symposium on Spatial User Interaction* (New York, NY, USA, Oct. 2019), SUI '19, Association for Computing Machinery, pp. 1–5.

[52] Cools, R., Venema, R., Esteves, A., and Simeone, A. L. The Impact of Near-Future Mixed Reality Contact Lenses on Users' Lives via an Immersive Speculative Enactment and Focus Groups. In *Proceedings of the 2024 ACM Symposium on Spatial User Interaction* (New York, NY, USA, Oct. 2024), SUI '24, Association for Computing Machinery, pp. 1–13.

[53] Cools, R., Zhang, X., and Simeone, A. L. CReST: Design and Evaluation of the Cross-Reality Study Tool. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (New York, NY, USA, Dec. 2023), MUM '23, Association for Computing Machinery, pp. 409–419.

[54] Coulton, P., Burnett, D., and Gradinar, A. Games as Speculative Design: Allowing Players to Consider Alternative Presents and Plausible Futures. https://www.drs2016.org/015, 2016. Accessed: 16-12-2022.

[55] CRUZ-NEIRA, C., SANDIN, D. J., DEFANTI, T. A., KENYON, R. V., AND HART, J. C. The CAVE: audio visual experience automatic virtual environment. *Commun. ACM 35*, 6 (June 1992), 64–72.

[56] DANYLUK, K., ENS, B., JENNY, B., AND WILLETT, W. A Design Space Exploration of Worlds in Miniature. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2021), Association for Computing Machinery (ACM), pp. 1–15.

[57] DE ARAÙJO, B. R., CASIEZ, G., AND JORGE, J. A. Mockup builder: direct 3D modeling on and above the surface in a continuous interaction space. In *Proceedings of Graphics Interface 2012* (CAN, May 2012), GI '12, Canadian Information Processing Society, pp. 173–180.

[58] DE BAUW, T., COOLS, R., AND SIMEONE, A. L. AdapTables: Using Conformal Mapping for Collaboration on Tables in Asymmetric Mixed Reality. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction* (New York, NY, USA, Dec. 2022), SUI '22, Association for Computing Machinery, pp. 1–2.

[59] DESAI, A. P., PEÑA-CASTILLO, L., AND MERUVIA-PASTOR, O. A Window to Your Smartphone: Exploring Interaction and Communication in Immersive VR with Augmented Virtuality. In *2017 14th Conference on Computer and Robot Vision (CRV)* (May 2017), pp. 217–224.

[60] DI LUCA, M., SEIFI, H., EGAN, S., AND GONZALEZ-FRANCO, M. Locomotion Vault: the Extra Mile in Analyzing VR Locomotion Techniques. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2021), CHI '21, Association for Computing Machinery, pp. 1–10.

[61] DI VERDI, S., NURMI, D., AND HOLLERER, T. ARWin - a desktop augmented reality Window Manager. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* (Oct. 2003), pp. 298–299.

[62] DREY, T., ALBUS, P., DER KINDEREN, S., MILO, M., SEGSCHNEIDER, T., CHANZAB, L., RIETZLER, M., SEUFERT, T., AND RUKZIO, E. Towards Collaborative Learning in Virtual Reality: A Comparison of Co-Located Symmetric and Asymmetric Pair-Learning. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), ACM, pp. 1–19.

[63] DUDLEY, J. J., SCHUFF, H., AND KRISTENSSON, P. O. Bare-Handed 3D Drawing in Augmented Reality. In *Proceedings of the 2018 Designing Interactive*

Systems Conference (New York, NY, USA, June 2018), DIS '18, Association for Computing Machinery, pp. 241–252.

[64] ELDER, A. Excellent online friendships: an Aristotelian defense of social media. *Ethics and Information Technology 16*, 4 (Dec. 2014), 287–297.

[65] ELSDEN, C., CHATTING, D., DURRANT, A. C., GARBETT, A., NISSEN, B., VINES, J., AND KIRK, D. S. On Speculative Enactments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2017), CHI '17, Association for Computing Machinery, pp. 5386–5399.

[66] EMMERICH, F., KLEMKE, R., AND HUMMES, T. Design Patterns for Augmented Reality Learning Games. In *Games and Learning Alliance* (Cham, 2017), J. Dias, P. A. Santos, and R. C. Veltkamp, Eds., Springer International Publishing, pp. 161–172.

[67] EMMERICH, K., KREKHOV, A., CMENTOWSKI, S., AND KRUGER, J. Streaming vr games to the broad audience: A comparison of the first-person and third-person perspectives. In *Conference on Human Factors in Computing Systems - Proceedings* (New York, NY, USA, May 2021), Association for Computing Machinery, pp. 1–14. arXiv: 2101.04449.

[68] ENS, B., BACH, B., CORDEIL, M., ENGELKE, U., SERRANO, M., WILLETT, W., PROUZEAU, A., ANTHES, C., BÜSCHEL, W., DUNNE, C., DWYER, T., GRUBERT, J., HAGA, J. H., KIRSHENBAUM, N., KOBAYASHI, D., LIN, T., OLAOSEBIKAN, M., POINTECKER, F., SAFFO, D., SAQUIB, N., SCHMALSTIEG, D., SZAFIR, D. A., WHITLOCK, M., AND YANG, Y. Grand Challenges in Immersive Analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2021), CHI '21, Association for Computing Machinery, pp. 1–17.

[69] FEINER, S., MACINTYRE, B., HAUPT, M., AND SOLOMON, E. Windows on the world: 2D windows for 3D augmented reality. In *Proceedings of the 6th annual ACM symposium on User interface software and technology - UIST '93* (Atlanta, Georgia, United States, 1993), ACM Press, pp. 145–155.

[70] FEINER, S., AND SHAMASH, A. Hybrid user interfaces: breeding virtually bigger interfaces for physically smaller computers. In *Proceedings of the 4th annual ACM symposium on User interface software and technology* (New York, NY, USA, Nov. 1991), UIST '91, Association for Computing Machinery, pp. 9–17.

[71] FELD, N., BIMBERG, P., WEYERS, B., AND ZIELASKO, D. Simple and Efficient? Evaluation of Transitions for Task-Driven Cross-Reality Experiences. *IEEE Transactions on Visualization and Computer Graphics* (2024), 1–18. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[72] Fender, A. R., and Holz, C. Causality-preserving Asynchronous Reality. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), ACM, pp. 1–15.

[73] Fischer, R., Lian, W.-X., Wang, S.-J., Hsu, W.-E., and Fu, L.-C. Seamless Virtual Object Transitions: Enhancing User Experience in Cross-Device Augmented Reality Environments. In *Extended Reality* (Cham, 2023), L. T. De Paolis, P. Arpaia, and M. Sacco, Eds., Lecture Notes in Computer Science, Springer Nature Switzerland, pp. 397–409.

[74] Frontier Developments plc. Elite dangerous. https://www.elitedangerous.com/, 2025. Accessed: 25-01-2025.

[75] Fröhler, B., Anthes, C., Pointecker, F., Friedl, J., Schwajda, D., Riegler, A., Tripathi, S., Holzmann, C., Brunner, M., Jodlbauer, H., Jetter, H.-C., and Heinzl, C. A Survey on Cross-Virtuality Analytics. *Computer Graphics Forum 41*, 1 (Feb. 2022), 465–494. Publisher: John Wiley & Sons, Ltd.

[76] Gallego Cascón, P., Matthies, D. J., Muthukumarana, S., and Nanayakkara, S. ChewIt. An Intraoral Interface for Discreet Interactions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2019), CHI '19, Association for Computing Machinery, pp. 1–13.

[77] Gamma, E., Helm, R., Johnson, R., and Vlissides, J. *Design Patterns: Elements of Reusable Object-Oriented Software*. Pearson Education, Oct. 1994. Google-Books-ID: 6oHuKQe3TjQC.

[78] Garcia, J. F., Simeone, A. L., Higgins, M., Powell, W., and Powell, V. Inside looking out or outside looking in? An evaluation of visualisation modalities to support the creation of a substitutional virtual environment. In *Proceedings of the Workshop on Advanced Visual Interfaces AVI* (New York, NY, USA, May 2018), Association for Computing Machinery, pp. 1–8.

[79] Garcia Estrada, J., and Simeone, A. L. Recommender system for physical object substitution in VR. In *2017 IEEE Virtual Reality (VR)* (Mar. 2017), pp. 359–360. ISSN: 2375-5334.

[80] Gasques, D., Johnson, J. G., Sharkey, T., Feng, Y., Wang, R., Xu, Z. R., Zavala, E., Zhang, Y., Xie, W., Zhang, X., Davis, K., Yip, M., and Weibel, N. ARTEMIS: A Collaborative Mixed-Reality System for Immersive Surgical Telementoring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2021), Association for Computing Machinery (ACM), pp. 1–14.

[81] GEORGE, C., TIEN, A. N., AND HUSSMANN, H. Seamless, Bi-directional Transitions along the Reality-Virtuality Continuum: A Conceptualization and Prototype Exploration. In *Proceedings - 2020 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2020* (Nov. 2020), Institute of Electrical and Electronics Engineers Inc., pp. 412–424.

[82] GHOSH, S., WINSTON, L., PANCHAL, N., KIMURA-THOLLANDER, P., HOTNOG, J., CHEONG, D., REYES, G., AND ABOWD, G. D. NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics 24*, 4 (Apr. 2018), 1447–1456. Publisher: IEEE Computer Society.

[83] GOOGLE. Google glass – introducing glass enterprise edition 2. https://www.google.com/glass/start/, 2023. Accessed: 09-01-2023.

[84] GOTTSACKER, M., BRUDER, G., AND WELCH, G. F. rlty2rlty: Transitioning Between Realities with Generative AI. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (Mar. 2024), pp. 1160–1161.

