Tangible World-in-Miniature Interaction in Virtual Reality

by

Junhyeok Kim

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Computer Science

Waterloo, Ontario, Canada, 2026

© Junhyeok Kim 2026

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Barrett Ens Professor, Deptartment of Computer Science, Mathematics, Physics and Statistics University of British Columbia
Supervisor(s):	Mark Hancock Professor, Deptartment of Management Science and Engineering, University of Waterloo
Internal Member:	Daniel Vogel Professor, David R. Cheriton School of Computer Science, University of Waterloo
	Craig S. Kaplan Professor, David R. Cheriton School of Computer Science, University of Waterloo
Internal-External Member:	Ville Mäkelä Assistant Professor, Stratford School of Interaction Design and Business, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Portions of the materials in this dissertation have appeared in the following peer-reviewed conference publications. The chapters which make use of the material are noted below.

Chapter 3 consists of

- Junhyeok Kim and Mark Hancock. 2025. Investigating Digital, Tangible, and Paper-Based Room Design at a Small Scale. In *Proceedings of the Graphics Interface 2025 (GI '25)*. ACM, New York, NY, USA, 11 pages.

Chapter 4 was submitted as

- Junhyeok Kim and Mark Hancock. 2027. Retargeting Scale: Investigating Visual–Proprioceptive Mismatch in Tangible Worlds-in-Miniature. In *Proceedings of the 2027 CHI Conference on Human Factors in Computing Systems*

Chapter 5 and chapter 6 were submitted as

- Junhyeok Kim and Mark Hancock. 2027. Laminar: Leveraging Intermediary World-in-Miniature Spaces to Make Haptic Retargeting More Flexible. In *Proceedings of the 2027 CHI Conference on Human Factors in Computing Systems*

I was the primary writer and researcher responsible for all materials presented, and I conducted all research studies under Mark Hancock's supervision.

Abstract

Virtual reality (VR) provides a spatial platform for working with miniature representations and tangible proxies across multiple scales and perspectives. However, current world-in-miniature (WiM)-based VR systems face persistent challenges in scale perception, interaction continuity, and the tradeoff between interaction precision and speed. These limitations collectively hinder intuitive multi-scale spatial interaction in VR.

This dissertation investigates how tangible user interfaces (TUIs) with WiM-based VR can support interaction across multiple scales and views using tangible proxies. The research examines how users perceive and manipulate miniature representations, how visual–proprioceptive scale discrepancies affect interaction, and how such discrepancies can be managed through interaction design.

To investigate this problem, I conducted a progressive, multi-stage research. First, I conducted an exploratory mixed-methods study examining three miniature design tools: online planner, paper-and-pencil sketching, and LEGO assembly. The study used observation, the Desirability Toolkit, the Flow Scale, and semi-structured interviews. Second, I conducted a controlled in-lab VR experiment examining visual–proprioceptive perceptual tolerance using tangible proxies paired with differently scaled virtual objects. Measures included interaction accuracy, task completion time, perceived scale estimation, and subjective discrepancy. Finally, I synthesized these empirical insights and prior literature into a multidimensional design framework, then instantiated one configuration of the framework as *Laminar*, a prototype interaction technique.

This dissertation presents empirical results on how people use and experience three miniature-based interior design tools. It provides quantified evidence of users' perceptual responses to visual-proprioceptive mismatch when manipulating tangible proxies against differently scaled virtual objects in VR. Finally, I present a design framework and instantiated interaction technique, *Laminar*, that integrates TUI and WiM within VR to support multi-scalar interaction.

Together, this dissertation delivers evidence-informed interaction techniques and design frameworks for tangible WiM in VR, addressing collective challenges in scale perception, multi-scale manipulation, and multi-view workflow fragmentation. This work establishes theoretical and practical foundations for seamless, intuitive, and scalable multi-scalar spatial interaction in VR.

Acknowledgements

First and foremost, I thank my supervisor, Mark Hancock, for his unwavering support. Through struggles, difficult moments, and joyous writing sessions, his support made this PhD journey possible. I genuinely appreciate his guidance, patience, and encouragement throughout my PhD. I also want to thank my committee members, Daniel Vogel and Craig Kaplan, for their thoughtful feedback on my research, and Jian Zhao for his support of my PhD proposal. I thank my examiners, Barrett Ens and Ville Mäkelä, for their valuable feedback on my thesis.

Whether through casual conversations or formal research discussions, I was fortunate to receive support and feedback from my labmates, Marco, Marvin, Ali, Ariel, Cayley, and Caroline.

The research I carried out also benefited from significant support from the Games Institute's facilities and equipment. I want to thank Neil and Agata for their continued support of my research.

Attending the University of Waterloo was an amazing experience, and I thank the School of Computer Science, including all its staff and faculty, for the supportive academic environment. I also acknowledge the continuous financial support I received from my supervisor, Mark Hancock, and the Natural Sciences and Engineering Research Council of Canada (NSERC). This research, especially the user studies, would not have been possible without the study participants. I want to thank them for their feedback and participation.

Countless encounters with research and development problems made me realize I could not complete this journey alone. I want to acknowledge everyone I have met, spoken with, worked with, or become acquainted with as I bring this journey to its close.

AID statement

Artificial Intelligence Tool(s): Microsoft Copilot; **Writing–Review & Editing:** Copilot was used for textual refinement and structuring paragraphs in chapter 1, chapter 2, chapter 5, and chapter 6; **Data Analysis:** Copilot was used for scripting R language for data analysis in chapter 4, chapter 6.; **Privacy and Security:** no identifiable data was shared in any use of the tools.

Dedication

To Mom, who believes in me.

Table of Contents

Examining Committee ii

Author's Declaration iii

Statement of Contributions iv

Abstract v

Acknowledgements vi

AID statement vii

Dedication viii

List of Figures xvii

List of Tables xxii

List of Abbreviations xxiii

1 Introduction 1

1.1 Overview 2

1.2 Thesis Statement 4

1.3 Scope of the Dissertation 4

1.4 Research Questions 5
1.5 Research Structure 7
1.5.1 Exploratory Study: Problem Identification 7
1.5.2 Experimental Study: Empirical Grounding 7
1.5.3 Design Space and Framework: Structuring 8
1.5.4 Research-through-Design: Integration 8
1.6 Contributions 8
1.7 Outline 9

2 Related Work **11**
2.1 Miniature Representations and Spatial Design 11
2.1.1 The Role of Miniature in Spatial Design 11
2.1.2 User Groups and Simplified Spatial Design Tasks 13
2.1.3 From Miniature Models to Interactive Interfaces 13
2.1.4 WiMs as Interactive Spatial Representations 15
2.2 Scale Perception in VR 16
2.2.1 Distance, Depth, and Size Perception 16
2.2.2 Absolute, Relative, and Familiar Scale 17
2.2.3 Interactivity and Scale Perception 17
2.3 Tangible Interaction and Visual–Proprioceptive Discrepancy 18
2.3.1 Tangible Interaction for Manipulation 18
2.3.2 Visual-Haptic Integration and Passive Haptics 20
2.3.3 Visual–Proprioceptive Discrepancy 20
2.3.4 Haptic Retargeting Techniques 21
2.3.5 Tangible user interfaces and Body-based Interaction 21
2.4 Multi-view, Multi-scale Interaction in VR 22
2.4.1 World-in-miniature as a Multi-Scale Interface 22
2.4.2 Multi-view, Multi-scale VR Systems 23

2.4.3 Interaction Challenges in Multi-Scale Systems 25
2.4.4 Design Space and Framework for Spatial Interaction 26
2.5 Challenges and Research Gaps 26
2.5.1 Scale Perception Challenges 27
2.5.2 Interaction and Workflow Challenges 27
2.5.3 Experience and Embodiment Limitations 28
2.5.4 Tangible WiM-VR as an Underexplored Intersection 28
2.6 Summary and Research Direction 29

3 Investigating Digital, Tangible, and Paper-Based Room Design at a Small Scale 31

3.1 Introduction 32
3.2 The Study 33
3.2.1 Participants 33
3.2.2 Method 35
3.2.3 Measures & Data Collection 37
3.3 Findings: Observations 38
3.3.1 The Disparity in Object Scaling 38
3.3.2 Freedom of Expression 40
3.3.3 Design Processes with Each Tool 40
3.4 Findings: Desirability Toolkit 42
3.4.1 Engaging 43
3.4.2 Satisfying 43
3.4.3 Empowering 44
3.4.4 Motivating 44
3.4.5 Organized 44
3.5 Findings: Flow Scale 44
3.6 Findings: Interview 46

3.6.1 Choosing the Right Tool for Each Stage . . . 46
3.6.2 Need for Different Perspectives . . . 47
3.7 Discussion . . . 47
3.7.1 Design Implications . . . 48
3.7.2 Envisioning Future Tools for Everyone . . . 50
3.7.3 Understanding the Users . . . 51
3.7.4 Implication of Full Remote Study . . . 51
3.8 Chapter Summary . . . 52

4 Retargeting Scale — Investigating Visual–Proprioceptive Mismatch in Tangible Worlds-in-Miniature 53
4.1 Introduction . . . 54
4.2 Methodology . . . 56
4.2.1 Participants . . . 56
4.2.2 Apparatus . . . 56
4.2.3 Experimental Design . . . 59
4.2.4 Main Task . . . 61
4.2.5 Measures . . . 62
4.2.6 Procedure . . . 62
4.3 Results . . . 64
4.3.1 Simulator Sickness (SSQ) . . . 64
4.3.2 Quantitative Findings . . . 64
4.3.3 Qualitative Findings . . . 68
4.3.4 Observed Interaction (Experimenter Notes) . . . 70
4.4 Discussion . . . 70
4.4.1 Phased Motor Control: The Limits of Visual Dominance . . . 71
4.4.2 Feedback: The Cognitive Reality Check . . . 72
4.4.3 The Exception: Asymmetry in Tolerance . . . 73

4.5 Implications for Tangible WiM Design 74
4.5.1 Scalability with Positive Bias 74
4.5.2 Tradeoff in Miniaturization 74
4.5.3 Informative Haptics 75
4.5.4 Ergonomics of Grasp 75
4.6 Chapter Summary 75

5 Design Space and Framework for Tangible WiM-VR Interaction 77
5.1 Introduction 78
5.2 Exploring the Haptic, Intermediary, Virtual Design Space 79
5.3 Design Space Foundations 81
5.3.1 Virtuality 82
5.3.2 Interaction Spaces 82
5.3.3 Interaction States 83
5.4 Design Framework 86
5.4.1 Dimensions for Handling Mismatch 86
5.4.2 Spatial Analysis 88
5.4.3 Transformation Interval Analysis 90
5.5 Design Implications 92
5.5.1 Direct Coupling vs. WiM Mediation 93
5.5.2 Local Alignment vs. World-scale Flexibility 94
5.5.3 Designing with Asymmetric Tolerance 95
5.6 Configurations 96
5.6.1 Spatially-Decoupled Retargeting with WiM 97
5.6.2 Bimanual Interaction Space Control 97
5.6.3 Scale-Mismatched Tangible Proxy 100
5.7 Chapter Summary 102

6 Laminar–WiM Retargeting for Multi-Scalar Spatial Design 104
6.1 Introduction . . . 104
6.2 Deriving Laminar . . . 106
6.3 Laminar Overview . . . 107
6.3.1 Transforming Embodiment Space . . . 108
6.3.2 Transforming Intermediate Space . . . 109
6.3.3 Hybrid Technique . . . 109
6.4 Research Through Design (RtD) . . . 110
6.4.1 Iterative Design . . . 110
6.4.2 External Evaluation . . . 111
6.5 Iterative Design . . . 111
6.5.1 Iteration 1: Haptic Anchor . . . 111
6.5.2 Iteration 2: Contextual Scale . . . 113
6.5.3 Iteration 3: Scale Mismatch Compensation . . . 114
6.5.4 Iteration 4: Seamless Workflow . . . 115
6.5.5 Iteration 5: Positional Retargeting . . . 117
6.5.6 Iteration 6: Combined Retargeting . . . 118
6.6 Preliminary Evaluation . . . 120
6.6.1 Participants . . . 120
6.6.2 Apparatus . . . 120
6.6.3 Evaluated Configurations . . . 121
6.6.4 Procedure and Measures . . . 123
6.7 Findings . . . 123
6.7.1 Qualitative Feedback . . . 124
6.7.2 Subjective Measure Analysis . . . 127
6.8 Discussion . . . 129
6.8.1 Laminar as a Configurable Technique . . . 130
6.8.2 Limitations . . . 131
6.8.3 Broader Implications . . . 132
6.9 Chapter Summary . . . 133

7 Synthesis **134**

7.1 Redistributing Spatial Discrepancies Across Time 134

7.1.1 Triad of Geometric Mismatch 135

7.1.2 Temporal Redistribution Across Interaction States 136

7.2 Worlds-in-miniature as Abstraction 138

7.2.1 Tangible Interaction as a Grounding Anchor 138

7.2.2 Embracing Mismatch: Two Spatial Frames, Two Profiles 139

7.3 Retargeting as a Tradeoff between Illusion and Performance 140

7.4 Chapter Summary 142

8 Conclusion and Future Work **143**

8.1 Contributions 143

8.1.1 Characterizing Miniature-Based Design Tools 144

8.1.2 Empirical Grounding for Visual–Proprioceptive Mismatch 144

8.1.3 Design Framework for Tangible WiM-VR Interaction 144

8.1.4 Laminar as a Prototype Interaction Technique 145

8.2 Limitations 145

8.2.1 Lab Disruptions and Participant Samples 145

8.2.2 Study and Data Collection 146

8.2.3 Scope of the Task 146

8.2.4 Validation 146

8.3 Future Work 147

8.3.1 Formal Evaluation and Optimization 147

8.3.2 Intent Detection and Personalization 147

8.3.3 Expanded Setup 147

8.3.4 Generalizing Knowledge 148

8.4 Revisiting the Thesis Statement 148

References **150**

APPENDICES **169**

A Technical Implementation Details **170**

A.1 Developing a Finger-Tracking Glove 170

A.1.1 Using the OptiTrack Algorithm 170

A.1.2 Developing a Simplified Solution 171

B Study Materials and Data **174**

B.1 Materials and Data for Chapter 3 174

B.1.1 Print-outs Grid Papers 174

B.1.2 Instruction and Interview Guide 177

B.1.3 Final Design Snapshots 180

B.2 Materials and Data for Chapter 4 203

B.2.1 Observation Notes 203

B.2.2 Semi-structured Interview Questions 204

B.2.3 Participants' Response 204

B.3 Materials and Data for Chapter 6 208

B.3.1 Subjective Measures Surveys 208

B.3.2 Interview Guidelines 212

List of Figures

1.1 Scope of the dissertation. 4

2.1 Two examples of Miniature Tools 12

2.2 Early WiM usage in VR; Mini-replica superimposed on viewer screen (Pausch, Burnette, Brockway, & Weiblen, 1995) 14

3.1 How I recorded the videos. Participants working on the design. 33

3.2 Sofa in three different design tools, highlighted with red 38

3.3 Size (length) comparison of participants' sofa sizes. 38

3.4 Boxplots for each dimension and FSS 45

4.1 The overview of the study 54

4.2 The experimental apparatus: A Cardboard mass model of the target (side table) shown with a Vive controller and the complete master set of 10 tangible proxies to illustrate relative scale. B Custom half-finger glove with active optical markers, and its representation in VR. C The master set of 10 tangible proxies. D A panoramic view of the virtual environment 57

4.3 Experimental task view: (Left) User's view of the virtual environment showing the world-in-miniature and the target docking receptacle. (Right) A participant performing the docking task with a tangible proxy. 61

4.4 Quantitative results of the visual–proprioceptive discrepancy evaluation across three tangible proxy sizes. A Total Trial Time reveals an overall interaction effect. Decomposing the task reveals that B Docking Time incurs a severe speed-accuracy penalty during high sensory conflict, whereas C ballistic Moving Time is completely unaffected by visual mismatches. D Cognitive Bias (Q1) demonstrates that visual dominance is modulated by the physical size of the grasped proxy. E Subjective Sensation scores (Q2) confirm that participants reliably detected the intended illusions across all physical scales. 65

4.5 Phased visual haptic interaction model for tangible WiM. Vision drives feedforward planning, proprioception dominates transport, and precision docking exposes visual–proprioceptive (V–P) conflict. Positive mismatch creates empty space that can be tolerated, whereas negative mismatch produces collision cues that drive rejection. 71

5.1 Five phases used and my review process. 79

5.2 Interaction states used in the temporal design framework: (A) browsing, (B) reaching, (C) grasping, (D) manipulating, and (E) releasing. The diagrams illustrate one tangible-WiM example. 84

5.3 A matrix of reviewed techniques across the four framework dimensions. Colored cells indicate the route, mismatch allocation, tangibility, and visibility assigned to each technique. 86

5.4 Spatially-Decoupled Retargeting Configuration: The WiM is placed away from the primary forward direction so that head/body movement between the WiM and the full-scale view creates a temporal buffer for retargeting before contacting the object. 98

5.5 Bimanual Interaction Space Control Configuration: One hand maintains coupling with a WiM object while the other hand controls the WiM container, allowing the WiM–full-scale world relationship to change. 99

5.6 Scale-Mismatched Configuration: tangible proxy can be coupled to differently scaled virtual WiM objects. The transformations of the WiM container or hand representation are dynamically adjusted based on the ratio between the tangible proxy and the selected object. 102

6.1 Overview of the chapter. Laminar: A tangible proxy provides passive haptic grounding while virtual objects are manipulated through a highlighted WiM interaction volume and verified in a concurrent full-scale contextual view. Laminar applies scale and positional retargeting to align selected virtual objects with the tangible proxy, while full-scale feedback preserves spatial context during manipulation. 105

6.2 Laminar mismatch allocation strategies. Embodiment scaling compensates for scale mismatch by transforming the virtual hand representation (left). Intermediate scaling compensates by transforming the WiM interaction volume (right). Hybrid scaling distributes compensation across both spaces (centre). All strategies aim to achieve congruent visual–tactile feedback when the user grasps the object. 107

6.3 Iteration 1: Haptic anchor. The design process begins by selecting the tangible proxy (left) as physical embodiment space over a standard VR controller (right), with the tabletop serving as a physical constraint for stable manipulation. This establishes passive haptic grounding before introducing WiM mediation or retargeting. 112

6.4 Iteration 2: Contextual scale. The design introduces a virtual WiM as the local interaction space and a concurrent full-scale world view as contextual feedback. This establishes the embodiment–WiM–world structure used by Laminar, but leaves scale mismatch between the fixed proxy and differently sized virtual objects unresolved. 113

6.5 Iteration 3: Scale mismatch compensation. Hand scaling assigns compensation to virtual embodiment space by changing the hand representation relative to the tangible proxy. WiM scaling assigns compensation to virtual WiM space by changing the interaction volume around the selected object. Both strategies preserve proxy–object contact alignment but redistribute discrepancy to different spaces. 115

6.6 Iteration 4: Proxemic workflow control. The user's approach toward the WiM is used as an implicit control signal: moving closer expands the WiM and supports finer manipulation, while moving away restores broader contextual control. This embeds scale and speed–accuracy control in body movement rather than explicit mode switching. 116

6.7 Iteration 5: WiM-localized positional retargeting. Compared with classic haptic retargeting, which transforms embodiment or world spaces directly, Laminar localizes positional retargeting inside the WiM. The WiM container pans so that the selected virtual object aligns with the tangible proxy while the full-scale world remains available as contextual feedback. 117

6.8 Iteration 6: Combined retargeting. Laminar combines scale compensation with WiM-localized positional retargeting. The top row shows WiM scaling combined with positional retargeting, while the bottom row shows hand scaling combined with positional retargeting. Proxemic movement provides a temporal window in which the WiM can scale and reposition so that differently sized virtual objects can be mapped to a single tangible proxy. . 119

6.9 Apparatus used in the study 121

6.10 Evaluated scale-mismatch compensation configurations. In the hand-scaling condition, the hand representation scales according to the selected virtual object size (a–b). In the WiM-scaling condition, the WiM container scales to compensate for mismatch between the selected object and tangible proxy (c–d). 122

6.11 NASA-Task Load Index (TLX) 128

6.12 Custom Experience Likert 129

7.1 Three discrepancy relationships among the referent, WiM, and tangible proxy.135

A.1 Early sketch and prototype exploring marker placement and wire routing for the finger-tracking glove. 171

A.2 Early working prototype glove for finger/hand tracking 172

A.3 Nail rings and prototypes using them. 173

B.1 P1 Final design 180

B.2 P2 Final design 181

B.3 P3 Final design 181

B.4 P4 Final design 182

B.5 P5 Final design 183

B.6 P6 Final design 183

B.7 P7 Final design 184
B.8 P8 Final design 185
B.9 P9 Final design 185
B.10 P10 Final design 186
B.11 P11 Final design 187
B.12 P12 Final design 188
B.13 P13 Final design 189
B.14 P14 Final design 190
B.15 P15 Final design 191
B.16 P16 Final design 192
B.17 P17 Final design 193
B.18 P18 Final design 194
B.19 P19 Final design 195
B.20 P20 Final design 196
B.21 P21 Final design 197
B.22 P22 Final design 197
B.23 P23 Final design 198
B.24 P24 Final design 198
B.25 P25 Final design 199
B.26 P26 Final design 200
B.27 P27 Final design 201
B.28 P28 Final design 202
B.29 P29 Final design 202
B.30 P30 Final design 203

List of Tables

3.1 The breakdown of participants by profession. While most had planning or design experience with architecture or interior design, this included simple tasks like laying out a room (which I consider common for non-experts). Additionally, six participants had experience with architectural and 3D modelling tools, respectively (5 of these were the same participants), though typically skill with architectural tools was rated lower than for 3D modelling tools. 34

3.2 Frequency tables (left: from initial 30-word list, right: top 5 words) of word choices by condition. 41

3.3 Summary of the results and recommended design stages for each tool . . . 48

4.1 Experimental Variables. The values represent the **Mismatch Ratio** (V_{scale}/T_{scale}). Ratios > 1 indicate Positive Mismatch (Visual > Physical), while ratios < 1 indicate Negative Mismatch (Visual < Physical). 59

5.1 My 2 (virtuality) $\times$ 3 (interaction space) model that I use as the spatial properties that provide the foundation for my design space framework. . . 81

5.2 Spatial analysis. Proposed Laminar variants are shown in bold. Laminar-Hand and Laminar-WiM occupy two intermediary-space retargeting configurations identified through the design-space analysis. Hybrid allocation techniques appear in multiple rows. 89

5.3 Temporal analysis showing when techniques introduce discrepancy management and when they reach their intended state during interaction. Proposed Laminar variants are shown in bold. 91

List of Abbreviations

AEC architecture, engineering, and construction 3, 4, 11, 13, 19, 29, 32, 78, 134, 148

AR augmented reality 24, 32, 48–50, 52

BIM building information modeling 50

CAD computer-aided design 19, 27, 32, 36, 50, 51, 64

DIY do-it-yourself 13, 32

DOF degree of freedom 58, 61, 75, 114

HCI human-computer interaction 4, 7, 13, 22, 23

LLM large language model 68

MLE maximum likelihood estimation 74

MR mixed reality 82

PLA polylactic acid 60

RBI Reality-Based Interaction 110, 111, 114

RtD research through design 8, 103, 104, 106, 110, 131–133, 143, 145, 146

TUI tangible user interface v, 4, 9, 14, 18–20, 29, 78, 138, 143, 149

VR virtual reality v, xvii, xix, 1–11, 13–26, 28, 29, 31, 32, 46, 48–50, 52–57, 59–64, 70, 72–78, 80, 82, 83, 93, 100–106, 112, 115, 120, 124, 125, 129, 132–134, 136, 138, 140–145, 148, 149, 171

WiM world-in-miniature v, xiii, xiv, xvii–xx, 1–6, 8–11, 14–29, 31, 32, 52–56, 59, 63, 70–85, 87–89, 91–109, 112–119, 121–145, 147–149

XR extended reality 13, 14, 140

Chapter 1

Introduction

In virtual reality (VR), world-in-miniature (WiM) interaction (Stoakley, Conway, & Pausch, 1995) allows users to manipulate large virtual environments through miniature representations. This approach offers flexibility for interacting with objects at different scales, but it introduces challenges in managing the scales and perspectives required for complex tasks such as design activities. It places a significant burden on users to interpret and coordinate multiple representations, viewpoints, and scale relationships. In particular, maintaining perceptual consistency and effective manipulation across different scales remains a significant challenge.

Existing WiM systems are primarily visual interfaces. While they provide effective contextual awareness and navigation, they often rely on controllers or indirect input mechanisms that provide limited physical grounding. Tangible interaction offers an alternative by introducing a physical reference that users can touch and manipulate, potentially improving precision, spatial understanding, and consistency across scales. Prior work has explored tangible interfaces and passive haptics in VR, but their integration with WiM interactions remains limited. This dissertation investigates whether combining tangible interaction with WiM can address limitations in scale perception, interaction continuity, and multi-scale manipulation.

This dissertation investigates how people engage with miniature-based tools in spatial design tasks and examines how scale, miniature representation, and tangible interaction influence user experience. It analyzes how these tools support room layout activities and explores the problem space and the role of multi-view interaction in VR. This dissertation also develops a design space and framework for tangible WiM-based interaction for VR, and presents a prototype interaction technique that integrates WiM interactive volume with a

tangible proxy. This approach explores how WiM can be integrated into VR interaction to support design activities with a tangible proxy while managing different scales and perspectives through users' natural body movements.

In this chapter, I begin by introducing the core concepts of WiM, tangible interaction, and VR as the foundations of this dissertation. I then explain the implications of miniature representation and the challenges of using it in design activities. This chapter also frames the research questions within this scope, explains how my research approaches them, and outlines its contributions.

1.1 Overview

WiM is a technique that scales down a large environment or object, allowing users to view, navigate, and manipulate within a miniaturized world (Pausch et al., 1995; Stoakley et al., 1995). Miniaturized models provide a comprehensive overview of a larger environment while maintaining spatial relationships between points of interest. This approach facilitates navigation and pathfinding by reducing and abstracting complex elements of a full-scale world (LaViola, Feliz, Keefe, & Zeleznik, 2001; Pausch et al., 1995). While other domains, such as entertainment and simulation (Horiuchi, Inoue, & Okada, 2012; Underkoffler & Ishii, 1999), leverage WiMs, this dissertation focuses on a particular case of spatial design. WiMs are particularly useful for spatial designs, such as architecture and interior design, where designers reason about spaces and the spatial relationships between multiple objects, within different views.

Although the concept of WiM (Stoakley et al., 1995) and its use in VR (Danyluk, Ens, Jenny, & Willett, 2021) is well established, existing research has largely examined scale perception, miniature representations, and interaction techniques independently. How these factors interact during multi-scale design activities, particularly when tangible interaction is introduced into WiM-based VR environments, remains underexplored.. An early WiM concept was introduced as an interface for navigating virtual worlds and manipulating virtual objects (Stoakley et al., 1995). Improvements, such as changing view scales and roaming around virtual environments, have been introduced to enhance the user experience with WiMs (LaViola et al., 2001; Pausch et al., 1995; Wingrave, Haciahmetoglu, & Bowman, 2006). While WiM research has focused more on new visualization techniques (Danyluk et al., 2021), the interactive value of WiMs has been underexplored, particularly how users manipulate scaled objects and perceive their relationships with tangible proxies.

Tangible interaction with miniature representations enables direct manipulation of

scaled objects mapped to a full-scale counterpart, while maintaining the user's spatial context within a compact visualization. Digital and physical interactions with scaled items are used in many applications (e.g., design, architecture, and interior design). Miniature models are common in architecture, engineering, and construction (AEC) for their capability to provide new dimensions for expression and better visualization options (Hosokawa, Takeda, Shioiri, Hirano, & Tanaka, 2008; H.-J. Kim, Kim, & Nam, 2016; Rauterberg et al., 1998). Miniature-based interaction has also been extended to other applications beyond spatial design, such as producing, animation, and virtual prototyping (Horiuchi et al., 2012; Kakehi, Yamaoka, Akatsuka, & Naemura, 2009; Nakanishi, 2012). In education and learning purposes, it serves as an immersive and engaging experience tool in VR with WiM (Marougkas, Troussas, Krouska, & Sgouropoulou, 2024). Likewise, serious gaming or gamification using miniature representations also offers a unique experience, making it a valuable tool for both education and entertainment (Oyelere et al., 2020).

Despite these benefits, using WiM in VR introduces significant challenges at perceptual and interaction levels. Miniaturization forces users to adapt their perception across multiple scales and perspectives, requiring them to continuously interpret the relationship between the WiM and the full-scale world (Creem-Regehr, Stefanucci, & Bodenheimer, 2022; Huk, 2006; Renner, Velichkovsky, & Helmert, 2013). This can lead to significant visual discrepancies, resulting in misperceptions on scales where accurate judgment is involved in design processes. At the same time, traditional VR controllers and mid-air interaction often lack support for precise manipulation (Teather & Stuerzlinger, 2011). As a result, these factors contribute to cognitive and physical fatigue when navigating and interacting with objects in WiM, which can detract from a seamless experience across multiple views.

To address these challenges, this dissertation follows a progressive research structure. First, I investigate how people use existing miniature-based design tools, including physical block methods, and identify difficulties related to scale perception and judgment. Second, I examine how WiM in VR affects user perception and manipulation performance across different discrepancy levels between the visual representation and embodied interaction via tangible proxies. Third, I incorporate these insights into a design framework for creating WiM interaction in VR, with a focus on managing these discrepancies. Lastly, I instantiate key design elements within a defined design space into a prototype of a tangible WiM-VR interaction technique.

1.2 Thesis Statement

This dissertation argues that although WiMs are a widely accepted paradigm in AEC and extensively used for visualization in VR, their full potential as a primary interactive mechanism remains limited by challenges in scale perception, physical grounding, and fragmented interaction across multiple views and scales. This research contends that integrating tangible interaction and WiM-based interfaces in VR can support a balance between interaction performance and perceptual consistency in WiM-based interaction.

1.3 Scope of the Dissertation

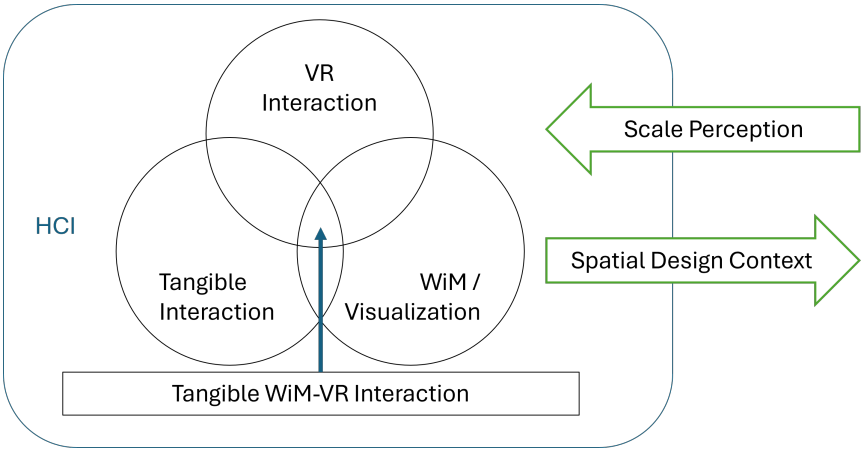


Figure 1.1: Scope of the dissertation.

This dissertation is situated at the intersection of human-computer interaction (HCI), VR interaction, visualization, tangible user interfaces (TUIs), WiMs, and perception in virtual environments. Although the work is motivated by architecture and interior design practices, this research provides evidence and empirical data about how users treat, experience, and expect their tools to perform during design activities. Specifically, while this work does not directly focus on design processes, a better understanding of how users perceive and perform such activities would help, and professionals can leverage the results of these insights to create more effective and engaging design tools. Therefore, the scope of this dissertation is the user experience of tangible WiM-based interaction in VR, especially when users interact with scaled elements and switch between view scales.

It aims to provide empirical and design-oriented evidence for user-centred development of VR systems, where spatial design practices intersect with human-computer interaction. Ultimately, this work aims to inform the development of design tools for broader user groups by exploring intuitive interaction methods with physical props. Although using such tools offers benefits for cooperative design practice, such as synchronous or asynchronous remote design sessions, this dissertation primarily focuses on perception, manipulation, and interaction with scaled virtual environments, rather than group use or collaborative workflows, which require different design and evaluation approaches.

This dissertation uses design activities as the main application context for WiM, as they provide several advantages as research materials. Although architecture and interior design are specialized professional domains, simplified design tasks (e.g., room layout and furniture placement) can be adapted for participants with varying levels of expertise and interests, making them suitable given the time required for controlled studies. Rather than replicating full-scale spatial design processes, this dissertation focuses on simplified spatial arrangement tasks as a way to examine interaction performance and perception. These tasks do not require extensive domain knowledge, while maintaining the key characteristics of '*design activity.*' Additionally, these miniature representations have been used for a long time in design activities and are in high demand among people with different expertise, expanding their applicable areas to virtual environments. Therefore, design processes in this work are simplified to accommodate a large participant pool and lower fatigue, including mental and physical load, while still serving the research purpose.

1.4 Research Questions

Prior work has identified multiple factors that influence user experience and performance in VR, including object scale, reference cues, and interaction technique. These studies examined user performance when adjusting a scaled item (Andrzejczak, Kozłowicz, Szrajber, & Wojciechowski, 2021; Combe, Posselt, & Kemeny, 2008) or explored the use of referent items to help perceive scaled items better (Hou, Cui, Chen, & Zeng, 2024). While this past work provides insight into factors that influence scale perception and interaction performance, it does not fully address how integrating tangible interaction, such as mapped physical props, affects WiM-based user experience in VR.

Scale perception with WiMs in VR interaction remains less understood. Additionally, how direct interaction with physical props representing miniaturized items affects user experience has been underexplored. Therefore, this dissertation investigates how users

perceive, manipulate, and reason about scaled representations across different miniature-based tools, and how these findings can inform tangible WiM-based interaction in VR.

This thesis is built around the following research questions:

RQ1. How do users perceive and interact with different miniature-based tools during room layout tasks?

RQ2. How do visual–proprioceptive scale discrepancies between tangible proxies and virtual representations affect scale perception and 3D manipulation performance in WiM-based VR?

RQ3. What are the key design elements and tradeoffs for managing discrepancies in tangible WiM-based VR interaction?

RQ4. How can scale discrepancies between WiM representations and tangible proxies be managed in VR to support haptic congruence and multi-view interaction?

RQ1 examines how users engage with miniature-based tools during room layout tasks. Miniature representations are common in spatial design, but how people perceive and interact with different sets of tools during these tasks remains unclear. This question motivates the exploratory study in chapter 3, which investigates how users work with different miniature-based design tools.

RQ2 focuses on scale discrepancies, especially when discrepancies arise between visual representations and their physical controls. Virtual objects and their miniature representations in WiM-based VR may not share the same scale. This question motivates the empirical study in chapter 4, which examines how visual–proprioceptive scale mismatch affects scale perception and 3D manipulation performance.

RQ3 builds on these findings by characterizing the design problem. Chapter 5 identifies key design elements, spatial relationships, and perceptual factors involved in tangible WiM-based interaction, and organizes them into a design space that captures the tradeoffs between design dimensions for such interaction.

RQ4 instantiates this design framework by exploring how design elements can be integrated into an interaction technique. This research question motivates chapter 6, which presents a prototype interaction technique that demonstrates how specific design decisions can be implemented to manage haptic congruence and multi-view interaction.

1.5 Research Structure

Since this dissertation investigates user experience in VR, the methodology draws primarily from HCI, and adapts methods from related disciplines for such purposes. The dissertation uses qualitative, quantitative, and design-oriented methods because the research problem involves both experiential and technical dimensions.

The dissertation follows a progressive and staged research design, allowing each subsequent research component to be grounded and motivated by insights from earlier research. This structure directly reflects the development of the research questions: first, identifying limitations in existing miniature-based design tools; second, empirically examining tangible scale mismatch in VR; and then, third, developing a design framework by analyzing interaction techniques within a design space; finally, synthesizing these insights and developing a prototype interaction technique from the design framework.

1.5.1 Exploratory Study: Problem Identification

The first study (chapter 3) used an exploratory approach using mixed methods to probe the problem space. These methods included semi-structured interviews, observation, think-aloud protocols, and self-report instruments such as the Desirability Toolkit (Benedek & Miner, 2002) and the Flow Scale (Jackson & Marsh, 1996). Qualitative analysis, including thematic analysis (Braun & Clarke, 2012, 2019), was used to understand participants' experiences and interaction strategies. These methods were paired with quantitative analysis where appropriate to identify patterns, such as scale misperception. This study addresses RQ1 of the dissertation by characterizing miniature-based tools for room layouts. It then reveals the limitations in existing miniature-based design methods and establishes the need for elevated tools to support interaction with miniature representations. The study informs the problem framing and design requirements that guide the subsequent research.

1.5.2 Experimental Study: Empirical Grounding

The second study (chapter 4) centers the problem defined in the earlier study to address the perceptual and proprioceptive gap when differently scaled virtual objects are paired with tangible proxies. This study addresses RQ2 concerned with tangible interaction, scale perception, and manipulation performance, grounding the research problem through empirical data. In a controlled study, I primarily used logged interaction data for performance

measures and behavioural analysis, as well as self-reporting perceptual responses for participants' perceived scales. These were used to examine how visual–proprioceptive scale mismatch affects scale perception and 3D manipulation performance in tangible WiM-based VR environments.

1.5.3 Design Space and Framework: Structuring

The scale mismatch and spatial discrepancies around WiMs are structured into a design framework. In chapter 5, I survey existing spatial interaction techniques in VR to probe how each technique distributes and manages these discrepancies. By constructing a design framework in temporal and spatial dimensions, the analysis exposes tradeoffs in choosing how mismatch is managed. In the analysis, I model these discrepancies in formulations as two *offsets* within the design space of this dissertation, tangible-WiM-VR. This design framework directly addresses RQ3 by showing how each design element leads to different tradeoffs.

1.5.4 Research-through-Design: Integration

The final research chapter (chapter 6) follows a research through design (RtD) approach, involving iterative prototyping, critique, reflection, and preliminary user feedback for an instantiated prototype interaction technique. This chapter focuses on analyzing design constraints by reviewing the literature and on system-level prototyping through iterative design processes, with preliminary user feedback on selected design alternatives within the design space.

1.6 Contributions

The objective of this dissertation is to understand and support interaction with scaled representations in VR, particularly in the context of tangible WiM-based interaction. This dissertation, therefore, investigates how users perceive and interact with miniature representations, examines how scale discrepancies affect perception and manipulation, and develops a structured design framework to inform the design of such interactions. Building on these insights, the dissertation produces the following contributions:

C1. An exploratory study of current design tools that examines the potential use of tangible interaction for small-scale objects.

C2. A user study that identifies what scales users can interact comfortably and confidently with in VR during design activities.

C3. A design framework for spatial interaction techniques that can be used when developing interaction techniques with scaled models and environments in VR.

C4. A prototype that demonstrates how a tangible proxy with WiM can be designed effectively in a full-scale virtual environment.

1.7 Outline

This dissertation begins by introducing the research problem and the key questions regarding WiM-based interaction in VR. It discusses how WiMs are used to support interaction across multiple scales and highlights the challenges introduced by discrepancies between the physical world and the virtual environment.

The following chapter, chapter 2, reviews existing literature on WiM, perception, TUI, and multi-view technique in VR. Then it positions the dissertation within these areas and identifies the gaps that motivate the research questions.

In chapter 3, an exploratory study examines how current miniature-based design tools are used in spatial layout tasks. The chapter analyzes both qualitative and quantitative data to characterize user workflows, interaction strategies, and difficulties related to scale perception and representation during room layout tasks.

In chapter 4, a controlled empirical study investigates how visual–proprioceptive discrepancies affect 3D manipulation performance and perceived object scales in WiM-based VR. This chapter provides empirical evidence of how scale mismatch influences both perception and interaction.

In chapter 5, the design space of tangible-WiM-VR in a spatial interaction context is explored. It identifies key design elements, relationships, and considerations, and organizes them into a structured framework. The design framework defines 5 interaction states that extend the insights from the previous study and provide key aspects of the design space.

In chapter 6, a prototype interaction technique is presented as an instantiation of the design framework. The chapter describes how this prototype technique addresses discrepancies between the real world and the virtual environment, promoting haptic congruence and continuous interaction.

In chapter 7, the findings from chapters are discussed and synthesized into higher-level insights and design implications for Tangible-WiM-VR interaction.

Finally, in chapter 8, I summarize the major contributions, discuss their implications, and outline the future direction of this research.

Chapter 2

Related Work

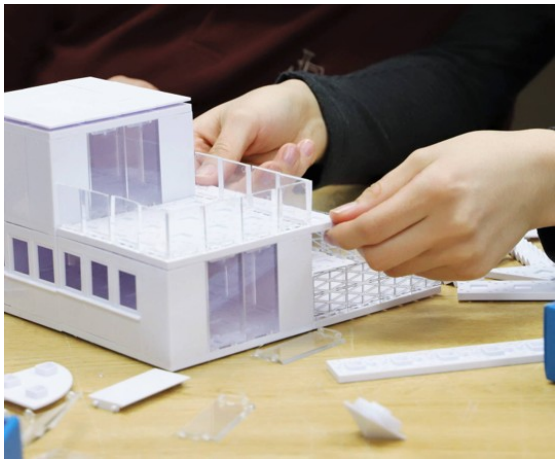
This chapter discusses how worlds-in-miniature (WiMs) have been used, and the current state of virtual reality (VR) and tangible interface development for such purposes.

2.1 Miniature Representations and Spatial Design

Miniature representations play a central role in spatial design, allowing designers to reason about large environments at a compact scale. Whether digital or physical, they support rapid iteration, overview, and comparison, making them essential tools in architecture, engineering, and construction (AEC). This section reviews the role of miniature representations and how they have evolved from static visualization to interactive systems. This establishes the conceptual foundation for understanding how WiMs have evolved and what characterizes WiM-based interaction.

2.1.1 The Role of Miniature in Spatial Design

Designing at scale, or working with miniature representations of larger physical spaces, is common practice in AEC and several other design areas (Morris, 2006). Miniature representations historically served as cognitive artifacts that allow practitioners and designers to reason about spatial relationships without the impracticality of full-scale implementation, which requires significant investment of time, labour, and material (Abruzzo, Ellingsen, & Solomon, 2007). For example, in architectural workflows, physical miniature models have



(a) ARCKit [a]



(b) LEGO [a]

[a](*Arckit 100 sqm. Architectural Model Building Kit*, 2026)

[a](*London 21034 | Architecture | Buy online at the Official LEGO® Shop CA*, 2026)

Figure 2.1: Two examples of Miniature Tools

been used for conceptual development and for communicating and presenting spatial concepts to clients, who can more readily interpret and understand three-dimensional models than technical drawings, which are mostly two-dimensional (Morris, 2006). These physical miniature representations help understand 3D models regardless of their actual scale and fill the gap between dimensional differences and spatial experiences.

Beyond their cognitive value, miniature prototypes reduce the need for full-scale testing during early design stages (Abruzzo et al., 2007; Morris, 2006). However, physical miniature prototyping still requires material and fabrication effort for each iteration. These factors can make physical miniatures resource-intensive, limiting the feasibility of repeated physical prototyping and underscoring the need for scalable, low-cost prototyping methods. Miniaturized architecture has already been supported, as in LEGO and ARCKit Figure 2.1, offering modular solutions for miniature physical prototyping, though these systems limit customization of geometry, surface, and material qualities. These constraints further highlight why practitioners and designers often rely on abstractions during prototyping, as abstractions provide the essential benefits of spatial reasoning while reducing the overhead associated with iterative design processes.

Overall, miniature representations highlight the practical value of abstraction and scaling for spatial design. They allow designers to reason about spaces and their relationships

across different scales while maintaining mappings between them, making them essential for spatial design tasks. Their roles form a foundation for interactive extended reality (XR) and VR miniature systems and motivate continued exploration of scale and spatial interaction in design contexts.

2.1.2 User Groups and Simplified Spatial Design Tasks

While miniature representations are widely accepted in AEC as practical tools because of their benefits, spatial design workflows often remain limited to a relatively small group of users. Prior work on design cognition shows fundamental differences in design approaches depending on domain knowledge and expertise (Ahmed, Wallace, & Blessing, 2003; Someren, Barnard, & Sandberg, 1994; Suwa & Tversky, 1997). These differences suggest that design tools must account for varying levels of expertise when design activities are made accessible to broader user groups.

Given movements such as Mass Architecture, and the increasing demand for related tasks (e.g., home do-it-yourself (DIY)) (Microsoft & Royal Institute of British Architects, 2018; Noghabaei, Heydarian, Balali, & Han, 2020; Vardouli & Buechley, 2014), making spatial design tasks accessible to a broader audience remains an important challenge, as AEC is still often considered closed to the general public (Vallerand, 2018). Simplified spatial design tasks provide a useful research context. Full professional AEC workflows involve lengthy timelines, external validation, and credentialing, making them difficult to use for studying interaction-level problems.

Room layout and 3D manipulation tasks preserve core aspects of spatial design while remaining accessible to participants without specialized training. Selecting, positioning, rotating, and evaluating objects are basic actions that design workflows frequently employ. In this dissertation, simplified room layouts and 3D manipulation tasks are used to isolate interaction problems related to miniature representations, scale perception, and tangible interaction. This down-toning of professional tasks still maintains design context, as these tasks involve reasoning about spatial relationships between objects. They preserve spatial reasoning and manipulation problems while making the studies suitable for controlled conditions and broader participant groups.

2.1.3 From Miniature Models to Interactive Interfaces

The human-computer interaction (HCI) community has explored the use of miniature and scaled objects as interactive elements, enabling interaction with complex and spatially

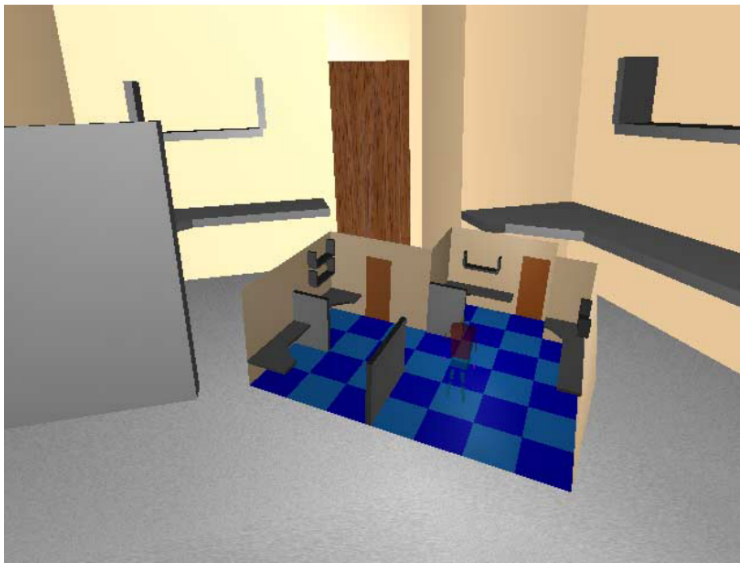


Figure 2.2: Early WiM usage in VR; Mini-replica superimposed on viewer screen (Pausch et al., 1995)

distant elements. The WiM metaphor (Stoakley et al., 1995) demonstrated that a WiM can function as an interactive tool rather than a static visualization in VR. In Stoakley et al.'s work, users manipulated a miniature replica of a virtual environment to navigate, select, and manipulate objects at a distance, with the ability to change points of view by rotating the WiM attached to a handheld device (physical clipboard and buttonball props). Their findings demonstrated that users can explore the virtual environment between miniature and full-scale views and minimize occlusion and context loss.

This WiM metaphor has been extended across various tangible user interface (TUI) and XR interfaces, leveraging the core benefit of miniature representations: they provide an overview of a large area and enable users to interact directly with scaled objects. Urp (Underkoffler & Ishii, 1999) leverages miniature representation for large-scale design contexts, such as urban planning, which is impractical at a one-to-one scale. The system illustrates how scaled physical models serve as both representations and physical controls for spatial simulation, such as analyzing shadows, wind velocity, proximity, and reflections. Similarly, a miniature design tool, BUILD-IT (Rauterberg et al., 1998), used physical blocks as controlling miniature representations in an architectural context, supporting natural interaction and rapid arrangement. Hosokawa et al. (2008) further showed that a tangible miniature-based interface supporting interior design tasks immediately displays its changes into computer-generated models. MiniStudio (H.-J. Kim et al., 2016) uses mini

objects to represent scaled-down versions of larger objects, which can be especially useful for large-scale projects, such as urban planning and architectural design. Collectively, these systems demonstrate how miniatures and their representations facilitate the design process, enabling rapid interaction and intuitive spatial reasoning in both professional practice and exploration.

Beyond architecture and design practices, the miniature metaphor has also been used in media production by researchers for character animation (Kakehi et al., 2009), which employed miniature objects for animation authoring. Horiuchi et al. (2012) also used physical miniatures to support theatrical staging and planning for multiple users. Prototyping using a miniature model on a public interactive display (Nakanishi, 2012) shows another example of the flexibility and adaptability of miniature metaphor across domains. Their effectiveness across diverse applications highlights the broader utility of miniature representations as a bridge between abstraction and embodied spatial experiences.

2.1.4 WiMs as Interactive Spatial Representations

WiMs build upon these advantages by transforming miniature representations into active interfaces. In a WiM system, a scaled-down model of an environment functions not only as a spatial overview but also as an interactive workspace through which users can perform tasks such as navigation, object selection, or environmental inspection. This provides a unified mechanism for switching across spatial scales, enabling users to maintain global awareness while performing detailed manipulations.

One of the primary reasons WiMs are widely used is that they excel at representing spatial relationships directly on the miniature, allowing users to interact with objects in ways that preserve their sense of spatial context. WiMs help mitigate common navigation challenges—such as disorientation, occlusion, or repeated viewpoint repositioning—by offering a stable external frame of reference. As spatial design tasks increasingly involve large or complex virtual environments, WiMs offer a practical means of bridging global and local viewpoints, supporting both exploration and controlled manipulation.

These systems demonstrate that WiMs can function as primary interfaces rather than passive visualizations. However, prior work often treats miniature interaction, physical control, and multi-scale navigation as separate design problems. This dissertation builds on these foundations by examining how miniature interaction, tangible proxies, and multi-view VR can be integrated into a coherent interaction framework.

2.2 Scale Perception in VR

Scale perception in VR is a complex interplay of factors, including depth, distance, and perspective (Creem-Regehr et al., 2022). This spatial ability directly affects users' experience and performance in virtual environments, as in design activities (Grechkin, Chihak, Cremer, Kearney, & Plumert, 2013; Ogawa, Narumi, & Hirose, 2018). This section reviews how scale perception affects the VR experience and why interaction with scaled objects is critical for understanding WiM-based VR.

2.2.1 Distance, Depth, and Size Perception

A general consensus in VR perception research is the compression of egocentric distance, where virtual spaces are perceived as smaller or more compressed than their real-world counterparts (Armbrüster, Wolter, Kuhlen, Spijkers, & Fimm, 2008; Waller & Richardson, 2008). This underestimation can limit design applications in VR, where spatial precision is important (Armbrüster et al., 2008). While distance underestimation is well established in numerous research studies (Kelly, Cherep, & Siegel, 2017; Thompson et al., 2004; Witmer & Kline, 1998), its specific impact on the perception of interactive scaled object size remains largely unknown.

The Size-Distance Invariance Hypothesis suggests that perceived size is interdependent with perceived distance (Gogel & Da Silva, 1987). Further research in VR suggests that object size may also be underestimated when distance is underestimated (Hornsey, Hibbard, & Scarfe, 2020). However, research has found that interaction within a virtual environment can recalibrate this perception, not only through walking in a virtual environment (Kelly, Donaldson, Sjolund, & Freiberg, 2013), but also through active interaction with virtual objects (Hornsey et al., 2020), warranting further investigation into interaction-dependent scale perception beyond controller-based interaction.

Depth cues also contribute to scale perception in VR. Widely used displays can limit certain depth cues that otherwise support accurate 3D object perception. The visual system processes binocular disparity as a major source of depth information (John, 2002; Rosahl, Gharabaghi, Hubbe, Shahidi, & Samii, 2006). While distance and depth perception are central to understanding 3D scenes, relatively less is known about how people perceive 3D size or volume, which is important for architecture, design, and 3D rendering.

Research also provides insight into how people use stereovision for spatial judgments. When determining a 3D object's distance, faster response times have been reported in

previous literature when stereoscopic depth cues were used (Yeh & Silverstein, 1992). Further research shows that parallax and stereovision affect perception of distance and scale, suggesting that binocular vision (disparity) can be a factor in determining distance and size. However, these effects can diminish as the distance increases between the object of interest and the viewer (Beall, Loomis, Philbeck, & Fikes, 1995).

2.2.2 Absolute, Relative, and Familiar Scale

Absolute size perception refers to perceiving the size of an object without additional contextual information, such as measurements of the surrounding environment or nearby objects (Gogel & Da Silva, 1987). Accurate absolute size perception is required in many applications, such as driving simulators or roadside simulations. Research shows that judgments of gaps between objects and distance perception in a virtual environment relate to performance in a real-world environment (Grechkin et al., 2013; Plumert & Kearney, 2014). This implies that design work in virtual environments also requires the same level of accurate size perception as in a real-world setting.

On the other hand, relative size perception occurs when an object of interest is interpreted in relation to another object or reference (Gogel & Da Silva, 1987; Ogawa et al., 2018). Similarly, various hand representations in VR affect how users perceive the scale of objects in or near their hands within a virtual environment (Ogawa et al., 2018). This is highly relevant to WiM-based interaction because users often judge scaled objects relative to other objects or their immediate surroundings.

A familiar scale refers to a mental reference based on prior knowledge of an object's typical size (Gogel & Da Silva, 1987). For example, a viewer is familiar with the size of a basketball in the real world and can tell the size approximately without seeing the real object. However, miniature and scaled designs can disrupt familiar scale perception because all objects are scaled by a specific ratio. Users therefore require a translated mental model based on the ratio provided in the miniature design, which could increase cognitive load (Huk, 2006). The use of familiarity in WiM settings can help users recognize the expected size or dimensions of objects. However, varied scales can still make accurate size perception difficult.

2.2.3 Interactivity and Scale Perception

As VR technology advances and becomes more interactive, users can manipulate 3D objects rather than only view them. Researchers have examined how users perceive objects and

how interaction affects perception in virtual environments. Previous literature shows that adding interaction to tasks helps people understand complex 3D objects, such as scientific images (Gagnier, Atit, Ormand, & Shipley, 2017). Together with stereovision, interactivity has also been reported to affect how users perceive their surroundings. For example, an interaction (e.g., walking or reaching) within a virtual environment recalibrates the scaling of perceived space (Kelly et al., 2013). Under similar conditions, interactivity within 3D models can make the model appear more interesting and encourage more active participation (Korakakis, Pavlatou, Palyvos, & Spyrellis, 2009).

However, providing interactivity does not necessarily improve performance, especially in spatial ability. Often, information relevant to the task is needed for improved performance (Keehner, Hegarty, Cohen, Khooshabeh, & Montello, 2008). Prior experience, such as 3D modelling expertise, can improve spatial reasoning regardless of gender or age (Šafhalter, Vukman, & Glodež, 2016).

These findings are important for WiM-based design tasks, where users interact with scaled objects while reasoning across multiple representations. Previous research shows that scale perception in VR depends on distance, depth cues, familiar and relative references, and interaction. However, less is known about how these factors shape perception and performance when users interact with scaled objects through tangible, physically grounded interfaces in WiM-based VR.

2.3 Tangible Interaction and Visual–Proprioceptive Discrepancy

Tangible user interfaces (TUIs) have long been recognized for their contributions to immersion and usability by connecting digital information with physical interaction (Bozgeyikli & Bozgeyikli, 2021; Ishii & Ullmer, 1997). However, integrating TUIs into complex virtual environments quickly becomes impractical due to physical constraints and the cost of fabricating a unique proxy for each virtual object. Consequently, research has increasingly focused on strategies for remapping and reusing a limited set of physical objects across multiple virtual objects.

2.3.1 Tangible Interaction for Manipulation

While design communities have shifted toward more computer-aided environments and adopted VR technology (Microsoft & Royal Institute of British Architects, 2018), physical

modelling remains an important basis for prototyping. This is especially relevant for spatial design, where physical representations of design elements allow users to directly manipulate objects and reason about their relationships at scale.

Computer-aided design (CAD) tools in AEC often require extensive training because they offer many features for complex design work. Tangible interfaces introduce alternative ways to interact with objects by integrating physical objects that represent digital counterparts. Bricks (Fitzmaurice, Ishii, & Buxton, 1995) introduced a graspable user interface with physical objects that can be mapped to virtual objects, establishing a baseline for TUIs research. Similarly, Urp (Underkoffler & Ishii, 1999) employs the concept of a graspable user interface in an urban planning tool suitable for tabletop settings. Schubert, Riedel, and Petzold (2013) used tangible interfaces for architectural design, specifically the early stage of design processes. The prototype used scaled, simplified physical props to represent individual buildings. As a similar concept in a tabletop setting, Salim et al. (2011) proposed a framework for urban design using physical models. TUIs align with WiMs because physical artifacts can function as both representations and interaction handles. As Ullmer and Ishii (2000) pointed out, these physical representations are used not only as input devices but also as output or display elements. However, VR separates input and output spaces, as users see the simulated environment on the display while still physically acting in the real world. This separation creates both an opportunity for passive haptics and a challenge for maintaining alignment between what users see and what they feel (Azmandian, Hancock, Benko, Ofek, & Wilson, 2016).

The lack of skills and design training is a significant barrier, making novices reluctant to be active participants (Booth, Taborda, Ramani, & Reid, 2016). Physical interfaces, such as LEGO, help users become more active participants and generate more elaborate and creative ideas during the design process (Bourdeau, Lesage, Couturier Caron, & Léger, 2020). Physical interfaces and props are used during prototyping and are standard practice in AEC, where they are considered essential across various design stages (Brereton & McGarry, 2000).

Tangible interfaces can bridge the gap between users and complex systems by providing intuitive, expressive, and physical controls without requiring extensive knowledge of professional design tools. Domains that require specific expertise, such as architecture, are in high demand for presentation methods that are easier for more users to understand, since reading complex technical drawings can be intimidating.

2.3.2 Visual-Haptic Integration and Passive Haptics

While virtual environments are visually dominant, interaction fidelity, including tangible interaction, plays a crucial role in spatial understanding (Feick, Bateman, Tang, Miede, & Marquardt, 2020; Lee-Cultura & Giannakos, 2020; Maher & Kim, 2005). Passive haptics have been used to increase presence and improve manipulation performance in VR (Medeiros, Wilson, Mcgill, & Brewster, 2023). Integrating a tangible interface into VR provides proprioceptive and haptic feedback (Choi, Culbertson, Miller, Olwal, & Follmer, 2017; Harley, Tarun, Germinario, & Mazalek, 2017). This feedback serves as a sensory anchor when users hold a physical object that corresponds to a virtual object (Cuijpers, Brenner, & Smeets, 2008; Swapp, Pawar, & Loscos, 2006). This can provide a strong feeling of coupling between the tangible proxy and the virtual object, complementing visual information during interaction (Feick et al., 2020; Ishii, 2008b).

Although integrating TUIs offers clear benefits, scaling this approach to complex virtual environments remains difficult. The repurposing of physical proxies mitigates this issue by reducing physical setup complexity, enhancing haptic feedback for richer interaction, and enabling larger and more dynamic virtual scenes. However, reusing physical proxies also introduces a significant challenge: the tangible proxy and virtual object do not always match in shape, position, or scale.

2.3.3 Visual–Proprioceptive Discrepancy

A recent study highlights discrepancies between tangible proxies and their virtual counterparts, as well as how these discrepancies depend on object size (Zheng et al., 2024). However, it remains unclear how discrepancies affect interactions in multi-scale environments. In spatial design applications such as WiM, the visual scale changes dynamically while the physical proxy remains constant. This creates a visual–proprioceptive discrepancy: users see an object at one scale or location while physically interacting with a proxy that may not match that representation.

This issue becomes more complex in WiM-based VR, where users may need to coordinate among the full-scale virtual environment, the WiM representation, and the tangible proxy simultaneously. Many existing remapping or retargeting systems typically operate within a single view or single scale, or focus on manipulating individual objects rather than supporting continuous transitions across different scales (Zenner & Krüger, 2019). As a result, visual–proprioceptive discrepancy remains an open problem for tangible WiM-based VR interaction.

2.3.4 Haptic Retargeting Techniques

Haptic retargeting addresses this issue by dynamically mapping a single tangible proxy or a constrained physical workspace to multiple virtual objects or larger virtual areas (Azmandian et al., 2016; Muender et al., 2019; Razzaque, Swapp, Slater, Whitton, & Steed, 2002). This technique leverages perceptual thresholds for visual–haptic discrepancies, allowing users to feel physical contact with virtual objects that are beyond their reach or outside the boundaries of the physical proxy itself. This expands the perceived tangibility and interaction fidelity of virtual environments while minimizing the required physical setup (Hettiarachchi & Wigdor, 2016).

More recent approaches further address the physical-virtual alignment by dynamically aligning virtual objects with physical proxies or by leveraging visual dominance to enhance the perceived tangibility of virtual content (Azmandian et al., 2016; Cheng, Ofek, Holz, Benko, & Wilson, 2017; Hettiarachchi & Wigdor, 2016; Z. Liu, Wu, Wang, Li, & Im, 2025; Matthews, Reichherzer, & Smith, 2023; Pohl, Lilija, McIntosh, & Hornbæk, 2021). Current research often relies on visual dominance, where the visual system can override or suppress small proprioceptive discrepancies to maintain a coherent experience (Ernst & Banks, 2002; Norman, Norman, Clayton, Lianekhammy, & Zielke, 2004). Haptic retargeting (Azmandian et al., 2016), for example, has leveraged this by introducing deliberate discrepancies between the physical and virtual hands, as well as between physical and virtual environments, to facilitate retargeting techniques that repurpose a single proxy across multiple virtual objects. However, the tolerance limits of this discrepancy, specifically when interacting with scaled virtual objects, remain underexplored.

2.3.5 Tangible user interfaces and Body-based Interaction

Techniques such as haptic retargeting often depend on users' body movement, hand motion, and attentional cues (gaze or head rotation) during interaction. These movements can introduce visual–proprioceptive discrepancies, but they can also provide signals for managing those discrepancies. This motivates implicit interaction approaches that use body movement and proxemic cues to adjust scale, focus, or mappings without requiring explicit mode switching.

Task continuity is essential for sustaining interactivity during multi-scale interaction. Proxemic interaction has been proposed as a means to reduce the need for explicit commands, and using proxemic cues can reduce redundant degrees of freedom in interaction design (Ballendat, Marquardt, & Greenberg, 2010). In tangible WiM-based VR, this is

highly relevant because body movement is already part of the interaction. Users already use their head, hands, and gaze to browse, inspect, reach toward objects, and coordinate between miniature and full-scale views. In WiM-based systems, proxemic and egocentric body-based cues can support interaction with both miniature and full-scale views. For example, head-distance-based zoom can use the user's distance from a representation as an implicit signal for changing scale or focus (Duchowski, 2002). Similarly, head motion can be translated into implicit interaction that reflects users' contextual focus (Harrison, Ishii, Vicente, & Buxton, 1995; Ware & Lewis, 1995). These examples show that interactions do not always need to be entirely controlled by explicit commands.

Proxemic and body-motion-based cues can be understood as proprioceptive signals already produced by users. Repurposing a tangible proxy across multiple virtual objects requires the system to determine when and how to adjust mappings between the user's body, the tangible proxy, and the virtual representation (Azmandian et al., 2016; Razzaque et al., 2002; Steinicke, Bruder, Jerald, Frenz, & Lappe, 2010). Because users naturally move, lean, reach, inspect, and reposition themselves during interactions, these movements can act as implicit indicators of interaction intent. For example, leaning in can signal the intent to inspect in more detail, as head movement and viewing distance have been used as implicit cues for focus-dependent interaction (Duchowski, 2002). As users lean in to engage with a specific object, this body movement can act as a precursor to a high-precision manipulation. This suggests that proxemics can be used as an implicit interaction paradigm (Ju & Leifer, 2008; Vogel & Balakrishnan, 2004), transforming natural body movement into continuous interaction commands. For tangible WiM-based VR, proxemic and body-based cues can help manage spatial discrepancies and proxy reuse without relying solely on explicit mode switching or controller input.

2.4 Multi-view, Multi-scale Interaction in VR

WiMs have been used to support interaction with large virtual environments by separating overview and detail views (Cockburn, Karlson, & Bederson, 2009; Stoakley et al., 1995). This section reviews how multi-view and multi-scale interaction has been integrated into VR, and why these systems motivate further research into tangible WiM-based interaction.

2.4.1 World-in-miniature as a Multi-Scale Interface

WiMs have long been used in HCI, especially as tools for navigation in virtual worlds (Stoakley et al., 1995). Early WiMs served as an interface to a large virtual world, pri-

marily for navigation, with small replica visualizations mapped to the represented world. A small user evaluation showed that users easily understood the mappings between virtual objects and WiM representations and adapted quickly for navigation (Stoakley et al., 1995). However, dual views (Figure 2.2) can confuse users when they move between the miniature and full-scale views, as attention shifts between the two (Pausch et al., 1995; Wingrave et al., 2006). Shortly after, "flying" was introduced as an alternative way to navigate in WiMs to mitigate this, where users slowly move an avatar in the miniature view (Pausch et al., 1995) to reduce confusion during WiM navigation.

Since their introduction to HCI, research has examined how WiMs should be displayed and controlled comfortably. WiM navigation naturally involves two views of the same world at different scales, requiring users to distribute attention between them, which can increase cognitive load (Pausch et al., 1995). The initial WiM interface only provided a fixed scale to the virtual world. As more interaction techniques were embedded into the virtual environment, the ability to control and switch between various scales became necessary for more flexible WiM interaction. For example, spatial design tasks may require switching between small-scale overviews and large-scale manipulation. Introducing the ability to switch between various scales has been proposed to address the problem (Wingrave et al., 2006). Additionally, hands-free interaction using feet in virtual environments (LaViola et al., 2001) has been introduced to keep the hands available for interaction in virtual worlds.

2.4.2 Multi-view, Multi-scale VR Systems

Multi-scale systems support interaction across different spatial representations. These systems combine a full-scale view for first-person to maximize embodiment and immersion, and miniature representations for abstract interaction, such as rapid placement and quick overview of large spaces. Previous work highlights the significance of these systems in the spatial design workflow, where users must transition between multiple perspectives without losing spatial context.

Several VR systems have explored multi-view and multi-scale interaction for spatial design and architectural contexts. For example, Racz and Zilizi (2018) developed a prototype that allows users to navigate and experience building interiors while providing a hologram of a miniature version of the space. Likewise, Hsu et al. (2020) presented an architectural design discussion system in VR, in which the prototype provides a first-person view with a minimap to visualize the user's position. Also, SeamlessDesign (Kiyokawa, Takemura, & Yokoya, 2000) provides a virtual design space, primarily for 3D object manipulation and creation in VR environments. Dreamscape Brick VR (Doma & Sener, 2022) uses LEGO-like bricks in a virtual environment. Users can scale *themselves* to life-size scale (1:1),

precision building scale (1:10), and LEGO figure-sized user scale (1:42.5). These systems leverage WiM related concepts by allowing users to move between different scales of view, such as first-person, overview, and bird's-eye view perspectives.

Beyond the foundational WiM paradigm, a variety of other multi-view and multi-scale techniques have been developed. These approaches include semantic zooming, focus+context visualizations, or hierarchical scene representations, to help users navigate and comprehend vast and intricate digital content (Cockburn et al., 2009; Perlin & Fox, 1993). Multi-viewport or view-multiplexing techniques are employed to present multiple scenes or views simultaneously within a single display, often for showing different perspectives or scales of data (Cockburn et al., 2009; Perlin & Fox, 1993; Schjerlund, Hornbæk, & Bergström, 2022). Even with sophisticated system designs, users may find it difficult to perceive and translate correct scales when working with miniature representations alone (Danyluk et al., 2021; J. Kim & Hancock, 2025; Wingrave et al., 2006). This suggests the value of multi-view systems, that can provide a perceptual anchor at full-scale view, while interacting at scale.

Commercial design tools and add-ons have increasingly incorporated VR and augmented reality (AR) features, including viewers, collaborative environments, and real-time modelling tools such as SketchUp (*SketchUp VR viewer*, n.d.), Revit (*Autodesk revit software*, n.d.), The Wild (*The wild integration | autodesk construction cloud*, n.d.), Gravity Sketch (*Gravity Sketch | 3D sketching and design software*, n.d.), Arkio (*Arkio*, n.d.), and Medium (*Medium by adobe*, n.d.). However, these tools primarily support viewing, presentation, modelling, or collaboration rather than integrated tangible interaction across multiple view scales. Although these systems help manage visual complexity across multiple views, they often remain visually driven. Integrating tangibility into such systems introduces additional complexity, including proxy reuse, haptic congruence, and visual–proprioceptive discrepancy.

To support multiple scales while minimizing loss of spatial context, visual systems often rely on multi-view interfaces. Zoomable interfaces such as Pad (Perlin & Fox, 1993) use spatially located information in parallel with nested, hierarchical graphic zoom in a plane. This approach expands how users access information by spatially organizing related data across nested zoom levels. Its concept is further developed in dual-view systems such as detail+context and focus+context visualizations (Cockburn et al., 2009), which aim to preserve context while providing access to multiple perspectives. In VR, maintaining spatial context becomes more difficult, as transitions between scales can cause disorientation (Bowman et al., 2003). Various interface techniques have been developed to mitigate such disorientation. Therefore, multi-scale and multi-view systems should employ spatially stable cognitive anchors to help users maintain orientation during transitions.

2.4.3 Interaction Challenges in Multi-Scale Systems

While multi-view and multi-scale systems provide access to both overview and detail, they introduce challenges for object manipulation and cognitive load. Users must adjust their movement strategies, interpret multiple spatial representations, and maintain context across changing scales and views.

Manipulating Objects

Switching between scales introduces not only perceptual demands but also motor-control challenges. These frequent changes can introduce inconsistencies in precision due to fluctuations in the Control-Display (C/D) ratio. Previous work in VR for multi-scale navigation highlights the need for consistency in C/D ratios (LaViola et al., 2001; Wingrave et al., 2006). Navigating WiM via flight introduces a different interaction scale dependent on users' cognitive anchor (Pausch et al., 1995; Wingrave et al., 2006). When users focus on the WiM rather than the full-scale view, small movements in the miniature can result in large, potentially confusing movements in the full-scale view. Conversely, physical movement, such as walking or reaching, affects the full-scale view and does not manipulate the WiM. This highlights the need for a consistent C/D ratio and for its dynamic application, depending on the situation and users' needs. These inconsistencies force users to adapt their movement strategies after every scale change, becoming a major source of frustration during detailed manipulation tasks.

Cognitive Load

Using multiple views to support multiple scales can compound users' cognitive load, especially when they must maintain multiple spatial mental models. While multi-view systems may increase cognitive load, prior work has explored strategies that can reduce mental effort. For instance, OVRlap (Schjerlund et al., 2022) demonstrates that overlapping multiple views can reduce the cognitive overhead associated with multiple transitions. At the same time, WiMs paired with full-scale views can reduce this load and support spatial reasoning when the views remain semantically consistent. The WiM provides a complementary overview, while the full-scale view supports a first-person spatial understanding. By maintaining consistent semantics across these perspectives, users minimize unnecessary re-orientation and reduce the cognitive load associated with scene transitions.

2.4.4 Design Space and Framework for Spatial Interaction

These challenges suggest that multi-scale interaction requires more than providing multiple visualizations. Users must coordinate perception, movement, and manipulation across multiple scales, especially in spatial design workflows that require both high-level planning and detailed refinement.

Design spaces and frameworks have been developed to make interaction design choices explicit, compare systems, and identify unexplored design opportunities. Prior frameworks show how conceptual models can provide shared vocabularies for describing interaction techniques and guide research beyond individual systems (Beaudouin-Lafon, 2000; Bowman et al., 2003). Frameworks for spatial and mixed-reality interfaces build on these. For example, Ethereal Planes (Ens, Hincapié-Ramos, & Irani, 2014) develops a framework for 2D information spaces in 3D environments by deriving design dimensions, mapping prior systems, and identifying usage patterns. These frameworks show how design dimensions can organize systems and reveal design opportunities. However, they primarily describe interaction structures, information spaces, or representations.

WiMs have also been analyzed using design space approaches that categorize their properties and relationships to the environments they represent (Danyluk et al., 2021). This work treats WiMs not as a single fixed technique, but as configurable spatial representations. However, existing WiM design spaces do not fully capture cases where the WiM becomes a physically grounded interaction volume integrated with tangible interaction in VR. In tangible WiM-based VR, the design problem is not only how the WiM represents a larger environment, but also how physical action is aligned with objects in the WiM and translated into effects in the full-scale virtual environment. Therefore, existing multi-view and multi-scale systems provide foundations, but they do not fully capture the combined design space of WiM, tangible interaction, and VR across perception, physical action, and representational relationships.

2.5 Challenges and Research Gaps

While WiMs provide substantial benefits for overview, navigation, and spatial reasoning, their use also comes with limitations related to scale perception, interaction continuity, and embodied feedback.

2.5.1 Scale Perception Challenges

Scale misperception is a central challenge for WiM interaction, as accurate scale judgment is fundamental to spatial design tasks and directly affects decision-making. While miniatures provide an effective overview of complex or large spaces and environments, scaling down can limit the accuracy of perceived spatial relationships. Familiar size cues can be disrupted by scale ratios, forcing users to rely on mental translation or external aids such as ratio notation, measurements, or reference objects. Even with these aids, the mental translation and semantics of representation remain major hurdles when dealing with miniatures.

When WiMs are scaled relative to the real world, this can affect judgements of spatial information, such as size and distance (Creem-Regehr et al., 2022; Loomis & Knapp, 2003; Renner et al., 2013). Although design tools such as CAD often include measurement aids, this addition also highlights the difficulty of perceiving scale directly in spatial design contexts.

Virtual reality often lacks the referent (in full-scale) of a scale model with familiar size cues (e.g., known size of an object), which helps calibrate scale perception. This is particularly true in scaled design scenarios (e.g., miniatures), where familiar size cues are cancelled out by the scaling factor, forcing users to mentally translate and calculate scaled models (Huk, 2006). Understanding how users perceive relative scale (ratios) rather than absolute size is therefore critical for evaluating VR interfaces in scaled scenarios.

2.5.2 Interaction and Workflow Challenges

WiM-based interactions introduce workflow fragmentation, disrupting continuity during iterative design tasks. WiMs allow users to switch between different perspectives and scales. However, this structure often disrupts work continuity. Users frequently switch between miniature and full-scale views for different purposes. Full-scale views support embodied inspection, whereas WiMs support rapid placement and the assessment of overall relationships between objects. This frequent switching forces users to re-establish context at every transition, disrupting workflow continuity required for highly creative work. This requires users to reconcile differences between perspectives and scenes, increasing cognitive load as they reorient their focus across scales (Cockburn et al., 2009; Perlin & Fox, 1993; Wang Baldonado, Woodruff, & Kuchinsky, 2000).

To counter this, WiMs adopted a minimap-style visual interface, placing miniatures atop full-scale views, creating a single-view system. This approach minimizes the switching required between multiple views and preserves spatial context in a single place (Pausch

et al., 1995; Stoakley et al., 1995). Further development of the concept suggests various techniques such as context zoom (Cockburn et al., 2009), avatar modifications and flying over (Piumsomboon, Lee, & Billinghurst, 2018; Wingrave et al., 2006). Despite these efforts, the underlying cost of maintaining mental models for multiple representations remains the largest burden for users using WiMs. For workflows involving multiple iterations, this further slows the progress and interrupts the workflow, especially when precise representations at scales are required simultaneously (Cockburn et al., 2009; Wang Baldonado et al., 2000). This fragmentation underscores a tension between overview and detail that current WiM interactions struggle with within a single, continuous interaction workflow.

2.5.3 Experience and Embodiment Limitations

WiMs can improve productivity when paired with complementary full-scale inspection. However, WiMs alone are limited in providing experience level compared to the full-scale version. Instead, they provide different perspectives specialized for an overview instead of a first-person experience. For example, miniature representations provide functional benefits such as layout planning in design workflow; however, they cannot reliably convey how design elements feel at full scale. This limitation relates to presence (as "being there") because users mentally imagine full-scale scenes while translating scales in VR (Cummings & Bailenson, 2016; Slater, 2009).

Additionally, manipulating objects at scales often requires motor precision beyond what standard input devices can offer, mouse sensitivity, VR controllers, and mid-air gestures, resulting in a reduced sense of control. This often leads to user frustration in detail-oriented tasks, especially spatial adjustment tasks. Thus, WiMs are better understood as complementary tools for spatial tasks rather than replacements for full-scale experience.

2.5.4 Tangible WiM-VR as an Underexplored Intersection

While WiMs provide contextual support, they often lack physical, tangible feedback during object manipulation. When users interact with virtual objects through a WiM, they often rely on visual representations rather than comprehensive haptic feedback, as noted by Azmandian et al. (2016); Hettiarachchi and Wigdor (2016); Muender et al. (2019). While other approaches address this by dynamically aligning virtual objects with physical proxies or leveraging visual dominance to enhance perceived tangibility (Azmandian et al., 2016; Hettiarachchi & Wigdor, 2016; Pohl et al., 2021), scaling such approaches to multi-view, multi-scale environments remains difficult. While visual dominance helps to create the

illusion in such systems, it does not eliminate the need for visual–proprioceptive alignment in multi-scale interfaces.

WiMs have been analyzed through design-space approaches that characterize their properties and relationships to larger represented environments (Danyluk et al., 2021). This framework treats WiMs not as a single technique, but as a configurable spatial representation. However, WiM can also serve as an interactive volume, rather than a passive viewport. Existing WiM design spaces do not focus on cases where the WiM becomes a physically grounded interaction volume accompanied by a tangible proxy. In tangible-WiM interaction, the design problem is not only how the WiM is structured, but also how physical action is aligned with objects in the WiM and translated into effects in the full-scale virtual environment.

Tangible and WiM-based interactions in VR remain largely disjoint, hindering a unified account of tangible-WiM-VR interaction. Tangible VR research primarily explains how tangible proxies can align with virtual objects, while WiM research explains how miniature spaces represent and control a larger full-scale environment. Tangible WiM interaction combines these relationships, but is not explicitly captured by existing frameworks.

2.6 Summary and Research Direction

This chapter reviewed previous work on spatial representation, scale perception in VR, WiM, TUI, haptic retargeting, and multi-view/scale systems. Collectively, this literature highlights the complexity of spatial design tasks that require seamless transitions between miniature overviews and full-scale views. Previous work has demonstrated the strength of WiM for navigation and context, but limitations in scale perception, manipulation precision, cognitive load, and physical integration remain unresolved challenges for tangible WiM-based VR experiences.

Mainly, four research gaps emerge from this review. First, while WiM is widely used in AEC, how users perceive and interact with different miniature-based tools during spatial design tasks remains largely unreported. Second, while topics in haptic retargeting, scale perception and visual–proprioceptive discrepancies are established individually, how they affect and interact with each other is unknown in the context of WiM in VR. Third, while prior research examined the design space of WiM, tangible interaction, VR, there is limited guidance for structuring the combined design problem of tangible WiM-based VR interaction. Lastly, few systems demonstrate how miniature representation, tangible proxies, and multi-view coordination can be integrated to manage spatial discrepancy. These gaps

motivate the dissertation's research questions across user interaction, perception, design space, and prototype instantiation.

Chapter 3

Investigating Digital, Tangible, and Paper-Based Room Design at a Small Scale

This chapter presents the first empirical study in the dissertation and establishes the foundational research problem that subsequent research addresses. This study examines how people use miniature-based design tools in a simplified spatial design task, comparing an online room planner, paper-based sketching, and tangible blocks (LEGO). This study reveals that current design methods with miniature representations do not reliably support accurate scale judgments between the representation and the expected experience. The findings highlight that digital tools provide ease of use, tangible-based tools provide engaging experiences through strong embodied interaction, warranting a combined approach of WiM-based tangible interaction in digital form. These findings motivate the next part of the dissertation research, centred around scale misperception while engaging with WiM. The core question is whether VR can preserve the benefits of physical, direct manipulation while improving perceptual and interaction support. While this study laid the foundation for the research problem space, the results directly inform fundamental problems in miniature representation and their use as design tools. This research has been published and presented at Graphics Interface (GI '25).

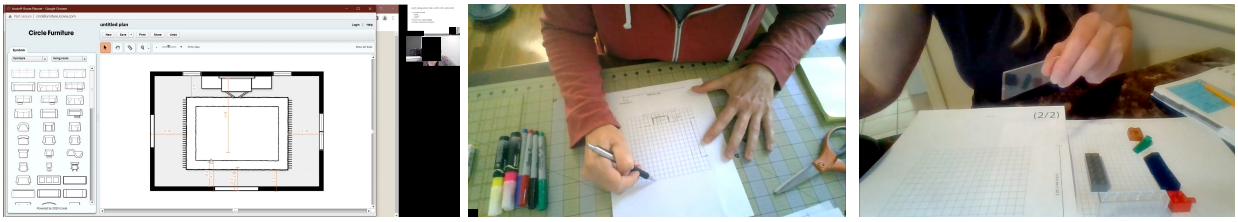
3.1 Introduction

Worlds-in-miniature (WiMs) have been around for a long while. These miniature models are common in architecture, engineering, and construction (AEC) (Hosokawa et al., 2008; H.-J. Kim et al., 2016; Rauterberg et al., 1998), but their use has been expanded into other applications (Horiuchi et al., 2012; Kakehi et al., 2009; Nakanishi, 2012). However, tools based on these concepts are mostly designed for and used by a small proportion of the general population, primarily experts, and often neglect any form of tangible experience.