[85] GOTTSACKER, M., CHEN, M., SAFFO, D., LU, F., AND MACINTYRE, B. Hybrid User Interface for Audience Feedback Guided Asymmetric Immersive Presentation of Financial Data. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Oct. 2023), pp. 199–204. ISSN: 2771-1110.

[86] GOTTSACKER, M., FURUYA, H., BATTISTEL, L., JIMENEZ, C. P., LAMONTAGNA, N., BRUDER, G., AND WELCH, G. F. Exploring Spatial Cognitive Residue and Methods to Clear Users' Minds When Transitioning Between Virtual Environments. In *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Oct. 2024), pp. 1000–1009. ISSN: 2473-0726.

[87] GOTTSACKER, M., NOROUZI, N., KIM, K., BRUDER, G., AND WELCH, G. Diegetic Representations for Seamless Cross-Reality Interruptions. In *ISMAR* (2021).

[88] GOTTSACKER, M., SYAMIL, R., WISNIEWSKI, P., BRUDER, G., CRUZ-NEIRA, C., AND WELCH, G. Exploring Cues and Signaling to Improve Cross-Reality Interruptions. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Oct. 2022), pp. 827–832. ISSN: 2771-1110.

[89] GRANDI, J. G., DEBARBA, H. G., AND MACIEL, A. Characterizing asymmetric collaborative interactions in virtual and augmented realities. In *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings* (Mar. 2019), Institute of Electrical and Electronics Engineers Inc., pp. 127–135.

[90] GRUBERT, J., LANGLOTZ, T., ZOLLMANN, S., AND REGENBRECHT, H. Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics 23*, 6 (June 2017), 1706–1724. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[91] GRUENEFELD, U., ALI, A. E., BOLL, S., AND HEUTEN, W. Beyond Halo and Wedge: visualizing out-of-view objects on head-mounted virtual and augmented reality devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services* (New York, NY, USA, Sept. 2018), MobileHCI '18, Association for Computing Machinery, pp. 1–11.

[92] GRUENEFELD, U., ALI, A. E., HEUTEN, W., AND BOLL, S. Visualizing out-of-view objects in head-mounted augmented reality. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (New York, NY, USA, Sept. 2017), MobileHCI '17, Association for Computing Machinery, pp. 1–7.

[93] GRUENEFELD, U., AUDA, J., MATHIS, F., SCHNEEGASS, S., KHAMIS, M., GUGENHEIMER, J., AND MAYER, S. VRception: Rapid Prototyping of Cross-Reality Systems in Virtual Reality. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), ACM, pp. 1–15.

[94] GRUENEFELD, U., ENNENGA, D., ALI, A. E., HEUTEN, W., AND BOLL, S. EyeSee360: designing a visualization technique for out-of-view objects in head-mounted augmented reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (New York, NY, USA, Oct. 2017), SUI '17, Association for Computing Machinery, pp. 109–118.

[95] GRØNBÆK, J. E. S., PFEUFFER, K., VELLOSO, E., ASTRUP, M., PEDERSEN, M. I. S., KJÆR, M., LEIVA, G., AND GELLERSEN, H. Partially Blended Realities: Aligning Dissimilar Spaces for Distributed Mixed Reality Meetings. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2023), CHI '23, Association for Computing Machinery, pp. 1–16.

[96] GRÜBEL, J., WEIBEL, R., JIANG, M. H., HÖLSCHER, C., HACKMAN, D. A., AND SCHINAZI, V. R. EVE: A Framework for Experiments in Virtual Environments. In *Spatial Cognition X* (Cham, 2017), T. Barkowsky, H. Burte, C. Hölscher, and H. Schultheis, Eds., Lecture Notes in Computer Science, Springer International Publishing, pp. 159–176.

[97] GUGENHEIMER, J. *Nomadic virtual reality: overcoming challenges of mobile virtual reality head-mounted displays*. PhD thesis, Universität Ulm, 2020.

[98] GUGENHEIMER, J., MAI, C., MCGILL, M., WILLIAMSON, J., STEINICKE, F., AND PERLIN, K. Challenges Using Head-Mounted Displays in Shared and Social Spaces. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2019), CHI EA '19, Association for Computing Machinery, pp. 1–8.

[99] GUGENHEIMER, J., STEMASOV, E., FROMMEL, J., AND RUKZIO, E. ShareVR: Enabling co-located experiences for virtual reality between HMD and Non-HMD users. In *Conference on Human Factors in Computing Systems - Proceedings* (May 2017), vol. 2017-May, Association for Computing Machinery, pp. 4021–4033.

[100] GUGENHEIMER, J., STEMASOV, E., SAREEN, H., AND RUKZIO, E. FaceDisplay: Towards asymmetric multi-user interaction for nomadic virtual reality. In *Conference on Human Factors in Computing Systems - Proceedings* (New York, New York, USA, Apr. 2018), vol. 2018-April, Association for Computing Machinery, pp. 1–13.

[101] GUGENHEIMER, J., TSENG, W.-J., MHAIDLI, A. H., RIXEN, J. O., MCGILL, M., NEBELING, M., KHAMIS, M., SCHAUB, F., AND DAS, S. Novel Challenges of Safety, Security and Privacy in Extended Reality. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), CHI EA '22, Association for Computing Machinery, pp. 1–5.

[102] GUZMAN, J. A., THILAKARATHNA, K., AND SENEVIRATNE, A. Security and privacy approaches in mixed reality: A literature survey. *ACM Computing Surveys 52*, 6 (Oct. 2019), 1–37. arXiv: 1802.05797 Publisher: Association for Computing Machinery.

[103] HAN, J., COOLS, R., AND SIMEONE, A. L. The Body in Cross-Reality: A Framework for Selective Augmented Reality Visualisation of Virtual Objects.

[104] HAN, J., MOERE, A. V., AND SIMEONE, A. L. Foldable Spaces: An Overt Redirection Approach for Natural Walking in Virtual Reality. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Mar. 2022), pp. 167–175. ISSN: 2642-5254.

[105] HARMS, C., AND BIOCCA, . A. Internal consistency and reliability of the networked minds social presence measure. Tech. rep., 2004.

[106] HART, S. G. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society* (2006), pp. 904–908. ISSN: 10711813.

[107] HARTMANN, J., HOLZ, C., OFEK, E., AND WILSON, A. D. RealityCheck: Blending Virtual Environments with Situated Physical Reality. In *Proceedings*

of the 2019 CHI Conference on Human Factors in Computing Systems (New York, NY, USA, May 2019), CHI '19, Association for Computing Machinery, pp. 1–12.

[108] HARTMANN, J., YEH, Y.-T., AND VOGEL, D. AAR: Augmenting a Wearable Augmented Reality Display with an Actuated Head-Mounted Projector. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (2020), ACM Press, pp. 1–14.

[109] HEILIG, M. L. Stereoscopic-television apparatus for individual use, Oct. 1960.

[110] HEILIG, M. L. El cine del futuro: the cinema of the future. *Presence Teleoperators Virtual Environ. 1*, 3 (1992), 279–294.

[111] HEINRICH, F., APILLA, V., LAWONN, K., HANSEN, C., PREIM, B., AND MEUSCHKE, M. Estimating depth information of vascular models: A comparative user study between a virtual reality and a desktop application. *Computers & Graphics 98* (Aug. 2021), 210–217.

[112] HENRIKSEN, I. M., SKAAR, M., AND TJORA, A. The Constitutive Practices of Public Smartphone Use. *Societies 10*, 4 (Dec. 2020), 78. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[113] HEPPERLE, D., DIENLIN, T., AND WÖLFEL, M. Reducing the Human Factor in Virtual Reality Research to Increase Reproducibility and Replicability. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Oct. 2021), pp. 100–105.

[114] HUBENSCHMID, S., WIELAND, J., FINK, D. I., BATCH, A., ZAGERMANN, J., ELMQVIST, N., AND REITERER, H. ReLive: Bridging In-Situ and Ex-Situ Visual Analytics for Analyzing Mixed Reality User Studies. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), ACM, pp. 1–20.

[115] HUBENSCHMID, S., ZAGERMANN, J., DACHSELT, R., ELMQVIST, N., FEINER, S., FEUCHTNER, T., LEE, B., REITERER, H., AND SCHMALSTIEG, D. Hybrid User Interfaces : Complementary Interfaces for Mixed Reality Interaction.

[116] HUSUNG, M., AND LANGBEHN, E. Of portals and ORBs: An evaluation of scene transition techniques for virtual reality. In *ACM International Conference Proceeding Series* (New York, NY, USA, Sept. 2019), vol. 19, Association for Computing Machinery, pp. 245–254.

[117] IBAYASHI, H., SUGIURA, Y., SAKAMOTO, D., MIYATA, N., TADA, M., OKUMA, T., KURATA, T., MOCHIMARU, M., AND IGARASHI, T. Dollhouse VR: A multi-view, multi-user collaborative design workspace with VR technology. In

SIGGRAPH Asia 2015 Emerging Technologies, SA 2015 (New York, New York, USA, Nov. 2015), Association for Computing Machinery, Inc, pp. 1–2.

[118] IKEA. IKEA - Apps on Google Play. https://play.google.com/store/apps/details?id=com.ingka.ikea.app&hl=en&gl=US. Accessed: 2022-08-11.

[119] INNOVEGA. Ar/vr needs a contact lens. we built it. https://innovega.io/, Oct 2022. Accessed: 09-01-2023.

[120] IRLITTI, A., LATIFOGLU, M., ZHOU, Q., REINOSO, M. N., HOANG, T., VELLOSO, E., AND VETERE, F. Volumetric Mixed Reality Telepresence for Real-time Cross Modality Collaboration. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2023), CHI '23, Association for Computing Machinery, pp. 1–14.