Trends like do-it-yourself (DIY) and democratizing design activities like software design and AR/VR/3D content authoring have gained popularity and are empowering non-experts to engage in design activities (Tanenbaum, Williams, Desjardins, & Tanenbaum, 2013; Vardouli & Buechley, 2014). Despite this popularity, digital tools for miniature representation (e.g., CAD tools) are often not designed for non-experts. As a result, non-experts are often left to use simplified design processes, like drawing on paper when planning their room layouts or using tools like the IKEA room planner. However, it is unclear if these simplified miniature representations are adequate and what the non-expert experience of using them is like.

Additionally, physical 3D representations (e.g., scaled architectural models) are currently inaccessible to non-experts and, while there have been efforts made to integrate haptic experiences (Azmandian et al., 2016; Hosokawa et al., 2008; Rauterberg et al., 1998; Underkoffler & Ishii, 1999), tangible interaction is still nascent. In this chapter, I address these gaps by investigating how non-experts use tools (both tangible and traditional) for design tasks.

This chapter presents findings from a mixed-method study where I observed participants designing a room layout using miniature representations of larger physical spaces with existing tools (paper and digital), as well as tangible LEGO blocks. I collected video data to observe the design process with each of these tools and used the Desirability Toolkit (Benedek & Miner, 2002) with a follow-up semi-structured interview to collect participants' affective experience with the tools. The analysis shows that while the online tool benefits from the constraint of having correctly scaled selections of furniture, it can limit creativity. On the other hand, paper and pen and physical LEGO bricks offer flexibility, promoting personalization and customization. These findings suggest that future designs could combine tangible interaction (e.g. LEGO blocks) with digital experiences, such as AR/VR, to further democratize design activities for a broader audience, including non-experts.



(a) Online (b) Paper (c) LEGO

Figure 3.1: How I recorded the videos. Participants working on the design.

3.2 The Study

The study investigates the use of physical and digital media for the creation and design of physical spaces. While architectural and interior design has a wide variety of miniature representations of physical spaces (e.g., architectural scale models, drafting documents, etc.), I focus on the application of room layout for several reasons: (1) it is a miniature representation of a physical space that many people are familiar with, (2) there are already online tools available for non-experts for this process, (3) it has both 2D and 3D aspects to the design process, and (4) there are examples already in HCI literature that use this task that this work can build on.

As this work is an early exploration of miniature representations of physical space, I opted for a mixed methods approach that leans more toward qualitative analysis when possible, because I did not have clear hypotheses. Nonetheless, I wanted specifically to contrast the experience of using physical 3D media (LEGO), physical 2D media (pen and paper), and digital tools. On the quantitative side, I used a validated scale to collect data about flow experience, but my primary data collection was a combination of the Desirability Toolkit (Benedek & Miner, 2002) and interview data.

3.2.1 Participants

I recruited 30 participants (21 identified as women, 9 as men, 0 as non-binary; age: 21–79, $Mdn = 42$) from UserInterviews.com[1], a third-party participant recruiting service. I did not set specific inclusion or exclusion criteria for background (knowledge, education, profession), though I aimed to recruit interested individuals for rich outcomes and analysis, as the study advertisement included the description of designing interiors and room layout.

[1] `https://www.userinterviews.com`

Employment	N	Design Exp	Arch Tools	Skill	3D Tools	Skill
Educational services	8	4				
Professional, scientific or technical services	8	6	3	6.0	4	5.7
Other services (except public administration)	5	4	2	4.0	1	3.0
Health care or social assistance	3	3				
Construction	2	2	1	5.0	1	5.0
Real estate or rental and leasing	1	1				
Utilities	1	1				
Retail trade	1	1				
Arts, entertainment or recreation	1	1				
	N=30	**N=23**	**N=6**	**M=4.3 (SD=2.3)**	**N=6**	**M=5.2 (SD=1.2)**

Table 3.1: The breakdown of participants by profession. While most had planning or design experience with architecture or interior design, this included simple tasks like laying out a room (which I consider common for non-experts). Additionally, six participants had experience with architectural and 3D modelling tools, respectively (5 of these were the same participants), though typically skill with architectural tools was rated lower than for 3D modelling tools.

As a result, most participants (23 out of 30) indicated they had been involved in these activities in some way, including basic room planning (arranging and moving furniture) or related gaming experiences like The Sims. Six participants had experience using 3D modelling tools like 3DS Max and Maya and described their skill level on a 7-point scale from beginner (1) to expert (7) with a mean of $M = 4.3$ ($SD = 2.3$). Five of those six also indicated experience with architectural tools and similarly rated their skill ($M = 5.2$, $SD = 1.2$), indicating a small subset of participants had related knowledge of and interests in interior and room layout activities (Table 3.1). All participants reported having at least a college credit or a Bachelor's degree, including 13 postgraduate degrees (Master's, Ph.D., professional degree). This study protocol (#41338) was reviewed and approved by the University of Waterloo Research Ethics Board and advertised in accordance with its guidelines.

A note on sample size

There are a number of methods for determining sample size in qualitative studies. I used a purposive sampling technique based on advice from Malterud, Siersma, and Guassora (2016) that uses information power rather than saturation. This model suggests five dimensions: study aim, sample specificity, established theory, quality of interview dialogue, and analysis strategy to determine information power. I determined that I had a narrow study aim (fewer participants), but the sample was quite general, with no exclusion criteria (more participants), with a theoretical foundation that was neither especially strong nor absent (neutral), a novice interviewer whose shyness might have affected the ability to establish a good dialogue (more participants), and the analysis crossed three conditions rather than focusing on one (more participants). Nonetheless, the interviewer's experienced supervisor has conducted many similar studies that have led to rich data with as few as 8 participants. Based on these considerations, I cautiously appraised my case to require about 30 participants (10 per condition).

3.2.2 Method

The study was conducted online using the video conferencing platform Zoom, due to the risk of COVID-19 at the time of the study. All participants were asked to prepare and set up a webcam prior to the session. After a brief introduction, ensuring the webcam and materials were set up properly, and obtaining informed consent, participants designed two rooms (a living room and a kitchen) for a family with two parents and two children (one teenager and one toddler). For each room, they were given a list of fixtures to place, and the dimensions and shape of the room, but no other constraints were suggested (i.e., specific furniture selection, size, colour, etc.). I also prompted participants with examples of what could be placed in each room and asked participants to choose and design features such as doors, windows, walls, and entryways. Participants were asked to think aloud while designing the rooms. Each session lasted about an hour, and participants were given a remuneration digital gift card valued at 20 USD.

Participants began with designing the living room, followed by the kitchen. The specifics of each room were as follows:

- The *living room* had a rectangular shape of 18 by 12 feet (5.49 m × 3.66 m) and participants were prompted to add a door, window, and lighting and furniture as they desired. I also provided the example of a play area for the toddler.

- The *kitchen* was 12 by 12 feet (3.66 m × 3.66 m), and participants were prompted to add a cabinet and major kitchen appliances (refrigerator, stove, etc.) and countertops/islands as desired.

Following the design activity, participants were asked to complete the FSS long version of the flow scale (Jackson & Marsh, 1996), and were then asked to complete the Desirability Toolkit (Benedek & Miner, 2002), followed by a short interview. After the interview, I collected demographics.

Conditions

I had participants design the layout of a room in one of three conditions:

- **Online**: I used a publicly available online planner[2], that was less likely to be used by my participants than other more popular tools (e.g., Ikea). This tool uses a 2D top-down view and provides comprehensive options, such as room shape, size, and lists of furniture and objects based on sections of a home (living room, bedroom, garden, etc.), essentially acting as a CAD tool for the non-expert. Before completing their design, I provided participants with a brief demonstration of the basic operations of the online planner, such as placing objects, rotating, scaling, and browsing.

- **Paper**: I included this option as a more traditional and accessible way to design the interior, using common household items, such as pen, pencil, paper, and ruler. While I did not limit the selection of objects used in each session, participants still needed to bring my base grid paper (provided as two printable digital documents, one for the living room and one for the kitchen) and a drawing tool (pen or pencil). Participants could also bring an eraser, ruler, and scrap paper.

- **LEGO**: I asked participants to bring at least 100 LEGO pieces (specifically, "Bricks"[3]). Participants would assemble a variety of pieces on the same piece of grid paper used in the paper condition.

I provided participants with two separate portable document format (PDF) files for the grids for the living room and kitchen. Each cell was approximately 8 mm × 8 mm and corresponded to 23 cm × 23 cm (9 inches × 9 inches) in real-world conversion. One cell

[2] `https://circlefurniture.icovia.com/`

[3] `https://www.lego.com/en-ca/pick-and-build/pick-a-brick?page=1&filters.i0.key=variants.attributes.subcategoryId&filters.i0.values.i0=1`

was designed to fit one bit of LEGO brick. In the online tool, I instructed participants to manually adjust the size of the room (by typing numbers). The online tool used a responsive design, so I could not easily control the physical size of the miniature representation. With the online tool, participants could also zoom in/out dynamically as needed. On a 1080p, 21-inch monitor, this would correspond to a 1:26.6 scale on the default view, at 100% zoom level (in contrast to the 1:28.7 scale for paper and LEGO).

3.2.3 Measures & Data Collection

I used several measures to gather qualitative and quantitative data covering both breadth and depth. I elaborate on collection methods, including the Flow Scale, the Desirability Toolkit, the follow-up interview, and observation:

- **Desirability toolkit**: The Desirability Toolkit (Benedek & Miner, 2002) was initially created to collect data about "desirability", but more generally, can be used to elicit a participant's experience using a tool. I included 30 descriptive words (e.g., innovative, time-saving, ineffective, irrelevant) and first asked participants to select as many words from this list as they wanted to describe their experience. Then, they were asked to select their top 5 words from that shorter list. The order of words in the list was randomized for each participant.

- **Interview**: An important part of the Desirability Toolkit is to follow this selection with an interview to discuss the experience. I used a semi-structured interview that began with, "How was your experience in general?" Whenever possible, I asked follow-up questions related to the previous question, even if the question was not on the question guideline. I also requested a specific rationale for their five words from the Desirability Toolkit.

- **Observations**: The design sessions by each participant were recorded, and the webcam was positioned so that the experimenter could observe the participant and the tool (online: screen capture + webcam on participant; paper and LEGO: webcam pointing down to see paper/LEGO during the design process).

- **Flow Scale**: The Flow Scale measures the responder's flow state, indicating the balance between their skills and the task's challenges, thus showing how deeply users are engaged in the task. Specifically, I used the FSS Long version (Jackson & Marsh, 1996), retrieving the original wording and question order for all participants.

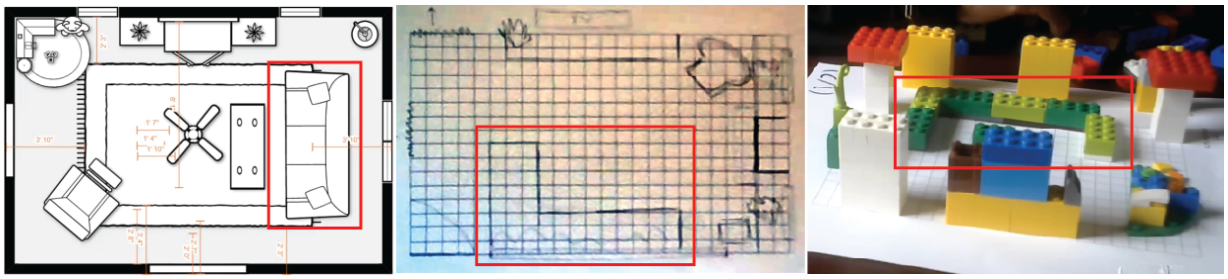


Figure 3.2: Sofa in three different design tools, highlighted with red

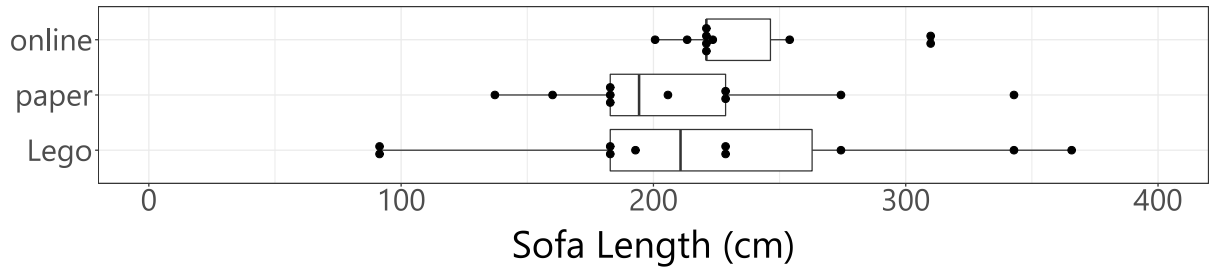


Figure 3.3: Size (length) comparison of participants' sofa sizes.

3.3 Findings: Observations

In addition to the interview data, I analyzed the video recordings from the design sessions, borrowing concepts from reflexive thematic analysis (Braun & Clarke, 2012, 2019), such as familiarizing ourselves with the dataset and using multiple passes to identify recurring patterns. Specifically, the experimenter did a first pass to review the videos, taking note of interesting patterns. In a second pass, the experimenter labelled interesting segments of the videos. In the third and subsequent passes, patterns of behaviour were identified.

3.3.1 The Disparity in Object Scaling

When representing furniture and other objects in their room layouts, there was a clear pattern of notable variation in scale representations in the paper and LEGO conditions, unlike the online condition, which had predefined sizes that were common in real-life scenarios. To demonstrate this discrepancy, I picked a sofa (which all participants represented) and converted the scaled representation to its actual length in centimetres. A boxplot and dot plot for each participant's sofa length is shown in Figure 3.3. When reading this plot, it

is useful to consider the online condition as "standard" in order to see the variety in the other two conditions.

To demonstrate how I determined these calculations, Figure 3.2 shows an example of sofa representation in each condition. The sofa in the paper condition (middle) was counted three deep as seating and 12 in length (longer, back of the couch). In the LEGO condition (Figure 3.2, right), I can count 2 in depth and 14 in length. As I provided the same grid paper for both the paper and LEGO conditions with the dimensions and conversion noted directly on the PDF (living room and kitchen), I can see the varied sizes of the same functional objects. In the grid paper, one cell converts to a 9-inch square (appx. 22.86cm on every side) in real-world size.

Online

In the online condition, participants used the supplied objects on the website I provided. All of the objects had dimensions (W $\times$ L $\times$ H) noted in a separate window, so the participants had access to the exact size. The default size of the objects was close to what is considered "standard" sizing in real life.

Paper and LEGO

In the paper and LEGO conditions, participants showed a large gap between their size and the standard size of objects. In the example in Figure 3.2, a participant in the online condition placed a three-seater sofa. In contrast, the paper and LEGO participants placed a sectional sofa but represented it as too small.

Because I did not ask how many people could be seated on the sofa they were placing, there is some room for interpretation of the intention of the participants. It seems clear that both when drawing or using physical props to design an interior, participants experienced difficulty converting their scale perception onto grid paper, even though the dimensions were provided. This scale perception difficulty might lead to other issues when designing. I occasionally see participants using an eraser in the paper condition and similar methods in LEGO, such as reassembling into a different size or unusual scaling relative to other objects in the same design. For example, in Figure 3.2 with LEGO (right), white blocks with red tops represent floor lamps (next to their sofa, red circled). Considering the "standard" floor lamp size, I can interpret that the scaling is difficult, or I can infer a limitation using LEGO bricks since LEGO bricks are provided "as-is" in their sizes and availability at hand, limiting flexibility.

3.3.2 Freedom of Expression

In the online condition, participants were limited to selecting only the objects available, but in both the paper and LEGO conditions, they had more freedom to express their own intentions. This was, in some cases, intimidating, but they spent time creating objects matching their real-world counterparts.

Paper

In the semi-structured interview, several participants mentioned they felt incompetent in drawing, even though I explicitly mentioned that I did not intend to test or evaluate their drawing skills or final drawing at the beginning of the study. In most cases, participants tried to draw the objects instead of using abstract geometrical shapes, such as rectangles or circles. This suggests that feeling less competent in drawing skills could impact the design process, making them passive rather than active participants.

LEGO

Participants with LEGO bricks showed a similar way of representing or replicating objects with their LEGO bricks. Since I asked the participants to bring enough pieces (100+ pieces) for the session, many participants brought pieces with different shapes and colours. Participants then assembled them into a similar shape or colour, resembling their real-life counterparts. For instance, they used green pieces to represent plants and clear (transparent) pieces for windows. I infer this as participants wanting to show higher fidelity so they could conceive the design better. Thus, high fidelity and instant feedback might be empowering, especially for non-experts.

3.3.3 Design Processes with Each Tool

I found that participants exhibited different behaviours when designing with each tool. In this section, I describe in detail how participants approached the design problem in each condition, looking both at how they set up the space around them as well as additional details about how they did their work.

Word	Online	Paper	Lego	Total
Engaging	7	10	10	27
Satisfying	5	8	7	20
Organized	7	5	7	19
Innovative	6	4	7	17
Intuitive	3	7	5	15
Empowering	3	9	2	14
Accessible	5	5	2	12
Efficient	2	4	6	12
Motivating	1	6	5	12
Comprehensive	5	3	3	11
Effortless	1	4	5	10
Time saving	5	3	2	10
Familiar	1	3	5	9
Confident	2	3	3	8
Unconventional	1	3	4	8
Complex	3	2	1	6
Collaborative	1	0	3	4
Intimidating	1	2	1	4
Time consuming	1	1	1	3
Difficult	1	1	0	2
Frustrating	2	0	0	2
Hard to use	2	0	0	2
Predictable	2	0	0	2
Too technical	2	0	0	2
Confusing	1	0	0	1
Irrelevant	0	0	1	1
Overwhelming	0	0	1	1
Unpredictable	0	1	0	1
Inconsistent	0	0	0	0
Ineffective	0	0	0	0
Total	70	84	81	235

Word	Online	Paper	Lego	Total
Engaging	7	9	8	24
Satisfying	5	7	5	17
Empowering	1	8	1	10
Motivating	1	5	4	10
Organized	5	2	3	10
Innovative	4	1	4	9
Intuitive	3	4	2	9
Comprehensive	5	1	2	8
Familiar	1	2	4	7
Efficient	1	1	4	6
Time saving	4	1	1	6
Accessible	2	2	1	5
Complex	3	1	1	5
Unconventional	1	1	3	5
Confident	0	1	2	3
Intimidating	0	2	1	3
Collaborative	0	0	2	2
Effortless	0	0	2	2
Frustrating	2	0	0	2
Confusing	1	0	0	1
Difficult	1	0	0	1
Predictable	1	0	0	1
Time consuming	1	0	0	1
Too technical	1	0	0	1
Hard to use	0	0	0	0
Inconsistent	0	0	0	0
Ineffective	0	0	0	0
Irrelevant	0	0	0	0
Overwhelming	0	0	0	0
Unpredictable	0	0	0	0
Total	50	48	50	148

Table 3.2: Frequency tables (left: from initial 30-word list, right: top 5 words) of word choices by condition.

Online

Participants frequently changed their minds about what furniture or features they wanted and replaced objects in their room layout. Since the online tool allows for easy editing, participants used this feature extensively. Once they decided what type of object to add (usually "landmark," the largest object), the participants navigated through lists of objects provided by the online tool. While browsing each object, they often placed them in the room to get more detail, such as exact measurements.

After initial browsing and navigating, participants decided on the object type (e.g., TV or Sofa for the living room) and its variations (e.g., loveseat or three-seater). Space dividers or landmarks were placed first, and minor features, such as tables and lighting, were decided later. I also noticed frequent relocation of objects as participants leveraged the benefit of having a digital tool. Participants repeatedly replaced their objects with the same type in a different size or variation. These actions were chained into other objects' replacements as they wanted to "match" their style or balance the available space.

Paper

Participants who used paper with drawing tools often began with light sketches or thin lines before committing to solid lines or noticeable colours. Non-committal drawing included hovering over the paper without making a thin line. Once they determined what objects to draw, the participants changed tactics and drew more solid lines by changing colour or using more pressure on the drawing tool. When the participants decided to add detail to an object, they either added a label to mark what the object represents or added details to the object, such as armrests on a sofa. Participants rarely replaced or removed the drawn objects, especially compared to online tool participants. Most participants used the eraser attached to the back of their pencil, if available. I also noticed that participants extended the line over their objects to make them bigger rather than using an eraser to remove and redraw them.

LEGO

LEGO participants were always actively thinking about what to build and how to collect the necessary pieces to achieve their design goals. Generally, I observed two types of builders during this study: batch builders and one-by-one builders. Batch builders decided on an object and then placed the necessary pieces in front of them for later assembly. One-by-one builders considered the basic structure of the object and connected each piece as they collected the necessary pieces. This one-by-one process is similar to the participants using pen and paper, as they needed to actively think about the required object and how they could achieve it. I rarely saw any detailing work from LEGO participants, adding small pieces to existing finished LEGO objects. This may be due to participants not having enough LEGO pieces or not feeling the need for a detailed model.

3.4 Findings: Desirability Toolkit

Table 3.2 shows word frequencies, with initial selections at the left and the top 5 words at the right. "Engaging" was the most frequent word from the list of 30 words across all conditions, though the rationale varied based on the condition. All participants in the Paper and LEGO conditions chose "engaging" as one of their initial words, but only 7 of 10 participants in the online condition did. Interestingly, all 7 of these online participants kept the word "engaging" as one of their top 5, but in both the paper and LEGO conditions, 1 and 2 participants, respectively, dropped the word "engaging" for their top 5 words. While

word frequencies provide a surface-level, holistic understanding of participants' experiences, I provide interview context for their top 5 words.

3.4.1 Engaging

Online participants selected "engaging" for two main reasons: their own motivation to continue or because of the tool's features. Participants felt compelled to "*keep going until it's complete*" (p5, online) or "*keep kind of doing it*" (p9, online). The online tool's features and ease of revising were seen as engaging to online participants: "*because it was kind of fun to do just moving things*" (p3, online). These experiences suggest that the online tool's instant feedback and user agency contributed to the engagement.

Participants in the paper condition found their thinking process engaging, emphasizing the need for planning and creative problem-solving. They noted that "*I have to think, you know, where I would place the items*" (p13, paper). This feeling stemmed from the permanency when drawing on paper, which encouraged thoughtful planning before committing to drawing. Participants also described "*thinking through the parameters*" (p11, paper) and said "*I activated the part of my brain. That's the creative part, the imagination*" (p16, paper), further demonstrating the root of their rationale.

In contrast to the other two methods, LEGO participants mentioned various engaging experiences, highlighting the physical, hands-on, play-like nature of the tool or the challenge of assembly and problem-solving. They described "*building things ... by hand*" (p22, LEGO) or similarly, "*to play around with it [LEGO]*" (p29, LEGO). The challenge of assembly, "*to think and get more in tune*" (p25, LEGO) and the "*process of elimination or, problem-solving*" (p30, LEGO), also contributed to the engaging experience.

3.4.2 Satisfying

Participants across all conditions found the design process "satisfying." They described the experience as "*a simple blank canvas to a room*" (p3, online), "*you get to see ... the end result*" (p15, paper) and "*seeing how things came together*" (p24, LEGO). LEGO participants also noted the tactile experience, as it *zones everything else out, and it's super satisfying*" (p21, LEGO).

3.4.3 Empowering

Paper participants predominantly selected "empowering," noting accomplishment in designing without professional experience. They felt they did "*better than [they] thought*" (p16, paper) and enjoyed "*seeing it on paper*" (p17, paper). They also felt empowered "*as opposed to hiring a professional*" (p11, paper) because they "*can be a part of that design process, which makes it empower*" instead of being "*subject to someone else's ideas*" (p18, paper).

3.4.4 Motivating

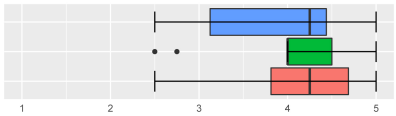
Participants with LEGO and Paper more frequently selected "motivating" than participants with the online condition. Paper participants found the design thinking process motivating with a similar remark: *making them think in a design way.* They described the experience "*to be more involved in design*" (p14, paper) and "*motivating [them] to want to make it happen*" (p16, paper). Whereas LEGO participants were motivated by the playful aspect of building and assembling to "*continue to do it*" (p25, LEGO) and "*to take on some projects*" (p26, LEGO), suggesting LEGO's potential as a playful introduction to the design activity.

3.4.5 Organized

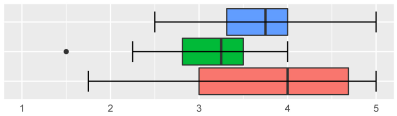
Online participants selected "organized" more frequently than participants with Paper or LEGO conditions. They appreciated the interface and the categorization of objects, found "*the program itself was very organized*" (p1, online) and that it made "*categories and like subcategories, ... easier to find*" (p10, online). LEGO participants, on the other hand, described "organized" in different ways, noting they "*can see exactly where each piece is going to fit*" (p23, LEGO) and "*the squares [the grid] ... you can fill however you want*" (p29, LEGO).

3.5 Findings: Flow Scale

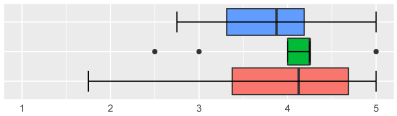
I used a one-way multivariate analysis of variance (MANOVA) with condition (paper, LEGO, online) as the independent variable and the nine dimensions of the flow scale as dependent variables. There was no significant difference found between the three conditions



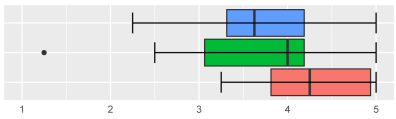
(a) Challenge-skill balance ($F_{2,27} = 0.203$, $p = .818$)



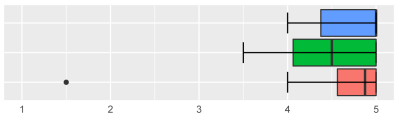
(b) Action-awareness merging ($F_{2,27} = 1.81$, $p = .183$)



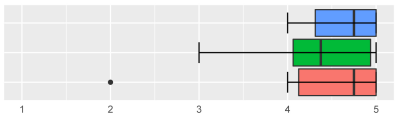
(c) Clear goals ($F_{2,27} = 0.222$, $p = .802$)



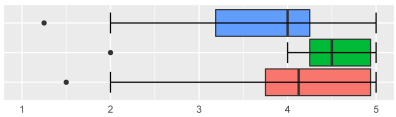
(d) Unambiguous feedback ($F_{2,27} = 2.025$, $p = .152$)



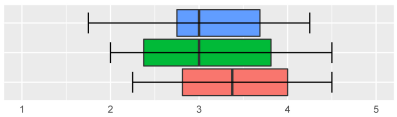
(e) Concentration on the task at hand ($F_{2,27} = 0.344$, $p = .712$)



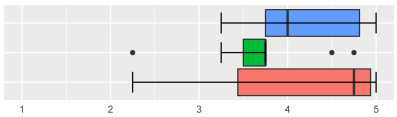
(f) Sense of control ($F_{2,27} = 0.533$, $p = .593$)



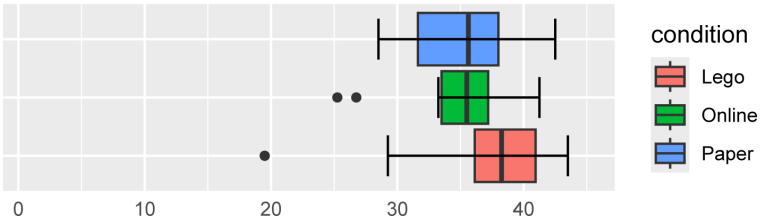
(g) Loss of self-consciousness ($F_{2,27} = 0.907$, $p = .416$)



(h) Time transformation ($F_{2,27} = 0.336$, $p = .718$)



(i) Autotelic experience ($F_{2,27} = 1.356$, $p = .275$)



(j) FSS total - overall flow ($F_{2,27} = 0.237$, $p = .790$)

Figure 3.4: Boxplots for each dimension and FSS

(Pillai's Trace = .84, $F_{2,27} = 1.6$, $p = .11$). I, therefore, cannot reject the null hypothesis. I show univariate ANOVA results and boxplots for each dimension and the overall flow score (FSS total) in Figure 3.4. This result is not surprising, given the sample size and the lack of reported differences in engagement between conditions in the Desirability Toolkit results.

3.6 Findings: Interview

Semi-structured interviews revealed two notable patterns beyond word choice comments and rationales. To identify these, I underwent multiple passes of the recordings: Initial familiarization followed by quote labelling and pattern grouping in the second pass. While the analysis resembles reflexive thematic analysis methodology (Braun & Clarke, 2012, 2019), my focus was on the methodology suggested by the Desirability Toolkit, resulting in topic summaries rather than themes.

3.6.1 Choosing the Right Tool for Each Stage

Participants favoured the paper condition for a quick draft to get a sense of the overall design, claiming it was useful for a "*very, very high level*" design stage (p13, paper) and "*rough sketch*" (p20, paper). Many wanted to transition from paper drafts to digital formats to "*visualize and make a really nice rendering*" (p20, paper). They valued digital convenience for instant visual feedback and precise measurements, noting paper's limitations in measurements and proportions: "*take measurements of all the like appliances or the cabinets, but on the piece of paper, it's, it's hard to do*" (p13, paper). Several participants mentioned they would use paper and pencil for a quick draft to get a sense of the overall design, noting that they "*would want to use pen and pencil to start*" (p2, online) or use the paper "*for the initial idea*" and for a "*very, very high level*" design stage (p13, paper). Similarly, It may be used "*in the beginning, actually as a first draft*" (p17, paper), and "*if I were to be, just do very quick things and being inspired to be* " and would "*use paper and pen as a rough sketch*" (p20, paper). Most participants in the online and paper conditions also agreed with the idea of sketching on paper as a draft and moving to another method. Participants "*would have liked for it to translate into some type of digital format*" (p14, paper) and "*transition to a digital version at the end after that [paper draft]*" (p17, paper). In the same vein, one participant "*may use an online tool to help me visualize and make a really nice rendering*" (p20, paper). They expressed the need "*to see what it's gonna look like and play with the items [objects]*" (p13, paper).

I suspect the main reason for wanting to move to other tools, such as VR, was to perform computations, such as calculating the object's size and seeing its proportion relative to the entire space, thus making the decision-making process much easier for users by providing direct visual feedback instantly. Notably, the participants commented that having a digital design could be beneficial because of its ability to measure everything and validate, removing unnecessary physical steps in order to measure their furniture manually: "*take*

measurements of all the like appliances or the cabinets, but on the piece of paper, it's, it's hard to do" (p13, paper).

3.6.2 Need for Different Perspectives

Participants desired different perspectives to judge object proportions, particularly as paper and online conditions offered only 2D top-down views. They expressed a need for their designs to be transformed into "*VR or 3D view*" (p1, online) to see "*how it looks within the room*" (p13, paper). Participants noted that 3D views would help those "*who sort of lack the vision of being able to see what something would look like*" (p11, paper). Despite the between-participants design, they recognized the benefits of other unseen methods, reinforcing the need for different perspectives. Participants wanted to look at their designs from different perspectives so that they could judge individual objects' proportions and adjust accordingly. This is especially true for this study where, in the paper and online conditions, participants only had a top-down view instead of 3D models of the objects. The participants would have liked to have a design transformed in "*like VR or 3D view of it. So then you can see like, oh wait, that's way too big. Or that's not big enough.*" (p1, online), or, "*like 3D where I could place objects and then it would show me exactly how it looked look within the room*" (p13, paper). One participant further commented that "*I think if the end product shows you more of a 3D image digitally, that can be really helpful for people who are, who sort of lack the vision of being able to see what something would look like*" (p11, paper). It is interesting that, even though the study was conducted in a between-participants manner, participants already discussed the need for additional (unseen) methods to continue and succeed in their design session. Furthermore, online participants emphasized that having different perspectives, such as in 3D or line of sight viewers, would be beneficial for checking their design.

3.7 Discussion

The study explored the current state of the design process and how non-experts perceive and use different tools in a familiar interior design scenario. I conducted the study to examine how people's experiences with miniature representations of large physical spaces vary based on tools both commonly used as non-expert tools (online planner, pen and paper) and physical objects with potential design utility (LEGO). The results revealed that the online tool offers convenient features, such as easy deletion and object measurement, but presented usability challenges for some participants. On the other hand, physical tools,

such as paper-pen and LEGO, can offer rich, engaging experiences, allowing participants to focus more on design than the tools. Furthermore, my observations yielded deeper insight into how people use each tool for design and factors that contribute to inaccuracies in scaled design, suggesting further investigation. These findings underscore the current state of such design tools and point to the potential for future solutions that combine the benefits of both digital and physical approaches. As technology advances, AR/VR solutions could be combined with tangible interaction to enable some of the benefits of physical miniature representations together with digital convenience.

3.7.1 Design Implications

Table 3.3 shows a summary of how the findings suggest what stage each of the tools can be used in and some of the advantages and disadvantages of each that I identified. I discuss how future tools can be designed to incorporate these insights.

	Pen and Paper	Online tool	LEGO
Design stages	(a) initial planning for quick sketch	(b) editing-heavy stages	(c) used as interface with online
Advantages	(d) complete freedom	(e) accessible easily modifiable	(f) intuitive, familiar
Disadvantages	(g) requires scale perception	(h) less freedom using predefined objects	(i) limited by availability of LEGO

Table 3.3: Summary of the results and recommended design stages for each tool

Different Workflow

The different workflows observed in the study highlight where each method excels at different design stages. This suggests that each method can be combined, complementing each other effectively.

For instance, a suitable interior design process for non-experts might begin with a quick sketch and ideation phase, allowing freedom of expression (Table 3.3d), which promotes personalization and customization. However, these methods can be intimidating for some users (Booth et al., 2016) and might encounter underlying requirements like scale perception (Table 3.3g). In contrast, online tools often provide a list of available objects which can

be placed. However, this heavily deters personalization and customization (Table 3.3h), which contradicts the recent DIY trend where users seek self-design opportunities for customization and personalization. This hints that freedom of expression is more important in the early design phase than producing accurate, high-fidelity designs. A digitalized tool can be developed to satisfy this expectation, similar to how current sophisticated CAD programs operate.

Once the initial ideation phase is completed, these sketches and ideas can be translated and transformed into a format that is accessible for detailed and heavy editing, taking advantage of *digital convenience* (Table 3.3e). In this stage, integrating a tangible interface can further assist in the design process to promote engagement and familiarity (Table 3.3f). One promising approach is integrating tangible interfaces to lower the barrier to entry for design activities. For example, using passive haptics, such as haptic retargeting (Azmandian et al., 2016), can compensate for the need for every object to be mapped to digital objects in physical forms (Table 3.3i).

Using Tangible Interfaces

As shown in previous work (Booth et al., 2016; Bourdeau et al., 2020), these findings also show that physical interaction and haptic feedback are great candidates for developing tools for non-experts. Responsive feedback and fidelity are both vital in the creation of supportive tools (Table 3.3f). The concept of "I'll Know It When I See It" (IKIWISI) resonates well within the design sector, particularly among less experienced users. The study participants' responses emphasized the demand for high fidelity in tasks such as LEGO assembly, colour matching to real-world objects, and the creation of realistic-looking drawings. This underscores the necessity for a tool that offers high fidelity while remaining user-friendly, nurturing creativity, and promoting ideation.

For urban planning, the interactive system Urp (Underkoffler & Ishii, 1999) demonstrates exceptional use of tangible user interfaces (Fitzmaurice et al., 1995; Ishii, 2008a). Bringing together technologies, such as traditional paper drawing with physical interfaces (McGee, Cohen, Wesson, & Horman, 2002) or augmented reality (Aliakseyeu, Martens, & Rauterberg, 2006), has the potential to improve the design experience. It is worth noting that both VR and AR have reached a stage where they can produce photorealistic representations with minimal differences in size and shape compared to real-world objects (Arnowitz, 2017; Hornsey et al., 2020; Ogawa et al., 2018). This suggests an opportunity to leverage these technologies to create a physical experience that enables miniature representations of larger physical spaces (Table 3.3c).

Self-Confidence in Sketching/Design

Despite the participants expressing intimidation from their lack of drawing ability, previous research suggests that sketching skills are not a significant factor in engineering design (M. C. Yang & Cham, 2006). While sketching skills are not directly relevant to the final design, the skills are often emphasized to make the individual participating in the design process more creative during ideation (Booth et al., 2016; Kudrowitz, Te, & Wallace, 2012).

Novices without related training often make inaccurate drawings with pen due to the misperception of the objects intended to be drawn (Cohen & Bennett, 1997) (Table 3.3g). As I saw in the study, the participants mentioned the need for a different perspective (3D or VR) to *check* their designs, and this suggests inaccuracies and *the disparity in scaling* may have the same underlying cause: misperception. This may be why participants felt the necessity for an extra validation tool such as a 3D viewer or digitally transformed design from using sketching on paper (Table 3.3a): the need for *the multiple design stages* (Table 3.3b,e).

3.7.2 Envisioning Future Tools for Everyone

The main tasks in the study centred on designing scaled worlds or, in other words, miniature worlds. Space design, such as this interior design task, requires an understanding of space and the relationship between objects. The birds-eye view benefits the participants by showing relative positions between objects, thus helping users further optimize the space. While miniaturization helps the planning phase of the space, experiencing these spaces at full scale remains a challenge.

Some of the findings, such as the "desire for 3D visualization after sketch" from the participants, underscore the need for further exploration. Future tools, especially incorporating immersive technology, hold promise for delivering highly interactive experiences with the convenience of digital workflows. AR/VR specifically offers tremendous potential as it allows for the seamless integration of digital information directly from existing software (CAD, building information modeling (BIM)) conveniently in an immersive and interactive environment.

This hybrid approach using AR/VR together with physical interaction aligns with a broader opportunity to combine the flexibility of digital tools with the intuitive control of physical interaction (Fitzmaurice et al., 1995; Ishii & Ullmer, 1997; Underkoffler & Ishii, 1999). The three methods I employed in this study—LEGO, paper sketching, and online interfaces—can serve as effective tools for initial sketches in low fidelity, then transition

to immersive technology for enhanced interaction and visualization. Combined with the tangible interface, this approach can be introduced for not only experts alike but everyday users as well, bridging physical and digital interfaces.

Integrating with advanced computing technology naturally addresses scalability in spatial design. While this study focused on limited spatial design tasks, such as room layout and object placements, transitioning and integration with computer-aided tools can enable users to design at any scale of spatial design, from room layout, interior, architecture, and facade to urban design. Interactive technology allows users to seamlessly navigate between multiple scales, views, and designs, ultimately fostering a dynamic and adaptable design process and enabling more informed design decisions.

3.7.3 Understanding the Users

While some participants reported some experience with architectural (e.g., CAD) and 3D modelling tools (7 in total), I consider the majority to still be non-experts (and even some of these 7 to have low expertise, given their self-ratings of skill with these tools). While many more participants rated some planning or design experience with architecture or interior design, I explicitly highlighted "planning your room layout or playing a game" as such an activity, so I expect this group to still be considered non-experts (and to be similar to participant pools from related studies), and thus think that the participants were representative of a general population (albeit a highly educated one).

Nonetheless, these experiences and education levels may have influenced participant experiences and responses during the interview. Despite having a different group than *typical* novices, the analysis and results still provide valuable insights. Because the advertisement for participation was sent broadly, I expect the nature of the study interested some participants more than others, so it is also reasonable to expect the participant pool to be a target audience of the types of tools, and the results can help inform developers and designers of related architecture and interior design tools.

3.7.4 Implication of Full Remote Study

I conducted a fully remote study due to the risk of COVID-19 at the time it was run. In addition to the typical tradeoffs between remote and in-person studies, this study required physical components (LEGO bricks and pen and paper), which limited the participant pool and necessitated individual differences in setups. This choice may have also affected the results, as participants that have LEGO at home, presumably, are already familiar with

those specific LEGO blocks, and similarly, a participant's regular use of their own pen and paper may have affected their word choices, for instance, the "familiar" word in the Desirability Toolkit. Although I instructed participants to bring certain types of LEGO and more than 100 pieces in total, an in-person study would have allowed providing a more consistent study setup, which may have influenced the outcomes of the study, such as participants' familiarity with the LEGO pieces and their final designs. Nonetheless, LEGO bricks are versatile in nature, and pen and paper are common household items, so I found the participants to have comparable setups, even if they were not identical. The remote setup also forced the research to be as simple as possible. While this research can be supported on applications with head-mounted displays (AR/VR), I decided not to use AR/VR in this study to streamline the sessions with participants.