[121] ISHII, A., TSURUTA, M., SUZUKI, I., NAKAMAE, S., SUZUKI, J., AND OCHIAI, Y. Let Your World Open: CAVE-based Visualization Methods of Public Virtual Reality towards a Shareable VR Experience. In *Proceedings of the 10th Augmented Human International Conference 2019 on - AH2019* (New York, New York, USA, 2019), ACM Press.

[122] ITOH, Y., LANGLOTZ, T., SUTTON, J., AND PLOPSKI, A. Towards Indistinguishable Augmented Reality: A Survey on Optical See-through Head-mounted Displays. *ACM Computing Surveys 54*, 6 (July 2021), 120:1–120:36.

[123] JAMES, R., BEZERIANOS, A., AND CHAPUIS, O. Evaluating the Extension of Wall Displays with AR for Collaborative Work. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2023), CHI '23, Association for Computing Machinery, pp. 1–17.

[124] JANSEN, P., FISCHBACH, F., GUGENHEIMER, J., STEMASOV, E., FROMMEL, J., AND RUKZIO, E. ShARe: Enabling co-located asymmetric multi-user interaction for augmented reality head-mounted displays. In *UIST 2020 - Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2020), Association for Computing Machinery, Inc, pp. 459–471.

[125] JERALD, J. *The VR Book: Human-Centered Design for Virtual Reality*. Morgan & Claypool, Sept. 2015. Google-Books-ID: ZEBiDwAAQBAJ.

[126] JOHNSTONE, R., MCDONNELL, N., AND WILLIAMSON, J. R. When Virtuality Surpasses Reality: Possible Futures of Ubiquitous XR. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), CHI EA '22, Association for Computing Machinery, pp. 1–8.

[127] JONES, B. R., BENKO, H., OFEK, E., AND WILSON, A. D. Illumiroom: Peripheral projected illusions for interactive experiences. In *ACM SIGGRAPH 2013 Emerging Technologies, SIGGRAPH 2013* (2013).

[128] JULIAN, A. Design Fiction: A Short Essay on Design, Science, Fact and Fiction. https://blog.nearfuturelaboratory.com/2009/03/17/design-fiction-a-short-essay-on-design-science-fact-and-fiction/, Mar. 2009. Accessed: 16-12-2022.

[129] JULIE. VR for Public and Social Settings | Future Immersive Interaction Group. http://juliericowilliamson.com/blog/vr-for-public-and-social-settings/. Accessed: 22-11-2022.

[130] JULIER, S., BAILLOT, Y., BROWN, D., AND LANZAGORTA, M. Information filtering for mobile augmented reality. *IEEE Computer Graphics and Applications 22*, 5 (Sept. 2002), 12–15. Conference Name: IEEE Computer Graphics and Applications.

[131] JUNG, T., TOM DIECK, M. C., LEE, H., AND CHUNG, N. Effects of Virtual Reality and Augmented Reality on Visitor Experiences in Museum. In *Information and Communication Technologies in Tourism 2016* (Cham, 2016), A. Inversini and R. Schegg, Eds., Springer International Publishing, pp. 621–635.

[132] KANDEL, J., DUPPEN, C., ZHANG, Q., JIANG, H., ANGELOPOULOS, A., NEALL, A. P.-A., WAGH, P., SZAFIR, D., FUCHS, H., LEWEK, M., AND SZAFIR, D. A. PD-Insighter: A Visual Analytics System to Monitor Daily Actions for Parkinson's Disease Treatment. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2024), CHI '24, Association for Computing Machinery, pp. 1–18.

[133] KAPTELININ, V., AND NARDI, B. *Activity Theory in HCI: Fundamentals and Reflections*. Synthesis Lectures on Human-Centered Informatics. Springer International Publishing, Cham, 2012.

[134] KARAOSMANOGLU, S., ROGERS, K., WOLF, D., RUKZIO, E., STEINICKE, F., AND NACKE, L. E. Feels like Team Spirit: Biometric and Strategic Interdependence in Asymmetric Multiplayer VR Games. Association for Computing Machinery (ACM), pp. 1–15.

[135] KARIM ISHIGAKI, S. A., ISMAIL, A. W., AND KAMRULDZAMAN, M. Q. MR-MEET: Mixed Reality Collaborative Interface for HMD and Handheld Users. In *2022 IEEE Global Conference on Computing, Power and Communication Technologies (GlobConPT)* (Sept. 2022), pp. 1–7.

[136] KASURINEN, J., PALACIN-SILVA, M., AND VANHALA, E. What Concerns Game Developers? A Study on Game Development Processes, Sustainability

and Metrics. In *2017 IEEE/ACM 8th Workshop on Emerging Trends in Software Metrics (WETSoM)* (May 2017), pp. 15–21. ISSN: 2327-0969.

[137] KENNEDY, R. S., LANE, N. E., BERBAUM, K. S., AND LILIENTHAL, M. G. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology 3*, 3 (July 1993), 203–220. Publisher: Lawrence Erlbaum Associates, Inc.

[138] KENT, L., SNIDER, C., GOPSILL, J., AND HICKS, B. Mixed reality in design prototyping: A systematic review. *Design Studies 77* (Nov. 2021), 101046.

[139] KESHAVARZI, M., YANG, A. Y., KO, W., AND CALDAS, L. Optimization and Manipulation of Contextual Mutual Spaces for Multi-User Virtual and Augmented Reality Interaction. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Mar. 2020), 353–362. Conference Name: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) ISBN: 9781728156088 Place: Atlanta, GA, USA Publisher: IEEE.

[140] KIM, D., JOSHI, N., AND VOGEL, D. Perspective and Geometry Approaches to Mouse Cursor Control in Spatial Augmented Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2023), CHI '23, Association for Computing Machinery, pp. 1–19.

[141] KIM, D., SHIN, J. E., LEE, J., AND WOO, W. Adjusting relative translation gains according to space size in redirected walking for mixed reality mutual space generation. In *Proceedings - 2021 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2021* (Mar. 2021), Institute of Electrical and Electronics Engineers Inc., pp. 653–660.

[142] KIM, D., AND VOGEL, D. Everywhere Cursor: Extending Desktop Mouse Interaction into Spatial Augmented Reality: Extended Abstract. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (New Orleans LA USA, Apr. 2022), ACM, pp. 1–7.

[143] KINETIC GAMES. Phasmophobia. https://www.kineticgames.co.uk/phasmophobia, 2025. Accessed: 25-01-2025.

[144] KISTER, U., KLAMKA, K., TOMINSKI, C., AND DACHSELT, R. GraSp: Combining Spatially-aware Mobile Devices and a Display Wall for Graph Visualization and Interaction. *Computer Graphics Forum 36*, 3 (2017), 503–514. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13206.

[145] KLEMMER, S. R., HARTMANN, B., AND TAKAYAMA, L. How bodies matter. Association for Computing Machinery (ACM), p. 140.

[146] KREIJNS, K., XU, K., AND WEIDLICH, J. Social Presence: Conceptualization and Measurement. *Educational Psychology Review 34*, 1 (Mar. 2022), 139–170.

[147] KREKHOV, A., PREUSS, D., CMENTOWSKI, S., AND KRÜGER, J. Silhouette Games: An Interactive One-Way Mirror Approach to Watching Players in VR. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Aug. 2020), 561–571. arXiv: 2008.02582 Publisher: ACM Place: New York, NY, USA.

[148] KRUSCHITZ, C., AND HITZ, M. The Anatomy of HCI Design Patterns. In *2009 Computation World: Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns* (Nov. 2009), pp. 202–207.

[149] KUDO, Y., TANG, A., FUJITA, K., ENDO, I., TAKASHIMA, K., AND KITAMURA, Y. Towards Balancing VR Immersion and Bystander Awareness. *Proceedings of the ACM on Human-Computer Interaction 5*, ISS (Nov. 2021), 1–22. Publisher: Association for Computing Machinery.

[150] KUMARAVEL, B. T., ANDERSON, F., FITZMAURICE, G., HARTMANN, B., AND GROSSMAN, T. Loki: Facilitating remote instruction of physical tasks using Bi-directional mixed-reality telepresence. In *UIST 2019 - Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (Oct. 2019), Association for Computing Machinery, Inc, pp. 161–174.

[151] KUMARAVEL, B. T., NGUYEN, C., DIVERDI, S., AND HARTMANN, B. TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators. ISBN: 9781450375146.

[152] LANGBEHN, E., LUBOS, P., AND STEINICKE, F. Evaluation of Locomotion Techniques for Room-Scale VR: Joystick, Teleportation, and Redirected Walking. In *Proceedings of the Virtual Reality International Conference (VRIC)* (2018), pp. 4–9.

[153] LANGNER, R., SATKOWSKI, M., BUSCHEL, W., AND DACHSELT, R. Marvis: Combining mobile devices and augmented reality for visual data analysis. In *Conference on Human Factors in Computing Systems - Proceedings* (May 2021), Association for Computing Machinery.

[154] LAZAR, J., FENG, J. H., AND HOCHHEISER, H. *Research Methods in Human-Computer Interaction*. Morgan Kaufmann, Apr. 2017. Google-Books-ID: hbkxDQAAQBAJ.

[155] LEBECK, K., RUTH, K., KOHNO, T., AND ROESNER, F. Towards Security and Privacy for Multi-user Augmented Reality: Foundations with End Users. In *2018 IEEE Symposium on Security and Privacy (SP)* (May 2018), pp. 392–408. ISSN: 2375-1207.

[156] LEDO, D., HOUBEN, S., VERMEULEN, J., MARQUARDT, N., OEHLBERG, L., AND GREENBERG, S. Evaluation Strategies for HCI Toolkit Research. In

Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (New York, NY, USA, Apr. 2018), CHI '18, Association for Computing Machinery, pp. 1–17.

[157] Lee, B., Cordeil, M., Prouzeau, A., Jenny, B., and Dwyer, T. A Design Space For Data Visualisation Transformations Between 2D And 3D In Mixed-Reality Environments. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), ACM, pp. 1–14.

[158] Lee, J., Kim, M., and Kim, J. RoleVR: Multi-experience in immersive virtual reality between co-located HMD and non-HMD users. *Multimedia Tools and Applications 79*, 1-2 (Jan. 2020), 979–1005. Publisher: Springer.

[159] Lee, J., Natarrajan, R., Rodriguez, S. S., Panda, P., and Ofek, E. RemoteLab: A VR Remote Study Toolkit. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2022), UIST '22, Association for Computing Machinery, pp. 1–9.

[160] Li, H., Zhang, X., Wang, H., Yang, Z., Liu, H., Cao, Y., and Zhang, G. Access to Nature via Virtual Reality: A Mini-Review. *Frontiers in Psychology 12* (Oct. 2021), 725288.

[161] Lubos, P., Bruder, G., Ariza, O., and Steinicke, F. Touching the Sphere: Leveraging Joint-Centered Kinespheres for Spatial User Interaction. In *Proceedings of the 2016 Symposium on Spatial User Interaction* (New York, NY, USA, Oct. 2016), SUI '16, Association for Computing Machinery, pp. 13–22.

[162] Luckerson, V. Google glass ends explorer beta program. https://time.com/3669927/google-glass-explorer-program-ends/, Jan 2015. Accessed: 09-01-2023.

[163] Martín-Gutiérrez, J., Luís Saorín, J., Contero, M., Alcañiz, M., Pérez-López, D. C., and Ortega, M. Design and validation of an augmented book for spatial abilities development in engineering students. *Computers & Graphics 34*, 1 (Feb. 2010), 77–91.

[164] Mcgill, M., Boland, D., Murray-Smith, R., and Brewster, S. A dose of reality: Overcoming usability challenges in VR head-mounted displays. In *Conference on Human Factors in Computing Systems - Proceedings* (New York, New York, USA, Apr. 2015), vol. 2015-April, Association for Computing Machinery, pp. 2143–2152.

[165] McGill, Mark. The IEEE Global Initiative on Ethics of Extended Reality (XR) Report–Extended Reality (XR) and the Erosion of Anonymity and Privacy. *Extended Reality (XR) and the Erosion of Anonymity and Privacy - White Paper* (Nov. 2021), 1–24. Conference Name: Extended Reality (XR) and the Erosion of Anonymity and Privacy - White Paper.

[166] McVeigh-Schultz, J., Kreminski, M., Prasad, K., Hoberman, P., and Fisher, S. S. Immersive Design Fiction: Using VR to Prototype Speculative Interfaces and Interaction Rituals within a Virtual Storyworld. In *Proceedings of the 2018 Designing Interactive Systems Conference* (New York, NY, USA, June 2018), DIS '18, Association for Computing Machinery, pp. 817–829.

[167] Medeiros, D., Anjos, R. D., Pantidi, N., Huang, K., Sousa, M., Anslow, C., and Jorge, J. Promoting reality awareness in virtual reality through proxemics. In *Proceedings - 2021 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2021* (Mar. 2021), Institute of Electrical and Electronics Engineers Inc., pp. 21–30.

[168] Medeiros, D., Dubus, R., Williamson, J., Wilson, G., Pöhlmann, K., and McGill, M. Surveying the social comfort of body, device, and environment-based augmented reality interactions in confined passenger spaces using mixed reality composite videos. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. 7*, 3 (sep 2023).

[169] Merino, L., Schwarzl, M., Kraus, M., Sedlmair, M., Schmalstieg, D., and Weiskopf, D. Evaluating Mixed and Augmented Reality: A Systematic Literature Review (2009-2019).

[170] Meta. Meta all-in-one sdk. https://developers.meta.com/horizon/downloads/package/meta-xr-sdk-all-in-one-upm/, 2025. Accessed: 25-01-2025.

[171] Microsoft. Microsoft-rocketbox. https://github.com/microsoft/Microsoft-Rocketbox, Dec 2020. Accessed: 27-05-2023.

[172] Microsoft. Microsoft flight simulator. https://www.flightsimulator.com/, 2025. Accessed: 25-01-2025.

[173] Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. Augmented Reality: A class of displays on the reality-virtuality continuum. Tech. rep., 1994. Publication Title: 282 / SPIE Volume: 2351.

[174] Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. Augmented reality: a class of displays on the reality-virtuality continuum. In *Telemanipulator and Telepresence Technologies* (Dec. 1995), vol. 2351, SPIE, pp. 282–292.

[175] Millette, A., and McGuffin, M. J. DualCAD: Integrating Augmented Reality with a Desktop GUI and Smartphone Interaction. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)* (Sept. 2016), pp. 21–26.

[176] MOGHADAM, K., BANIGAN, C., AND RAGAN, E. D. Scene Transitions and Teleportation in Virtual Reality and the Implications for Spatial Awareness and Sickness. *IEEE Transactions on Visualization and Computer Graphics 26*, 6 (June 2020), 2273–2287. Publisher: IEEE Computer Society.

[177] MOJO VISION. What if tomorrow you could... https://www.mojo.vision/, 2023. Accessed: 09-01-2023.

[178] MONTERO, C. S., ALEXANDER, J., MARSHALL, M. T., AND SUBRAMANIAN, S. Would you do that? understanding social acceptance of gestural interfaces. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services* (New York, NY, USA, Sept. 2010), MobileHCI '10, Association for Computing Machinery, pp. 275–278.

[179] MORTEZAPOOR, S., VASYLEVSKA, K., VONACH, E., AND KAUFMANN, H. CoboDeck: A Large-Scale Haptic VR System Using a Collaborative Mobile Robot. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)* (Mar. 2023), pp. 297–307. ISSN: 2642-5254.

[180] MOSSEL, A., VENDITTI, B., AND KAUFMANN, H. DrillSample: precise selection in dense handheld augmented reality environments. In *Proceedings of the Virtual Reality International Conference: Laval Virtual* (New York, NY, USA, Mar. 2013), VRIC '13, Association for Computing Machinery, pp. 1–10.

[181] NEBELING, M., LEWIS, K., CHANG, Y.-C., ZHU, L., CHUNG, M., WANG, P., AND NEBELING, J. XRDirector: A Role-Based Collaborative Immersive Authoring System. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2020), CHI '20, Association for Computing Machinery, pp. 1–12.

[182] NEBELING, M., SPEICHER, M., WANG, X., RAJARAM, S., HALL, B. D., XIE, Z., RAISTRICK, A. R. E., AEBERSOLD, M., HAPP, E. G., WANG, J., SUN, Y., ZHANG, L., RAMSIER, L. E., AND KULKARNI, R. MRAT: The Mixed Reality Analytics Toolkit. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2020), CHI '20, Association for Computing Machinery, pp. 1–12.

[183] NGUYEN, B. V. D., HAN, J., AND VANDE MOERE, A. Towards Responsive Architecture that Mediates Place: Recommendations on How and When an Autonomously Moving Robotic Wall Should Adapt a Spatial Layout. *Proceedings of the ACM on Human-Computer Interaction 6*, CSCW2 (Nov. 2022), 467:1–467:27.

[184] NGUYEN, B. V. D., SIMEONE, A. L., AND VANDE MOERE, A. Exploring an architectural framework for human-building interaction via a semi-immersive

cross-reality methodology. In *ACM/IEEE International Conference on Human-Robot Interaction* (Mar. 2021), IEEE Computer Society, pp. 252–261. ISSN: 21672148.

[185] Nguyen, B. V. D., and Vande Moere, A. The Adaptive Architectural Layout: How the Control of a Semi-Autonomous Mobile Robotic Partition was Shared to Mediate the Environmental Demands and Resources of an Open-Plan Office. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2024), CHI '24, ACM.

[186] Nguyen, P., Le, L. K.-D., Nguyen, D., Gao, L., Dunstan, D. W., and Moodie, M. The effectiveness of sedentary behaviour interventions on sitting time and screen time in children and adults: an umbrella review of systematic reviews. *International Journal of Behavioral Nutrition and Physical Activity 17*, 1 (Sept. 2020), 117.

[187] Norman, D. A. Stages and levels in human-machine interaction. *International Journal of Man-Machine Studies 21*, 4 (Oct. 1984), 365–375.

[188] Nreal Air. Give the gift of ar. https://www.nreal.ai/air/, 2023. Accessed: 09-01-2023.

[189] Oberdörfer, S., Fischbach, M., and Latoschik, M. E. Effects of VE transition techniques on presence, illusion of virtual body ownership, efficiency, and naturalness. In *SUI 2018 - Proceedings of the Symposium on Spatial User Interaction* (New York, NY, USA, Oct. 2018), Association for Computing Machinery, Inc, pp. 89–99.

[190] Odom, W., Zimmerman, J., Davidoff, S., Forlizzi, J., Dey, A. K., and Lee, M. K. A fieldwork of the future with user enactments. In *Proceedings of the Designing Interactive Systems Conference* (New York, NY, USA, June 2012), DIS '12, Association for Computing Machinery, pp. 338–347.

[191] O'Hagan, J., Gugenheimer, J., Bonner, J., Mathis, F., and McGill, M. Augmenting People, Places & Media: The Societal Harms Posed by Everyday Augmented Reality, and the Case for Perceptual Human Rights. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (New York, NY, USA, Dec. 2023), MUM '23, Association for Computing Machinery, pp. 225–235.

[192] O'Hagan, J., Saeghe, P., Gugenheimer, J., Medeiros, D., Marky, K., Khamis, M., and McGill, M. Privacy-Enhancing Technology and Everyday Augmented Reality: Understanding Bystanders' Varying Needs for Awareness and Consent. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 6*, 4 (Jan. 2023), 177:1–177:35.