3.8 Chapter Summary

In this chapter, I presented a mixed methods study comparing room layout design experiences with a digital online tool, paper and pen, and physical LEGO blocks. The study shows that miniature-based design tools support different aspects of design activity regardless of background and experiences the users might have, but also reveals difficulties in scale judgment.

The findings show the benefits of each tool. The digital experience allows quick editing and precise measurements, while tangible tools (paper-pen combo and LEGO) are familiar and intuitive, offering the flexibility and familiarity of physical media. Moreover, participants experienced difficulties properly scaling objects within a room when using physical media, motivating further investigation of scale perception with WiM. These results highlight that miniature-based design tools are not interchangeable. Each medium introduces different benefits and constraints for spatial reasoning and interaction.

This chapter supports the first argument of the dissertation: existing miniature-based design tools provide valuable benefits in interaction and representation, but they do not fully support accurate scale judgment. While they reveal limitations, the positive experience reported with each tool used in this study suggests that tangible interaction with WiM remains a promising direction for design tools. These findings motivate chapter 4, which examines how tangible proxies in VR support scale perception and manipulation performance under different scale-mismatch conditions within a WiM environment.

Chapter 4

Retargeting Scale — Investigating Visual–Proprioceptive Mismatch in Tangible Worlds-in-Miniature

Building on the exploratory findings from chapter 3, this study examines whether tangible interaction can support scale perception and 3D manipulation performance in a controlled WiM-based VR environment. While chapter 3 showed that miniature-based design tools can be engaging but susceptible to scale misperception, this chapter focuses on scale mismatch as an experimental variable and evaluates how users interact and perceive when visual and physical scales do not fully match.

The role of this chapter in the dissertation is to empirically examine the core mechanism that motivates the later design framework. Specifically, the study investigates how visual–proprioceptive mismatch between a tangible proxy and a virtual object affects object size perception, subjective discrepancy, and manipulation performance. By using tangible proxies at different scales within a WiM environment, this chapter examines whether users can maintain performance and perceptual stability even when physical and virtual representations are mismatched.

This chapter supports the second argument of the dissertation: tangible proxies can provide physical grounding for WiM-based interaction in VR, but their effectiveness depends on the direction and magnitude of visual–proprioceptive mismatch. The findings show that tangible interaction can preserve functional manipulation performance under some mismatch conditions, while also revealing perceptual and performance tradeoffs that must be accounted for in system design.

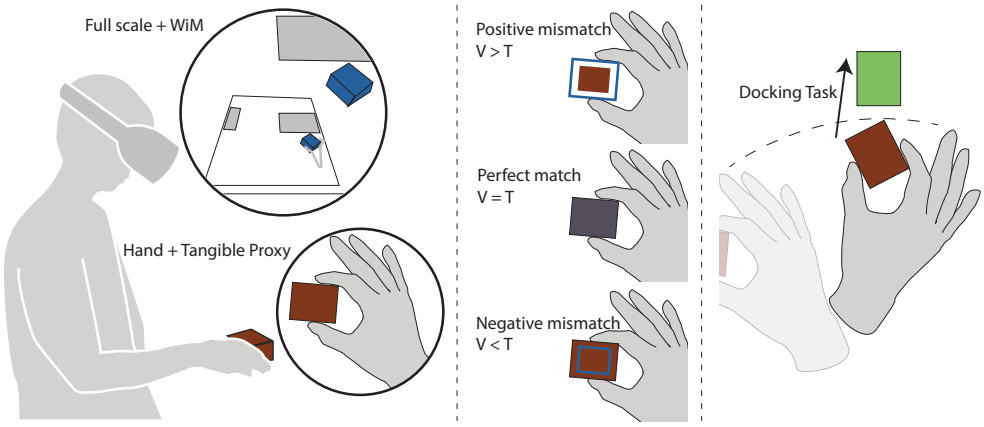


Figure 4.1: The overview of the study

These findings motivate chapter 5. If tangible proxies can support manipulation and scale perception despite mismatch, then the next step is to examine how this mechanism is managed in existing interaction techniques and how this can be integrated into a broader tangible WiM-VR design framework.

4.1 Introduction

Chapter 3 showed that miniature-based design tools can support engaging room layout activities in both digital and physical forms. Physical tools provided intuitive control, while digital tools offered flexibility. However, the findings suggest persistent difficulties in accurately judging and expressing scale through miniature representations. Previous literature suggests that such difficulties may stem from the mental translation required between scaled representations and their full-scale referents (Cutting & Vishton, 1995; Gogel & Da Silva, 1987; Huk, 2006). In WiM-based VR, this translation becomes both interactive and embodied. Users not only view a scaled virtual model but also manipulate it through a physical proxy that may differ from the model's visual scale. In a virtual system, a single proxy may need to represent multiple virtual objects at different sizes, creating a visual–proprioceptive scale discrepancy between the physical object being grasped and the virtual objects being manipulated. Understanding how users perceive and perform under this discrepancy is necessary to determine whether the mismatch can be tolerated, compensated for, or deliberately leveraged in later interaction designs.

In contrast to extensive studies on distance and depth perception (Armbrüster et al.,

2008; Creem-Regehr et al., 2022; Gogel & Da Silva, 1987; Hornsey et al., 2020; Kelly et al., 2013), scale perception under visual–proprioceptive mismatch with tangible proxies remains comparatively underexplored. Although haptic retargeting demonstrates that users can interact with a tangible proxy despite such discrepancies (Azmandian et al., 2016), few studies have focused on scale-based manipulation and tolerance to scale mismatches. Previous research has explored how small size differences affect judgment (Combe et al., 2008; Thomas, 2020), but the accuracy of scale judgment under conditions of sensory conflict remains largely unknown. I hypothesize that while users will consciously detect significant visual–proprioceptive mismatches, the visual dominance of VR will allow them to maintain scale judgements without degrading interaction performance. This lack of empirical data hinders developers and designers aiming to create VR applications where accurate spatial understanding is essential.

In order to address this gap and understand how users perceive scaled objects in virtual environments, I conducted a within-participants user study. Participants were immersed in a full-scale virtual room containing a world-in-miniature (WiM) representation positioned in front of them (Figure 4.1). Their primary task was to estimate the ratio of a miniaturized virtual object to its full-scale object, both of which were visible simultaneously. The study employed a within-participants design evaluating three tangible scales against three mismatch types: negative (tangible proxy larger than visual scale), congruent, and positive (proxy smaller than visual scale). This design allowed me to isolate the effects of scale discrepancies on user performance and perception. The findings provide empirical insights that performance is constrained by the tangible proxy and its physical dimensions, with significant degradation at the smallest scales. Furthermore, the results reveal a perceptual asymmetry: users are highly sensitive to "negative" mismatches where a larger tangible proxy causes physical collisions with fingers, whereas they are tolerant of "positive" mismatches where the proxy is smaller than the virtual object. I present empirical results, followed by a discussion of future research directions and consolidated design recommendations.

This chapter contributes to a deeper understanding of scale perception in VR by examining how visual–proprioceptive mismatches affect users' perception of 3D scaled models. This work expands the knowledge of human perception within virtual environments by extending haptic retargeting beyond positional warping to include scale-based manipulation. By revealing the directionality of mismatch tolerance and the conditionality of visual dominance with tangible interaction in VR, I provide novel insights into how tangible proxies interact with scaled virtual objects. These empirical findings provide grounding for design space and framework explorations in chapter 5 and for the prototype interaction technique in chapter 6, where spatial discrepancy is treated as a design variable to be managed.

4.2 Methodology

This study is grounded in the optimization of sensory processing of the brain (Ernst & Banks, 2002), which suggests the brain resolves sensory conflict by weighting the most reliable sense. In VR, the most reliable and direct source of 'reality' is vision. I designed this experiment to identify the threshold of visual dominance in a VR environment with a WiM. The study employed a within-participants factorial design to investigate the tolerance limits of visual-proprioceptive mismatch in virtual reality.

4.2.1 Participants

A total of 12 participants[1] were recruited for an in-person study (approx. 60 minutes per participant) through university email listings and recruitment posters. Participants' ages ranged from 18 to 33 years ($M = 26.08, SD = 4.38, Mdn = 25.00$). Of those, 3 identified as women, 9 as men, and none as non-binary. Participants also self-reported their virtual reality (VR) experience levels: 6 were beginners, 5 were intermediates, and 1 was advanced. Educational backgrounds included 1 with a high school diploma, 5 with a bachelor's degree, 4 with a master's degree, and 2 with a PhD.

Exclusion criteria included uncorrected visual impairments, clinically diagnosed balance disorders, and the use of medical devices or implants, such as pacemakers and neurostimulators, due to potential risks associated with electromagnetic interference from the VR devices and setup.

All participants provided informed consent prior to the study and were compensated with a $20 CAD digital gift card. This study protocol (#45474) was reviewed and approved by the University of Waterloo Research Ethics Board and was advertised in accordance with its guidelines.

4.2.2 Apparatus

The VR system was powered by a PC equipped with an AMD Ryzen 7 5800 processor, an Nvidia RTX 3060 graphics card, and 16 GB of RAM. The experiment was conducted using

[1] I had originally planned for the study to have 24 participants, but the lab experienced a theft of the equipment needed to run the study. I replaced the equipment, and this replaced equipment was also burgled. I have conducted a post-hoc power analysis to validate that this sample is sufficient to draw interesting conclusions from the data

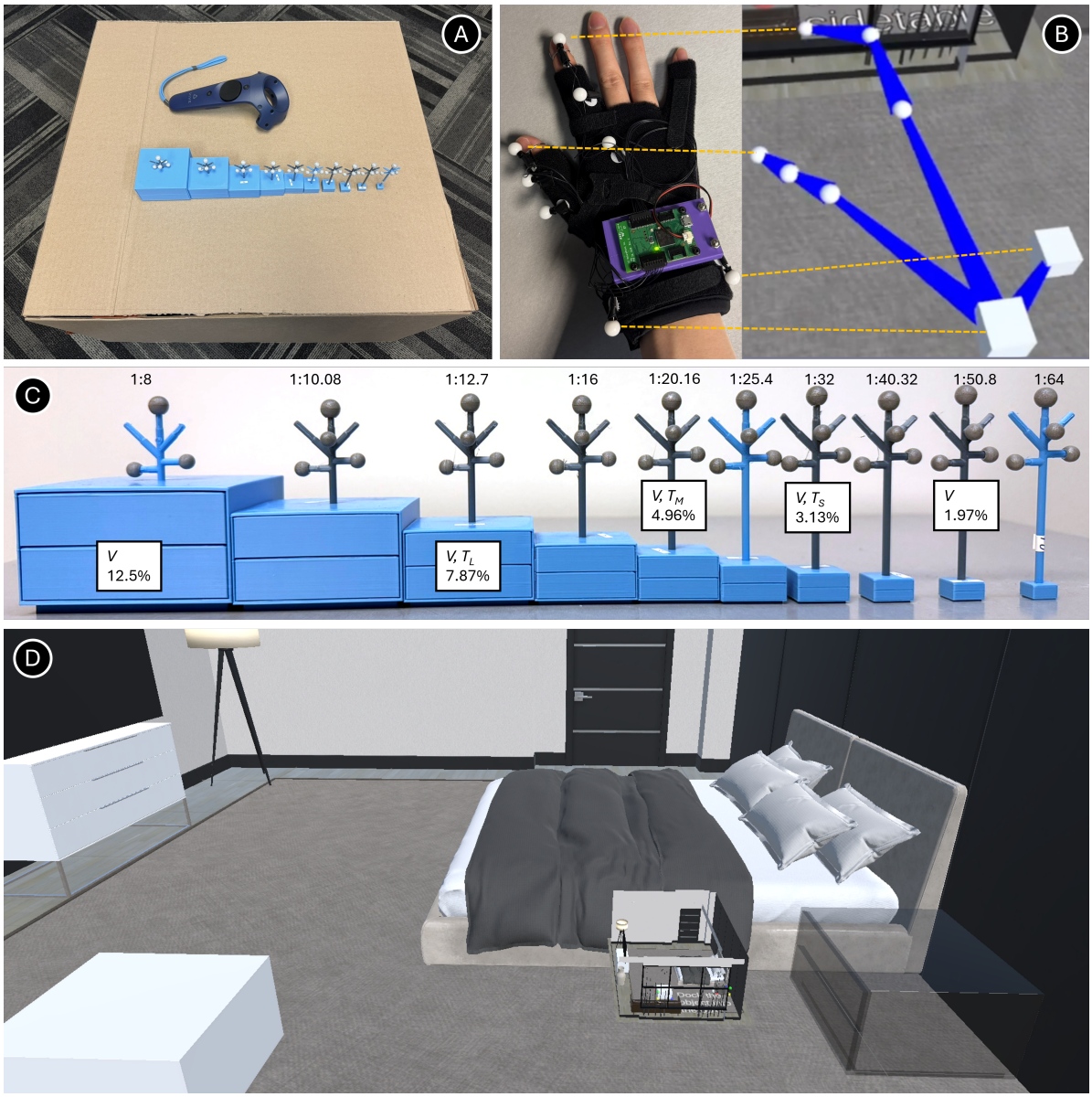


Figure 4.2: The experimental apparatus: (A) Cardboard mass model of the target (side table) shown with a Vive controller and the complete master set of 10 tangible proxies to illustrate relative scale. (B) Custom half-finger glove with active optical markers, and its representation in VR. (C) The master set of 10 tangible proxies. (D) A panoramic view of the virtual environment

an HTC Vive Pro Virtual Reality Headset, which rendered visuals at a refresh rate of 90 Hz. An OptiTrack system with eight Prime 13 cameras, operating at 100 Hz, provided 6-degree of freedom (DOF) tracking for the headset, tangible proxies (Figure 4.2C), and the user's hand through markers on a custom half-finger glove (Figure 4.2B).

Virtual Environment

The virtual environment was developed using Unity, depicting a realistic, full-scale bedroom furnished with standard objects, including a king-size bed, lamp, TV and a couch, to provide spatial context (Figure 4.2D). A side table (64 cm (W) x 38 cm (H) x 59 cm (D)) served as the primary interactive target. Simultaneously, a miniaturized replica of this room, containing a scaled-down version of the target object, was positioned directly in front of the participant, approximately 50 cm away from the designated location for the tangible proxy.

The Custom Glove and Hand Representation

I used a half-finger glove to maximize the tactile feedback directly on their fingers while maintaining the robust placement of the active LED markers (Figure 4.2B). For each finger (thumb and index), a nail ring was used to attach an active marker. I then strategically placed the markers to represent anatomical landmarks: the fingertips (distal phalanges), the two main finger joints (proximal interphalangeal, or PIP), the two knuckle joints (metacarpophalangeal, or MCP), and two points on the wrist.

I visualized the user's hand as an abstract skeletal model to minimize body-centric comparison as indicated in previous literature, specifying that the realism of the 3D hand model affects the perceived other 3D models around (Ogawa et al., 2018). By implementing minimal representation of a hand, I reduced the cognitive weight of the virtual body, encouraging the participants to treat the representation as a tool, rather than their direct hand avatar. The model consisted of small spheres rendered at the tracked marker positions and connected by thin lines to form a simplified skeletal structure. This abstract representation facilitated the dynamic scaling of the hand's geometry. I avoided standard scaling of the hand applying at the wrist root, which would unnaturally displace the virtual fingertips. Instead, I dynamically set the scaling pivot to the functional centre of the grasp. The system calculated the midpoint between the tracked index fingertip and thumb tip. The hand was then scaled relative to this contact-centric pivot point, adjusting automatically based on the scale difference between the virtual object and the tangible proxy, ensuring visual congruency with the virtual object size.

4.2.3 Experimental Design

The study employed a within-participants factorial design to systematically evaluate user performance and perception across varying degrees of visual-proprioceptive mismatch. In the context of worlds-in-miniature (WiMs), there are three relevant representations of artifacts that are experienced by participants: the tangible proxy, the WiM representation of the artifact in VR, and the full scale representation of the artifact in VR. The ratios between sizes of each of these representations are all relevant to the investigation, which I manipulated by considering two primary variables: the *tangible scale* (T), which is the ratio between the tangible proxies and the full scale model, and the *visual scale* (V), which is the ratio between the WiM representation and the full scale model. Rather than evaluating absolute sizes in an unbalanced matrix, I categorized the variables into a fully crossed 3×3 within-participants design, which is defined by tangible scale (T) (small, medium, large) and mismatch type, negative ($V < T$), congruent ($V = T$) or positive ($V > T$). To ensure consistent mismatch intensity across different sizes, I utilized a logarithmic spacing with a constant step factor of 1.6. The levels of scales used in the study were motivated by the Weber-Fechner Law (Fechner, 1860), which suggests that noticeable differences may depend on proportional rather than absolute differences. Logarithmic spacing allowed the study to apply a fixed proportional mismatch across proxy sizes, paired with visual scales.

Table 4.1: Experimental Variables. The values represent the **Mismatch Ratio** (V_{scale}/T_{scale}). Ratios > 1 indicate Positive Mismatch (Visual > Physical), while ratios < 1 indicate Negative Mismatch (Visual < Physical).

Visual Scale (V)		**Tangible Scale (T)**		
Ratio	% Full Size	T_L (1:12.7)	T_M (1:20.2)	T_S (1:32)
1 : 8	12.50%	**1.58**	-	-
1 : 12.7	7.87%	1.00*	**1.58**	-
1 : 20.2	4.96%	**0.63**	1.00*	**1.58**
1 : 32	3.13%	-	**0.63**	1.00*
1 : 50.8	1.97%	-	-	**0.63**

**Note: cells indicate congruent conditions ($V = T$).*

Tangible Scales

I fabricated a master set of 10 tangible proxies, logarithmically spaced from ratios of 1:8 to 1:64 relative to the full scale VR model, using rigid polylactic acid (PLA) (Figure 4.2C). These proxies were precise miniature replicas of the target object (side table), fitted with passive reflective markers for optical tracking. While the "tree-like" tracking mounts added height to the proxies, the height of the marker structure was minimized to lower the centre of gravity as much as possible, and pilot testing confirmed that this design balanced the tradeoff between maintaining optical tracking visibility and minimizing occurrences of the proxies tipping during manipulation. From this master set, I selected a subset of three tangible scales: $T_L = 1:12.7 = 7.87\%$, $T_M = 1:20.16 = 4.96\%$, and $T_S = 1:32 = 3.13\%$.

Visual Scales

I added two visual scales to this set of three tangible scales, for a total of five visual scales (V): 1:8, 1:12.7, 1:20.16, 1:32, and 1:50.8. The addition of these two visual scales allowed for negative, congruent, and positive mismatches for each tangible scale (Table 4.1).

Study Conditions

To create a robust study design while preserving the effect of mismatch types, I paired each tangible scale (small, medium, large) with three specific visual scales. For a congruent match ($V = T$), the exact match scale was used. The immediately adjacent smaller and larger visual scales were used for the negative ($V < T$) and positive ($V > T$) mismatches, respectively. This pairing ensures a complete within-participants design, isolating the directional effect of the mismatch. Extreme mismatch conditions (e.g., T_L vs $V_{1:50.8}$) were excluded because they represent a considerable discrepancy that offers little insight into the tolerance thresholds of visual-proprioceptive mismatch. This focused approach maintained a constant mismatch ratio across all conditions, minimized participant fatigue, and isolated the core variables. All participants completed trials for every selected combination in a randomized order.

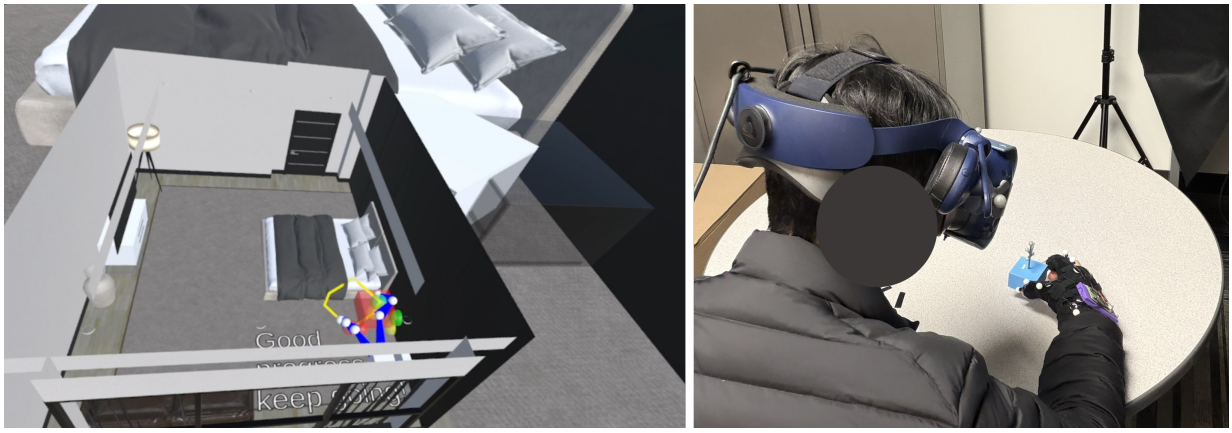


Figure 4.3: Experimental task view: (Left) User's view of the virtual environment showing the world-in-miniature and the target docking receptacle. (Right) A participant performing the docking task with a tangible proxy.

4.2.4 Main Task

6-DOF Docking Task

I employed a 6-DOF precision docking task to ensure consistent, active interaction. Docking is a standard task for assessing 6-DOF manipulation performance in VR (Zhai & Milgram, 1998). For each trial, participants used the assigned tangible proxy to grasp and align the virtual miniature object with a corresponding semi-transparent target receptacle ('ghostimage') within the miniature room beside a bed (Figure 4.3, left).

Alignment and Success Criteria

To guide alignment, four colour-coded spheres marked the corners of both the target object (the side table) and its receptacle. The system then continuously monitored the spatial intersection between these spheres. The receptacle provided real-time visual feedback, appearing red when misaligned, and turning green only when the target object achieved at least 70% volumetric overlap over all four independent spheres simultaneously. The threshold was established during pilot testing to enforce high positional (x, y, z) and rotational (pitch, yaw, roll) precision without inducing fatigue from impossible tolerances. Successful docking was registered when the object was in the *green* state (i.e., achieving $>$ 70% overlap on all four spheres) and the participant lifted their fingers from the proxy. A

3-second visual inspection time was provided to avoid a sudden flash/screen change after their docking phase. The full-scale room and its miniature room then disappeared, and a screen appeared to administer perceptual questions (subsection 4.2.5).

4.2.5 Measures

The dependent variables included task completion time and docking accuracy, as well as the responses to the following questions asked after each trial:

Q1. **(Visual Estimation)**: "Based on what you saw, which scale do you think the miniature object was?" They selected from five fixed percentage options corresponding to the visual scale levels. This forced-choice method isolates the cognitive judgment of the visual scale.

Q2. **(Sensation of Discrepancy)**: "Compared to the miniature you saw, the physical object you held was:" Responses were recorded on a 5-point Likert scale ranging from 'Way too small' to 'Way too big,' with 'Just right' indicating perceived congruence.

4.2.6 Procedure

The study was conducted in a quiet laboratory setting to minimize external distractions. Upon arrival, participants were greeted and briefed on the study's purpose, before providing informed consent. The session consisted of a training phase followed by the main experimental phase. The institutional research ethics board reviewed and granted ethics clearance for all components of this study.

Training Phase

Participants first underwent a guided training session to familiarize themselves with the VR equipment, the VR environment, and the entire trial procedure. This phase included:

- **Task Explanation:** Detailed instructions were provided regarding the main task: the docking task and the two post-trial perceptual questions. The concept of estimating size as a percentage of the full-scale object was explicitly clarified through its mass model, fabricated using a cardboard box at full scale (Figure 4.2A) and corresponding scales with 3D printed proxies (Figure 4.2C).

- **VR Environment Exploration:** A brief period to acclimate to the VR headset and the full-scale virtual room. Participants took approximately 3 minutes to fit the headset and adjust the straps, and were briefed on virtual environments, including a full-scale and WiM- interactive view, before practice trials.
- **Practice Trials:** Participants completed a minimum of 3 practice trials, mimicking the full procedure of a trial with detailed explanation in each sequence to ensure full comprehension and comfort with the task, response reporting, and trial progression. After a minimum number of trials, participants were asked to confirm their full comprehension of the trial procedure and were prompted to continue either with more practice or with data collection.

Experimental Trials

The main experiment consisted of 45 trials, structured into 5 blocks (repetitions). Within each block, the unique 9 conditions were presented exactly once in a randomized order. This randomized block design ensured that all conditions were evenly distributed across the duration of the experiment, mitigating the impact of fatigue or learning effects on any specific condition.

For each trial:

1. Pre-trial: The system was in a 'ready' paused state, allowing the experimenter to place the tangible proxy for the condition in a designated position. The experimenter carefully placed the proxy, minimizing impact noise to reduce unintended auditory cues. The system displayed the tangible proxy ID for the experimenter only, ensuring zero configuration error. The participant could use this paused status to rest if needed.
2. Docking: Once the proxy was placed and the participant was ready, the trial began. The participant grasped the tangible proxy, which was highlighted with a yellow outline when picked up, and aligned the virtual object with the target.
3. Transition: Upon successful docking, the participant waited 3 seconds, then the scene disappeared.
4. Response: The first question appeared on a virtual panel. Participants verbally selected an option (1-5), which the experimenter logged. The second question followed immediately.

After completing all 45 trials, participants immediately completed the Simulator Sickness Questionnaire (SSQ) to assess their motion sickness level (Kennedy, Lane, Berbaum, & Lilienthal, 1993), followed by their demographic information and prior experience related to virtual environments and 3D-related tasks (CAD, 3D modelling). Upon completing the survey on Qualtrics, a short semi-structured interview was conducted, with four guiding questions about their overall experience.

4.3 Results

Data from all 12 participants (540 trials) were analyzed using repeated measures analysis of variance (RM-ANOVA) for quantitative measures. The analysis is organized into three components: objective performance (accuracy and efficiency), perceptual metrics (estimation bias and subjective discrepancy), and qualitative findings from interviews and research notes.

4.3.1 Simulator Sickness (SSQ)

Simulator sickness was assessed using the standard SSQ. The Total Score was ($M = 20.88$ $SD = 24.52$). Breakdown by subscale revealed the following scores: Nausea ($M = 15.90, SD = 20.47$), Oculomotor ($M = 16.42, SD = 17.05$), and Disorientation ($M = 24.36, SD = 38.52$). Note that the high standard deviations suggest a small subset of participants reported mildly significant symptoms related to motion sickness, while the majority experienced negligible effects. No participants reported significant discomfort or nausea during or after the study, and all sessions were completed. The experimenter did not notice any discomfort during interaction with the participants, except that some of them requested water or wanted to take off their jackets. This discrepancy may be explained by the central heating settings affecting the room temperatures (winter season settings), which likely elevated scores on thermal items, such as sweating, independently of the VR task. Overall, these scores suggest that the VR interaction was generally well-tolerated, with symptoms remaining within an acceptable range for active VR tasks.

4.3.2 Quantitative Findings

To evaluate the effects of multi-sensory discrepancy, I aggregated trials by participants to ensure the independence of observations. I structured the analysis around three distinct

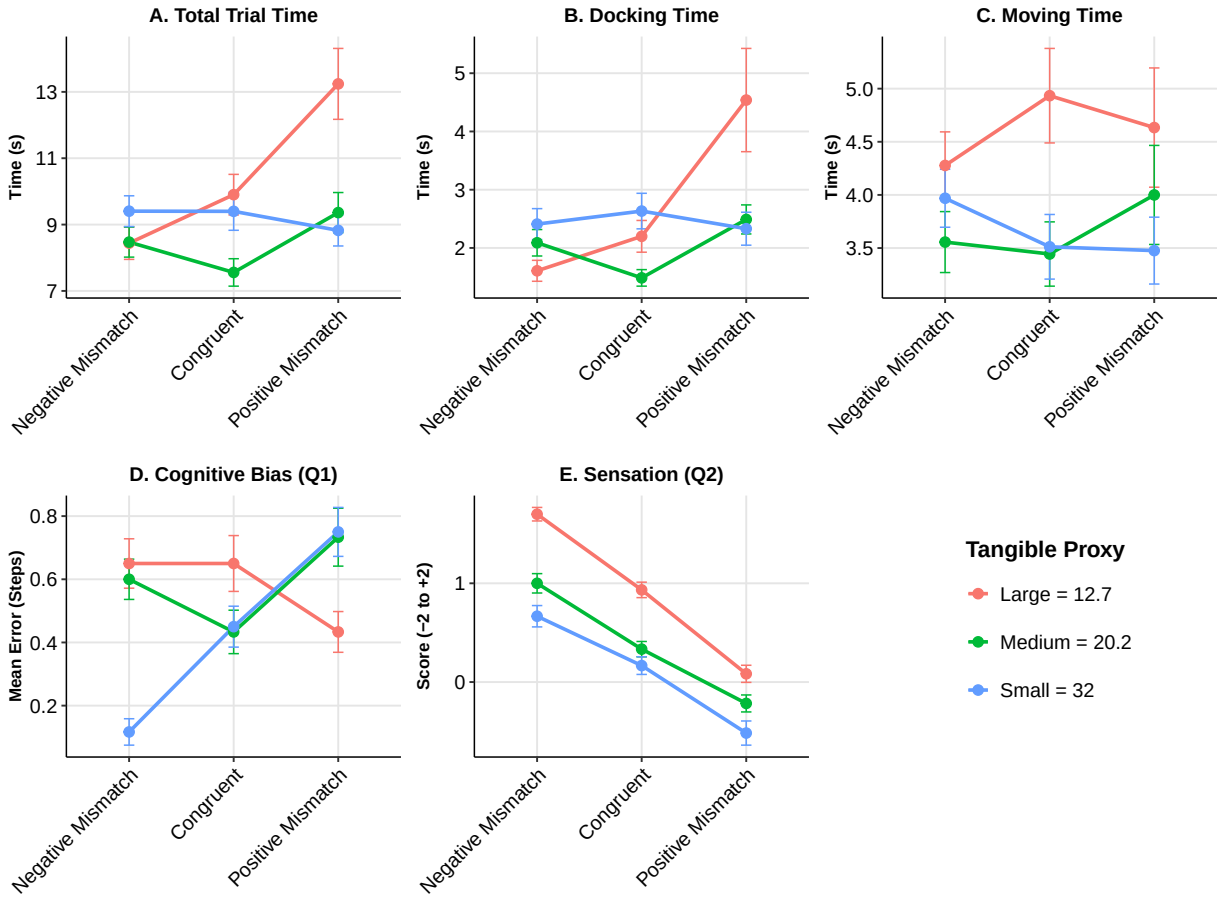


Figure 4.4: Quantitative results of the visual–proprioceptive discrepancy evaluation across three tangible proxy sizes. (A) Total Trial Time reveals an overall interaction effect. Decomposing the task reveals that (B) Docking Time incurs a severe speed-accuracy penalty during high sensory conflict, whereas (C) ballistic Moving Time is completely unaffected by visual mismatches. (D) Cognitive Bias (Q1) demonstrates that visual dominance is modulated by the physical size of the grasped proxy. (E) Subjective Sensation scores (Q2) confirm that participants reliably detected the intended illusions across all physical scales.

visual–proprioceptive alignment conditions: **Positive Mismatch** (visual object (V) > tangible proxy (T)), **Congruent** (visual object (V) = tangible proxy (T)), and **Negative Mismatch** (visual object (V) < tangible proxy (T)). I analyzed the dependent variables using a repeated-measures ANOVA with Mismatch Type and Tangible Size as within-participants factors. Greenhouse-Geisser corrections were applied where sphericity was violated, and generalized eta-squared (η^2_G) effect sizes are reported.

Total Task Time

Analysis of the total trial duration for completing the docking task revealed significant main effects of both Mismatch Type ($F_{1.69,18.59} = 9.52, p = .002, \eta^2_G = .056$) and Tangible Scale ($F_{1.87,20.55} = 10.16, p = .001, \eta^2_G = .070$, Figure 4.4A). Furthermore, a significant interaction (mismatch type × tangible scale) was found ($F_{2.00,22.00} = 8.45, p = .002, \eta^2_G = .090$), indicating that the impact of the visual-proprioceptive mismatches on overall completion time depended on the tangible proxy used. To further examine the root cause of these delays, I decomposed task completion time into two segments: the ballistic phase (moving time) and the fine-tuning phase (docking time).

Docking Time

Analysis showed a significant interaction between Mismatch Type and Tangible Scale ($F_{1.66,18.27} = 5.78, p = .015, \eta^2_G = .171$, Figure 4.4B). While the visual illusion did not slow down the hand in motion, attempting to precisely align a mismatched object caused a speed-accuracy trade-off. The penalty incurred by a visual mismatch was highly dependent on the tangible proxy being grasped. Specifically, the most severe penalty occurred when participants manipulated the largest proxy under a Positive Mismatch condition; docking time spiked to 4.61 seconds, compared to 2.20 seconds for congruent trials. In contrast, when manipulating the smallest proxy, docking performance remained relatively flat and unaffected by visual mismatches, fluctuating only slightly between 2.33 and 2.63 seconds across all conditions.

Moving Time

The RM-ANOVA revealed a significant main effect of Tangible Scale ($F_{1.61,17.72} = 5.67, p = .017, \eta^2_G = .040$), but no effect of Mismatch Type ($F_{1.27,13.94} = 0.06, p = .870$, Figure 4.4C). The interaction was also not significant ($p = .211$). This indicates that participants were

not influenced by the visual mismatch between the virtual object and the tangible proxy during rapid, feed-forward motor actions; participants adjusted their ballistic velocity based on the size of the physical object they were holding.

Cognitive Bias: Scale Estimation

To quantify perceptual accuracy (Q1), I calculated the absolute step error for each trial. Participants reported their estimation by selecting one of five displayed options representing the visual scales (12.5%, 7.9%, 5.0%, 3.1%, and 2.0%). Because tangible and visual scales were logarithmically spaced at discrete intervals of 1.6, this 5-point response interface directly mapped to the sequence of tangible proxy sizes in the study. A step error of 0 indicates perfect estimation of the tangible proxy size, while an error of 1 indicates estimation was shifted by exactly one scale step (e.g., estimating medium as a large proxy). To evaluate the extent to which the visual scale captured the user's cognitive estimation of scale, I analyzed the step error using RM-ANOVA.

The RM-ANOVA revealed a significant Mismatch Type × Tangible Scale interaction ($F_{2.53,27.85} = 4.97, p = .010, \eta^2_G = .146$). This demonstrates that "visual dominance" is not a static phenomenon; the participants' susceptibility to visual scale was strongly modulated by the physical dimensions of the proxy they were holding (Figure 4.4D). When participants grasped the smallest proxy, their estimation accuracy varied by mismatch direction: they were highly accurate during Negative Mismatches ($M = 0.12$ steps, 95% CI $[-0.04, 0.27]$) but exhibited strong visual dominance and poor accuracy during Positive Mismatches ($M = 0.75$ steps, 95% CI $[0.49, 1.01]$). However, when grasping the largest physical proxy, this visual capture effect flattened out, resulting in moderate and relatively consistent error rates across Negative ($M = 0.65$ steps, 95% CI $[0.35, 0.95]$), Congruent ($M = 0.65$ steps, 95% CI $[0.41, 0.89]$), and Positive Mismatches ($M = 0.43$ steps, 95% CI $[0.17, 0.70]$).

Subjective Detection: Sensation of Discrepancy

To confirm whether the system successfully induced the intended haptic illusions, I analyzed the perceived mismatch score (Q2). Ratings were scored from -2 ("Way too small") to +2 ("Way too big"), with 0 representing "Just right" (perceived congruence).

The results revealed a significant main effect of Mismatch Type ($F_{1.54,16.91} = 91.78, p < .001, \eta^2_G = .620$). Post-hoc pairwise comparisons confirmed that participants successfully distinguished between all three states ($p < .01$ for all comparisons). Positive Mismatches

were rated significantly higher (feeling "smaller") than Congruent trials, which in turn were rated significantly higher than Negative Mismatches (feeling "larger") (Figure 4.4E).

Interestingly, I also found a significant main effect of the Tangible Scale ($F_{1.39,15.34} = 37.71, p < .001, \eta_G^2 = .376$). The physical size of the proxy shifted the baseline of the illusion: holding the smallest proxy inherently biased participants toward feeling a Positive Mismatch, whereas holding the largest proxy biased them toward a Negative Mismatch. Importantly, there was no significant interaction ($p = .125$), demonstrating that the relative illusion of discrepancy is highly robust and functions similarly, regardless of the physical anchor's absolute size.

4.3.3 Qualitative Findings

Qualitative data from the semi-structured interviews were typed by the experimenter during the interviews. Given the focused nature of the four guiding questions, I employed an iterative review process rather than a formal thematic analysis. The experimenter conducted an initial pass of the interview notes to aggregate and identify recurring patterns and establish data saturation. To validate these initial observations and assist with conceptual synthesis, I utilized a large language model (LLM) as an assistive tool. Post-study interviews revealed five distinct themes related to sensory strategy, adaptation, and awareness of mismatch. The interview primarily consisted of four questions: difficulty, sensory reliance, strategy, and perception, with follow-up questions to probe for deeper insights.

Sensory Prioritization

Feedback suggested a reversal of the expected sensory processing order for their response strategy. Three participants (P_3, P_8, P_{11}) indicated that they assessed the proprioceptive mismatch (Q2) before formulating their visual estimation (Q1) during the docking task. Rather than relying solely on visual sensation, they first determined how the object "felt" relative to what they saw, and then used that sensation of discrepancy to deduce their visual scale estimate. This suggests that the sensory conflict was often the most salient signal.

Rapid Motor Recalibration

Participants frequently described the mismatch in terms of a change in sensitivity or "mouse sensitivity" (P_9). Their responses indicated that motor adaptation to the mismatch con-

ditions was rapid, often described as almost instant, or *"less than a second"* (P_5) upon grasping the tangible proxy. One participant (P_9) shared a specific cognitive loop: (a) forming a visual expectation (before touching the proxy), (b) sensing the physical discrepancy, and (c) immediately recalibrating their motor skill to recalibrate for the discrepancy. This aligns with the quantitative finding that task completion time remained stable; the cognitive cost of adaptation was paid upfront and quickly resolved, compensating quickly not to affect their overall completion time.

Anchoring Strategies

To resolve the scale ambiguity, participants employed specific anchoring strategies. First, body-centric strategies involved using the participant's own hand or finger span as a fixed "ruler" to gauge the tangible proxy's size (P_2, P_4). On the other hand, world-centric strategies relied on environmental cues in the virtual environment, such as the object's proportion relative to the miniature room boundary or background elements. One participant (P_9) reported shifting their strategy based on the perceived visual scale, relying more on global context when dealing with smaller objects because it provided less distinct proprioceptive feedback.

Perception of Difficulty

While extreme scales on both sides, largest and smallest, were often described as obvious or distinct (P_9, P_{11}), participants reported that the intermediate scale ratios (e.g., 3.1% and 5%) were the most cognitively challenging to differentiate (P_1, P_2). This suggests that differentiating scale was subjectively more difficult in the middle of the logarithmic range, whereas the extremes provide unambiguous sensory signals.

Tolerance of Conscious Mismatch

Consistent with the quantitative results, participants were highly aware of the scale mismatches between the visual and tangible proxies. A recurring theme was functional resilience. Despite being consciously aware of, or "bothered" by the discrepancy (P_5, P_7), participants agreed that it did not interfere with their ability to physically complete the docking task. This confirms that users can functionally adapt to significant proprioceptive discrepancies even when the sensory illusion breaks.

4.3.4 Observed Interaction (Experimenter Notes)

Complementing the formal measures, the experimenter documented recurring interaction behaviours. These observations are reported as complementary results that reinforce the quantitative and qualitative results.

Depth Perception

Participants consistently struggled with the z-axis (depth, forward direction) alignment, often '*undershooting*' the target for their docking task. This suggests that depth perception in the specific VR setup (e.g., the WiM was placed ~50 cm from their seated position) presented a challenge separate from the scale mismatch.

Grasp Ergonomics

Although participants were instructed to use their thumb and index fingers as a pinch grip, participants frequently adopted a third finger, often the middle finger. One participant (P_{12}) particularly noted the use of the third finger in a supportive role to stabilize the tangible proxy and distribute its weight, as well as in delicate manipulation (e.g., rotation of the tangible proxy).

Finger collision

Participants' fingers, mostly the index finger, as it was the most forward position toward the tangible proxy, often collided with the tangible proxy in mismatch conditions where the tangible proxy was bigger than its visual (virtual) object (negative mismatch). This finger 'undershooting' is a direct physical manifestation of the sensory conflict, yet participants swiftly managed to adapt to it.