[193] O'Hagan, J., and Williamson, J. R. Reality aware VR headsets. In *Proceedings - Pervasive Displays 2020: 9th ACM International Symposium on Pervasive Displays, PerDis 2020* (New York, NY, USA, June 2020), Association for Computing Machinery, Inc, pp. 9–17.

[194] Oikonomou, K., Chatzilari, E., Nikolopoulos, S., Kompatsiaris, I., Gavilan, D., and Downing, J. Snapwear: A Snapchat AR filter for the virtual tryon of real clothes. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Oct. 2021), pp. 46–51.

[195] Olin, P. A., Issa, A. M., Feuchtner, T., and Grønbæk, K. Designing for Heterogeneous Cross-Device Collaboration and Social Interaction in Virtual Reality. In *ACM International Conference Proceeding Series* (New York, NY, USA, Dec. 2020), Association for Computing Machinery, pp. 112–127.

[196] Olson, J. A., Sandra, D. A., Chmoulevitch, D., Raz, A., and Veissière, S. P. L. A nudge-based intervention to reduce problematic smartphone use: Randomised controlled trial.

[197] Onishi, Y., Takashima, K., Higashiyama, S., Fujita, K., and Kitamura, Y. WaddleWalls: Room-scale Interactive Partitioning System using a Swarm of Robotic Partitions. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2022), UIST '22, Association for Computing Machinery, pp. 1–15.

[198] Ouverson, K. M., and Gilbert, S. B. A Composite Framework of Co-located Asymmetric Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction 5*, CSCW1 (Apr. 2021), 1–20. Publisher: Association for Computing Machinery.

[199] Ouverson, K. M., Scherber, C., Oldham, E., and Gilbert, S. B. What Does "Asymmetric VR" Mean? A Directed Content Analysis of Co-Located Use of VR by Users on Reddit. *Frontiers in Virtual Reality 2* (Dec. 2021), 154. Publisher: Frontiers Media SA.

[200] Ozaki, H., Matoba, Y., and Siio, I. Can I GetToyIn? A Box Interface Connecting Real and Virtual Worlds. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2018), CHI EA '18, Association for Computing Machinery, pp. 1–4.

[201] Palmer, M. G., and Johnson, C. M. Experimenter presence in human behavior analytic laboratory studies: Confound it? *Behavior Analysis: Research and Practice 19* (2019), 303–314. Place: US Publisher: Educational Publishing Foundation.

[202] PARADISO, J. A., AND LANDAY, J. A. Guest Editors' Introduction: Cross-Reality Environments. *IEEE Pervasive Computing 8*, 3 (July 2009), 14–15. Conference Name: IEEE Pervasive Computing.

[203] PAUSCH, R., BURNETTE, T., BROCKWAY, D., AND WEIBLEN, M. E. Navigation and locomotion in virtual worlds via flight into hand-held miniatures. In *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics* (New York, New York, USA, 1995), ACM, pp. 399–400.

[204] PHAM, T., VERMEULEN, J., TANG, A., AND VERMEULEN, L. M. D. Scale impacts elicited gestures for manipulating holograms: Implications for AR gesture design. In *DIS 2018 - Proceedings of the 2018 Designing Interactive Systems Conference* (New York, NY, USA, June 2018), Association for Computing Machinery, Inc, pp. 227–240.

[205] PHOTON ENGINE. Photon. https://www.photonengine.com/, 2023. Accessed: 27-05-2023.

[206] PIUMSOMBOON, T., LEE, G. A., HART, J. D., ENS, B., LINDEMAN, R. W., THOMAS, B. H., AND BILLINGHURST, M. Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2018), ACM.

[207] PIUMSOMBOON, T., LEE, G. A., IRLITTI, A., ENS, B., THOMAS, B. H., AND BILLINGHURST, M. On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. In *Conference on Human Factors in Computing Systems - Proceedings* (New York, NY, USA, May 2019), Association for Computing Machinery, pp. 1–17.

[208] PIUMSOMBOON, T., ONG, G., URBAN, C., ENS, B., TOPLISS, J., BAI, X., AND HOERMANN, S. Ex-Cit XR: Expert-elicitation and validation of Extended Reality visualisation and interaction techniques for disengaging and transitioning users from immersive virtual environments. *Frontiers in Virtual Reality 3* (Dec. 2022). Publisher: Frontiers.

[209] PLASSON, C., BLANCH, R., AND NIGAY, L. Selection Techniques for 3D Extended Desktop Workstation with AR HMD. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Oct. 2022), pp. 460–469. ISSN: 1554-7868.

[210] POINTECKER, F., FRIEDL, J., SCHWAJDA, D., JETTER, H.-C., AND ANTHES, C. Bridging the Gap Across Realities: Visual Transitions Between Virtual and Augmented Reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Oct. 2022), pp. 827–836. ISSN: 1554-7868.

[211] Pooryousef, V., Cordeil, M., Besançon, L., Bassed, R., and Dwyer, T. Collaborative Forensic Autopsy Documentation and Supervised Report Generation Using a Hybrid Mixed-Reality Environment and Generative AI. *IEEE Transactions on Visualization and Computer Graphics 30*, 11 (Nov. 2024), 7452–7462. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[212] Qian, Y., Deng, J., and Chen, S. Application of Multiple Design Patterns in Virtual Reality Game Development. In *Proceedings of the 2023 7th International Conference on Big Data and Internet of Things* (New York, NY, USA, Nov. 2023), BDIOT '23, Association for Computing Machinery, pp. 103–108.

[213] Radianti, J., Majchrzak, T. A., Fromm, J., and Wohlgenannt, I. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers and Education 147* (Apr. 2020), 103778. Publisher: Elsevier Ltd.

[214] Ramsey Buchanan, L., Rooks-Peck, C. R., Finnie, R. K. C., Wethington, H. R., Jacob, V., Fulton, J. E., Johnson, D. B., Kahwati, L. C., Pratt, C. A., Ramirez, G., Mercer, S. L., and Glanz, K. Reducing Recreational Sedentary Screen Time: A Community Guide Systematic Review. *American Journal of Preventive Medicine 50*, 3 (Mar. 2016), 402–415.

[215] Rauh, S. F., Bogdan, C., Meixner, G., and Matviienko, A. Navigating the Virtuality-Reality Clash: Reflection and Design Patterns for Industrial Mixed Reality Applications. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (New York, NY, USA, July 2024), DIS '24, Association for Computing Machinery, pp. 2247–2266.

[216] Razzaque, S., Kohn, Z., and Whitton, M. C. Redirected Walking. *EUROGRAPHICS* (2001), 2–7.

[217] Reipschlager, P., Flemisch, T., and Dachselt, R. Personal Augmented Reality for Information Visualization on Large Interactive Displays. *IEEE Transactions on Visualization and Computer Graphics 27*, 2 (Feb. 2021), 1182–1192. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[218] Reipschläger, P., and Dachselt, R. DesignAR: Immersive 3D-Modeling Combining Augmented Reality with Interactive Displays. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (New York, NY, USA, Nov. 2019), ISS '19, Association for Computing Machinery, pp. 29–41.

[219] Reski, N., Alissandrakis, A., Tyrkkö, J., and Kerren, A. "Oh, that's where you are!" – Towards a Hybrid Asymmetric Collaborative Immersive

Analytics System. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (New York, NY, USA, Oct. 2020), NordiCHI '20, Association for Computing Machinery, pp. 1–12.

[220] RICO, J., AND BREWSTER, S. Usable gestures for mobile interfaces: evaluating social acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2010), CHI '10, Association for Computing Machinery, pp. 887–896.

[221] RIVU, R., ABDRABOU, Y., PFEUFFER, K., ESTEVES, A., MEITNER, S., AND ALT, F. Stare: Gaze-assisted face-to-face communication in augmented reality. In *ACM Symposium on Eye Tracking Research and Applications* (New York, NY, USA, 2020), ETRA '20 Adjunct, Association for Computing Machinery.

[222] ROBERTSON, J., AND KAPTEIN, M., Eds. *Modern Statistical Methods for HCI*. Human–Computer Interaction Series. Springer International Publishing, Cham, 2016.

[223] ROESNER, F., KOHNO, T., AND MOLNAR, D. Security and privacy for augmented reality systems. *Communications of the ACM 57*, 4 (Apr. 2014), 88–96.

[224] ROO, J. S., AND HACHET, M. One reality: Augmenting how the physical world is experienced by combining multiple mixed reality modalities. In *UIST 2017 - Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2017), Association for Computing Machinery, Inc, pp. 787–795.

[225] RUVIMOVA, A., KIM, J., FRITZ, T., HANCOCK, M., AND SHEPHERD, D. C. "Transport Me Away": Fostering Flow in Open Offices through Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2020), CHI '20, Association for Computing Machinery, pp. 1–14.

[226] RZAYEV, R., MAYER, S., KRAUTER, C., AND HENZE, N. Notification in VR: The effect of notification placement, task, and environment. In *CHI PLAY 2019 - Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (New York, NY, USA, Oct. 2019), Association for Computing Machinery, Inc, pp. 199–211.

[227] SALOVAARA, A., OULASVIRTA, A., AND JACUCCI, G. Evaluation of prototypes and the problem of possible futures. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2017), CHI '17, Association for Computing Machinery, p. 2064–2077.

[228] SCHJERLUND, J., HORNBÆK, K., AND BERGSTRÖM, J. OVRlap: Perceiving Multiple Locations Simultaneously to Improve Interaction in VR. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), CHI '22, Association for Computing Machinery, pp. 1–13.

[229] SCHMANDT, C. Input and display registration in a stereoscopic workstation. *Displays 5*, 2 (Apr. 1984), 89–92.

[230] SCHMIDT, A. Implicit human computer interaction through context. *Personal Technologies 4*, 2 (June 2000), 191–199.

[231] SCHMIDT, S., NUNEZ, O. J. A., AND STEINICKE, F. Blended Agents: Manipulation of Physical Objects within Mixed Reality Environments and Beyond. 1–10.