4.4 Discussion

The results expand on how retargeting techniques can be used and adapted to different scales of objects for effective VR haptic interaction. While the sample size ($N = 12$) limits broad generalization, the large effect sizes observed in the analysis indicate the robustness of the trends. This discussion synthesizes both quantitative and qualitative

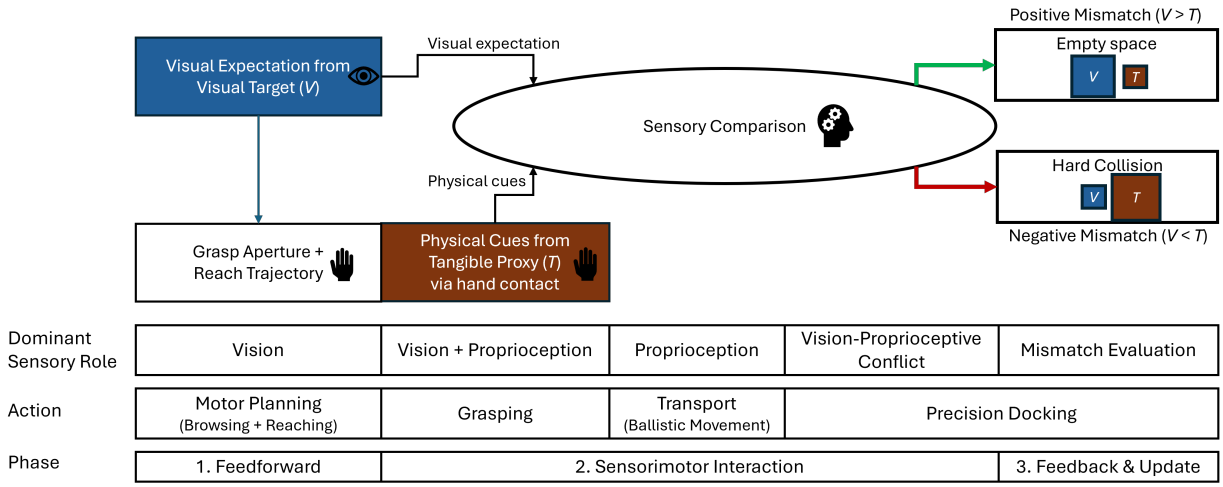


Figure 4.5: Phased visual haptic interaction model for tangible WiM. Vision drives feedforward planning, proprioception dominates transport, and precision docking exposes visual–proprioceptive (V–P) conflict. Positive mismatch creates empty space that can be tolerated, whereas negative mismatch produces collision cues that drive rejection.

insights into a phased model of the visual-proprioceptive cycle (Figure 4.5), explaining how vision and haptics collaborate across different stages of movement to maintain highly functional performance in the presence of retargeted sensory mismatch.

4.4.1 Phased Motor Control: The Limits of Visual Dominance

The primary finding is that visual dominance does not persist throughout the entire duration of tangible interaction in virtual reality. Rather, it is dependent on the phase of the movement. Prior literature establishes that during the initial motor planning phase, vision provides a critical feedforward signal, assessing the target's size to control the pre-shape of the grasp aperture (Marino, Stucchi, Nava, Haggard, & Maravita, 2010), effectively setting up the expectation. The participants then successfully start interacting with the tangible proxy. The empirical data from this study extend this motor planning phase into how a sensory *handshake* between vision and haptics occurs after this initial feedforward phase.

During the initial ballistic transport phase (*Moving Time*), the efficacy of visual dominance was severely limited. The results of this study show that visual scale had no significant impact on transport speed ($p = .870$). The physical dimensions of the tangible proxy were the primary driver of how fast participants moved their hands through space

($p = .017, \eta_G^2 = .040$). This indicates a heavy reliance on proprioceptive feedback rather than visual feedforward during macro-movement.

However, as users transitioned into the fine-motor control phase (*Docking Time*), the findings suggest that visual dominance emerged as the leading factor in the precision movement, with discrepancies between vision and proprioceptive feedback. This had a significant performance impact when the visual expectations for docking conflicted with the physical dimensions of the proxy in hand, especially the largest proxy under positive mismatch.

Furthermore, the qualitative findings regarding 'contextual grounding' suggest that the initial visual feedforward signal may rely on the WiM's dual-context rather than the scaled model alone. Relying on both full-scale and scaled models implies that the tangible WiM works on a dual-track perception to estimate and plan the eventual motor actions. The visual system plans ballistic movement based on the scaled model, while simultaneously semantically anchoring the scale judgment in a full-scale referent. This dual-track processing validates the core architecture of WiM, which states that the miniature and the full-scale environment must be visible simultaneously to support accurate spatial understanding.

4.4.2 Feedback: The Cognitive Reality Check

While the completion time analysis indicates that the motor system adapted to the sensory discrepancy during the transport phase, the qualitative analysis suggests that haptics drove cognitive judgement. After contacting the tangible proxy, the proprioceptive feedback served as a 'reality check', injecting physical evidence into the cognitive system.

This strongly suggests that, while vision provided feedforward control for the motor planning phase, the body acted as the cognitive anchor for evaluation. In the perceptual-physical cycle, the haptic feedback provided an 'error signal,' which the brain compares against the visual expectation, answering the feedforward command. As evidenced in the results, the participants could easily detect the mismatch (Q2), while they struggled to accurately estimate how much the mismatch was (Q1). In post-study interviews, participants noted that their decision-making process sometimes reversed, as they felt the discrepancy first, and then used that proprioceptive anchor to deduce the visual scale. (*Sensory Prioritization*)

This clearly distinguishes the role of haptic feedback from that of visual information in determining the mismatch, serving as a binary validator of reality. This underscores why multimodality in VR is essential not only for better immersion but also for delivering critical information as a channel for situational awareness.

4.4.3 The Exception: Asymmetry in Tolerance

However, participants' tolerance to sensory mismatch was not symmetric. The subjective analysis indicates that a positive mismatch ($V > T$) is accepted, but a negative mismatch ($V < T$) is widely rejected. This asymmetry in tolerance may be due to the physical constraint of collision between the body (the participants' hand) and the expectation (the visual information). This finding is consistent with previous literature, which demonstrates that VR users are significantly less sensitive to visual-proprioceptive discrepancy (offset) than to visual interpenetration of a virtual object (physics conflict) (Burns et al., 2005).

In the positive mismatch condition, the participants' fingers close earlier than expected relative to the large virtual object when grasping the smaller proxy. The wider aperture of fingers (thumb and index) detects empty space, and the brain resolves this discrepancy through visual dominance. This trend of tolerance for positive mismatch provides strong justification for tangible WiMs as a robust interaction paradigm. Because tangible WiMs inherently involve manipulating a scaled tangible proxy to control a larger, full-scale virtual counterpart, this natural bias toward 'empty space' effectively masks the scale discrepancy required for the technique to work. However, the results reveal a trade-off: while positive mismatches are subjectively stealthy, they incur an objective performance cost. With the large proxy, *docking time* doubled during positive mismatch ($4.61s$) compared to congruent trials ($2.20s$). This suggests that while the empty space masks the illusion, it simultaneously increases the difficulty of a precision task due to the control-display gain.

Conversely, in the negative mismatch condition, the tangible proxy (T) is larger than the visual expectation (V). This causes participants' fingers to collide with the proxy early, triggering a strong signal that breaks the illusion, leading to rejection.

Although the hand representation provided to the participants had an abstract skeletal structure, this phenomenon can be further explained by the 'hand-as-tool' metaphor in VR. The intention to reduce reliance on body-centric perception worked exactly as intended: participants treated the virtual hand representation as a tool rather than literal hands. This bypassed the aperture of the virtual hand representation and instead focused more on their proprioception (of their actual hands). This reveals a critical nuance to the theory of visual dominance–while vision dominates the spatial trajectory of the tool, proprioception dictates the physical state of the hand (grasp). Ultimately, this suggests that abstract representations facilitate functional flexibility, but cannot fully mask the physical constraints.

4.5 Implications for Tangible WiM Design

The results offer empirical guidance for designing tangible WiM-based VR . I distill these findings into four actionable recommendations for practitioners.

4.5.1 Scalability with Positive Bias

The findings on asymmetry suggest that if mismatch is unavoidable, biasing toward a smaller physical handle for different sized virtual objects may provide more natural experience. The results suggest that miniaturizing the tangible proxy (positive mismatch, $V > T$), relative to the virtual object, is supported and accepted by users for larger virtual objects, thereby minimizing breaks in immersion during VR tasks. This minimizes the risk of "hard collision" conflicts associated with negative mismatch ($V < T$) and leverages the brain's natural tolerance for visual capture in empty space. Furthermore, using smaller proxies leads to practical advantages, such as reduced manufacturing costs and material usage. However, this miniaturization strategy is not infinitely scalable; rather, it comes with a trade-off between speed and precision.

4.5.2 Tradeoff in Miniaturization

While miniaturization offers a robust interaction paradigm, the results reveal a complex trade-off between perceived mismatch and task precision. Statistical analysis confirms a significant main effect of tangible scale on *Total Task Time* (section 4.3.2) and an interaction effect on *Scale Estimation Error* (section 4.3.2). Interestingly, I observed that the large proxy served as a "perceptual buffer"; participants were unable to consciously detect positive mismatches (score ≈ 0.08). However, this "tactile stealth" came at a cost to performance: the *docking time* for this condition spiked to 4.61 seconds, doubling the duration of congruent trials and resulting in the slowest *Total Task Time* in the study ($13.03s$). Conversely, the small proxy assisted to produce fast trial times ($8.8s$) overall, but resulted in high *Scale Estimation Error* (0.75) during positive mismatches. With the large proxy, the negative mismatch was highly detectable (score ≈ 1.70) and physically rejected by participants.

Following maximum likelihood estimation (MLE) models (Ernst & Banks, 2002), I suspect that at the smallest scales, the brain's reliance on visual dominance is compromised by hardware noise (e.g., tracking jitter and display resolution). While the small proxy

hides the sensation of collision, the distinct tactile cues provide a conflicting physical signal. Therefore, I recommend that designers target the medium tangible scale for the best balance of speed and precision, using miniaturization as a tool to mask mismatch only when high-accuracy docking is not the primary task requirement.

4.5.3 Informative Haptics

Haptics act as a critical medium for cognitive sensemaking, rather than a secondary confirmation channel. While traditional VR design often treats non-visual senses as subordinate to visual dominance, the results suggest that haptics play a significant role in cognitive decision-making, particularly under conditions of ambiguity (the mismatch). The visual isolation of VR increases the user's reliance on physical cues for establishing cognitive spatial mappings. Therefore, designers can leverage haptic mismatch as an intentional semantic signal to highlight data in physical display form by carefully adjusting the proxy sizes, such as a small haptic proxy for sparse data, by contrasting the mismatch.

4.5.4 Ergonomics of Grasp

The observations suggest that using additional fingers (e.g., middle, ring, or pinky) for the tangible object is a natural behaviour, even for lightweight proxies, despite the instruction to use pinching with index and thumb. Biomechanically, this preference suggests that the middle finger offers superior planar opposition to the thumb for stable 6-DOF manipulation, whereas the index finger's higher position in a side-grip configuration may induce greater strain or instability. While traditional ergonomic models often treat grasping as an index+thumb interaction, the observations suggest that the middle+thumb grasp combination may be optimal for certain proxy geometries. These insights provide the recommended design for the tangible proxy as passive haptics for better affordances in coordination with mapped virtual objects. In conjunction, virtual objects can be designed to increase immersion and manipulative accuracy. Prioritizing these considerations is essential to reduce fatigue for prolonged tangible interaction in VR.

4.6 Chapter Summary

In this chapter, I presented an empirical study examining how visual–proprioceptive scale mismatch affects perception and manipulation performance in a tangible WiM-based VR

environment. Building on the scale judgment difficulties identified in chapter 3, this study isolated visual and tangible scale as experimental factors and examined how users responded when a physical proxy and its virtual representation differed in size.

The findings show that tangible proxies can support functional 3D manipulation even under visual–haptic mismatch, but this support is conditional. Participants were able to complete the docking task across mismatch conditions, suggesting that tangible interaction can preserve basic manipulation performance. However, the results also revealed that visual dominance was not uniform across the interaction. Physical proxy size influenced both task performance and scale judgment, indicating that users relied on both visual information and proprioceptive feedback when interpreting scale.

A central finding from this chapter is the asymmetry between positive and negative mismatch. Participants tolerated cases where the virtual object was larger than the tangible proxy more readily than cases where the tangible proxy was larger than the virtual object. This suggests that visual–haptic mismatch is not simply a matter of magnitude, but also of direction. In particular, physical collision and grasp constraints can break the illusion more strongly than empty space, revealing limits in how far visual dominance can support tangible WiM interaction.

This chapter supports the second argument of the dissertation: tangible proxies can provide a stable interaction reference in WiM-based VR, but their use requires careful management of scale mismatch, perceptual consistency, and manipulation performance. The study therefore provides empirical grounding for the later design framework by identifying conditions under which tangible proxies remain effective and conditions where mismatch becomes disruptive.

These findings motivate chapter 5, which explores how the tangible proxy mechanism examined in this chapter can be integrated into a broader tangible WiM-VR interaction framework. Rather than treating visual–haptic mismatch as a problem to eliminate entirely, chapter 5 examines how mismatch can be managed through design principles, haptic retargeting, and multi-view interaction.

Chapter 5

Design Space and Framework for Tangible WiM-VR Interaction

This chapter addresses RQ3 by developing a design space and framework for managing discrepancies in tangible WiM-based VR interaction. Chapter 4 empirically examined how visual–proprioceptive mismatch is perceived and affects manipulation performance when pairing tangible proxies with differently scaled virtual objects. The phased model from chapter 4 motivates the temporal structure of this chapter: discrepancy may be described across interaction states. The observed asymmetry between positive and negative mismatch further suggests that discrepancy should not be treated as directionally neutral. This chapter transitions from empirical observation into a design framework for describing how interaction techniques distribute and manage spatial discrepancies temporally and spatially.

To develop this framework, I examine representative spatial interaction techniques and organize them across interaction spaces, virtuality, and interaction states. By analyzing these techniques, this chapter shows how WiM space changes the management of these discrepancies by introducing an intermediate interaction space between local embodied action and world-scale virtual effect. The analysis reveals major tradeoffs in deciding whether to preserve direct physical–virtual alignment, introduce WiM-based mediation, or redistribute mismatch across interaction states. This chapter lays the groundwork for chapter 6, outlining specific design implications and potential configurations.

5.1 Introduction

Virtual reality (VR) interaction is valuable in domains where users' primary task involves different perspectives and viewpoints at varying scales, such as AEC. A fundamental discrepancy between the virtual environment and the real world limits how users interact with virtual objects and spaces in VR. Existing haptic retargeting techniques (Azmandian et al., 2016; Clarence et al., 2024; Matthews, Reichherzer, Thomas, & Smith, 2023; Matthews, Thomas, Von Itzstein, & Smith, 2022) address how a limited set of physical props can be used for multiple virtual objects. However, virtual objects may be diverse in many characteristics, with objects of various sizes and shapes, and fabricating physical controls for every virtual object becomes quickly impractical.

Previous work has addressed this discrepancy in different ways. For example, haptic retargeting (Azmandian et al., 2016) leverages perceptual thresholds for change blindness by reusing and remapping spatially displaced tangible proxies and virtual objects, aligning the embodied object with the virtual object. Other approaches, such as OVRlap (Schjerlund et al., 2022), change how user actions are mapped into the virtual world, by explicitly remapping through alternate spaces and multi-view systems. WiMs provide another approach by offering a scaled representation of a larger environment (Stoakley et al., 1995). While these techniques have been heavily researched, their combination remains underexplored. In this chapter, I position tangible WiM interaction as a special case, introducing three coordinated spaces: the embodiment, the WiM space, and the full-scale world. These then can be independently transformed, allowing distribution of discrepancies across all three spaces.

To identify design opportunities, I explore the design space of TUIs, WiMs, and retargeting techniques and characterize how transformations of space are used to accommodate the discrepancy between virtual and physical for object manipulation through a tangible proxy in VR. Little work has examined how an intermediate space can be used to absorb the discrepancy across interaction spaces. Introducing an intermediate space allows independent transformations of position and scale across different interaction spaces. This broadens the discrepancy management opportunity not only in positional differences but also in scale.

The literature suggests that intermediary interaction spaces provide an underexplored design opportunity for managing spatial discrepancies between physical and virtual interaction. By mediating user interaction through an intermediate space, discrepancies, such as position displacement and scale differences, can be distributed across embodiment, intermediate, and world spaces.

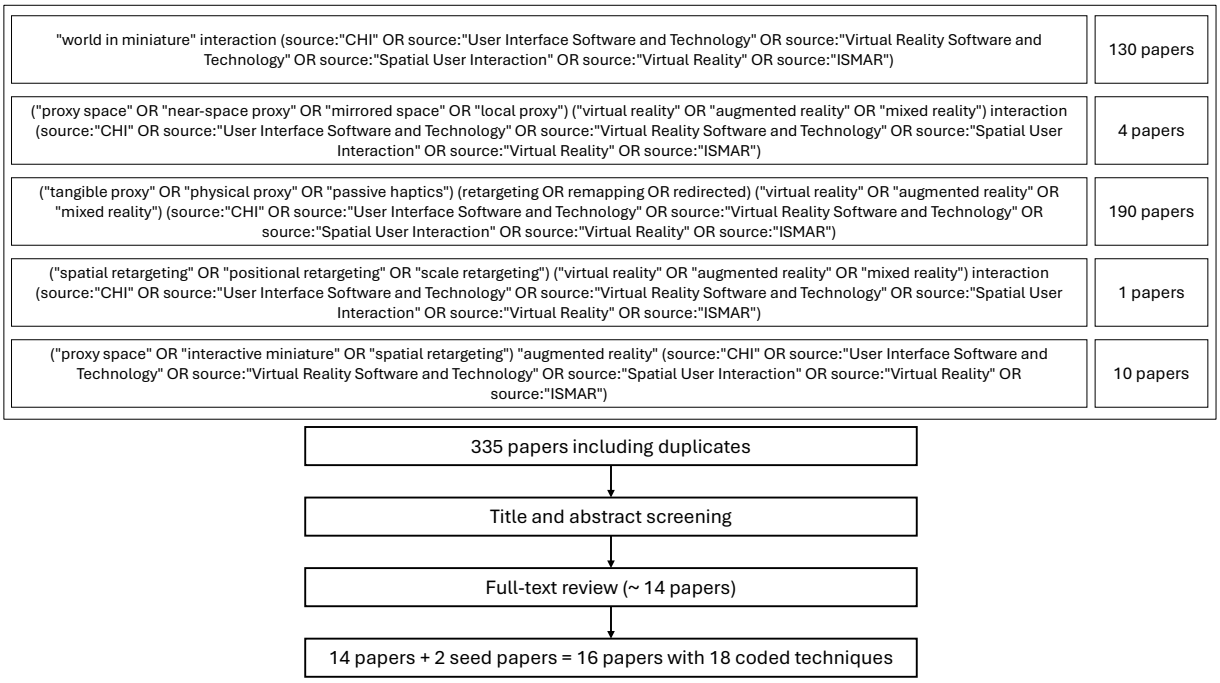


Figure 5.1: Five phases used and my review process.

This chapter's contributions are twofold. First, I introduce a design vocabulary that provides a consistent language for describing and categorizing interaction techniques into four design dimensions, namely Interaction Route, Mismatch Allocation, Tangibility and Visibility. Second, I apply this framework to analyze representative techniques from haptic retargeting and WiM-based interaction, identifying how WiM changes the management of spatial discrepancies between physical embodiment and the full-scale virtual world. The analysis leads to design implications and identifies an underexplored design opportunity, using intermediary spaces for retargeting. This motivates the Laminar interaction technique presented in chapter 6.

5.2 Exploring the Haptic, Intermediary, Virtual Design Space

To establish a design space for exploring haptic/physical experiences in a mixed reality environment via an intermediary WiM space, I first reviewed relevant interaction techniques

from the literature. My review focused primarily on recent publications, specifically from 2022–2025, and in related venues, such as ACM CHI, ACM VRST, ACM UIST, ACM SUI, IEEE VR, and IEEE ISMAR. My corpus is not an exhaustive list of techniques; papers were selected based on their relevance to the research context, such as tangibility, retargeting, and alternate space usage for object manipulation. I also included a small set of 'seed' papers, including haptic retargeting (Azmandian et al., 2016), classic WiM (Stoakley et al., 1995) to build on my design dimensions of interest.

Papers were included if they described interaction techniques involving object-level manipulation, retargeting mechanisms, use of proxies, intermediary spaces, or transformed spatial representations (Figure 5.1). Papers focused on locomotion, navigation, rendering, visualization, or tracking capabilities were excluded. Techniques without tangibility support were retained when they introduced interaction through space management such as transformations or intermediary representations that altered how users manipulated virtual content. This screening intentionally favoured object manipulation through spatial transformations, which addresses the interaction problem, rather than surveying a broader corpus including all WiM, VR, or haptic-related publications.

From the papers I reviewed, I derived a set of properties that distinguish interaction techniques from one another. I began with haptic retargeting as a seed technique because it presents techniques that overcome certain physical-digital barriers, specifically position and orientation differences between real-world and VR experiences. My goal was to identify cases to cover a wide breadth of designs rather than to conduct an exhaustive review of all research publications. I included techniques focused on the spatial transformation of objects and spaces, and how their relationship changes during interaction.

I selected representative techniques based on their relevance to the research problem and their position within my design space, when the literature contains multiple variants. I then coded each technique using provisional dimensions, iteratively refining their definitions and values when existing design dimensions could not meaningfully distinguish the reviewed techniques.The initial coding focused on identifying where interaction occurred, where transformations were applied, what physical grounding was available (Tangibility), and whether transformations were exposed or hidden from users (Visibility). All reviewed techniques could be described using these initial codes, suggesting that they captured recurring characteristics across the corpus. Dimensions were retained when they grouped multiple techniques and were removed or merged when they did not provide meaningful variation.The resulting dimensions were then used to construct a coded matrix of all reviewed techniques (Figure 5.3), allowing common patterns to be identified.

I established a design vocabulary whose values were used to describe where interactions

Table 5.1: My 2 (virtuality) × 3 (interaction space) model that I use as the spatial properties that provide the foundation for my design space framework.

	Physical	**Virtual**
Embodiment (E)	hand/finger, controller, tangible proxy	virtual hand, controller model, cursor/tool
Intermediate (I)	miniature physical model	WiM container and objects
World (W)	room, desk, wall	full-scale virtual environment

occur, where representations exist, and when discrepancies are managed during interactions. Using this vocabulary, I iteratively coded the reviewed techniques and refined a set of dimensions that ultimately formed the design framework presented in the following sections.

5.3 Design Space Foundations

In my review of existing interaction techniques that leverage some combination of enabling tangibility and interaction through intermediate (WiM) spaces, the concept of mismatch required an understanding of *virtuality* (i.e., what objects/artifacts existed virtually vs. physically) and in what *interaction space* (i.e., a person's body, the full-scale environment, and/or the world-in-miniature) action occurred in. These provide a consistent vocabulary used in section 5.4 to derive a set of design dimensions and analyses that characterize how interaction techniques manage physical–virtual discrepancy.

To describe this coordination, I use a 2 × 3 model (Table 5.1) that combines virtuality and interaction spaces to define the spatial properties of interaction techniques. In my review of literature, I also noted that movement and change over time significantly impacted each interaction technique's instantiation, and so also considered these *interaction states* as a third temporal property.

In this section, I provide descriptions of these spatial and temporal properties, which provide the foundation for my design framework. The design framework (section 5.4) then leverages these properties to describe how interaction techniques from the existing literature map across virtuality (e.g., to handle mismatch), leverage interaction spaces (e.g., directly in full-scale space or through an intermediate WiM), and what interaction states are used to handle any mismatch (e.g., when reaching or when browsing).

5.3.1 Virtuality

I distinguish between a *physical* layer and a *virtual* layer. This distinction is informed by the reality–virtuality continuum, ranging from physical environments to fully virtual environments (Milgram & Kishino, 1994). In this framework, this is collapsed into two analytical layers to identify where each interaction entity exists in the coupling model, rather than treating it as a continuous measure. This simplification is adopted because the framework focuses on how embodiment, WiM, and world spaces are coordinated across physical and virtual entities, rather than on measured degrees of mixed reality (MR).

Physical

refers to entities, actions, or environmental features that can be touched or tracked. This includes the user's body, gaze/head direction, physical controllers, tangible proxies, physical miniature models, and the surrounding physical environment.

Virtual

refers to rendered or simulated entities in VR. This includes virtual hands, cursors, controller models, WiM containers, miniature objects, and full-scale virtual environments.

The presence of an additional interface alone does not establish an intermediate interaction space. The additional space literally mediates user action and flows it into a full-scale world space.

5.3.2 Interaction Spaces

I define three interaction spaces to describe tangible-WiM-VR interaction: embodiment space, intermediate (WiM) space, and world space. These spaces describe where action originates, where intermediate interaction occurs, and where interaction effects are realized.

Embodiment space

Embodiement space refers to the user's means of action. This includes the physical hand, fingers, head or gaze, VR controller, and tangible proxies held in hand. It also includes their virtual representations, such as a virtual hand, controller model, cursor, or virtual tool. A

tangible proxy belongs to embodiment space when it is used as an interaction instrument. This space is motivated by previous work on body-relative and direct interaction in virtual environments, which leverage the user's body and proprioception as central resources for interaction (Mine, Brooks, & Sequin, 1997; Poupyrev, Billinghurst, Weghorst, & Ichikawa, 1996).

Intermediate space

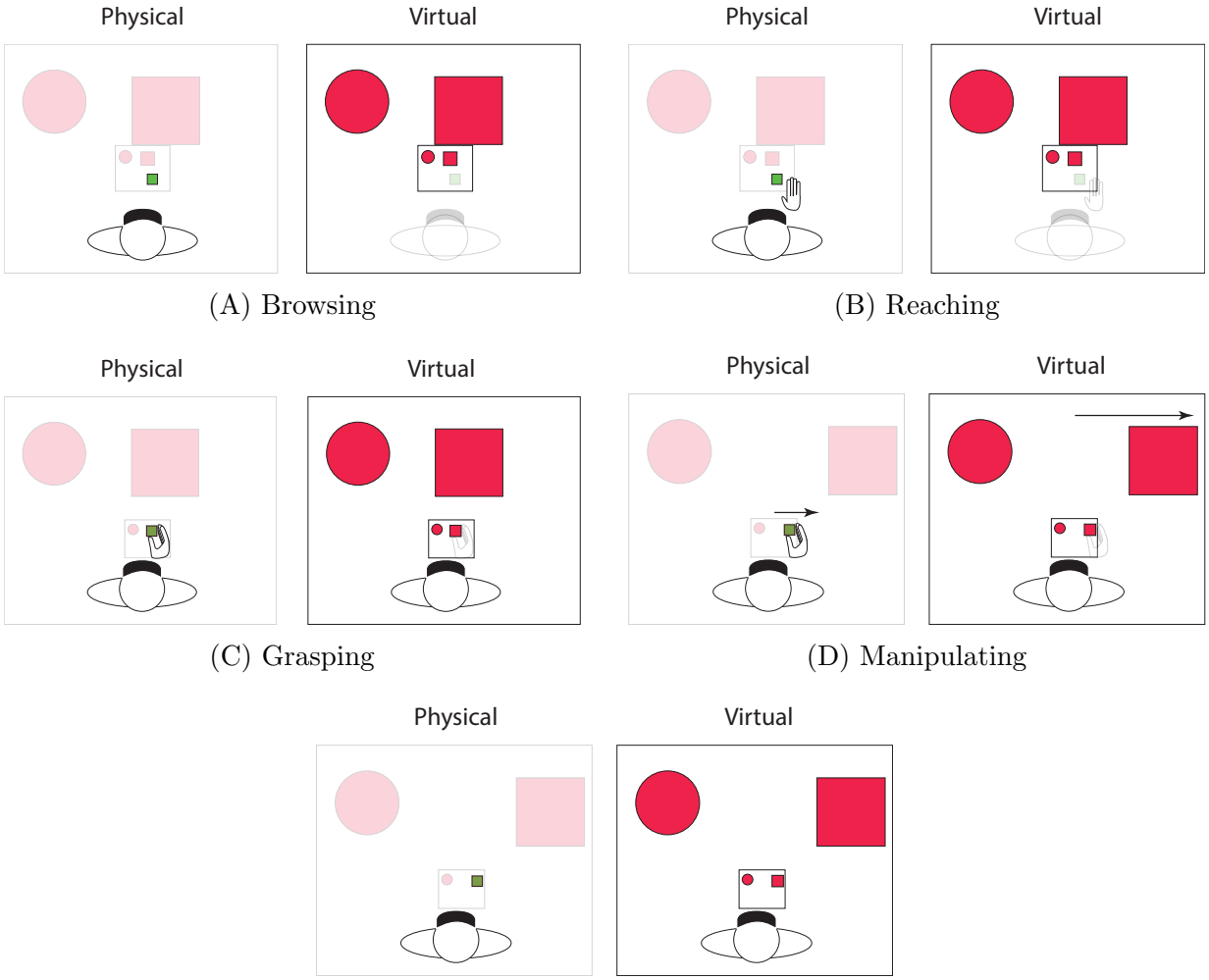
Intermediate (WiM) space refers to an intermediate miniature interaction space. Previous work introduced WiM as a miniature representation of the full-scale world in VR (Stoakley et al., 1995). This includes virtual WiM containers and miniature objects in the virtual layer. Although WiMs often indicate virtual WiM space, a WiM space may also exist physically, for example, as a physical miniature model, tabletop model, or tangible scaled representation.

World space

World space refers to the full-scale environment where user interaction effects are ultimately realized. This is distinguished from WiM space because world space is the full-scale referent environment that the WiM represents, controls, or mediates, and contains the objects and environments that users ultimately intend to observe or manipulate. Like WiM space, world space may exist physically, virtually, or both. In some cases, the physical world may be absent when the system does not explicitly model or render the surrounding physical environment. When present, the physical world includes the physical space (e.g., room, desk, floor, and walls), and the virtual world includes the full-scale virtual environment and world-scale virtual objects.

5.3.3 Interaction States

Interaction states represent common stages in object-based 3D interaction. Previous 3D UI work identifies selection and manipulation as central tasks in 3D user interfaces (Bowman, Kruijff, LaViola, & Poupyrev, 2004), and reach-to-grasp studies show that visual representations of the body affect reaching and grasping behaviour, including hand pre-shaping during grasping (Marino et al., 2010). Based on these foundations, I define five analytical states for tangible-WiM-VR interaction: browsing, reaching, grasping, manipulating, and releasing. These states are interaction-space agnostic. Depending on the technique,



(A) Browsing (B) Reaching

(C) Grasping (D) Manipulating

(E) Releasing

Figure 5.2: Interaction states used in the temporal design framework: (A) browsing, (B) reaching, (C) grasping, (D) manipulating, and (E) releasing. The diagrams illustrate one tangible-WiM example.

interaction may occur in embodiment space, WiM space, world space, or across multiple spaces.

Browsing

Browsing (Figure 5.2A) refers to the stage where users explore the environment and identify a potential target.

Reaching

Reaching (Figure 5.2B) describes directed movement toward the intended object or interaction space.

Grasping

Grasping (Figure 5.2C) is the state when the user establishes a coupling with the target through contact, selection, grasp, or intended action to interact.

Manipulating

Manipulating (Figure 5.2D) is the stage where the embodiment, object, or interaction space is coupled and continuously transformed through interaction.

Releasing

Releasing (Figure 5.2E) is the stage where the user decouples from the target and the technique resolves the spatial relationship. The transformed mapping may return to its initial state, persist, or change its transformation to prepare for the subsequent interaction.

These states provide a common analytical sequence rather than a strict interaction phase that every technique fits. These states temporally determine when relationships among embodiment, intermediate, and world spaces become active, transformed, or terminated.

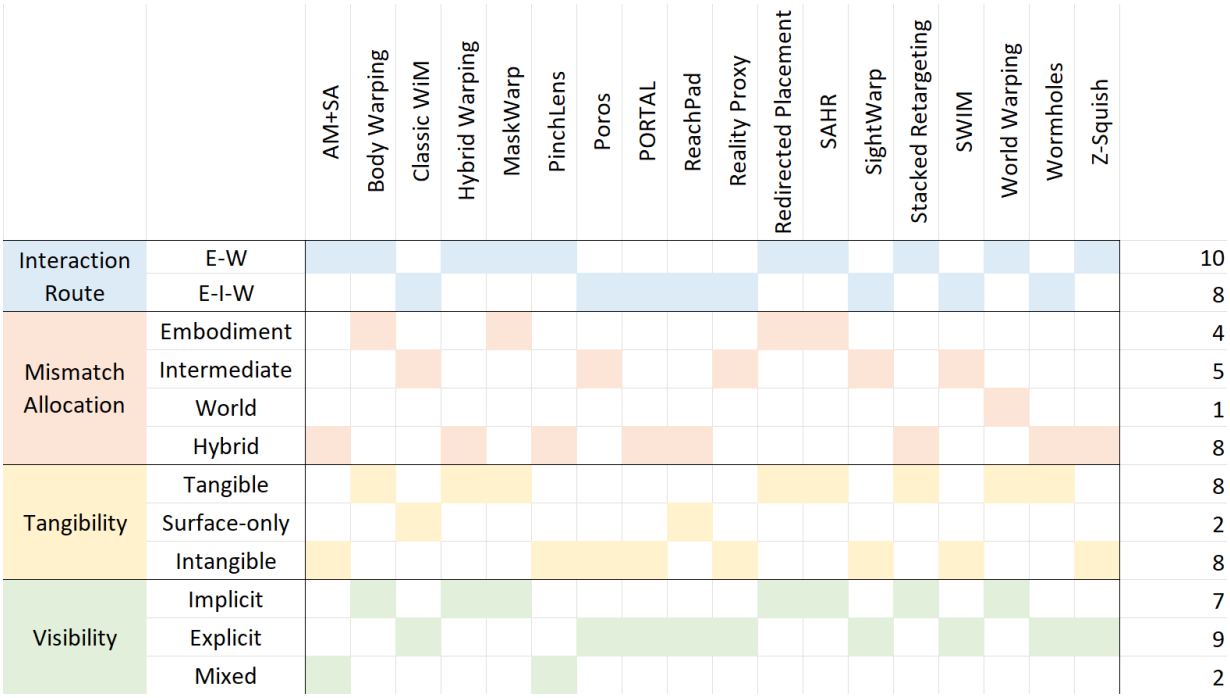


Figure 5.3: A matrix of reviewed techniques across the four framework dimensions. Colored cells indicate the route, mismatch allocation, tangibility, and visibility assigned to each technique.

5.4 Design Framework

Interaction spaces, virtuality, and interaction states provide a vocabulary for comparing interaction techniques. Using this framework, I examine how techniques distribute transformations across embodiment, intermediate, and world spaces, when those transformations occur, and how they are coordinated throughout interaction.

5.4.1 Dimensions for Handling Mismatch

I identified four dimensions which allow us to characterize each technique in my literature review (section 5.2) with respect to how they allocate and organize mappings to overcome mismatches between physical interaction and virtual content. These dimensions are intended to support the classification of existing techniques by describing where discrepancy

is introduced, how it is distributed, what physical grounding is available, and whether transformations are expected to be perceptually hidden.

Interaction Route

The interaction route describes the sequence of interaction spaces where the user action propagates to the full-scale world space. I identified two routes: techniques that resolve the discrepancy directly, and techniques that use intermediate space. The *direct route* from embodiment to world space (E-W) omits the intermediate interaction space. All haptic retargeting techniques, for instance, direct user actions from their embodiment to a full-scale virtual environment. Body warping (Azmandian et al., 2016) demonstrates that user action is realized without an intermediate step. PinchLens (Zhu, Sidenmark, Sousa, & Grossman, 2023) zooms in on the world space itself instead of introducing a separate visual interface. An *intermediate route* (E-I-W) passes user action from embodiment space through an intermediate space before affecting the full-scale world space. Classic WiM (Stoakley et al., 1995) leverages a scaled interface of a full-scale virtual environment to enable interaction across a larger space. A significant number of techniques in my corpus follow a similar mechanism of mediating user action through an intermediate step (Ban, Matsumoto, Narumi, & Kuzuoka, 2022; Pohl et al., 2021; H. Shi et al., 2025).

Mismatch Allocation

The mismatch allocation describes which interaction spaces accommodate the spatial discrepancy between user action and the final effect. This is determined by identifying where mismatch is being managed. Although multiple spaces may be involved in an interaction technique, mismatch allocation is assigned to the interaction spaces that absorb the discrepancy through transformations of representations, mappings, or both. Techniques that distribute across multiple interaction spaces are classified as hybrid allocations. *Embodiment allocation* modifies the virtual hand representation, such as body warping, Redirected Placement, and MaskWarp (Azmandian et al., 2016; Matthews, Reichherzer, Thomas, & Smith, 2023; X. Yang et al., 2023). *Intermediate allocation* instead absorbs the discrepancy within its intermediate space, such as classic WiM (Pivovar, DeGuzman, & Rosenberg, 2022; Pohl et al., 2021; Stoakley et al., 1995). *World space allocation* transforms its full scale environment either implicitly or explictly, such as world warping, Z-Squish (Azmandian et al., 2016; Jennings & Matviienko, 2025). *Hybrid allocation* uses a mix of allocation approaches, transforming multiple interaction spaces; this includes hybrid warping, Worm-

holes, ReachPad, Portal, PinchLens (Azmandian et al., 2016; Ban et al., 2022; Han, Kim, & Cho, 2022; H. Shi et al., 2025; Zhu et al., 2023).

Tangibility

Tangibility describes the degree of physical grounding available during interaction. I distinguish between tangible, surface-only, and intangible techniques. *Tangible* techniques use tangible proxy mapped to a virtual object, providing tactile feedback, such as Wormholes, MaskWarp, and Redirected Placement (Ban et al., 2022; Matthews, Reichherzer, Thomas, & Smith, 2023; X. Yang et al., 2023). Some techniques only provide a *surface-only* tactile feedback, for example, ReachPad and Classic WiM (H. Shi et al., 2025; Stoakley et al., 1995) provide tactile surface (e.g., a clipboard) to accommodate interaction instead of mapping a tangible proxy to a virtual object. While tangibility provides valuable feedback, some used retargeting mechanisms without haptics (*intangible*), such as Poros, Reality Proxy, and SightWarp (X. Liu et al., 2025; Y. Liu et al., 2025; Pohl et al., 2021).

Visibility

Visibility describes whether transformations are perceptually hidden from the user or intentionally exposed. Techniques may be implicit, mixed, or explicit. Haptic retargeting techniques and their relatives are *implicit*, since they use change blindness, increasing the degree of transformations (Azmandian et al., 2016; Clarence et al., 2024; Matthews, Reichherzer, Thomas, & Smith, 2023), while current techniques that use intermediate space are *explicit*, as they do not necessarily rely on the user's perceptual threshold in most cases, since this space may be exposed already (Han et al., 2022; Y. Liu et al., 2025; Pohl et al., 2021; Stoakley et al., 1995). Some edge cases, such as PinchLens and AM+SA, (Shin et al., 2022; Zhu et al., 2023) suggest *mixed* visibility, in which some aspects of the transformation are exposed while other aspects are hidden during interaction. PinchLens transforms world space explicitly, and its gain of cursor movement is changed during interaction.

5.4.2 Spatial Analysis

The interaction route primarily describes two groups of interactions. Without intermediate space, techniques are forced to use embodiment space, world space or both for retargeting. Techniques like haptic retargeting variants, Z-squish, Redirected Placement, MaskWarp (Azmandian et al., 2016; Jennings & Matviienko, 2025; Matthews, Reichherzer, Thomas,

Table 5.2: Spatial analysis. Proposed Laminar variants are shown in bold. Laminar-Hand and Laminar-WiM occupy two intermediary-space retargeting configurations identified through the design-space analysis. Hybrid allocation techniques appear in multiple rows.

		Interaction Route	
		E–W	**E–I–W**
Mismatch Allocation	Embodiment	body, hybrid warping (Azmandian et al., 2016), Z-squish (Jennings & Matviienko, 2025), AM+SA (Shin et al., 2022), Stacked retargeting (Clarence et al., 2024), Redirected Placement (X. Yang et al., 2023), MaskWarp (Matthews, Reichherzer, Thomas, & Smith, 2023), PinchLens (Zhu et al., 2023), SAHR (Matthews et al., 2022)	Portal (Han et al., 2022), Wormholes (Ban et al., 2022), ReachPad (H. Shi et al., 2025)
	Intermediate	—	classic WiM (Stoakley et al., 1995), Poros (Pohl et al., 2021), SWIM (Pivovar et al., 2022), Reality Proxy (X. Liu et al., 2025), SightWarp (Y. Liu et al., 2025), Portal (Han et al., 2022), Wormholes (Ban et al., 2022), ReachPad (H. Shi et al., 2025)
	Full-scale world	World, Hybrid warping (Azmandian et al., 2016), Z-Squish (Jennings & Matviienko, 2025), AM+SA (Shin et al., 2022), Stacked retargeting (Clarence et al., 2024), PinchLens (Zhu et al., 2023)	—

& Smith, 2023; X. Yang et al., 2023) distribute discrepancy by modifying the virtual hand, the virtual world, or a combination of both. The two spaces limit where the discrepancy can be distributed. Hybrid allocation techniques appear in multiple cells because they distribute discrepancy across more than one interaction space. On the other hand, with intermediate space, it becomes possible to use this in a retargeting mechanism. The most common pattern I observed from existing techniques (Ban et al., 2022; Han et al., 2022; Y. Liu et al., 2025; Pohl et al., 2021; H. Shi et al., 2025; Stoakley et al., 1995) uses transformation through a variety of the three spaces, but primarily relies on intermediary space.