[232] SCHOMAKER, L. A taxonomy of Multimodal Interaction in the Human Information Processing System.

[233] SCHRÖDER, J.-H., SCHACHT, D., PEPER, N., HAMURCULU, A. M., AND JETTER, H.-C. Collaborating Across Realities: Analytical Lenses for Understanding Dyadic Collaboration in Transitional Interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2023), CHI '23, Association for Computing Machinery, pp. 1–16.

[234] SCHWIND, V., KNIERIM, P., HAAS, N., AND HENZE, N. Using Presence Questionnaires in Virtual Reality. 12–2019. ISBN: 9781450359702.

[235] SCHWIND, V., REINHARDT, J., RZAYEV, R., HENZE, N., AND WOLF, K. Virtual reality on the go? a study on social acceptance of VR glasses. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (New York, NY, USA, Sept. 2018), MobileHCI '18, Association for Computing Machinery, pp. 111–118.

[236] SCOTT, S. D., BESACIER, G., AND MCCLELLAND, P. J. Cross-device transfer in a collaborative multi-surface environment without user identification. In *2014 International Conference on Collaboration Technologies and Systems, CTS 2014* (2014), IEEE Computer Society, pp. 219–226.

[237] SERAJI, M. R., PIRAY, P., ZAHEDNEJAD, V., AND STUERZLINGER, W. Analyzing User Behaviour Patterns in a Cross-Virtuality Immersive Analytics System. *IEEE Transactions on Visualization and Computer Graphics 30*, 5 (May 2024), 2613–2623. Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[238] SERAJI, M. R., AND STUERZLINGER, W. HybridAxes: An Immersive Analytics Tool With Interoperability Between 2D and Immersive Reality Modes. In

2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) (Oct. 2022), pp. 155–160. ISSN: 2771-1110.

[239] SERAJI, M. R., AND STUERZLINGER, W. XVCollab: An Immersive Analytics Tool for Asymmetric Collaboration across the Virtuality Spectrum. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Oct. 2022), pp. 146–154. ISSN: 2771-1110.

[240] SERENO, M., BESANÇON, L., AND ISENBERG, T. Point specification in collaborative visualization for 3D scalar fields using augmented reality. *Virtual Reality 26*, 4 (Dec. 2022), 1317–1334.

[241] SERENO, M., GOSSET, S., BESANÇON, L., AND ISENBERG, T. Hybrid Touch/Tangible Spatial Selection in Augmented Reality. *Computer Graphics Forum 41*, 3 (2022), 403–415. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14550.

[242] SERENO, M., WANG, X., BESANCON, L., MCGUFFIN, M. J., AND ISENBERG, T. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* (Oct. 2020), 1–1. Publisher: Institute of Electrical and Electronics Engineers (IEEE).

[243] SHIGEYAMA, J., NARUMI, T., HASHIMOTO, T., TANIKAWA, T., YOSHIDA, S., AND HIROSE, M. Transcalibur: A Weight Shifting Virtual Reality Controller for 2D Shape Rendering based on Computational Perception Model. *Conference on Human Factors in Computing Systems - Proceedings* (2019), 1–11. ISBN: 9781450359702.

[244] SIMEONE, A., KHAMIS, M., ESTEVES, A., DAIBER, F., KLJUN, M., PUCIHAR, K., ISOKOSKI, P., AND GUGENHEIMER, J. *International Workshop on Cross-Reality (XR) Interaction*. Nov. 2020. Pages: 114.

[245] SIMEONE, A., WEYERS, B., BIALKOVA, S., AND LINDEMAN, R. W., Eds. *Everyday Virtual and Augmented Reality*. Human–Computer Interaction Series. Springer International Publishing, Cham, 2023.

[246] SIMEONE, A. L. The VR motion tracker: Visualising movement of non-participants in desktop virtual reality experiences. In *2016 IEEE 2nd Workshop on Everyday Virtual Reality, WEVR 2016* (Feb. 2017), Institute of Electrical and Electronics Engineers Inc., pp. 1–4.

[247] SIMEONE, A. L., COOLS, R., DEPUYDT, S., GOMES, J. M., GORIS, P., GROCOTT, J., ESTEVES, A., AND GERLING, K. Immersive Speculative Enactments: Bringing Future Scenarios and Technology to Life Using Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), CHI '22, Association for Computing Machinery, pp. 1–20.

[248] Simeone, A. L., Mavridou, I., and Powell, W. Altering User Movement Behaviour in Virtual Environments. *IEEE Transactions on Visualization and Computer Graphics* (2017).

[249] Simeone, A. L., Speicher, M., Molnar, A., Wilde, A., and Daiber, F. LIVE: the Human Role in Learning in Immersive Virtual Environments. In *Symposium on Spatial User Interaction* (New Orleans LA USA, Oct. 2019), ACM, pp. 1–11.

[250] Simeone, A. L., Velloso, E., and Gellersen, H. Substitutional reality: Using the physical environment todesign virtual reality experiences. In *Conference on Human Factors in Computing Systems - Proceedings* (Apr. 2015), vol. 2015-April, Association for Computing Machinery, pp. 3307–3316.

[251] Skarbez, R., Smith, M., and Whitton, M. C. Revisiting Milgram and Kishino's Reality-Virtuality Continuum. *Frontiers in Virtual Reality 2* (Mar. 2021), 647997. Publisher: Frontiers.

[252] Slater, M., Gonzalez-Liencres, C., Haggard, P., Vinkers, C., Gregory-Clarke, R., Jelley, S., Watson, Z., Breen, G., Schwarz, R., Steptoe, W., Szostak, D., Halan, S., Fox, D., and Silver, J. The Ethics of Realism in Virtual and Augmented Reality. *Frontiers in Virtual Reality 1* (2020).

[253] Slater, M., Usoh, M., and Steed, A. Depth of presence in virtual environments. *Presence: Teleoper. Virtual Environ. 3*, 2 (Jan. 1994), 130–144.

[254] Snap Inc. Spectacles by snap inc. the next generation of spectacles. https://www.spectacles.com/, 2023. Accessed: 09-01-2023.

[255] Snyder, H. Literature review as a research methodology: An overview and guidelines. *Journal of Business Research 104* (Nov. 2019), 333–339.

[256] Souza, M. T. d., Silva, M. D. d., and Carvalho, R. d. Integrative review: what is it? How to do it? *einstein (São Paulo) 8* (Mar. 2010), 102–106. Publisher: Instituto Israelita de Ensino e Pesquisa Albert Einstein.

[257] Speicher, M., Hall, B. D., and Nebeling, M. What is Mixed Reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, May 2019), CHI '19, Association for Computing Machinery, pp. 1–15.

[258] Speicher, M., Hall, B. D., and Nebeling, M. What is mixed reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), CHI '19, Association for Computing Machinery, p. 1–15.

[259] SPEICHER, M., HALL, B. D., YU, A., ZHANG, B., ZHANG, H., NEBELING, J., AND NEBELING, M. XD-AR: Challenges and Opportunities in Cross-Device Augmented Reality Application Development. *Proceedings of the ACM on Human-Computer Interaction 2*, EICS (June 2018), 7:1–7:24.

[260] SPEICHER, M., HELL, P., DAIBER, F., SIMEONE, A., AND KRÜGER, A. A virtual reality shopping experience using the apartment metaphor. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces* (New York, NY, USA, May 2018), AVI '18, Association for Computing Machinery, pp. 1–9.

[261] SPINDLER, M., BÜSCHEL, W., AND DACHSELT, R. Use your head: Tangible windows for 3D information spaces in a tabletop environment. In *ITS 2012 - Proceedings of the ACM Conference on Interactive Tabletops and Surfaces* (New York, New York, USA, 2012), ACM Press, pp. 245–254.

[262] SPINDLER, M., STELLMACH, S., AND DACHSELT, R. PaperLens: Advanced magic lens interaction above the tabletop. In *ITS 2009 - The ACM International Conference on Interactive Tabletops and Surfaces, Proceedings* (New York, New York, USA, 2009), ACM Press, pp. 69–76.

[263] SRA, M., MOTTELSON, A., AND MAES, P. Your Place and Mine: Designing a Shared VR Experience for Remotely Located Users. In *Proceedings of the 2018 Designing Interactive Systems Conference* (New York, NY, USA, June 2018), DIS '18, Association for Computing Machinery, pp. 85–97.

[264] STEED, A., IZZOUZI, L., BRANDSTÄTTER, K., FRISTON, S., CONGDON, B., OLKKONEN, O., GIUNCHI, D., NUMAN, N., AND SWAPP, D. Ubiq-exp: A toolkit to build and run remote and distributed mixed reality experiments. *Frontiers in Virtual Reality 3* (2022).

[265] STEFANIDI, H., LEONIDIS, A., KOROZI, M., AND PAPAGIANNAKIS, G. The ARgus Designer: Supporting experts while conducting user studies of AR/MR applications. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (Oct. 2022), pp. 885–890. ISSN: 2771-1110.

[266] STEINICKE, F., BRUDER, G., HINRICHS, K., LAPPE, M., RIES, B., AND INTERRANTE, V. Transitional environments enhance distance perception in immersive virtual reality systems. In *Proceedings - APGV 2009: Symposium on Applied Perception in Graphics and Visualization* (2009), pp. 19–26.

[267] STEINICKE, F., BRUDER, G., HINRICHS, K., AND STEED, A. Gradual transitions and their effects on presence and distance estimation. *Computers & Graphics 34*, 1 (Feb. 2010), 26–33.

[268] SUGIURA, Y., IBAYASHI, H., CHONG, T., SAKAMOTO, D., MIYATA, N., TADA, M., OKUMA, T., KURATA, T., SHINMURA, T., MOCHIMARU, M., AND IGARASHI,

T. An asymmetric collaborative system for architectural-scale space design. In *Proceedings - VRCAI 2018: 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry* (New York, NY, USA, Dec. 2018), Association for Computing Machinery, Inc, pp. 1–6.

[269] SUTHERLAND, I. E. Sutherland - The Ultimate Display. In *In Proceedings of IFIPS Congress*, vol. 2. 1965, pp. 506–508. arXiv: astro-ph.IM/1601.03459 ISSN: 0018-9162.

[270] SUTHERLAND, I. E. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I* (New York, NY, USA, Dec. 1968), AFIPS '68 (Fall, part I), Association for Computing Machinery, pp. 757–764.

[271] SUZUKI, R., HEDAYATI, H., ZHENG, C., BOHN, J. L., SZAFIR, D., DO, E. Y.-L., GROSS, M. D., AND LEITHINGER, D. RoomShift: Room-scale Dynamic Haptics for VR with Furniture-moving Swarm Robots. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2020), CHI '20, Association for Computing Machinery, pp. 1–11.

[272] SZALAVÁRI, Z., SCHMALSTIEG, D., FUHRMANN, A., AND GERVAUTZ, M. "Studierstube": An environment for collaboration in augmented reality. *Virtual Reality 3*, 1 (1998), 37–48. Publisher: Springer London.

[273] TCHAPMI, L., CHOY, C., ARMENI, I., GWAK, J., AND SAVARESE, S. SEGCloud: Semantic Segmentation of 3D Point Clouds. In *2017 International Conference on 3D Vision (3DV)* (Oct. 2017), pp. 537–547. ISSN: 2475-7888.

[274] THANYADIT, S., PUNPONGSANON, P., AND PONG, T. C. ObserVAR: Visualization system for observing virtual reality users using augmented reality. In *Proceedings - 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2019* (Oct. 2019), Institute of Electrical and Electronics Engineers Inc., pp. 258–268.

[275] THORAVI KUMARAVEL, B., AND WILSON, A. D. DreamStream: Immersive and Interactive Spectating in VR. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), ACM, pp. 1–17.

[276] TIDWELL, J. *Designing Interfaces: Patterns for Effective Interaction Design.* "O'Reilly Media, Inc.", Nov. 2005. Google-Books-ID: oCTSeKfMaWUC.

[277] TSENG, W.-J., BONNAIL, E., MCGILL, M., KHAMIS, M., LECOLINET, E., HURON, S., AND GUGENHEIMER, J. The Dark Side of Perceptual Manipulations in Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), CHI '22, Association for Computing Machinery, pp. 1–15.

[278] Turner, P., and Walle, G. V. D. Familiarity as a basis for universal design. *Gerontechnology 5*, 3 (July 2006), 150–159.

[279] Ugwitz, P., Šašinková, A., Šašinka, C., Stachoň, Z., and Juřík, V. Toggle toolkit: A tool for conducting experiments in unity virtual environments. *Behavior Research Methods 53*, 4 (Aug. 2021), 1581–1591.

[280] Unity. Xr interaction toolkit. https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@3.0, 2025. Accessed: 25-01-2025.

[281] Unity Technologies. Xr interaction toolkit: Xr interaction toolkit: 2.3.2. https://docs.unity3d.com/Packages/com.unity.xr.interaction.toolkit@2.3/manual/index.html, May 2018. Accessed: 27-05-2023.

[282] Unity Technologies. Unity game engine. https://unity.com/, 2023. Accessed: 27-05-2023.

[283] Usoh, M., Catena, E., Arman, S., and Slater, M. Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments 9*, 5 (2000), 497–503. ISBN: 1054-7460.

[284] Valkov, D., and Flagge, S. Smooth immersion: The benefits of making the transition to virtual environments a continuous process. In *SUI 2017 - Proceedings of the 2017 Symposium on Spatial User Interaction* (New York, NY, USA, Oct. 2017), vol. 2017-October, Association for Computing Machinery, Inc, pp. 12–19.

[285] Van Delm, N., Cools, R., and Simeone, A. L. Investigating the Effect of Material and Shape Abstraction on Object Recognition for Private Objects in Augmented Reality Environments. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia* (New York, NY, USA, Dec. 2023), MUM '23, Association for Computing Machinery, pp. 486–488.

[286] Vasser, M., Kängsepp, M., Magomedkerimov, M., Kilvits, K., Stafinjak, V., Kivisik, T., Vicente, R., and Aru, J. VREX: an open-source toolbox for creating 3D virtual reality experiments. *BMC Psychology 5*, 1 (Feb. 2017), 4.

[287] Vock, K., Hubenschmid, S., Zagermann, J., Butscher, S., and Reiterer, H. IDIAR: Augmented Reality Dashboards to Supervise Mobile Intervention Studies. In *Proceedings of Mensch und Computer 2021* (New York, NY, USA, Sept. 2021), MuC '21, Association for Computing Machinery, pp. 248–259.

[288] Von Willich, J., Funk, M., Müller, F., Marky, K., Riemann, J., and Mühlhäuser, M. You invaded my tracking space! Using augmented virtuality for spotting passersby in room-scale virtual reality. In *DIS 2019 - Proceedings*

of the 2019 ACM Designing Interactive Systems Conference (New York, NY, USA, June 2019), Association for Computing Machinery, Inc, pp. 487–496.

[289] WADHWA, M., AND SHRIVASTAVA, U. Ubiquitous Computing: A Comprehensive Review. In *Innovations in Electrical and Electronic Engineering* (Singapore, 2024), R. N. Shaw, P. Siano, S. Makhilef, A. Ghosh, and S. L. Shimi, Eds., Springer Nature, pp. 591–600.

[290] WANG, C.-H., CHEN, B.-Y., AND CHAN, L. RealityLens: A User Interface for Blending Customized Physical World View into Virtual Reality. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2022), UIST '22, Association for Computing Machinery, pp. 1–11.

[291] WANG, C. H., TSAI, C. E., YONG, S., AND CHAN, L. Slice of light: Transparent and integrative transition among realities in a multi-HMD-user environment. In *UIST 2020* (New York, NY, USA, Oct. 2020), Association for Computing Machinery, Inc, pp. 805–817.

[292] WANG, C.-H., YONG, S., CHEN, H.-Y., YE, Y.-S., AND CHAN, L. HMD Light: Sharing In-VR Experience via Head-Mounted Projector for Asymmetric Interaction. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2020), ACM, pp. 472–486.

[293] WANG, D., OHNISHI, K., AND XU, W. Multimodal Haptic Display for Virtual Reality: A Survey. *IEEE Transactions on Industrial Electronics 67*, 1 (Jan. 2020), 610–623.

[294] WANG, J., AND LINDEMAN, R. Coordinated 3D interaction in tablet- and HMD-based hybrid virtual environments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction* (New York, NY, USA, Oct. 2014), SUI '14, Association for Computing Machinery, pp. 70–79.

[295] WANG, N., AND MAURER, F. A Design Space for Single-User Cross-Reality Applications. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces* (New York, NY, USA, June 2022), AVI 2022, Association for Computing Machinery, pp. 1–5.

[296] WANG, N., ZIELASKO, D., AND MAURER, F. User Preferences for Interactive 3D Object Transitions in Cross Reality - An Elicitation Study. In *Proceedings of the 2024 International Conference on Advanced Visual Interfaces* (New York, NY, USA, June 2024), AVI '24, Association for Computing Machinery, pp. 1–9.

[297] WANG, X., BESANÇON, L., ROUSSEAU, D., SERENO, M., AMMI, M., AND ISENBERG, T. Towards an Understanding of Augmented Reality Extensions for

Existing 3D Data Analysis Tools. In *Conference on Human Factors in Computing Systems - Proceedings* (Apr. 2020), Association for Computing Machinery.

[298] WARE, C., ARTHUR, K., AND BOOTH, K. S. Fish tank virtual reality. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems* (New York, NY, USA, May 1993), CHI '93, Association for Computing Machinery, pp. 37–42.

[299] WARE, C., AND JESSOME, D. Using the bat: a six-dimensional mouse for object placement. *IEEE Computer Graphics and Applications 8*, 6 (Nov. 1988), 65–70. Conference Name: IEEE Computer Graphics and Applications.

[300] WEICHEL, C., HARDY, J., ALEXANDER, J., AND GELLERSEN, H. ReForm: Integrating Physical and Digital Design through Bidirectional Fabrication. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (New York, NY, USA, Nov. 2015), UIST '15, Association for Computing Machinery, pp. 93–102.

[301] WEISER, M. The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun. Rev. 3*, 3 (July 1999), 3–11.

[302] WELLER, S. C. Cultural Consensus Model. In *Encyclopedia of Social Measurement*. Elsevier Inc., Jan. 2004, pp. 579–585.

[303] WELSFORD-ACKROYD, F., CHALMERS, A., KUFFNER DOS ANJOS, R., MEDEIROS, D., KIM, H., AND RHEE, T. Spectator View: Enabling Asymmetric Interaction between HMD Wearers and Spectators with a Large Display. *Proceedings of the ACM on Human-Computer Interaction 5*, ISS (Nov. 2021). Publisher: Association for Computing Machinery.

[304] WHITE, M., GAIN, J., VIMONT, U., AND LOCHNER, D. The case for haptic props: Shape, weight and vibro-tactile feedback. In *Proceedings - MIG 2019: ACM Conference on Motion, Interaction, and Games* (Oct. 2019), Association for Computing Machinery, Inc.

[305] WHITLOCK, M., SMART, S., AND SZAFIR, D. A. Graphical Perception for Immersive Analytics. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Mar. 2020), pp. 616–625. ISSN: 2642-5254.

[306] WILLIAMSON, J. Being Social in XR with Stable Interpersonal Realities | Future Immersive Interaction Group. http://juliericowilliamson.com/blog/being-social-in-xr-with-stable-interpersonal-realities/. Accessed: 22-11-2022.