While intermediary techniques (E-I-W) involve both embodiment and intermediate spaces, they use each space for a mutually exclusive retargeted space. While hybrid retargeting (Azmandian et al., 2016) distributes discrepancies across both embodiment and world spaces, little research examines how a hybrid approach using embodiment and intermediate space would work for retargeting.

5.4.3 Transformation Interval Analysis

I also analyze how the interaction techniques I reviewed leverage the five interaction states temporally to achieve retargeting. I construct Table 5.3 to characterize when transformations are introduced (*transformation start*) and when they reach their intended state (*transformation peak*) to eventually match the discrepancy.

Pre-contact alignment

Pre-contact alignment techniques introduce and resolve discrepancies before the user makes contact with objects. All haptic retargeting variants (Body warping, World warping, Hybrid warping), SAHR, MaskWarp, PinchLens, and Stacked Retargeting (Azmandian et al., 2016; Clarence et al., 2024; Matthews, Reichherzer, Thomas, & Smith, 2023; Matthews et al., 2022; Zhu et al., 2023) reach the transformation peak at the grasping of the object. Of these techniques, the biggest distinction is when the execution of compensation starts (rows of Table 5.3). World warping and hybrid warping start the transformation when browsing (Browsing-Grasping), whereas other techniques begin transformations when the user starts reaching out to the target object (Reaching-Grasping). These prioritize congruent visual–tactile feedback at the moment of contact, while most of them introduce transformation implicitly, benefiting from change blindness.

Table 5.3: Temporal analysis showing when techniques introduce discrepancy management and when they reach their intended state during interaction. Proposed Laminar variants are shown in bold.

		Peak State				
		Browsing	**Reaching**	**Grasping**	**Manipulating**	**Releasing**
Starting State	**Browsing**	—	—	World warping, Hybrid warping	—	Classic WiM, Poros, SWIM, Reality Proxy, SightWarp
	Reaching	—	—	Body warping, MaskWarp, SAHR, Stacked Retargeting, PinchLens	—	Portal, Wormholes, ReachPad, Z-Squish, AM+SA
	Grasping	—	—	—	—	Redirected Placement
	Manipulating	—	—	—	—	—
	Releasing	—	—	—	—	—

Post-contact adjustment

Redirected Placement (X. Yang et al., 2023) demonstrates a rare edge case in terms of the transformation in my design framework. It shows that discrepancy management is still possible after contact has been made, heavily prioritizing the final destination-oriented instead of stabilizing manipulation.

Passive workspace

Techniques with intermediary space show a clear trend in maintaining their transformation at a persistent level regardless of interaction states, actively leveraging their third space. Classic WiM, Poros, SWIM, Reality Proxy, and SightWarp (X. Liu et al., 2025; Y. Liu et al., 2025; Pivovar et al., 2022; Pohl et al., 2021; Stoakley et al., 1995) establish a workspace that remains active across interaction states, rather than necessarily prioritizing congruent contact alignment.

Active workspace

Another pattern shows a group of techniques that actively transform throughout the interaction states. Portal, Wormholes, ReachPad, Z-Squish, and AM+SA (Ban et al., 2022; Han et al., 2022; Jennings & Matviienko, 2025; H. Shi et al., 2025; Shin et al., 2022) begin transforming before the interaction and maintain their transformations until the interaction concludes. These techniques usually require user intervention to trigger and terminate the mechanism.

The spatial and temporal analysis of related techniques reveals a distinction between haptic retargeting and intermediary space techniques. Techniques that use the direct approach (E-W) mainly perform discrepancy management through pre-contact, reaching their peak transformation before object contact, whereas intermediary space techniques (E-I-W) primarily maintain persistent mediated workspaces or actively invoked workspaces. This suggests an unexplored combination: techniques that employ an intermediary interaction space while using pre-contact alignment to achieve congruent haptic feedback.

5.5 Design Implications

The primary value of modelling interaction techniques using the framework (section 5.4) is that it exposes systematic design patterns that can serve as guidelines. The analysis shows

a shift from techniques that manage discrepancy through direct interaction routes (E-W), such as haptic retargeting, to techniques that introduce WiM as an intermediate interaction space (E-I-W). This shift creates two core tradeoffs: whether to manage discrepancy through direct physical–virtual correspondence or WiM-based mediation, and whether to prioritize local alignment or world-scale interaction flexibility. The spatial and temporal analyses (table 5.2 and table 5.3) suggest that interaction techniques differ not only in where discrepancies are managed in interaction spaces, but also when transformations are introduced and peaked throughout interaction (temporally).

5.5.1 Direct Coupling vs. WiM Mediation

The spatial analysis reveals a fundamental distinction between techniques that manage discrepancy through direct physical-virtual correspondence (E-W) and those that leverage intermediary (WiM) interaction spaces (E-I-W).

Direct manipulation can be considered a baseline for 3D interaction in VR. In direct manipulation, physical embodiment and virtual representation are tightly coupled, offering intuitive interaction with matching visual–proprioceptive expectations. However, this configuration introduces a bounded ceiling, as interaction in VR remains limited by the user's physical reach, available workspace, and physical proxy configuration.

To mitigate this, interaction techniques introduce mediated spatial relationships. Haptic retargeting manages this by altering physical–virtual correspondences in embodiment or world space. For example, body warping transforms the relationship between physical and virtual embodiment, while world warping transforms the relationship between physical and virtual world space (Azmandian et al., 2016). These techniques preserve the possibility of local physical contact by adjusting position alignment, but rely on implicit transformations to adjust the correspondence between physical and virtual spaces.

WiM-based techniques offer another strategy. Introducing an intermediate interaction volume, such as a WiM (Stoakley et al., 1995), or mirrored and multiplexed views (Pohl et al., 2021; Schjerlund et al., 2022), allows users to interact with objects through contained, mirrored, or multiplexed environments from a limited physical workspace. In this case, the design choice is not simply whether to preserve or break direct manipulation, but where mediation and abstraction should occur. The WiM provides a way to explicitly contain abstraction, rather than hiding all discrepancies inside body or world warping.

Implicit warping techniques are bounded by human perceptual thresholds. For example, users can tolerate only a limited amount of translational or rotational gain before consciously detecting transformation changes (Azmandian et al., 2016; Steinicke et al.,

2010). Explicit interaction spaces, such as WiM, offer another strategy: instead of hiding the discrepancy, they expose and contain it within a visible representation. This makes WiM mediation useful when discrepancy cannot be hidden reliably within body or world correspondences.

When discrepancies exceed physical limits or perceptual detection thresholds, interaction can fail in different ways. First, a perceptual break occurs when users detect the mismatch, and the illusion breaks. An implicit warping technique may shift into an explicit sliding effect, becoming noticeable to users and breaking the illusion of embodiment. Second, the interaction loop can break because users under- or overshoot the target, miss contact, or experience discomfort from severe visual–vestibular conflict (Azmandian et al., 2016; Razzaque et al., 2002; Steinicke et al., 2010).

The key temporal difference is where discrepancy management is concentrated. Haptic retargeting leverages visual dominance during pre-contact interaction states, whereas WiM-based interaction uses a distributed strategy across temporal states to achieve stable interaction. This makes haptic retargeting primarily an alignment strategy. The transformation is concentrated in states before or around grasping, and the goal is to make later manipulation feel stable and direct. In contrast, WiM-based techniques shift this structure. Rather than treating discrepancies to be hidden before contact, WiM-based interaction provides a mediating space where discrepancies can be contained, distributed, and made useful during interaction.

5.5.2 Local Alignment vs. World-scale Flexibility

The design framework suggests a systemic tension between local alignment and world-scale flexibility. Local alignment refers to the user's ability to maintain a stable relationship between the body, proxy, or tool and the object being manipulated. World-scale flexibility refers to the ability for that local action to affect distant, scaled, or reconfigured world content. Preserving local alignment often requires physical movement, system-driven adjustment, or both. In contrast, leveraging WiM-based mediation can reduce physical effort, but shifts the burden onto mental interpretation: users must understand how local action inside the WiM relates to world-scale effect.

This tension appears differently in techniques without and with WiM. Techniques such as haptic retargeting preserve local contact by transforming embodiment or world spaces before the user makes contact. The system attempts to make the tangible proxy and virtual target align so that the subsequent manipulation phase can proceed as if the interaction

were direct. This supports local alignment, but depends on prediction, perceptual masking, and tangible proxy arrangements.

With WiM, systems can preserve local interaction while allowing the full-scale effect to differ from the local action. The user may manipulate a miniature object, or a WiM container, while the full-scale world space receives the mediated result. This supports larger-scale, distant, or cross-scale interaction, making WiM interaction versatile under larger spatial discrepancies, but requires the user to interpret the relationship between the local WiM action and the resulting full-scale transformation.

This distinction becomes critical for tangible-WiM interaction. A tangible proxy can provide physical grounding, but the WiM determines how that local action becomes a full-scale world effect. Therefore, the design problem is not only how to align the proxy with a virtual object, but also how to make the proxy–WiM–world relationship understandable across interaction states.

5.5.3 Designing with Asymmetric Tolerance

The findings in chapter 4 suggest that scale discrepancy should not be treated as directionally neutral. Positive mismatch, where the virtual object is larger than the tangible proxy, and negative mismatch, where the tangible proxy is larger than the virtual object, create different perceptual and interaction consequences. Within the interaction spaces $\times$ virtuality framework, this asymmetry narrows the design choices for how discrepancy should be redistributed across time and space.

Temporally, scale discrepancy should be managed before contact whenever possible. The browsing and reaching states provide opportunities to adjust the relationship between the tangible proxy, virtual WiM object, and full-scale referent before grasping establishes physical contact. This is especially important because negative mismatch can produce premature collision cues once the user reaches or grasps the proxy. Therefore, when scale compensation is required, designers should prioritize transformations during pre-contact states, such as rescaling the WiM, adjusting the virtual hand representation, or preparing proxy–object alignment before the user enters the manipulating state.

Spatially, WiM provides the intermediate space where this redistribution can occur. Rather than forcing direct physical–virtual alignment between the tangible proxy and the full-scale virtual object, the WiM can mediate the scale relationship between local object action and full-scale world effect. Designers can use WiM space to expose the proxy as a scaled handle, rescale the WiM container around the selected object, or adjust the visual

representation of the hand or object to manage the mismatch without producing premature collision cues.

This turns asymmetric tolerance into a directed design choice. When scale mismatch is unavoidable, designers should bias the configuration toward cases where the tangible proxy remains smaller than the virtual WiM object it controls. Positive mismatch should not be treated as unlimited free scaling, because it can still introduce performance costs during precision docking or fine manipulation. However, it provides a safer direction for redistribution than negative mismatch. This reframes WiM as a space where scale discrepancy can be biased toward more tolerable forms, redistributed before contact, and made usable during manipulation rather than simply hidden.

Ultimately, preserving local alignment and increasing world-scale flexibility require different kinds of design effort. Algorithmic alignment hides transformation inside physical–virtual correspondence, while WiM mediation exposes transformation through an explicit interaction space. The former can reduce visible abstraction but may require continuous perceptual adaptation; the latter can reduce physical effort but may require users to understand how local WiM actions relate to full-scale world effects.

This suggests that the role of WiM is not only to provide an overview or scaled representation. It can manage spatial discrepancies by separating local tangible interaction from world-scale transformation. This makes WiM a design resource for redistributing discrepancy across embodiment, WiM, and world spaces.

5.6 Configurations

The implications suggest that WiM can be used not only as a scaled representation of world space, but as a design element for managing spatial discrepancies between embodiment and world spaces. Based on this, I outline several potential interaction configurations that explore how WiM space can be positioned, controlled, or scaled across interaction states and across physical and virtual layers.

These configurations are not intended as complete systems, but as examples of how introducing intermediate space can expand design possibility for interaction techniques leveraging them within the design framework. Each configuration foregrounds a different design variable: WiM placement, WiM container control, and proxy–object scale alignment.

5.6.1 Spatially-Decoupled Retargeting with WiM

Existing retargeting techniques rely on user movement to mask spatial warping. I propose a configuration where retargeting is explicitly coordinated with the spatial placement of the interaction space (Figure 5.4). In this configuration, both the WiM interaction volume and full-scale view are available, but the WiM interaction volume is deliberately positioned and oriented away from the user's primary forward axis, for example, at approximately 90 degrees (Figure 5.4A). As a result, the user rotates their head or body to inspect the full-scale view, then returns to the WiM interaction volume to manipulate objects (Figure 5.4B).

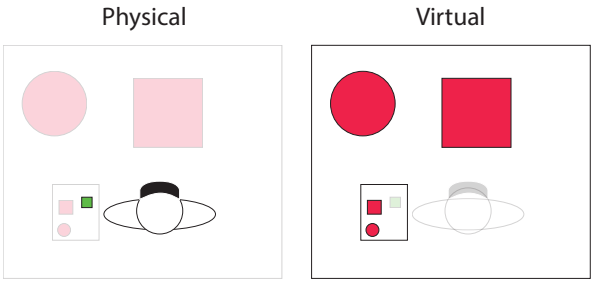
During this movement, the system applies retargeting transformations to the WiM, body, or world representation, possibly leveraging body- or world-warping. For example, when the user turns away from the WiM to validate how their manipulation relates to the full-scale environment, the system can adjust the position or scale of the WiM object so that the target object is aligned with the tangible proxy when the user returns.

This effectively creates a temporal and spatial buffer for retargeting, integrating implicit transformation with explicit interaction space design. In terms of the interaction spaces $\times$ virtuality framework, this configuration uses the spatial relation between WiM and world spaces to create an opportunity for adjusting physical–virtual correspondences in embodiment or world space. While such a configuration may introduce additional physical fatigue from frequent head/body rotation for a full-scale view, it may enable more aggressive retargeting without noticeable perceptual warping. This configuration demonstrates how control cues, interaction space placement, and implicit transformation can be designed to dynamically manage object alignment.

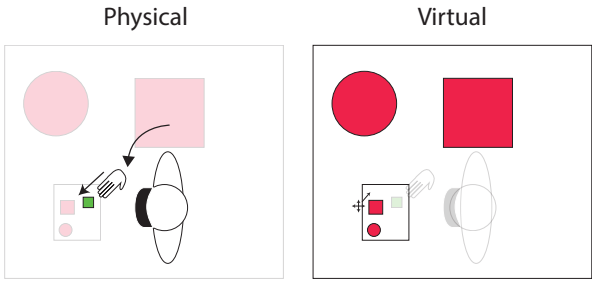
5.6.2 Bimanual Interaction Space Control

Most WiM-based systems assume a static interaction volume during manipulation to avoid confusion from moving objects. I propose a configuration where the interaction volume is the WiM container and moves while the user is grasping objects (Figure 5.5).

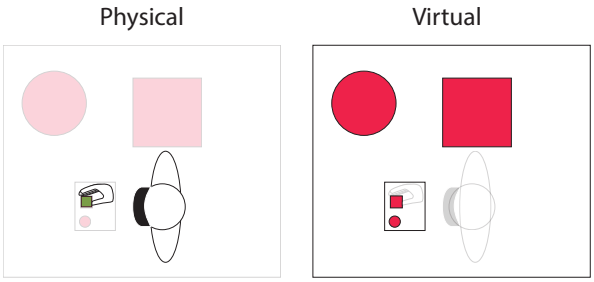
In this configuration, one hand grabs and manipulates the miniature object with a tangible proxy, while the other hand manipulates the WiM container (Figure 5.5D). The WiM container translates or rotates relative to the full-scale world. This keeps the local object-proxy relationship stable within the WiM space because the target object remains mapped to the tangible proxy. At the same time, the relationship between the WiM and the world becomes dynamic as the container moves relative to the full-scale world. As a



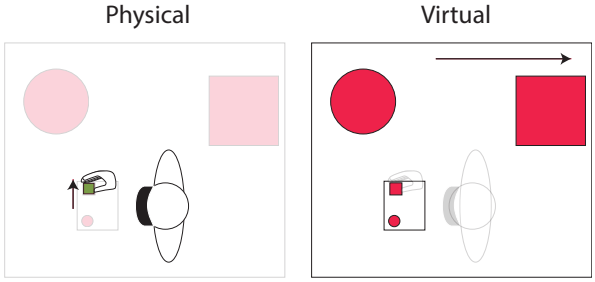
(A) Browsing: WiM interactive volume is positioned away from the user.



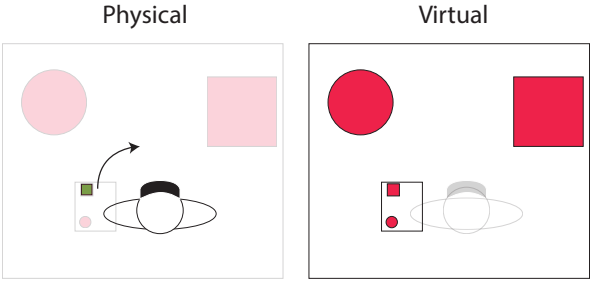
(B) Reaching: As the user rotates and reaches toward WiM space, the transformation occurs under the spatial buffer, and becomes proxy-object aligned.



(C) Grasping: The user grasps the object via the tangible proxy.

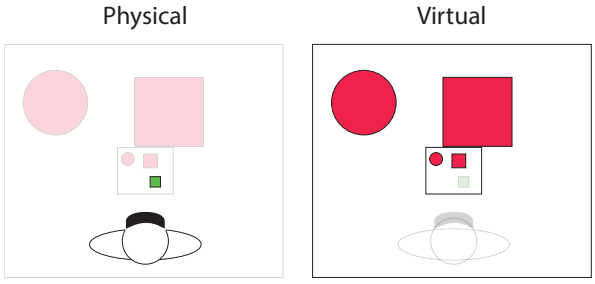


(D) Manipulating: Manipulation action is translated into world space effect via the WiM space.

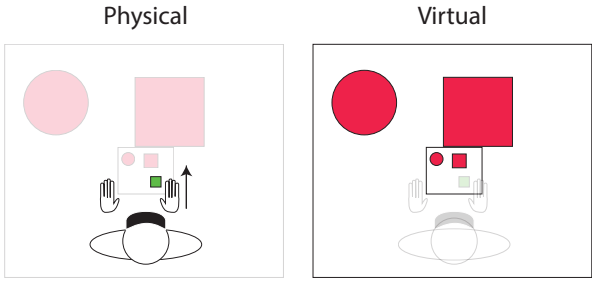


(E) Releasing: the user releases the object and inspects the full-scale world space.

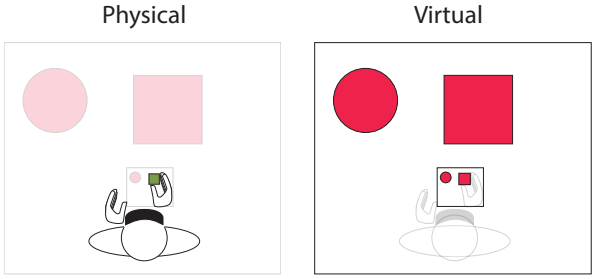
Figure 5.4: Spatially-Decoupled Retargeting Configuration: The WiM is placed away from the primary forward direction so that head/body movement between the WiM and the full-scale view creates a temporal buffer for retargeting before contacting the object.



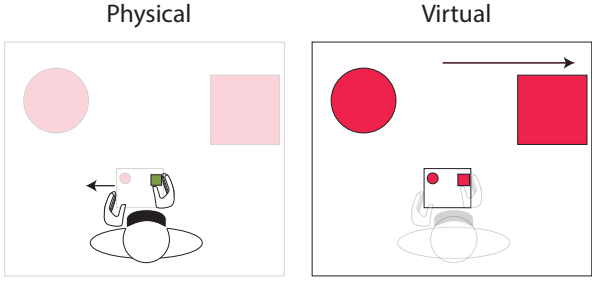
(A) Browsing: WiM and world space are visible to the user.



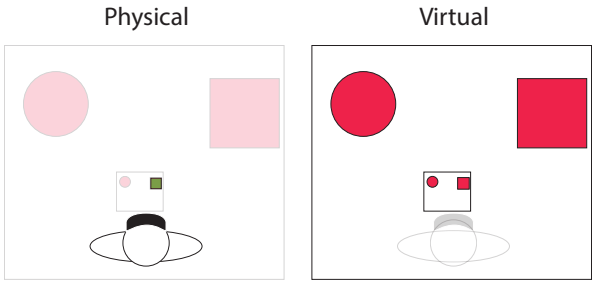
(B) Reaching: As the user reaches toward WiM space, the transformation occurs.



(C) Grasping: The user grasps the object via the tangible proxy; the other hand is locked to the WiM container.



(D) Manipulating: The object-hand stays still while the WiM container-hand moves the container. Moving the container to the left causes the selected object to move to the right relative to the space.



(E) Releasing: the user releases the object and inspects the full-scale world space.

Figure 5.5: Bimanual Interaction Space Control Configuration: One hand maintains coupling with a WiM object while the other hand controls the WiM container, allowing the WiM–full-scale world relationship to change.

result, the full-scale object can move even when the local object-proxy relationship remains stable.

This configuration shifts discrepancy from the proxy–object relationship to the WiM–world relationship. The tangible proxy can remain locally aligned with the selected WiM object, while the WiM container changes how that local action maps into the full-scale world. As a result, the user can preserve stable local manipulation with one hand while using the other hand to reposition the interaction space itself. This effectively elevates the interaction technique from object manipulation to world-scale space manipulation. This may support large-scale translation or navigation while preserving local proxy–object alignment.

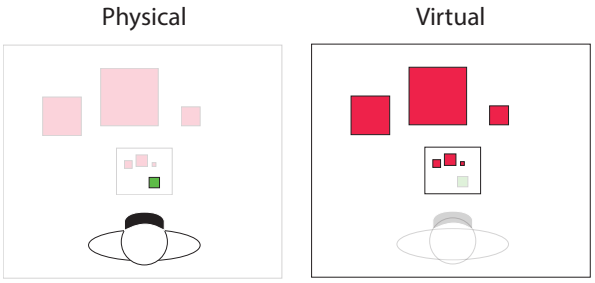
5.6.3 Scale-Mismatched Tangible Proxy

This configuration explores discrepancy caused by scale mismatch rather than positional mismatch. A tangible proxy has fixed physical dimensions and cannot match every virtual object. However, a WiM may contain multiple miniature virtual objects at different sizes. In this configuration, the system can positionally align a single tangible proxy with different miniature objects by warping or shifting the WiM container. However, the scale relationship between the proxy and each virtual object may still differ.

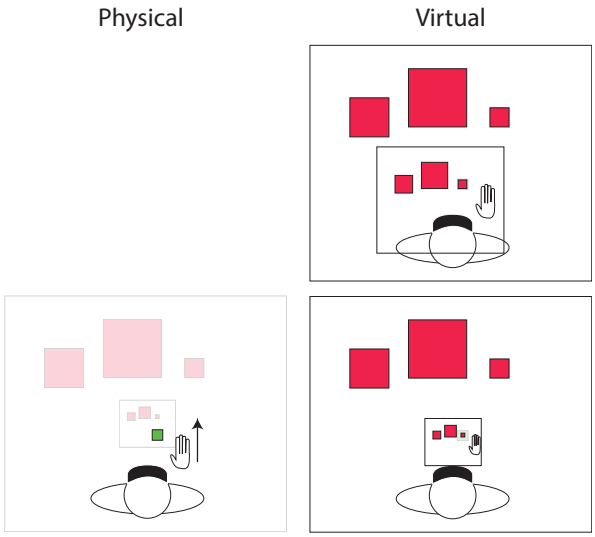
For example, the same proxy may be mapped to manipulate a small chair, a large piece of furniture, or a building inside the WiM. If the WiM container does not compensate for the scale mismatch, the proxy may remain positionally co-located with the target object while becoming scale-mismatched. This example highlights that physical–virtual discrepancy can emerge from scale discrepancy, not only from positional displacement.

A system could respond to this mismatch in several ways based on the framework. It could dynamically rescale the WiM container around the selected object, visually indicate that the proxy is acting as a scaled handle (Figure 5.6C, Top), or adjust the virtual body representation, such as the virtual hand, to match the scale ratio between the proxy and virtual target object (Figure 5.6C, Bottom). These choices determine whether scale compensation is hidden, exposed, or used as an explicit interaction resource.

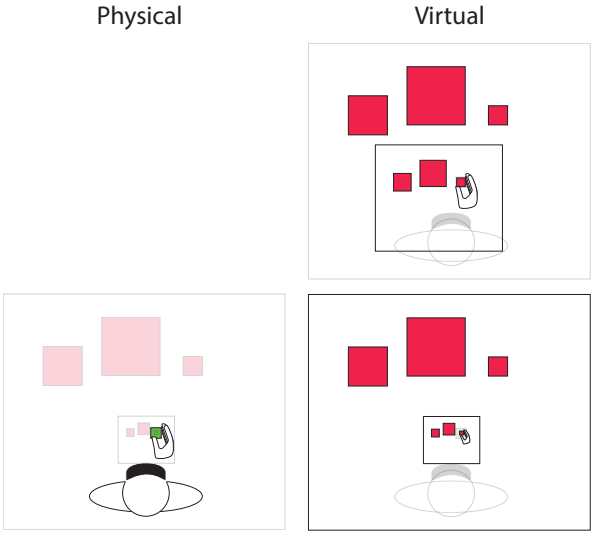
This configuration is particularly important for tangible-WiM-VR because it shows how a fixed physical proxy can interact with multiple differently scaled virtual objects and directly addresses the scale discrepancy challenges identified in chapter 4. In this case, the WiM becomes the place where scale discrepancy is either corrected, exposed, or repurposed for interaction.



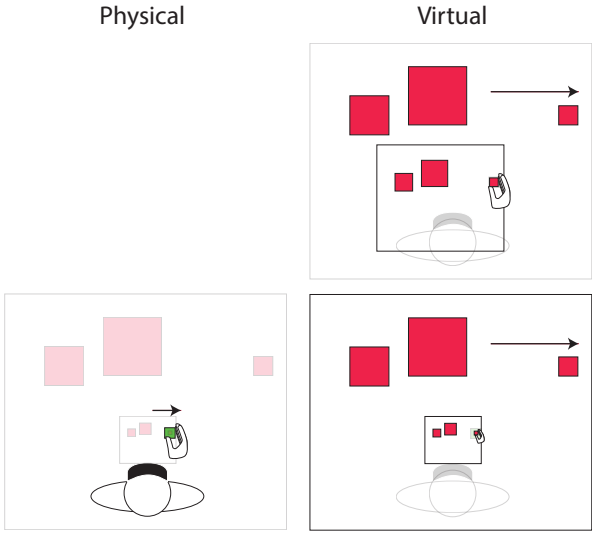
(A) Browsing: WiM and world space are visible to the user.



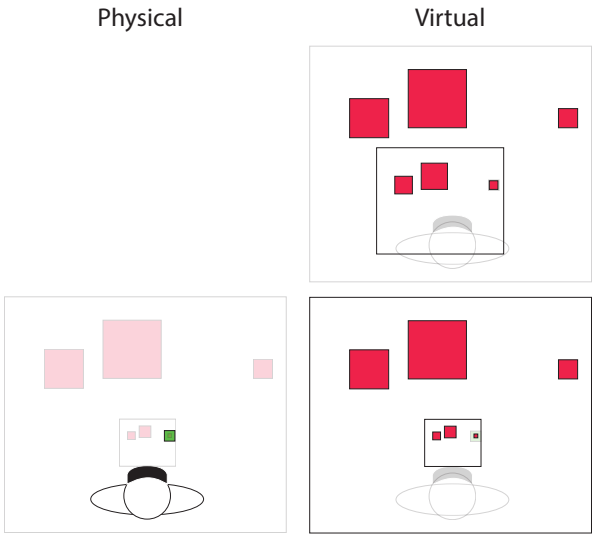
(B) Reaching: As the user reaches toward WiM space, the transformation occurs in either WiM (top), or hand representation (bottom).



(C) Grasping: WiM container expands (top), or hand representation (bottom) shrinks in VR environment to maintain hand contact to the virtual object.



(D) Manipulating: The user manipulates in the transformed WiM container (top), or with the transformed hand representation (bottom).



(E) Releasing: the user releases the object and inspects the full-scale world space.

Figure 5.6: Scale-Mismatched Configuration: tangible proxy can be coupled to differently scaled virtual WiM objects. The transformations of the WiM container or hand representation are dynamically adjusted based on the ratio between the tangible proxy and the selected object.

This final configuration motivates Laminar in chapter 6. Laminar instantiates this configuration by using a single tangible proxy to control differently scaled virtual WiM objects, while evaluating two ways of assigning scale compensation: to virtual embodiment through hand scaling, or to virtual WiM space through WiM scaling.

5.7 Chapter Summary

This chapter presented a design framework for tangible-WiM-based spatial interaction techniques in VR. The framework focuses on how interaction techniques manage mismatches and redistribute them across interaction spaces and interaction states. I model interaction as coordination between interaction spaces, virtuality, and interaction states.

To support design space analysis, I introduced design space foundations, which provide a consistent vocabulary for classifying interaction techniques. This design framework allowed me to compare techniques with and without WiM space. Through the analysis, I showed

that techniques without WiM, such as haptic retargeting, primarily manage discrepancies through direct interaction route (E-W) in embodiment and world space, while WiM-based techniques introduce an intermediate interaction space that mediates discrepancies between local action and world-scale effects (E-I-W).

Within the framework, spatial interaction techniques are understood to manage discrepancies rather than eliminate them entirely. This leads to design tradeoffs between direct physical–virtual coupling and WiM-based abstraction, as well as between local alignment and world-scale flexibility. Building on these insights, I proposed configurations that explore alternative ways to use WiM placement, WiM container control, and scale alignment to mitigate discrepancies. These configurations show how the design framework can be used not only to analyze existing systems, but also to discuss and generate new interaction designs.

This chapter establishes the conceptual foundation for designing tangible WiM-based VR interactions that manage spatial discrepancies, specifically scale mismatch between tangible proxies and virtual representations. Chapter 6 builds on this framework and analysis by instantiating the scale-mismatched tangible proxy configuration as *Laminar*. The next chapter functions as the second part of the research through design (RtD) process: it takes the design variables identified here—WiM mediation, proxy-object scale alignment, and temporal redistribution across interaction states—and operationalizes them in a prototype interaction technique.

Chapter 6

Laminar–WiM Retargeting for Multi-Scalar Spatial Design

This chapter addresses RQ4 by presenting a prototype interaction technique that instantiates one configuration of the design space developed in chapter 5. This chapter should be read as the second part of the RtD process developed in chapter 5, which identified tangible proxy coupling into virtual WiM as an underexplored configuration, particularly when a passive tangible proxy must control multiple virtual objects at different scales. The prototype exemplifies how spatial discrepancies between tangible proxies, WiM, and full-scale representations can be managed across interaction states. It is implemented with two different mismatch compensation strategies, which are evaluated in a preliminary user study. This chapter demonstrates how tangible interaction can support haptic congruence, multi-view coordination through WiM space and more continuous interaction across scales despite spatial discrepancies.

6.1 Introduction

The analysis presented in chapter 5 provided valuable insights into tangible WiM-based VR interaction design. Although leveraging intermediate interaction spaces, such as WiM, may provide a mechanism for resolving discrepancies in tangible proxies and multiple virtual objects, a significant challenge remains when those virtual objects differ in spatial properties such as scale and position. This chapter further investigates this challenge and realizes the design opportunity identified in chapter 5. Existing solutions typically utilize

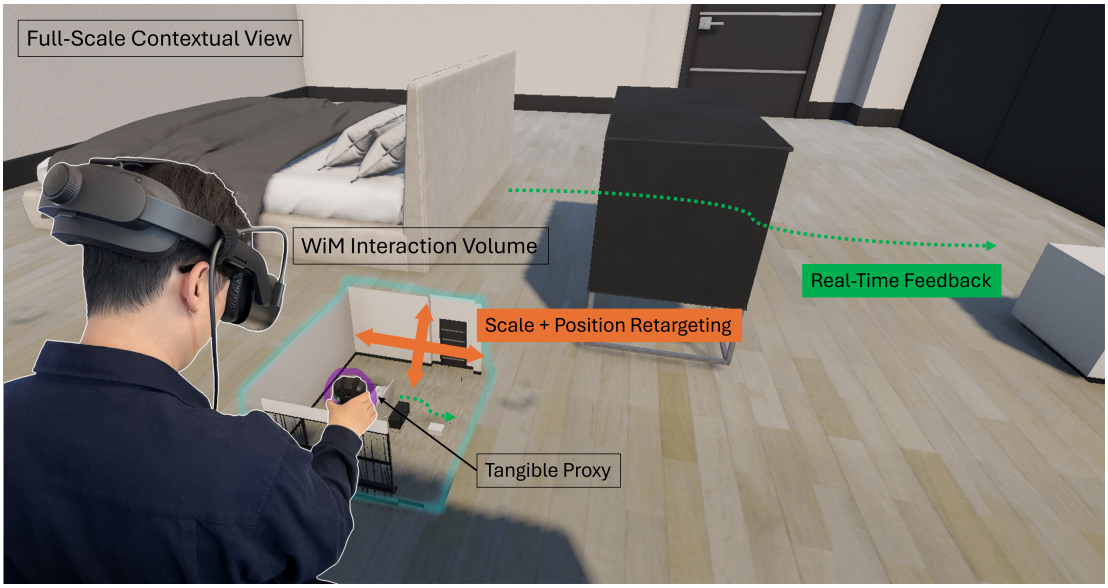


Figure 6.1: Overview of the chapter. Laminar: A tangible proxy provides passive haptic grounding while virtual objects are manipulated through a highlighted WiM interaction volume and verified in a concurrent full-scale contextual view. Laminar applies scale and positional retargeting to align selected virtual objects with the tangible proxy, while full-scale feedback preserves spatial context during manipulation.

dedicated physical proxies for each object (Muender et al., 2019) or rely on one-to-one mappings between everyday objects and virtual counterparts (Hettiarachchi & Wigdor, 2016). However, these tangible solutions do not fully address the physical–virtual discrepancy that arises when one physical proxy is used to manipulate multiple virtual objects with different spatial properties, including size and position. A tangible proxy may provide physical grounding, but cannot reproduce every virtual object's spatial properties due to physical limitations. As a result, systems with tangible proxies must determine where these mismatches are absorbed and managed, as such discrepancies can fragment workflows by forcing users to switch between interactions and system management, or tolerate unstable visual–proprioceptive alignment.

To address these interdependent discrepancies, I introduce Laminar, a tangible interaction technique for managing physical–virtual mismatch in multi-scale VR. Laminar distributes mismatch across three coupled mechanisms: scale compensation, WiM-localized retargeting, and body-cue-based implicit control. Central to Laminar is a retargeting mechanism that leverages body-centric cues and dynamic warping of the WiM view to align a

single tangible proxy with multiple virtual objects. By utilizing a multi-view paradigm, Laminar enables concurrent dual views, providing a combination of global context and local precision.

The design process was guided by an iterative research through design (RtD) methodology to ensure that the approach effectively addresses the architectural constraints of multi-scale tangible interaction. This chapter makes the following contributions:

- Laminar, a multi-scale tangible interaction technique for spatial design in VR that maps multiple virtual objects to a single tangible proxy through WiM-localized retargeting and concurrent multi-scale views.
- A configurable implementation of scale mismatch compensation strategies for scaled object manipulation, demonstrating how spatial discrepancies can be distributed across different spaces.
- A preliminary evaluation examining how these mismatch compensation strategies affect perceived alignment, predictability, and effort.

6.2 Deriving Laminar

The analyses in chapter 5 reveal an underexplored interaction design approach: integrating intermediate interaction space (WiM) to allocate mismatches and transformations for haptic retargeting before the user makes contact with the object. While existing haptic retargeting techniques primarily manage discrepancy through direct interaction routes (E-W), intermediate space techniques typically rely on persistent or user-invoked separate interaction spaces. The spatial analysis (table 5.2) shows that existing retargeting techniques primarily use either embodiment or full-scale world space for transformations, and the temporal analysis (table 5.3) shows when transformations occur and how they relate to visibility. These findings suggest an interaction technique that combines intermediate space mediation with pre-contact discrepancy management.

Furthermore, the findings from chapter 4 indicate that scale mismatch plays a vital role in how users perceive scaled objects in VR. These observations motivate Laminar, a tangible interaction technique that leverages a WiM as a workspace for haptic retargeting. Laminar explores how scale and positional discrepancies can be redistributed across embodiment, intermediary, and world spaces while maintaining physically grounded interaction through a single tangible proxy.

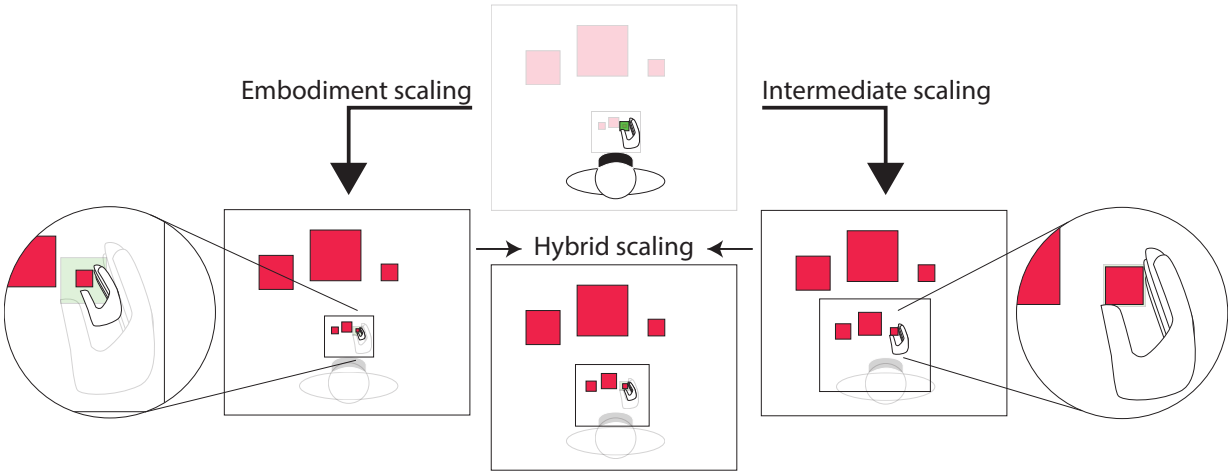


Figure 6.2: Laminar mismatch allocation strategies. Embodiment scaling compensates for scale mismatch by transforming the virtual hand representation (left). Intermediate scaling compensates by transforming the WiM interaction volume (right). Hybrid scaling distributes compensation across both spaces (centre). All strategies aim to achieve congruent visual–tactile feedback when the user grasps the object.

6.3 Laminar Overview

The analysis reveals a design opportunity to use a WiM that replicates full-scale world space at a small scale and provides retargeted haptic interaction in that space. Laminar, therefore, is an instantiation of configurations proposed in chapter 5, using a WiM interaction volume as an intermediary workspace for retargeting. I name the technique *Laminar* (Figure 6.2) after the laminar flow property in fluid dynamics, where fluid particles move seamlessly in layers around objects, reflecting how discrepancy can be distributed across layered interaction spaces while maintaining a coherent interaction relationship.

Central to Laminar is how both position and scale mismatches are compensated after object selection. Laminar computes the discrepancy between the tangible proxy and the selected virtual object using their positions and dimensions. This discrepancy determines the amount of compensation required to achieve visual-tactile alignment at grasping. Once the user starts to move and reaches toward the selected target, the embodiment space (hand representation), the intermediate space (WiM), or both begin to transform. The transformation progresses continuously as the user approaches the target and reaches its intended state at object grasping, completing the transformation for visual-tactile congruency.

Laminar can be described directly using the design framework (chapter 5). Laminar occupies an **intermediate route (E-I-W)** (*Interaction Route*) and uses a tangible proxy to support **tangible** (*Tangibility*) interaction. Depending on its variant, Laminar allocates its transformation to compensate for a mismatch in either the **embodiment** or **intermediate** spaces (*Mismatch Allocation*) (table 5.2). Unlike existing intermediate route techniques, Laminar adopts a pre-contact alignment strategy that reaches its intended peak transformation state at the grasping state (Reaching-Grasping) (table 5.3) rather than a user-invoked or trigger-based approach (Reaching-Releasing). These transformations remain **explicitly visible** (*Visibility*) to the user.

I chose a WiM container because it provides a locally and physically grounded interface when paired with a tangible proxy that mimics direct manipulation in a separate space. Although world space mismatch allocation could be achieved, I focus on leveraging WiM containers because they directly maximize the advantage of separate interaction space provided by WiM. World space allocation, like the world warping version of haptic retargeting, primarily transforms the represented environment rather than the local interaction volume. I identified two primary ways to compensate for discrepancies between tangible proxies and virtual objects by transforming embodiment and intermediate spaces: hand representation-based and intermediate WiM volume-based strategies.

Conceptually, Laminar variants mirror haptic retargeting variants (Azmandian et al., 2016). Laminar-hand follows body warping by allocating compensation to embodiment space, whereas Laminar-WiM provides an intermediary space alternative to world warping by allocating compensation to the WiM rather than the full-scale environment. Laminar, therefore, combines the spatial structure of many intermediary space techniques with the temporal characteristic of pre-contact alignment techniques identified in the design space analysis.

6.3.1 Transforming Embodiment Space

Laminar-hand follows an embodiment allocation strategy, similar to body warping. However, unlike body warping, which uses a direct route (E-W), Laminar-hand operates through an intermediary interaction route (E-I-W). This variant hence combines embodiment-based mismatch allocation with an intermediate interaction structure.

The discrepancy is compensated for by transforming the virtual hand representation while the other two interaction spaces remain untouched (Figure 6.2, left). When users approach a target object, the virtual hand gradually scales to match the relationship between the tangible proxy and the selected miniature object. The system requires the size

of the tangible proxy to calculate the difference between the selected target and the proxy used for compensation. While scaling the hand representation, the centre of scaling is also adjusted, allowing the technique to compensate for position discrepancy in addition to scale discrepancy. The transformation reaches its peak state at object contact, allowing tactile feedback from the proxy to remain spatially and visually consistent.

6.3.2 Transforming Intermediate Space

Laminar-WiM follows an intermediate mismatch allocation, similar to WiM driven techniques. Discrepancy is absorbed mainly within the intermediary space rather than the embodiment or world spaces. Unlike trigger-based intermediary techniques, Laminar-WiM uses pre-contact alignment behaviour identified in the temporal analysis (table 5.3) and reaches its intended state at object grasping, resembling haptic retargeting variants.

Instead of scaling the embodiment (hand representation), Laminar-WiM entirely allocates the mismatch compensation into an intermediary space. The discrepancy is compensated by transforming the WiM itself (Figure 6.2, right). The difference is calculated by the size of the tangible proxy and the visual scale of the selected target. As in the embodiment-based variant, scaling is achieved via a corresponding positional adjustment, ensuring that the selected target remains spatially consistent during the transformation. The WiM container is then scaled to match the scale of the selected target to the tangible proxy to provide congruent tactile feedback, while the user's virtual embodiment remains unchanged. As a result, alignment is achieved through intermediary representation rather than the embodiment of the user.

6.3.3 Hybrid Technique

It is also possible to combine the two and distribute how much compensation each interaction space absorbs from the compensating mechanism (Figure 6.2, centre). This hybrid approach distributes compensation across both embodiment and intermediary spaces. The hybrid approach corresponds to the hybrid allocation category, in which discrepancy compensation is distributed across multiple interaction spaces for the transformation (Figure 5.3). Rather than relying on a single interaction space, the hybrid approach shares the amount of transformation needed between the virtual hand representation and the WiM. This creates a highly configurable technique that can be adjusted to determine how much compensation each space absorbs.

6.4 Research Through Design (RtD)

Laminar was developed through a research through design (RtD) methodology (Zimmerman, Forlizzi, & Evenson, 2007), treating the design process itself as a method of research. Evaluation of each iteration occurs through cycles of analysis, critique, and reflective synthesis rather than the summative nature of user studies. Each iteration, therefore, serves as an evaluation probe for design decisions and as a means of examining the design space, exploring constraints among potential design elements, and informing subsequent refinements through iterative processes.

6.4.1 Iterative Design

To structure design decisions, I grounded the iterations in the Reality-Based Interaction (RBI) framework (Jacob et al., 2008). RBI argues that effective interfaces should leverage users' innate understanding of the world, specifically naive physics and body awareness, as the foundation for interaction. The RBI framework explicitly permits deviations from reality only when necessary to overcome physical limitations, as evidenced in haptic retargeting (Azmandian et al., 2016), which prioritizes the physical constraints (the passive prop) and introduces virtual deviations (warping) to accommodate them.

Guided by the RBI framework, my design iterations followed a hierarchical progression: from immutable physical constraints (the reality) to virtual augmentations, and then to the interaction techniques that bridge them.

1. Input Modality (Reality): I began by prioritizing the input method that leverages naive physics. Physical limitations, such as users' reach, the tangible proxy's friction on surfaces, and its properties, are immutable constraints.
2. View Management (Augmentation): I designed the visual interface to overcome these physical limitations, deploying virtual augmentations, such as concurrent multiviews, to expand the workspace.
3. Flow Control (Bridge): Finally, I designed the interaction technique to bridge the gap between the static physical proxy and the dynamic virtual augmentation, supporting a continuous transition.

In the RtD process, each iteration surfaced distinct breakdowns and insights that progressively clarified how each design choice addressed the underlying challenges. This iterative process helped expose and navigate the interdependent constraints and tradeoffs,

informing the design decisions carried into subsequent iterations. The following section outlines this iterative progression and provides important details on how the final technique emerged.

6.4.2 External Evaluation

I also conducted a preliminary evaluation with a small group of participants as a design probe rather than a summative validation of Laminar. The goal was to surface experiential consequences of design decisions around scale mismatch compensation strategies. This external evaluation grounded the design iterations and decisions by observing the participants' interactions.

6.5 Iterative Design

The iterative process unfolded across two phases: **Phase 1**, comprising iterations 1–4, which establish the foundational components of the system. This phase focuses on eliminating unsuitable design alternatives to surface and address the collective constraints. In **Phase 2**, iterations 5 and 6 focused on developing retargeting mechanisms, refining how the system utilizes body-centric cues to map multiple virtual objects to a single tangible proxy. Each iteration followed three phases: (a) design, (b) critique, and (c) reflection. In the *design* phase, I formulated hypotheses based on the constraints identified in RBI analysis, translating requirements into prototypes. In the *critique* phase, I conducted internal critiques to identify interaction bottlenecks. This allowed me to assess stability, haptic congruence, and task continuity against the established physical constraints. In the *reflection* phase, I distilled the successes and failures of each prototype into design knowledge, which refined the constraints for the subsequent iteration.

6.5.1 Iteration 1: Haptic Anchor

Motivation: The goal of this iteration was to establish a robust haptic baseline as a reality. A specific objective for the spatial design workflow was to ensure a stable and consistent haptic experience. I hypothesized that direct physical manipulation of a tangible proxy would provide superior cognitive grounding compared to mid-air gestures (Ishii & Ullmer, 1997).



Figure 6.3: Iteration 1: Haptic anchor. The design process begins by selecting the tangible proxy (left) as physical embodiment space over a standard VR controller (right), with the tabletop serving as a physical constraint for stable manipulation. This establishes passive haptic grounding before introducing WiM mediation or retargeting.

Design: Alternative approaches included controller-based interaction, mid-air gesture interaction, self-haptics, and passive tangible proxies. Considering the requirements, I immediately discarded mid-air gestures due to the absence of tactile feedback. Although self-haptic techniques (Fang & Harrison, 2021) were considered, their reliance on bi-manual interaction made them highly incompatible with the type of interaction and workflow required for tabletop-based spatial interaction workflows. I compared standard controller-based input with direct haptic experience provided by passive tangible proxies. I implemented a scenario where virtual furniture (a side table) had to be aligned against a wall and a bed.

Critique: Standard controllers with vibration feedback does not provide tangible feedback in *Tangibility* in the design framework. Additionally, operating in mid-air poses a high risk of instability, leading to tremor due to a lack of physical support and Gorilla Arm fatigue. While custom-built haptic devices provide higher fidelity, their reliance on a one-to-one physical-virtual mapping approach introduces a scalability bottleneck, requiring high fabrication costs. Furthermore, shape-changing haptic devices were also excluded because they are highly complex in larger spaces, which is not the main use case in spatial design, such as architecture and interior design tasks.

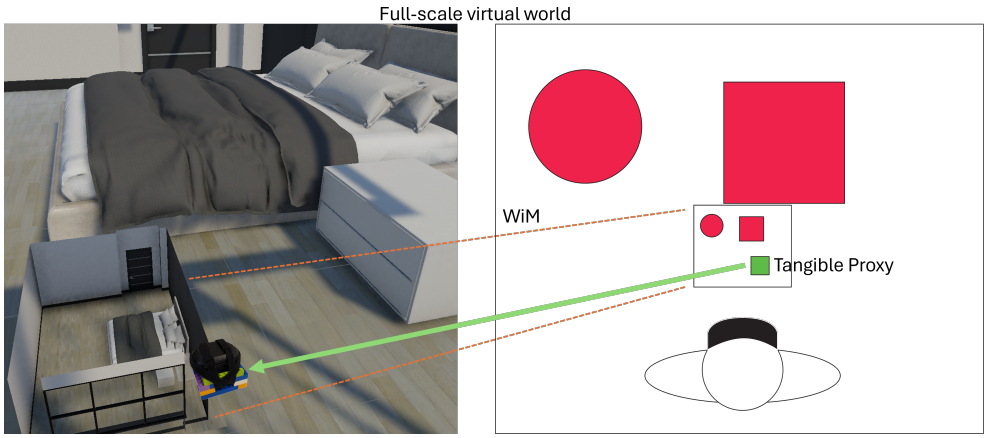


Figure 6.4: Iteration 2: Contextual scale. The design introduces a virtual WiM as the local interaction space and a concurrent full-scale world view as contextual feedback. This establishes the embodiment–WiM–world structure used by Laminar, but leaves scale mismatch between the fixed proxy and differently sized virtual objects unresolved.

Reflection: I concluded that tangible proxy-based interaction on a tabletop surface is necessary for higher fidelity interaction. Since tangible proxies operate on a desk or tabletop setting, they essentially constrain the y-axis (height), allowing precise refinement with friction on the desk. Furthermore, this setting enables richer and more intuitive interaction (such as grasping), which leverages the user's naive physics knowledge of traditional model-making. However, committing to passive haptics introduces a scalability constraint, which necessitates a retargeting approach (Azmandian et al., 2016). This scalability constraint became the primary focus of the next iteration, which explored how various objects and environments can be mapped to limited physical workspace and resources.

6.5.2 Iteration 2: Contextual Scale

Motivation: Having established the input modality for using a tangible proxy, I encountered a scalability bottleneck in mapping a large-scale virtual environment within a constrained physical workspace.

Design: Potential approaches included teleportation-based locomotion, overlapping views, minimap representations, and WiM-based interaction.I explored paradigms to support a large-scale environment within the user's field of view. Literature proposed notable use of multiview paradigms such as (1) superimposition/overlapping (Schjerlund et al.,

2022), (2) teleportation/switching (Bozgeyikli, Raij, Katkoori, & Dubey, 2016), and (3) scaling views (Abtahi, Gonzalez-Franco, Ofek, & Steed, 2019; Stoakley et al., 1995).

Critique: While locomotion techniques, such as teleportation, can provide access to large-scale context, teleportation failed the continuous workflow ; the abrupt context switches broke the user's spatial mental model. Overlapping or superimposing multiple views maintained the context but introduced visual clutter that interfered with manipulation. I investigated a simultaneous multiple-view paradigm, such as a minimap-like (2D) and a WiM (3D) as a static secondary viewport. Using a 2D minimap was rejected as it lacked the 3D mapping required for 6DOF object manipulation.

Reflection: This iteration established the embodiment-WiM-world space relationship that forms Laminar's design: intermediary interaction route (E-I-W). While WiM serves as a secondary viewport for contextual access, it introduces a new design question regarding scale mismatch between the fixed tangible proxy and virtual objects of varying sizes. At this stage, two conceptual strategies were possible: compensating scale mismatch at the level of either the user's body (e.g., scaling representation of the hand) or the interaction space (scaling the WiM space itself). Both approaches preserve reachability while implying different impacts on embodiment and interaction stability, warranting further exploration.

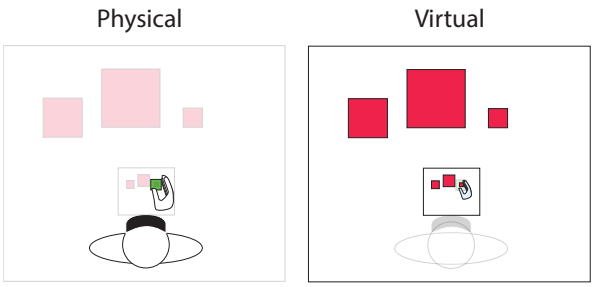
6.5.3 Iteration 3: Scale Mismatch Compensation

Motivation: Multiple views, with the full-scale environment as a primary view and the WiM as a secondary view, introduced a design fork in how scale mismatch could be addressed. As the main challenge remains with haptic retargeting, I strategically designed two mismatch compensation strategies while preserving the contact alignment between visual (virtual) and physical sensation.

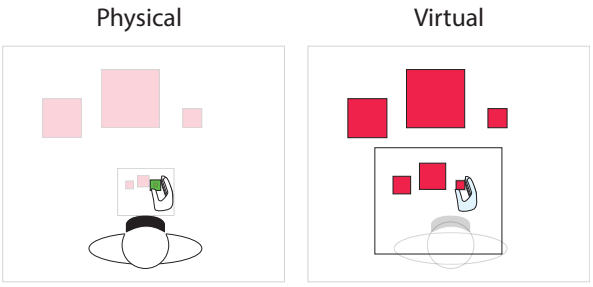
Design: I first explored a space-centric solution by expanding the flexible WiM view. This implementation takes the geometrical dimension of the tangible proxy to calculate how much scale is needed for the WiM view, which contains the selected object mapped to the proxy. Another possible design was to scale the hand representation. Scaling the hand representation would deviate from the RBI-based framework; however, I included it as a possible alternative to compensate for scale mismatch.

Critique: While both strategies visually and haptically provide reasonable contact alignment, I observed that they were viable in different ways, preserving contact alignment.

Reflection: I left both options as "viable" options to be evaluated in later stages. These correspond to the mismatch allocation dimension identified in the framework, assigning discrepancy either to embodiment or intermediary (WiM) spaces. These two strategies



(A) Hand scaling strategy (Laminar-hand): hand representation is transformed to compensate for the scale discrepancy, while maintaining contact alignment.



(B) WiM scaling strategy (Laminar-WiM): WiM container is transformed to compensate for the scale discrepancy, while maintaining contact alignment.

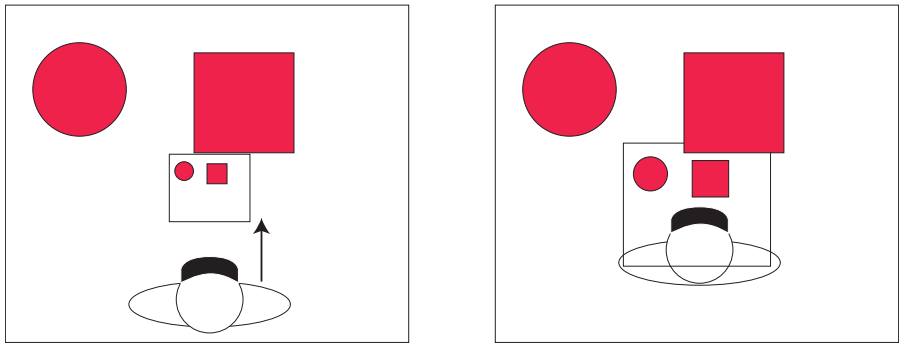
Figure 6.5: Iteration 3: Scale mismatch compensation. Hand scaling assigns compensation to virtual embodiment space by changing the hand representation relative to the tangible proxy. WiM scaling assigns compensation to virtual WiM space by changing the interaction volume around the selected object. Both strategies preserve proxy–object contact alignment but redistribute discrepancy to different spaces.

preserve reachability but reassign stability and embodiment differently, motivating further exploration in subsequent iterations. These alternatives exposed a workflow challenge, motivating exploration of seamless scale transitions in the subsequent iteration.

6.5.4 Iteration 4: Seamless Workflow

Motivation: This iteration addressed the integration of the tangible proxy with the WiM paradigm in VR. WiM is used as an interactive volume, and the full-scale environment supports the contextual verification made in the WiM view. Because the WiM represents a large virtual space, small physical movements translate to massive virtual displacements (a high C/D ratio). This integration created a fundamental tradeoff between speed and accuracy .

Design: The primary candidate design approach is to include explicit UI widgets and implicit ways through users' body movement. I established that the system must use the WiM view as the primary interactive space, where the miniature movement translates into rapid, large-scale movement in the full-scale virtual environment. The scale of the miniature view directly translates into the C/D ratio, supporting rapid movement across a large space (speed). The challenge was designing a mechanism to support fine move-



(A) Leaning in: The user's approach acts as a body cue that expands the WiM container.

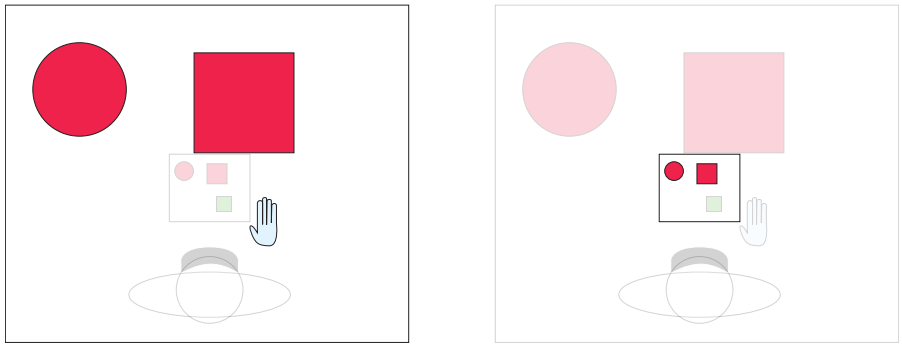
(B) After leaning-in: Moving closer expands the WiM, lowering the effective control–display ratio for finer manipulation.

Figure 6.6: Iteration 4: Proxemic workflow control. The user's approach toward the WiM is used as an implicit control signal: moving closer expands the WiM and supports finer manipulation, while moving away restores broader contextual control. This embeds scale and speed–accuracy control in body movement rather than explicit mode switching.

ment for small adjustments (precision) without adding cognitive load. I hypothesized that adding separate input devices or bimanual operation would also increase cognitive overhead. Instead of explicit UI widgets, such as toggles, buttons, and sliders, an implicit mode-switching mechanism based on proxemics (Ballendat et al., 2010) was implemented.

Critique: Explicit interaction would add cognitive load to users. Making a conscious decision to switch modes can disrupt the workflow. It can turn a continuous design thought into a fragmented sequence of system management, increasing interaction overhead .

Reflection: I rejected explicit switching in favour of implicit body cue-based scaling of the WiM. I leveraged the body awareness cue of "leaning in" to examine an object in detail. The distance between the user's head and the object determines the scale of the WiM view, also allowing for the minimization of visual clutter when examining the full-scale view. Crucially, I leveraged this dynamic scaling window to mask retargeting the tangible proxy. As the user leans in to zoom, the visual changes dominate their perception (change blindness) (Luo, Tang, Xu, Tong, & Zhang, 2024), allowing the system to subtly pan the miniature view and realign the proxy unnoticed. Having established a base interaction flow, the next iteration examined how retargeting could be integrated into this interaction paradigm.



(A) Classic haptic retargeting: discrepancy is managed through body or world warping.

(B) WiM retargeting: discrepancy is localized by panning the WiM container.

Figure 6.7: Iteration 5: WiM-localized positional retargeting. Compared with classic haptic retargeting, which transforms embodiment or world spaces directly, Laminar localizes positional retargeting inside the WiM. The WiM container pans so that the selected virtual object aligns with the tangible proxy while the full-scale world remains available as contextual feedback.

6.5.5 Iteration 5: Positional Retargeting

Motivation: After establishing the core interaction components, I focused on integrating the retargeting technique into the system.

Design: I explored how existing haptic retargeting techniques, including body warping, world warping, and hybrid warping (using both), align with multi-view, multi-scalar workflows with proxemics. I used the suggested values from the existing literature, along with optimized values for the retargeting illusion (Azmandian et al., 2016), which guided the initial probe for retargeting the proxy positionally. However, instead of applying these techniques directly to the full-scale world, I opted to test a miniature version of world warping, *WiM warping*, by panning/sliding the WiM container to retarget the position.

Critique: Each traditional retargeting technique exhibits a critical limitation in the current design context. Body warping results in noticeable distortion of hand reaching, with a narrow range of illusion, even when applied within the WiM interactive volume. World warping is designed to work with a full-scale first-person view in mind. It produced substantial disorientation as environmental elements (i.e., multiple objects with complex material patterns) compounded the warping. In contrast, WiM warping produced minimal illusion breaks since the warping was localized within a small viewing angle. Initial probes

showed that WiM warping could be integrated and used as a retargeting technique to (re)position to the tangible proxy without disrupting workflow or causing disorientation.

Reflection: This iteration explored how retargeting could be achieved leveraging intermediary space (WiM). The WiM warping allowed repositioning of WiM to match the position of the desired object with the tangible proxy. Although the choice might have been different if I had tested those existing techniques in a more abstract environment with primitive 3D shapes and simpler colours, they conflicted with the detail-oriented setting of workflows. The WiM warping provided flexibility through proxemic-based WiM scaling, enabling positional retargeting for multiple objects. This led to the exploration of simultaneous retargeting, both positional and scale, in the next iteration.

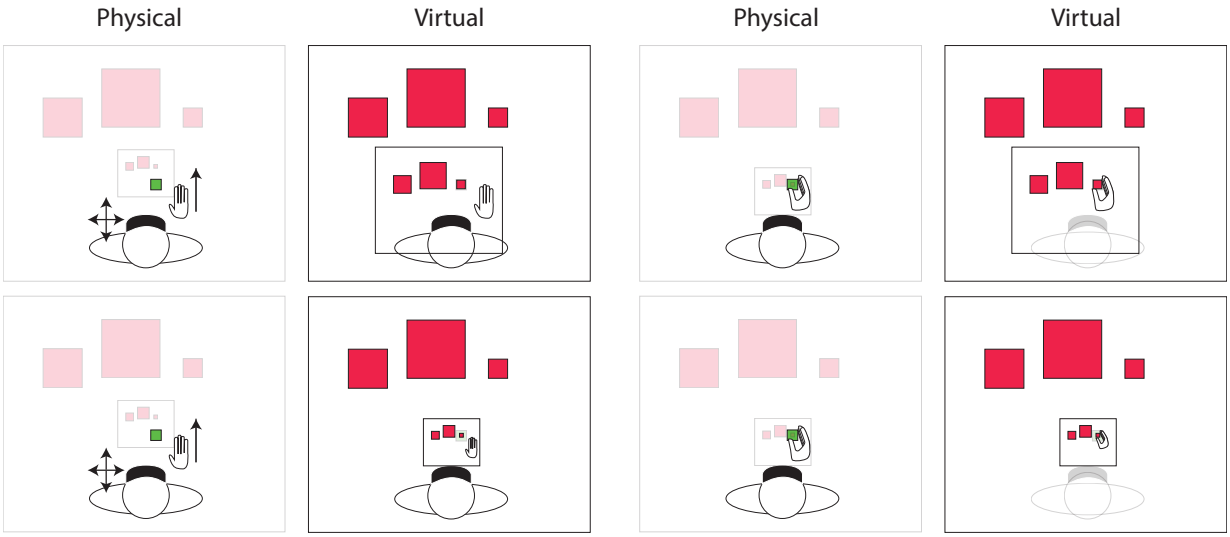
6.5.6 Iteration 6: Combined Retargeting

Motivation: I observed that WiM scaling is largely disregarded over broad ranges, as visual attention becomes dominated by the user's own body movement and the natural perceptual effect that objects appear larger when viewed up close (the "looming" effect). This suggested that proxemic-based WiM scaling could provide a masking window for other retargeting operations to enhance object mappings and improve haptic congruence.

Design: Building on iteration 5, I explored whether positional retargeting could be combined with WiM scaling based on proxemics, allowing the system to realign and scale the WiM during proxemic masking windows. I hypothesized that these adjustments with WiM could be masked by the user's natural attention during their own movements (e.g., leaning in) for scaling WiM.

Critique: Initial probing of panning the WiM suggested that panning during approach conflicts with the reaching plans already in place, eventually leading the hand to move farther to chase the target. When panning (warping) occurs concurrently with proxemic-based WiM scaling, it appeared less salient, suggesting that dynamic scaling may provide a perceptual masking for small adjustments. Although WiM scaling did not explicitly match the footprint of the tangible proxy, the scaling (up or down) naturally resulted in a more comfortable and congruent manipulation experience. However, this led to a "chasing" behaviour, where the user began reaching toward a selected object, but the panning shifted the object's position, causing the user to repeatedly correct the hand's reaching trajectory. Hence, chasing the object remains a challenge beyond the current probe.

Reflection: These findings revealed that body-based cues provide an effective channel for both scale and positional retargeting. This combined retargeting operation, by coupling WiM panning and scaling, formed a candidate mechanism for mapping multiple virtual



(A) Reaching: Proxemic movement provides a temporal window for combined scale compensation and positional retargeting.

(B) Grasping: The selected virtual object becomes aligned with the tangible proxy after scale and positional compensation.

Figure 6.8: Iteration 6: Combined retargeting. Laminar combines scale compensation with WiM-localized positional retargeting. The top row shows WiM scaling combined with positional retargeting, while the bottom row shows hand scaling combined with positional retargeting. Proxemic movement provides a temporal window in which the WiM can scale and reposition so that differently sized virtual objects can be mapped to a single tangible proxy.

objects to a single tangible proxy while maintaining workflow continuity. At the same time, the chasing behaviour highlights an unresolved tension between retargeting and natural interaction (reaching), suggesting an opportunity for future refinement, such as an intuitive and intelligent feedforward visual interface.

6.6 Preliminary Evaluation

I conducted a preliminary evaluation to examine how key aspects of Laminar's scale mismatch compensation strategies are experienced, focusing on how different configurations of Laminar shape embodied perception, trust, and workflow continuity. I designed this study as an exploratory probe into how participants perceive its core design decisions, rather than validating Laminar as a complete system.

6.6.1 Participants

I recruited 10 participants for an in-person study in a controlled lab environment. Participants' ages ranged from 22 to 33 years ($M = 27.70, SD = 3.50, Mdn = 28.00$). Of those, 6 identified as women, 4 as men, and none as non-binary.

Participants also self-reported their VR experience levels: 7 were beginners, 3 were intermediates, and none reported advanced. Educational backgrounds included 6 with a Bachelor's degree, 2 with a Master's degree, and 2 with a PhD.

They were all recruited from a mailing list. The exclusion criteria included uncorrected visual impairments, balance disorders, and the use of medical devices or implants that were at risk of electromagnetic interference, such as pacemakers and neurostimulators.

All participants provided informed consent prior to the study and were compensated with a $20 CAD digital gift card. This study protocol (#48304) was reviewed and approved by the University of Waterloo Ethics Board and was advertised in accordance with its guidelines.

6.6.2 Apparatus

The study was conducted using an HTC Vive Focus Vision headset, its controllers, and an HTC Vive Ultimate Tracker attached to a tangible proxy (LEGO assembled) (Figure 6.9A). The tangible proxy (Figure 6.9B) provided passive haptic feedback and physical grounding



(A) left hand controller, VR headset, and proxy (B) Proxy close up: LEGO assembled bottom

Figure 6.9: Apparatus used in the study

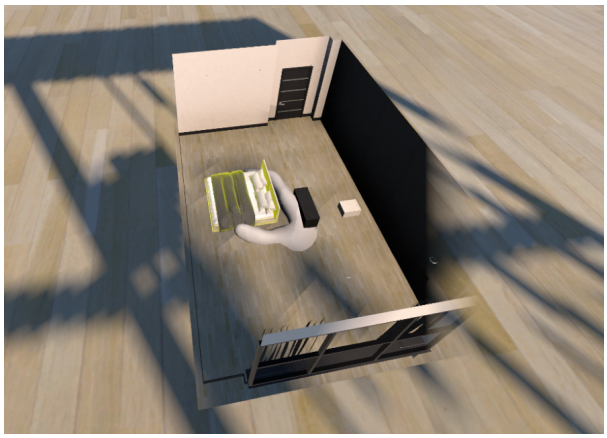
during manipulation on the surface (a desk). All interaction conditions were implemented in Unity 6 and incorporated positional retargeting, scale compensation mechanisms, and head motion-based control derived from the Laminar prototype.

Participants interacted with a set of furniture rendered concurrently in a WiM view and a full-scale contextual view. The WiM view functioned as the primary interactive volume while the full-scale view allowed participants to verify spatial outcomes in real time as they manipulated objects in the WiM. Participants used the left-hand controller's trigger button to switch the active target among the three virtual objects.

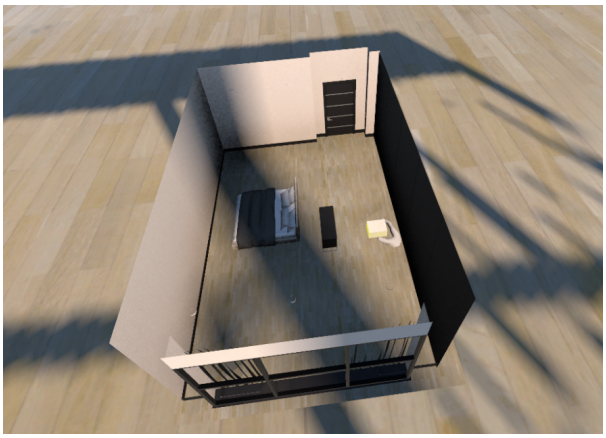
6.6.3 Evaluated Configurations

I examined two configurations of Laminar that embody different approaches to managing scale mismatch.

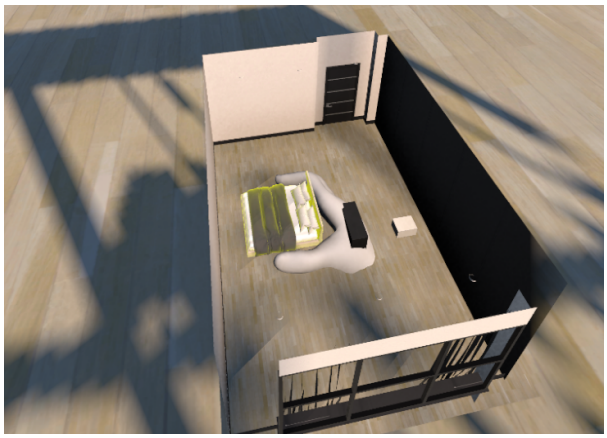
- Hand Scaling strategy: the hand representation scaled relative to the tangible proxy (Figures 6.10A and 6.10B).
- WiM Scaling strategy: WiM view (primary interactive volume) scaled to match the currently selected object's size with the tangible proxy (Figures 6.10C and 6.10D).



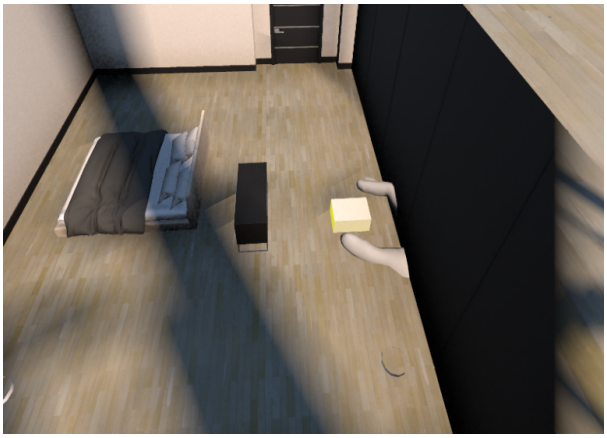
(A) (Hand Scaling) When grabbing *Bed*



(B) (Hand Scaling) When grabbing *Sidetable*



(C) (WiM scaling) When grabbing *Bed*



(D) (WiM scaling) When grabbing *Sidetable*

Figure 6.10: Evaluated scale-mismatch compensation configurations. In the hand-scaling condition, the hand representation scales according to the selected virtual object size (a–b). In the WiM-scaling condition, the WiM container scales to compensate for mismatch between the selected object and tangible proxy (c–d).

Training: I included a version for training purposes which had no scaling effect. This served as a reference in the study, as well as a familiarization with the equipment and the motivation for the study.

Hybrid Scaling (Post-study probe): While hand and WiM scaling strategies were evaluated as main conditions, I probed an alternative mismatch compensation strategy, where both hand and WiM scale to compensate. In the post-study probe, I included equal contributions from hand and WiM.

6.6.4 Procedure and Measures

All participants were briefed about the study prior to giving consent to participate. They were given approximately 10 minutes for equipment fitting and to learn how the tangible proxy works with the system.

Participants then completed a spatial manipulation task (arranging furniture) under two evaluated configurations (hand scaling and WiM scaling, respectively), presented in counterbalanced order. During each configuration, participants were encouraged to think aloud while engaging with the system.

After each strategy, I collected subjective measures (NASA-TLX and a custom 7-point Likert for 6 statements (section B.3.1)), followed by semi-structured interviews for deeper discussion (section B.3.2). Upon completing an interview for evaluating strategies, I presented an alternative, hybrid compensation strategy and conducted a post-study interview. Finally, participants completed the demographic and experience questionnaire to complete the study.

6.7 Findings

The results are primarily qualitative, as the goal of the study was to evaluate the consequences of the design iteration and probe unresolved design choices. I focus on how different mismatch compensation strategies affect embodied experience within Laminar.

Simulator Sickness

Participants completed the Simulator Sickness Questionnaire (SSQ) at the end of the study session. The mean total score was 7.11 ($SD = 10.20$). Breakdown by subscale revealed the

following scores: Nausea ($M = 6.68, SD = 11.06$), Oculomotor ($M = 4.55, SD = 7.32$), and Disorientation ($M = 8.35, SD = 13.45$). These scores suggest that the implementation was generally well-tolerated and did not induce significant motion sickness.

6.7.1 Qualitative Feedback

I gathered qualitative feedback through think-aloud protocols during individual interactions and post-session semi-structured interviews. Each scaling strategy session captured immediate reactions to visualization and scaling strategies, often in comparison to previous interactions with Laminar. The semi-structured interviews were designed to elicit detailed reflections on the primary design probes and other aspects of Laminar.

Hand vs. WiM

There was a clear distinction between preferences in scaling strategies. Participants articulated the differences between the two strategies and their rationales for one over the other in terms of a stable reference frame, which served as a strong contextual anchor during interaction.

With the hand scaling strategy, a majority of participants (6/10) treated their virtual hand as a reliable point of reference. As P5 noted, the interaction felt *'very natural, even though [the] hand [was] small'*, and P4 emphasized that seeing a hand *'exact[ly] the same size [as] the one I am seeing... feels like I am actually grabbing it'*. This suggests that leveraging body representation stabilizes interactions and maintains a connection to the real world. These participants preferred hand scaling because it preserved environmental consistency. They adapted to the dynamic scaling of their hand representation by adopting an instrumental view of their hands as *'[the] hand was [like a] computer mouse, more like [a] proxy to my hand'* (P3).

On the other hand, under WiM scaling, participants (4/10) often described the dynamic resizing of the WiM as responsive. For example, P9 noted that if the miniature room scales up enough, *'it can substitute the environment (the full-scale view)'*. Because their virtual hand size remained consistent across different objects, the interaction naturally aligned with their physical proprioception. P6 rejected the dynamic resizing of the hand, mentioning that *'I know my hand sizes already, it didn't really make sense seeing my hand being so small'*. Similarly, P8 found the changing size of the hand representation in VR *'annoying'*, explicitly pointing out, *'I don't know how to hold things with my small hand'*.

Although preferences diverged, their responses were centred on the hand representation as a reference frame and the WiM as a functional interface. This suggests that balancing proprioception with a stable visual interface is a critical design consideration in VR.

Trust, Predictability and Mismatch

Participants generally showed higher trust in the hand scaling strategy because it provided less discrepancy of control (virtual hand grabbing virtual object) and a stable frame of reference (WiM view's size remained relatively constant). Notably, P1 stated that the *'hand scaling version was more trustworthy. [I] feel more control, information from hand scaling version'*, which indicates a preference for a predictable frame of reference. Some participants noted that discrepancies between virtual objects and the physical proxy felt *'unnatural'* due to the inherent shape mismatch; for example, P4 described *'grabbing [the] chest feel[s] unnatural because of the shape mismatch'*.

With hand scaling, participants (6/10) felt a more consistent and accurate alignment between their hand representation and virtual objects. P2 noted that *'contact alignment helped trust'*, P5 confirmed that *'contact alignment is exactly adjusting to the virtual object depending on the size of the object'*, and P10 added that hand scaling was *'visually convincing'*. Participants described this as intuitive and natural despite the hand (representation) size. This predictability contributed to a sense of control. In contrast, participants reported that with WiM, contact alignment was inconsistent, leading to less predictability, particularly when the WiM view expands dramatically.

Despite the training phase, some participants struggled with positional retargeting when aligning with the proxy. This triggered noticeable offsets between the virtual object and proxy, described as *'offset a lot though, chest is here, but the thing [proxy] is here'* (P8). This system-based limitation occasionally caused frustration due to its unpredictable behaviour.

Tradeoff of Head Movement

Participants described head movement as a control for scaling the WiM in the WiM scaling condition. While this enabled flexible exploration, participants differed in how they experienced this mechanism.

Several participants (4/10) noticed that head movement functioned as an implicit control for zooming and retargeting, allowing them to transition between full-scale and WiM views. P4 liked this mechanic, stating that the *'intuitive head movement to zoom in and*

out... makes things easier'. In these cases, head-based control supported a continuous workflow by reducing the need for explicit mode switches. For these participants, this implicit control made the WiM scaling strategy highly advantageous.

However, other participants experienced head movement as an additional cognitive and physical burden, especially when repeated retargeting was required. P3 commented that they *'have to move a lot, arms, hand, [and] head'* because the WiM view dynamically changed and *'exploded'* when selecting a smaller virtual object. Furthermore, P6 noted confusion around how head movement worked, having a *'[tendency] to do the side way [horizontal rotation of head movement] to align'* before having to actively *'learn [to] move forward and backward'*. Similarly, P10 described this movement as *'unnecessary'* and needing *'exceeded head movement until the room (WiM)... stop[s]'* to align with the proxy. These participants found the head movement to be an unnecessary step for aligning the object with the proxy, suggesting that the system should automatically align upon object selection. While positional haptic retargeting is mainly designed to reuse a single proxy, specific use cases reveal a strong tradeoff between preserving spatial context and achieving functional efficiency.

Hybrid and Laminar Tradeoffs

As a post-study probe, I gauged a hybrid method that dynamically scaled both the hand and WiM views. Some participants (4/10) noticed it as a hybrid, while others felt it was similar to their earlier experience.

Participants perceived the hybrid as an improved version overall, calling it *'super easy'* (P1) and a *'major improvement'* (P3) over the two mono strategies. P1 appreciated the fluid transition: *'once [I] switch object, it [WiM] scales, and getting close to the object, [and] hand scales. [I] really like it'*. This mitigated a primary issue of the WiM room getting *'dramatically big'* (P3). They noted that the system felt responsive to their movement, linking this directly to enhanced predictability.

On the other hand, other participants (6/10) found the simultaneous dynamic resizing confusing and inconsistent. Because the visual scales were constantly changing, participants experienced altering movement gains, leading targeting errors when guiding their hand trajectory to the target. P5 explicitly mapped this visual-motor disconnect: *'[when] I want to grab small object, it [WiM] gets big, so I overshoot, [and] when I want [to grab a] big object, it gets small, so I undershoot'*. This frustration also appeared on another occasion, with P6 calling it *'awkward'* and noting it *'doesn't feel [like it has] consistency'*.

This dynamic scaling of both representations was confusing because, as P10 stated, users *'do not know what rule does the zooming.'*

These mixed responses suggest that hybrid scaling should be treated as an adjustable parameter for scaling retargeting rather than a fixed 50/50 blend.

6.7.2 Subjective Measure Analysis

While the analysis relied primarily on qualitative research notes, I also examined the quantitative measures, including NASA-TLX and responses to a custom 7-point Likert questionnaire. Given the small sample size ($N = 10$) and paired within-participants comparison between hand scaling and WiM scaling, I used two-sided paired Wilcoxon signed-rank tests to compare each NASA-TLX subscale and each Likert item. These tests were treated as exploratory and used to contextualize the qualitative findings rather than to provide confirmatory statistical evidence.

The Workload Assessment

The overall workload scores were relatively low and similar between the two strategies ($V = 18, p = 1.000$), with the hand scaling strategy ($Mdn = 5.58, IQR = 4.71$) scoring slightly lower than the WiM scaling strategy ($Mdn = 6.75, IQR = 5.04$) on a 21-point scale. Paired Wilcoxon tests did not reveal statistically significant differences between hand scaling and WiM scaling: Effort ($V = 5.5, p = .344$), Frustration ($V = 8.5, p = .393$), Mental Demand ($V = 11, p = 1.000$), Performance ($V = 13, p = .674$), Physical Demand ($V = 16, p = .799$), and Temporal Demand ($V = 7.5, p = 1.000$).