[307] WILLIAMSON, J. R., CROSSAN, A., AND BREWSTER, S. Multimodal mobile interactions: usability studies in real world settings. In *Proceedings of the 13th*

international conference on multimodal interfaces (New York, NY, USA, Nov. 2011), ICMI '11, Association for Computing Machinery, pp. 361–368.

[308] WILLIAMSON, J. R., MCGILL, M., AND OUTRAM, K. PlaneVR: Social acceptability of virtual reality for aeroplane passengers. In *Conference on Human Factors in Computing Systems - Proceedings* (New York, NY, USA, May 2019), Association for Computing Machinery, pp. 1–14.

[309] WOJCIECHOWSKI, R., WALCZAK, K., WHITE, M., AND CELLARY, W. Building Virtual and Augmented Reality museum exhibitions. In *Proceedings of the ninth international conference on 3D Web technology* (New York, NY, USA, Apr. 2004), Web3D '04, Association for Computing Machinery, pp. 135–144.

[310] WONG, R. When no one can go to school: does online learning meet students' basic learning needs? *Interactive Learning Environments 31*, 1 (Jan. 2023), 434–450. Publisher: Routledge _eprint: https://doi.org/10.1080/10494820.2020.1789672.

[311] WOODWORTH, J. W., BROUSSARD, D., AND BORST, C. W. Redirecting Desktop Interface Input to Animate Cross-Reality Avatars. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (Mar. 2022), pp. 843–851. ISSN: 2642-5254.

[312] WÖLFEL, M., HEPPERLE, D., PURPS, C. F., DEUCHLER, J., AND HETTMANN, W. Entering a new Dimension in Virtual Reality Research: An Overview of Existing Toolkits, their Features and Challenges. In *2021 International Conference on Cyberworlds (CW)* (Sept. 2021), pp. 180–187. ISSN: 2642-3596.

[313] XIA, H., HERSCHER, S., PERLIN, K., AND WIGDOR, D. SpaceTime: Enabling fluid individual and collaborative editing in virtual reality. In *UIST 2018 - Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA, Oct. 2018), Association for Computing Machinery, Inc, pp. 853–866.

[314] XIE, B., LIU, H., ALGHOFAILI, R., ZHANG, Y., JIANG, Y., LOBO, F. D., LI, C., LI, W., HUANG, H., AKDERE, M., MOUSAS, C., AND YU, L.-F. A Review on Virtual Reality Skill Training Applications. *Frontiers in Virtual Reality 2* (Apr. 2021). Publisher: Frontiers.

[315] YANG, K. T., WANG, C. H., AND CHAN, L. ShareSpace: Facilitating shared use of the physical space by both VR head-mounted display and external users. In *UIST 2018 - Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA, Oct. 2018), vol. 18, Association for Computing Machinery, Inc, pp. 499–509.

[316] YANG, Y., DWYER, T., WYBROW, M., LEE, B., CORDEIL, M., BILLINGHURST, M., AND THOMAS, B. H. Towards Immersive Collaborative Sensemaking. *Proc. ACM Hum.-Comput. Interact. 6*, ISS (Nov. 2022), 588:722–588:746.

[317] YAREMYCH, H. E., AND PERSKY, S. Tracing physical behavior in virtual reality: A narrative review of applications to social psychology. *Journal of Experimental Social Psychology 85* (Nov. 2019), 103845.

[318] YIXIAN, Y., TAKASHIMA, K., TANG, A., TANNO, T., FUJITA, K., AND KITAMURA, Y. ZoomWalls: Dynamic Walls that Simulate Haptic Infrastructure for Room-scale VR World. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, Oct. 2020), UIST '20, Association for Computing Machinery, pp. 223–235.

[319] YU, K., ECK, U., PANKRATZ, F., LAZAROVICI, M., WILHELM, D., AND NAVAB, N. Duplicated Reality for Co-located Augmented Reality Collaboration. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–1.

[320] ZHANG, X., AND SIMEONE, A. L. Using the Think Aloud Protocol in an Immersive Virtual Reality Evaluation of a Virtual Twin. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction* (New York, NY, USA, Dec. 2022), SUI '22, Association for Computing Machinery, pp. 1–8.

[321] ZHANG, Y., LIU, H., KANG, S.-C., AND AL-HUSSEIN, M. Virtual reality applications for the built environment: Research trends and opportunities. *Automation in Construction 118* (Oct. 2020), 103311.

[322] ZHOU, Q., FITZMAURICE, G., AND ANDERSON, F. In-Depth Mouse: Integrating Desktop Mouse into Virtual Reality. In *CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2022), ACM, pp. 1–17.

[323] ZHU, F., AND GROSSMAN, T. BISHARE: Exploring Bidirectional Interactions between Smartphones and Head-Mounted Augmented Reality. In *Conference on Human Factors in Computing Systems - Proceedings* (Apr. 2020), vol. 20, Association for Computing Machinery.

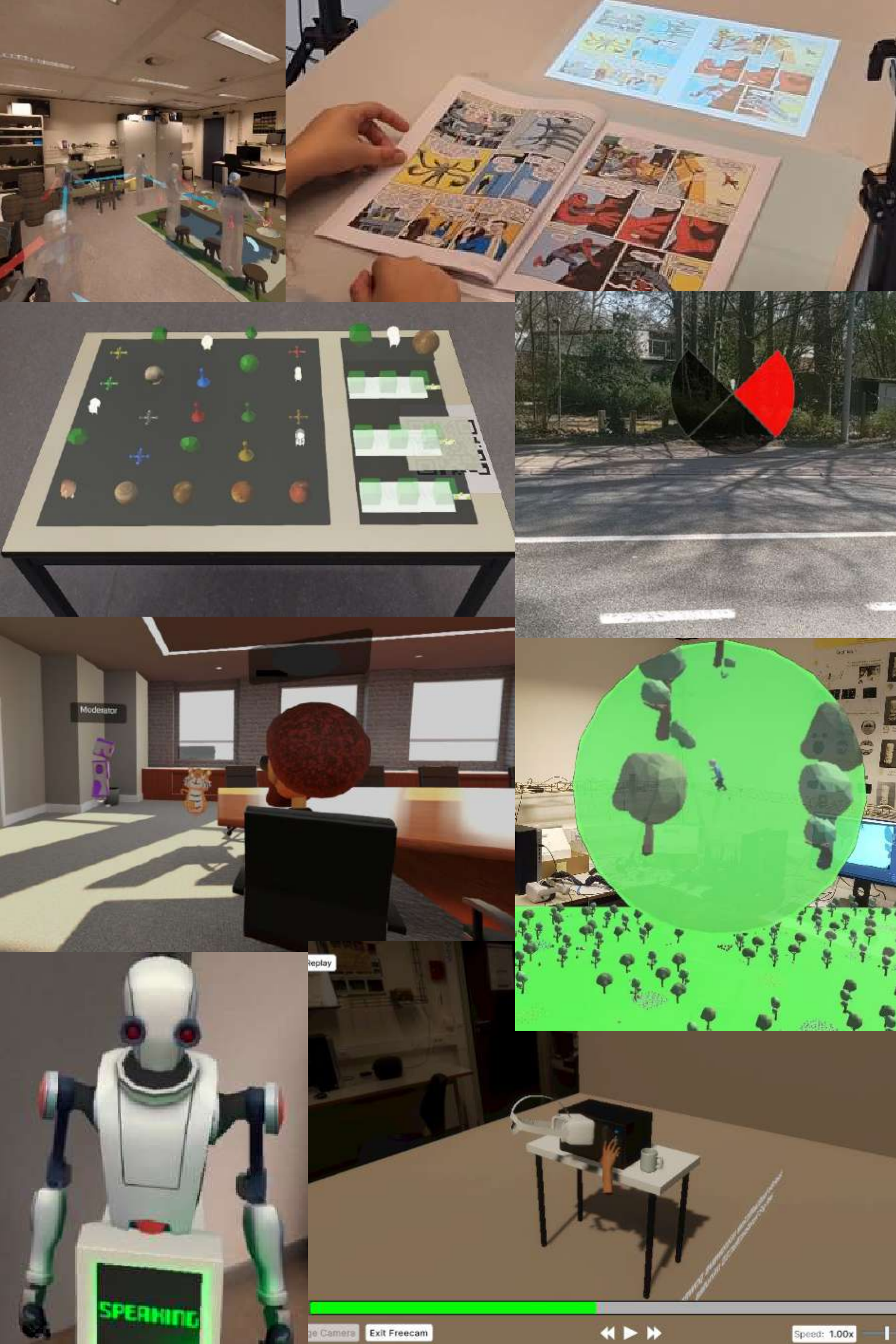
[324] ZIMMERMANN, L. "Your Screen-Time App Is Keeping Track": Consumers Are Happy to Monitor but Unlikely to Reduce Smartphone Usage. *Journal of the Association for Consumer Research 6*, 3 (July 2021), 377–382. Publisher: The University of Chicago Press.

Statement on the use of Generative AI

I did not use generative AI assistance tools during the research/writing process of my thesis, except for mere language assistance.
The text, code, and images in this thesis are my own (unless otherwise specified). Generative AI has only been used in accordance with the KU Leuven guidelines and appropriate references have been added. I have reviewed and edited the content as needed and I take full responsibility for the content of the thesis.

The following is a collage of supervised master students, collaborations, and other research that I was involved in that were not included of this thesis.


12:00
explore


Moderator
SPEAKING
Replay
Exit Freecam
Speed: 1.00x

FACULTY OF ENGINEERING SCIENCE
DEPARTMENT OF COMPUTER SCIENCE
Celestijnenlaan 200A box 2402
B-3001 Leuven
robbe.cools@kuleuven.be
https://aria.cs.kuleuven.be/