Despite similar overall scores, the WiM scaling strategy showed a slightly higher median Effort score ($Mdn = 5.00, IQR = 6.25$) than the hand scaling strategy ($Mdn = 3.00, IQR = 5.25$). Although this difference was not statistically significant, it aligns with the qualitative feedback, where participants reported that the WiM scaling required more physical arm and head movement when the WiM view changed dynamically.

The Alignment and Movement Assessment

Results from the custom experience Likert scale also showed no statistically significant differences between hand scaling and WiM scaling across the evaluated items: Cognitive Load (reversed) ($V = 13.5, p = .590$), Coherence ($V = 21, p = .240$), Contact Alignment

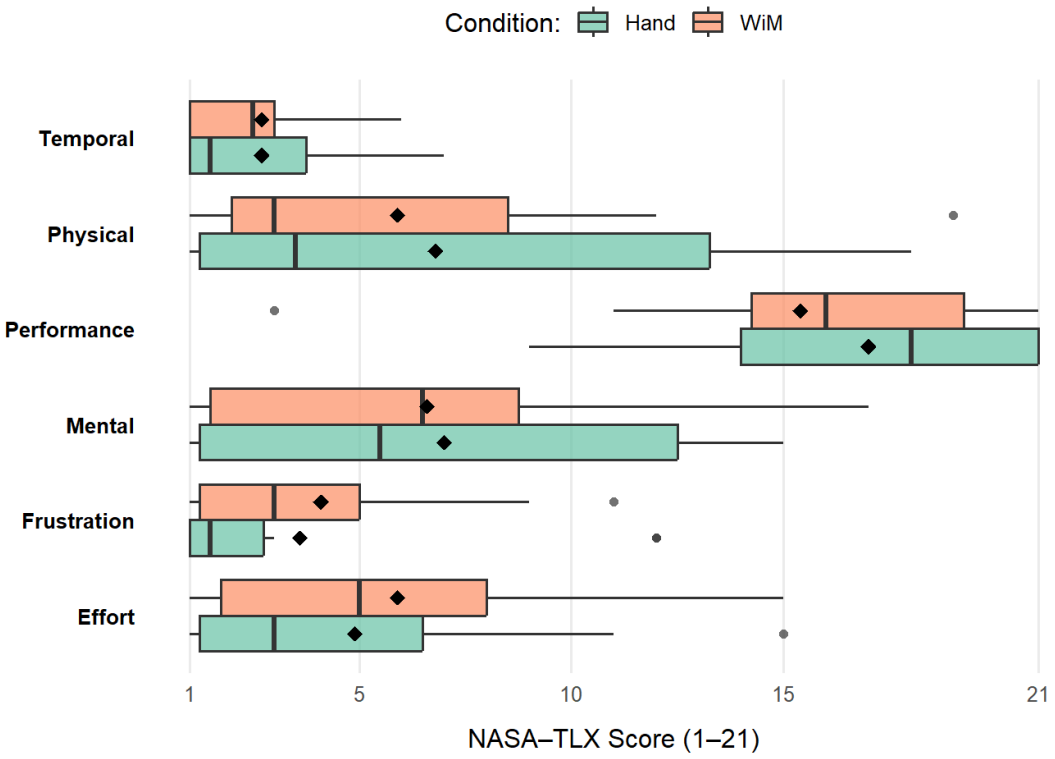


Figure 6.11: NASA-Task Load Index (TLX)

($V = 18, p = .138$), Embodiment ($V = 15.5, p = .863$), Predictability ($V = 5, p = .281$), and Workflow ($V = 15, p = .931$). Most items were similar across the two strategies. Cognitive Load (reversed), Coherence, Embodiment, and Predictability had equal or near-equal median ratings across conditions, while Workflow showed a small descriptive advantage for hand scaling (Hand scaling: $Mdn = 6.00, IQR = 2.25$; WiM scaling: $Mdn = 5.00, IQR = 1.00$).

The largest descriptive difference between the two strategies appeared in Contact Alignment, where hand scaling received a slightly higher median rating ($Mdn = 6.00, IQR = 1.50$) than WiM scaling ($Mdn = 4.50, IQR = 2.75$), suggesting that participants felt the physical-to-virtual alignment was more accurate and intuitive when their virtual hand representation scaled rather than the environment (WiM view). Although this difference was not statistically significant, it helps contextualize the qualitative findings that participants often perceived hand scaling as providing more stable contact, whereas the WiM scaling sometimes introduced uncertainty when the interaction volume changed dynamically.

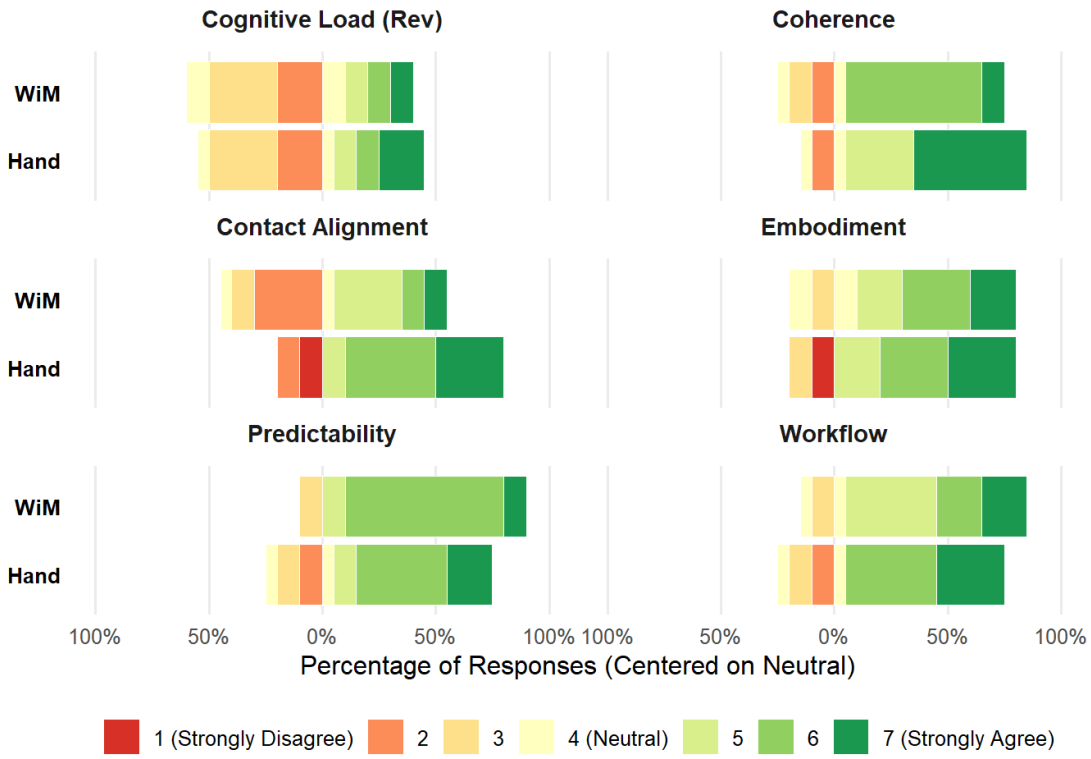


Figure 6.12: Custom Experience Likert

6.8 Discussion

Through the design framework developed in chapter 5, Laminar demonstrates how WiM mediation can turn scale mismatch from a static physical–virtual conflict into a temporally managed interaction design resource. The technique uses reaching and approach motions as opportunities for compensation, while maintaining the WiM as the space where the relationship between tangible proxy, miniature object, and full-scale referent remains visible. Having instantiated Laminar as a configurable technique, this section discusses broader design implications, limitations, and future directions for tangible, multi-scalar interaction for VR spatial design.

6.8.1 Laminar as a Configurable Technique

Building on the iterative development process and preliminary evaluation, I summarize Laminar as a configurable interaction technique designed for distributing physical–virtual discrepancy across scale and positional retargeting. Laminar supports navigation and manipulation as a single continuous interaction across a WiM, a full-scale contextual view, and a tangible proxy. As the user approaches, Laminar increases detail in the WiM by scaling up the viewport and dampens the control–display ratio for precise control. The system preserves alignment between the selected object and the tangible proxy by coupling retargeting with body-centric cues, such as head movement. The WiM and full-scale views are concurrent and spatially consistent. Users manipulate in the WiM while verifying immediate feedback in the full-scale view.

I position Laminar as a configurable interaction technique. Rather than prescribing a single optimal configuration, Laminar organizes the design decisions revealed through the iterations into adjustable interaction parameters. The reference instantiation emphasizes interaction behaviours that balance stability, predictability, and effort in participants' responses, while leaving configuration choices adaptable to task demands, hardware setup, and user preferences.

Parameterized Interaction Model

The findings can also be interpreted as evidence for the framework proposed in chapter 5. Participants' responses show that assigning discrepancy to different spaces produces different experiential consequences. Hand scaling stabilized virtual embodiment for some users, while WiM scaling preserved proprioceptive consistency but sometimes increased physical effort and reduced predictability. This distinction motivates treating hand scaling and WiM scaling as configurable parameters rather than as competing fixed solutions. This supports the framework's claim that discrepancy is not eliminated, but redistributed across embodiment, WiM, and world spaces. The findings suggest two primary parameter groups for configuring Laminar: scale retargeting and positional retargeting. These parameters define how scale and position mismatch are distributed across the user's body representation, the WiM interaction volume, and the tangible proxy. These parameters are presented as design guidance rather than optimized values.

Scale Retargeting

In this model, scale retargeting is bounded by two complementary limits. WiM size imposes an upper bound because excessive expansion of the WiM container requires greater physical movement and can reduce spatial predictability. This reflects participants' reports that WiM scaling sometimes made the room feel dramatically enlarged or required more arm and head movement. Hand representation size imposes a lower bound because excessive body-centric scaling can disrupt the legibility of the hand as an embodied control. Participants' responses suggest that the hand representation can shift between being experienced as part of the body and being treated as an indirect instrumental cursor. When hand scaling becomes too aggressive, this shift can reduce grasp predictability and weaken visual–proprioceptive expectation. A bounded hybrid configuration therefore distributes mismatch between WiM scaling and hand scaling rather than relying exclusively on either strategy.

Positional Retargeting

The findings also motivate two approaches to positional retargeting: head-driven and hand-driven alignment. Head-driven retargeting couples head movement with a proxemic approach, allowing retargeting to be masked during view scaling and shifts in visual attention. However, participants' responses also revealed that repeated head movement could become physically demanding or feel unnecessary when the goal was simply to align the object with the proxy. While the preliminary evaluation involved head-driven retargeting, hand-driven retargeting is included as a design variant motivated by participants' desire for more direct and less effortful alignment. Hand-driven retargeting, therefore, represents a configurable variant that prioritizes the reaching hand as the primary cue for alignment. In this configuration, the WiM container moves in response to the hand's trajectory, velocity, and distance to align the selected virtual object with the tangible proxy. This approach prioritizes functional haptic access and motor intent, rather than relying primarily on masking the retargeting process through head movement.

6.8.2 Limitations

While Laminar was developed through an iterative RtD methodology with the defined design principles, the current artifact has several limitations that narrow the scope of its immediate generalizability.

Passive Haptic Constraints

Since the system relies on passive haptics using a single tangible proxy on a tabletop (or desk), the interaction is limited to local, arms-reachable space. However, the WiM mitigates this limitation by allowing local tabletop manipulation to affect a larger environment in VR. Laminar does not support room-scale locomotion via physical walking, but it does support large-scale manipulation beyond the immediate physical workspace and contextual verification through WiM interaction. This surface computing metaphor is therefore appropriate for detailed design work but may be less suitable for tasks that require full-body navigation, large physical movement, or direct physical traversal of the virtual world. Furthermore, the reliance on passive haptics brings several inherent limitations, notably limited kinesthetic feedback (force or resistance) and inherent shape mismatch, such as irregular sculptures or curved lamps. While the system supports positional congruence, it does not provide geometric congruence.

Generalization of Implicit Control Parameters

The parameters used to complete the dynamic haptic retargeting are currently based on a non-personalized model. This assumption was made to accommodate rapid iterations for the RtD process. The implicit control relies heavily on a comfortable range of (head) motion from its users. It may not be suitable for every user, or the system might introduce long-term cognitive or ergonomic fatigue for some users, which warrants further investigation on a personalized solution or an adaptive algorithm to prevent ergonomic fatigue or disorientation.

6.8.3 Broader Implications

Beyond specific applications such as interior design, Laminar reinforces the value of proximity-based interaction (Ballendat et al., 2010) in VR tool design. While implicit, distance-based interactions have been explored for public displays and proxemic interfaces, Laminar demonstrates their critical utility in managing the scale of WiM interaction in VR. Current VR paradigms largely rely on explicit commands for menus and modes, which imposes a high cognitive tax. My research instead suggests that systems can leverage natural body language, such as head motion (R. Shi, Zhu, Liang, & Zhao, 2021). This implies that the solution to complex, multi-variable interaction problems, such as retargeting, may be achieved by leveraging existing relationships between the user and system, rather than

relying on hardware-heavy development. This approach could be extended to other domains that require multi-scale, multi-view manipulation, where the workflow should not be interrupted for optimal results, such as 3D artwork or medical visualization.

6.9 Chapter Summary

This chapter presented Laminar, a configurable tangible interaction technique that manages physical–virtual discrepancy in multi-scale VR spatial design. By promoting continuous interaction workflows with reduced overhead, the technique addresses the friction often found in VR by exploring the tradeoffs between body- and space-centric (WiM) scale mismatch compensation strategies. The findings reveal critical design implications for haptic congruence, contextual continuity, and implicit control. Laminar was iteratively designed using a RtD approach based on the design framework in chapter 5, and the technique was examined through a preliminary user study. Ultimately, this technique provides foundational insights for future systems seeking to manage spatial discrepancies while combining the stability and precision of tangible interaction with the flexibility and scale of virtual environments.

In relation to the design framework developed in chapter 5, Laminar demonstrates how one generated configuration—scale-mismatched tangible proxy coupling into virtual WiM—can be instantiated, probed, and refined through RtD. The findings show that different compensation strategies redistribute discrepancy to different spaces: hand scaling assigns compensation to virtual embodiment, while WiM scaling assigns compensation to virtual WiM space. This supports the framework's central claim that discrepancy is not eliminated, but managed across embodiment, WiM, and world spaces over time. Laminar therefore serves as a concrete instance of the design framework, showing how WiM mediation can support haptic congruence, contextual continuity, and continuous multi-scale interaction.

Chapter 7

Synthesis

Across the four research chapters, a common pattern emerges: tangible WiM-based VR interaction is less about eliminating or masking spatial discrepancy than about deciding where, when, and how discrepancy should be distributed across representation, embodiment, and interaction state. This pattern appears first in chapter 3, where users struggled to translate scale across miniature-based tools; it is then examined experimentally in chapter 4, where visual–proprioceptive mismatch affected perception and manipulation depending on mismatch direction and proxy size. Chapter 5 formalizes this problem as a temporal design framework, and chapter 6 instantiates one configuration of the framework through Laminar. The thesis synthesis can be understood through three themes: the temporal redistribution of spatial discrepancy, the evolution of WiMs from miniature overviews into active interaction volumes, and retargeting as a tradeoff between perceptual illusion and interaction performance.

7.1 Redistributing Spatial Discrepancies Across Time

Prior work often uses controlled discrepancies between physical and virtual representations to increase perceived immersion, interaction possibilities, or tangible experiences. Scale mismatch, for example, can be a persistent discrepancy in VR, especially because cost-effective and scalable tangible proxies cannot always match every virtual object. In VR, this discrepancy is exacerbated by physical-digital asymmetry. While scaling digital form may be trivial, changing the spatial properties of physical form requires complex, specialized hardware. While this discrepancy exists, many design sectors, including AEC, already incorporate scale differences into their design process.

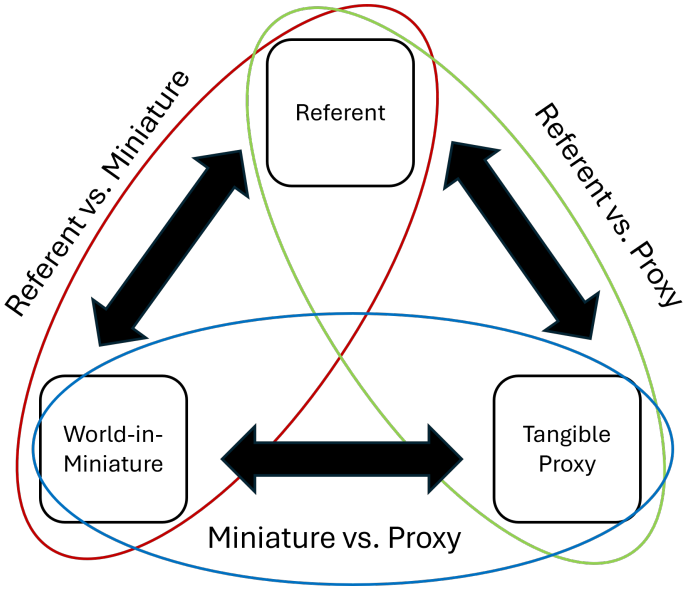


Figure 7.1: Three discrepancy relationships among the referent, WiM, and tangible proxy.

In this dissertation, mismatch is treated as a manageable design choice rather than as a factor to be hidden. The goal is not necessarily to eliminate this mismatch or mask it entirely, but to keep it within an acceptable, usable range, parameterize its operational tradeoffs, and understand how mismatch can be actively redistributed temporally, to support a continuous design workflow.

7.1.1 Triad of Geometric Mismatch

Miniature representations are often incorporated into design processes to compress and represent large-scale environments into compact, manageable forms. Scale mismatch appears in this process. Designers develop practices to address *mismatch*, or discrepancies, in order to perceive scale correctly, sometimes with the help of external aids. Frequent switching between multiple scale views may introduce noise and misperception, making it hard to express and implement design decisions without appropriate tools, as evidenced in chapter 3.

In this dissertation, mismatch appears across three relationships among the referent object, its miniature representation, and the tangible proxy Figure 7.1.

1. **Referent vs. Miniature:** First, there is a mismatch between the referent object

and the miniature representation of the referent. This exists wherever a miniature or scaled model is used. For example, blueprints and tabletop gaming both use scaled representations naturally. However, in VR, this becomes more dynamic because the virtual representation can change dynamically.

2. **Referent vs. Proxy:** Second, there is a discrepancy between the referent object and the tangible proxy: when integrating a tangible proxy to control significantly different-sized objects, users often treat it as an instrument rather than as a direct representation, especially when there is no clear visual mapping. It is often unclear whether a user treats the proxy as a local counterpart to the miniature representation inside the WiM, or as a direct tool for the macro referent in the full-scale world space. While this does not pose a significant problem in the real world, the separation between physical touch and virtual representation in VR makes it difficult to interact with virtual referent objects.

3. **Miniature vs. Proxy:** Third, there is a mismatch between miniature representations and the tangible proxies controlling them. This relates to the first mismatch category but becomes more complex when tangible control is introduced. Although the sizes may not differ significantly, having multiple miniature representations of different sizes can be an obstacle.

The studies in this dissertation examine these discrepancies and provide evidence for this framing. Chapter 3 examines the mismatch between full-scale referent and miniature representation in different media. While this involves a tangible representation (the LEGO condition) as a means of directly controlling the assembly and placement of objects, it deliberately removes other layers to consider (e.g., LEGO condition: no VR; online condition: no tangibility). The insights show that the spatial mismatch between these two does not break the interaction; rather, it introduces perceptual noise as misperception that must be managed. This framing continues in chapter 4, where scale mismatch is isolated and explained between the virtual object and tangible proxies, and in chapter 5, where mismatch becomes part of the design choices.

7.1.2 Temporal Redistribution Across Interaction States

This dissertation demonstrates that spatial discrepancies, especially scale mismatch, can be managed dynamically by distributing them across interaction states: browsing, reaching, grasping, manipulating, and releasing. By decomposing an interaction technique into these

temporal states, designers can determine whether discrepancies should be implicitly masked in pre-contact phases, explicitly exposed, or managed through hybrid combinations across interaction spaces.

Techniques such as haptic retargeting illustrate one temporal strategy: concentrating discrepancy management before contact through implicit transformation. During the early phase of an interaction loop, especially browsing and reaching states, the human motor system relies primarily on visual feedback to plan trajectories and estimate the target for the grasp aperture. This temporal window offers an opportunity for a system to execute implicit coordinate transformations, such as body or world warping. By leveraging visual dominance described in chapter 4 during this window, the interface can subtly transform to guide user embodiment from physical to virtual while physically guiding the hand toward the tracking coordinate of the tangible proxy. In this strategy, discrepancy is redistributed temporally so that grasping and manipulation can proceed stably.

WiM-based interaction shows a second temporal strategy: distributing discrepancy across the interaction loop through a persistent mediating space. Rather than focusing on transformation during pre-contact states, WiM allows discrepancy to be managed during browsing, reaching, grasping, manipulating, and releasing. Intentionally exposing these discrepancies through WiM space allows the user's local visual–proprioceptive expectations to remain stable within a broader range, while the system translates those local actions into full-scale environment effects in world space.

This persistent mediation becomes important as the user moves into the contact phases, such as when grasping and manipulating. During these states, the interaction loop encounters the perceptual tolerance boundaries between vision and haptics (chapter 4). If visual alignment exceeds tolerance during this phase, the visual–proprioceptive illusion breaks.

Chapter 6 provides the realization of this design framework by implementing scale mismatch compensation. Laminar demonstrates that the scale mismatch can be managed temporally and mediated through WiM space. By coupling positional and scale retargeting to the user's natural reaching and approaching motions, Laminar uses movement as a window for adjustment. At the same time, the WiM keeps the relationship between the tangible proxy, miniature object, and full-scale referent explicit during interaction. Ultimately, the physical-virtual mismatch is transformed from a static interaction hindrance into a dynamic, temporal mechanism that manages and drives multi-scalar design.

7.2 Worlds-in-miniature as Abstraction

The WiM system is traditionally used as a viewpoint for a localized, separate overview of the full-scale VR. Using a WiM as an interaction space offers strong embodiment and high spatial confidence for direct manipulation within a separate viewpoint. My thesis builds on this paradigm, transforming the WiM into a primary interactive volume with tangible interaction rather than just a visual container.

Because passive haptic proxies are immutable, other components of the system should be changed to make interaction as flexible as possible. In this case, the WiM interactive volume can change its form, as it is already a scaled view of the full-scale referent. Transforming it minimizes the impact on the user, as long as it stays within thresholds.

7.2.1 Tangible Interaction as a Grounding Anchor

Tangible user interfaces (TUIs) have been recognized as a major contributor for enhanced immersion, realism, and usability in VR. However, integrating TUIs into VR often introduces scalability bottlenecks, as it is impractical to fabricate a unique tangible proxy for every virtual object. Therefore, in this dissertation, a tangible proxy is primarily used as a stability tool and physical control for virtual objects.

I used a tangible proxy to ground perception and stabilize physical reference during interaction. Surface computing often uses this grounding to provide stable, predictable interaction. Operating on a surface with a passive haptic proxy allows precise manipulation by leveraging the surface friction and effectively stabilizes the y-axis (height). This physical anchoring grounds proprioception and fine motor skills.

In chapter 3, findings suggested that participants strongly agreed that tangibility with a miniature representation was engaging. The study in chapter 4 deliberately introduces scale mismatches to examine how participants perceive them and how they affect manipulation performance. The findings support a coordinated, segmented sensation when 3D manipulation, visual information, and haptic sensation cooperate. The tangible proxy provides a strong grounding effect, maintaining functional performance during specific interaction phases.

This grounding is then used to begin retargeting and to provide stable, predictable interaction with multiple virtual objects. Although physical grounding helps build stable interaction techniques, geometric differences stemming from a fundamental mismatch between physical and virtual dimensions make the interaction inevitably rely on virtual transformations.

7.2.2 Embracing Mismatch: Two Spatial Frames, Two Profiles

Suggested by the findings in chapter 3, users in miniature interactions diversify their choices, making the results for divergent preferences in tools dependent on design stages and functional objectives. Participants did not converge on a single tool for the specific task (e.g., room layout) across different media. Instead, there were two possible passages of functional choice. The digital form was valued for its easy editing capabilities and external aids, such as rulers and measurements provided, whereas the tangible form (i.e., LEGO bricks) was valued for its direct, tactile feedback, which led to higher engagement. This clear separation can be understood in terms of the directions in which participants diverged, rather than a simple preference for tools. This dissertation articulates this separation of choice in terms of different cognitive preferences regarding spatial ability and trust when encountering a fundamental mismatch between physical and virtual spaces.

The diversification of tool preferences can be mapped onto the user behaviours observed during the evaluation of the scale mismatch compensation strategies in chapter 6. Participants who preferred hand-scaling strategies anchored their cognitive model in the stability of the surrounding workspace in WiM. They tolerated visually distorted (scaled) virtual hand representation, depending on the active virtual object, to compensate for visually aligned contact simulation, despite losing embodied visual feedback on hand representation. This approach reflects traditional indirect proxy interactions, such as computer mouse input, where the physical hand functions as an externalized instrument or cursor navigating a stable, exocentric reference frame. For this profile of users, a static and stable workspace offers better predictability and control, allowing the visual dominance thresholds in chapter 4 to safely override proprioceptive drift.

On the other hand, participants who preferred the WiM scaling strategy commented on their embodiment for their controllable representation. This group of participants anchored their cognitive referent in their virtual representation, and generally disliked working with a distorted referent body. The visually scaled hand representation broke their internal cognition for physical grasping, generating sensory frustration when they encountered physics-breaking visuals. They require the environment to absorb the entirety of the spatial discrepancy. They experience the dynamically scaling WiM interaction volume as a highly responsive functional interface. For this profile of users, responsiveness and body-centric direct manipulation are favoured over breaking the egocentric reference frame.

These two passages suggest that there is no universal design paradigm for tangible WiM-based interaction. Instead, user behaviours separate into distinct profiles: instrument-driven, which treats the proxy as a tool directly on the macro referent in world space, and proprioception-driven, which anchors the proxy to their physical body as a $1:1$ local coun-

terpart to the visual miniature. This dissertation shows that the tangible proxy must act as a stable physical grounding anchor, while the WiM interaction volume manages spatial discrepancies, maintaining contextual continuity without breaking haptic congruence.

However, dynamically altering these reference frames to achieve spatial integrity does not eliminate the systemic costs. It shifts the type of burden placed on the user. Specifically, the hybrid strategy used in chapter 6 was not received well by all participants due to the complex initial experience. When warping both the hand and the space simultaneously, participants experienced fluctuating movement gains that corrupted their trajectory planning, leading to visual-motor performance that broke down into over- and undershooting. This interaction establishes a design knowledge for XR scaling that shows how scale transformations can be mapped to dynamic transitions like reaching, but it is best to avoid simultaneous warping across different ego- and exo-centric frames, which requires heavy visual-motor coordination.

7.3 Retargeting as a Tradeoff between Illusion and Performance

Prior work often emphasizes and uses the retargeting technique as a perceptual illusion to repurpose a single tangible proxy for multiple objects. This dissertation reframes retargeting as an enabling mechanism for multi-view, multi-scale interaction rather than as a means to hide warping. Retargeting techniques provide opportunities for interaction with multiple virtual objects by relying on users' tolerance of visual–proprioceptive discrepancy and by intentionally warped representations of body or environment in VR. Optimizing it within the context of tangible WiM interaction requires careful balancing between cognitive and physical efforts from users instead of eliminating them.

This tradeoff between illusion and performance forces designers and developers to balance the two in implementations. Retargeting benefits from perceptual illusions and visual dominance, but relying on illusions can become fragile when the conditions for their effectiveness are not met or when proprioception, physical contact, or precise manipulation make discrepancies noticeable. While retargeting enables interaction with multiple virtual objects, certain applications can demand extra effort from users, such as spatial design or 3D manipulation. In chapter 3, the study highlights that participants struggled to judge the accurate scaling ratio of miniature representations and to express them properly across different media. This suggests the need for interaction techniques that support scale reasoning, not only visually expressed ones, but also other forms of support. This directly

connects to the study in chapter 4, which investigates scale perception across three conditions. The findings show that users can maintain functional performance despite mild discrepancies and detect mismatches. Its outcome also suggests that scale mismatch affects perception. This informs how a systemic solution can maintain functional performance regardless of noticeable discrepancy.

Chapter 5 builds on these insights by reframing retargeting as a strategy for redistributing discrepancy across interaction spaces and interaction states. In the framework, retargeting is not only a way to mask discrepancy, but also a design choice about where discrepancy should be managed, when it should be adjusted, and how it affects the relationship between embodiment, WiM, and world spaces. Based on the framework, chapter 6 instantiates this argument through Laminar: hand scaling assigns compensation to virtual embodiment space, WiM scaling assigns scale compensation to virtual WiM space, and WiM-localized positional retargeting aligns the selected virtual object with the tangible proxy. This synthesis suggests that haptic retargeting for tangible WiM-VR should be treated as a trade-off rather than only as a solution to perceptual illusion. While prior retargeting often emphasizes perceptual masking, Laminar instead emphasizes practical value by balancing illusion, performance, and workflow continuity. This dissertation highlights that retargeting is not only about shifting perceived positions, but also managing tradeoffs for multi-view, multi-scalar VR systems, including how burden shifts between system and users.

Chapter 3 suggests that participants struggled with scale misperception, likely due to the lack of an external guide, such as automatic scaling within digital media (e.g., an online tool) and no explicit references. Preliminary evaluation in Chapter 6 demonstrates how users interact with dynamic scaling of two different scale mismatch compensation strategies. Rather than providing conclusive evidence, these insights suggest a descriptive pattern: WiM scaling strategies may increase perceived effort when the interaction volume changes substantially, while assigning compensation to virtual embodiment can affect how users interpret their hand as either a body representation or an instrument.

The empirical data from chapter 4 suggests that users can maintain functional interaction performance even when there is a noticeable discrepancy between the tangible proxy and the virtual object across a range of differences. Chapter 6 also demonstrates that participants could engage and interact with scale compensation strategies, indicating that scale discrepancy does not necessarily invalidate and break the interaction loop.

Overall, this dissertation demonstrates that spatial discrepancy and retargeting are closely related, and that discrepancy can be managed across embodiment representation, WiM scale, world mapping, or user movement. This dissertation therefore reframes re-

targeting techniques as a balancing mechanism to distribute spatial discrepancies across interaction spaces and interaction states.

7.4 Chapter Summary

This chapter presents a comprehensive synthesis of this dissertation, covering empirical, theoretical, and system-level aspects. The synthesis describes my perspective on geometric mismatch as a design element that can be managed through temporal and spatial dimensions.

This synthesis identifies three distinct mismatches among the referent object, the miniature representation, and the tangible proxy and provides a deeper understanding of how they occur across interaction spaces and states. This thesis demonstrates how these discrepancies appear in existing interaction techniques and how they can be managed. During the early interaction states, such as browsing and reaching, discrepancies can be managed implicitly by leveraging visual dominance and movement. However, WiM space extends this strategy by providing a persistent mediating space across the interaction loop. Rather than only hiding discrepancies before the user makes contact with the object, WiM can contain, expose, or redistribute them across different interaction states, allowing local visual–proprioceptive expectations to remain stable while local actions are translated into world-scale effects.

This chapter also examines how a virtual interactive WiM volume functions as an active abstraction rather than a passive visualization tool. The dissertation demonstrates how WiM space absorbs spatial discrepancies by mediating immutable physical limitations through augmentation, leveraging a flexible WiM scaled volume to compensate for these discrepancies, while providing spatially aligned proxy–virtual object contact.

Finally, the synthesis identifies two distinct user profiles for tangible-WiM-VR interaction when spatial discrepancies are compensated between user embodiment and WiM space. User behaviours are split into distinct operational profiles based on how they anchor their cognitive spatial reasoning. Users who favour the hand scaling strategy show an exocentric cognitive style, choosing to preserve the stability of the visual WiM space by tolerating dynamic changes within the virtual embodiment. On the other hand, users who favour the WiM scaling strategy show an egocentric style, requiring their virtual embodiment to remain unaltered.

Collectively, these insights provide a theoretical foundation for future spatial user interfaces and interaction design to balance human spatial cognition with physical limitations.

Chapter 8

Conclusion and Future Work

This thesis centres on the concept of worlds-in-miniature (WiMs) in VR, integrating a single tangible proxy to control virtual objects. My research has been motivated to enhance the usability of the WiM-VR system through a tangible proxy, especially by supporting spatial design workflows and exposing tradeoffs. The first study explores how current tools are used, identifies perceptual issues, and empirically measures how people perceive scale mismatch and perform under such conditions. Then, I derive a design framework based on how WiM contributes to managing discrepancies as an intermediate interaction space. The last research chapter of the dissertation instantiates a prototype interaction technique using the RtD methodology and examines how two scale mismatch compensation strategies affect experiences through a preliminary evaluation.

This chapter summarizes the research contributions, identifies the thesis's limitations, discusses future research, and finally presents closing remarks.

8.1 Contributions

Building on prior work in VR, TUI, and WiM interaction, this dissertation contributes interaction design knowledge for developing tangible WiM-based interaction in VR, with empirical, system-level findings. This dissertation characterizes existing miniature-based design tools and identifies underlying cognitive limitations. It then examines how visual–proprioceptive mismatch between tangible proxy and virtual counterparts affect interaction states, and constructs a design framework and an instantiated prototype interaction technique for multi-scale, multi-view spatial interaction. The contributions are, as described in chapter 1, are as follows:

C1. An exploratory study of current design tools that examines the potential use of tangible interaction for small-scale objects.

C2. A user study that identifies what scales users can interact comfortably and confidently with in VR during design activities.

C3. A design framework for spatial interaction techniques that can be used when developing interaction techniques with scaled models and environments in VR.

C4. A prototype that demonstrates how a tangible proxy with WiM can be designed effectively in a full-scale virtual environment.

8.1.1 Characterizing Miniature-Based Design Tools

This dissertation research begins by characterizing miniature-based tools through the context of layout design workflows. By comparing a digital online tool, paper and pen, and physical LEGO blocks, this work identifies how people experience each tool differently and in which aspects each excels and exposes weaknesses, signalling a potential approach to different design stages. This contribution (C1) establishes the dissertation's initial problem space by showing that miniature-based tools can support engagement and flexibility. In contrast, this work demonstrates that inaccurate scale judgment may appear without external help.

8.1.2 Empirical Grounding for Visual–Proprioceptive Mismatch

Depth perception and distance perception have been extensively studied, though the concept of scale perception remains relatively underexplored, especially within tangible and miniature interaction. This dissertation formalizes and investigates scale mismatch across three different conceptual levels: the referent object, its miniature virtual representation, and the tangible proxy. This contribution (C2) provides empirical grounding for understanding how visual and proprioceptive feedback interact, revealing perceptual asymmetries in thresholds when users manipulate scaled objects in a WiM environment.

8.1.3 Design Framework for Tangible WiM-VR Interaction

The design framework (C3) contributes design knowledge to tangible-WiM-VR interaction by analyzing parameters for spatial discrepancies. It organizes the design space for tangible

WiM-VR interaction by identifying how spatial properties, such as WiM, the full-scale environment, and body representation, transform to mitigate spatial discrepancies. Rather than presenting a single solution, the framework explains how designers and developers can reason about tradeoffs, manage and distribute spatial discrepancies across explicit interaction spaces and states, tailored to the desired workflow and applications.

8.1.4 Laminar as a Prototype Interaction Technique

The final contribution of the dissertation is Laminar (C4), an instantiated prototype interaction technique that operationalizes a specific design path through the developed design framework. Laminar demonstrates how tangible proxy manipulation, WiM interaction, haptic retargeting, and multi-view VR can be synthesized in practice to manage spatial discrepancies. The interaction technique is not presented as a final, optimized system, but rather as a research artifact that illustrates the feasibility and consequences of the proposed design framework, demonstrating that interaction technique can successfully distribute discrepancies across different design dimensions, both temporally and spatially.

8.2 Limitations

While this dissertation provides empirical and research artifact contributions, these contributions are bound by the scope of the studies, tasks, participants, and technical implementations used throughout the research.

8.2.1 Lab Disruptions and Participant Samples

The studies presented in chapter 4 and chapter 6 were heavily impacted by the theft of lab equipment. All research computers and VR headsets were stolen during active research, leading to a major pause in active research for technical recovery. This forced the premature conclusion of the research study, resulting in smaller sample sizes than originally planned. The final prototype evaluation was inevitably conducted with a different VR headset and technical setup than originally planned. This structural shift led to additional delays due to necessary implementation changes and protocol adaptation. Although the study still served its purpose as an exploratory research through design (RtD) probe, the change in apparatus limits direct continuity across the technical implementations used in the dissertation.

As a result of these events, the controlled experiments concluded with final samples of 12 participants in chapter 4 and 10 participants in chapter 6. Although the effect size and robustness of the analysis yielded clear outcomes, the generalizability of the claims may be limited by the smaller sample size. While their study designs and nature did not require a large sample size, a larger sample would be needed to support broader claims about population-level performance or usability.

8.2.2 Study and Data Collection

The first study was conducted remotely via video conferencing. This allowed me to recruit from a broader pool of participants rather than relying only on participants from the local institution. However, the remote sessions may still have introduced variability in participants' environments and in the availability of the tools required for participation, as they were asked to prepare several study materials, such as LEGO bricks, printed papers, and writing tools. These differences in setup may have affected the study outcomes.

8.2.3 Scope of the Task

The dissertation leverages spatial design tasks, specifically room layouts. These tasks were selected because they involve scale perception, object placement, and 3D manipulation, while being more accessible to participants without specialized training. Focusing on a specific class of spatial design tasks may limit the scope of the research outcomes. While this interaction technique and knowledge generated may be applicable to other tasks or domains, this dissertation does not directly evaluate or validate them in alternative or adjunct areas, such as macro-scale urban planning, individual product design or collaborative workflows.

8.2.4 Validation

While Laminar, an instantiated interaction technique, is grounded in an iterative RtD and a preliminary evaluation, this dissertation does not provide summative and comparative performance benchmarks. Instead, the preliminary study presents an exploratory probe to capture fundamental differences in spatially distinct scale-mismatch compensation strategies, rather than to validate the system against existing techniques. Additionally, this dissertation does not provide full coverage of all parameter configurations in the design framework, while multiple configurations are possible.

8.3 Future Work

The limitations of this dissertation suggest several directions for future research.

8.3.1 Formal Evaluation and Optimization

Future work should focus on formal, full-scale studies of the design framework. Further investigation should involve comparing Laminar to baseline techniques for empirical performance benchmarks, such as direct manipulation or controller-based mid-air manipulation. These evaluations would test the full set of parameters established in the framework, strengthening its generalizability. Further studies should also optimize balanced parameter values and propose thresholds for various tasks based on tradeoffs. Ultimately, these can be formalized mathematically as offset formulations.

8.3.2 Intent Detection and Personalization

While the current evaluation used a left-hand controller to switch the target among three virtual objects for selection, in future work, an intelligent system could detect which object the user intends to select without requiring explicit user input. For example, machine learning models could incorporate various body cues, such as head movement, hand poses, reaching trajectory, and gaze tracking, to distinguish the target and improve prediction.

An intelligent component will also help shape the system's personalization. Since Laminar is derived from the design framework rather than a generalized interaction technique, future systems could adapt retargeting parameters to individual users. For example, the intelligent component could take users' unique cognitive reference style by analyzing behavioural logs against a different set of virtual objects, such as hand and head movements. Based on findings in chapter 3 and chapter 6, there were divided user profiles about which transformations should scale, whether WiM scales or hand representation scales. This suggests that future work can build on the technique to make it more personalized for individual users.

8.3.3 Expanded Setup

Future systems can expand this paradigm by incorporating multiple proxies. This dissertation focused on a single proxy to examine how retargeting across multiple virtual objects

can be designed and to isolate its impact on interaction design. Future work should investigate the technique with additional dimensions to the design framework, along with a complex algorithm to optimize its retargeting behaviour. For example, a WiM container could slide toward the closest tangible proxy, or a more proximate one to the user, depending on the situation.

Adding multiple proxies also opens opportunities for collaborative settings. Each tangible proxy could be assigned to a different user or shared among them, and the system could intelligently retarget proxies based on the task and the proximity of elements. This should be thoroughly investigated through group-based studies, identifying specific possible scenarios to optimize its system behaviour.

8.3.4 Generalizing Knowledge

Future work should investigate how the design framework and empirical results can be applied beyond spatial design. Domains other than AEC should be investigated, especially domains involving 3D manipulation that benefits from dual views, various scales, and physical grounding. This future investigation could identify which parts of the framework are specific to a domain and how they can be generalized to broader multi-scalar interactions in VR.

8.4 Revisiting the Thesis Statement

This dissertation has argued that although WiMs are widely used for visualization in VR, their full potential as the primary tangible interactive mechanism has been limited by visual–proprioceptive tensions. This research has distilled these into a progressive research arc.

The first study (chapter 3) showed that current design tools that leverage miniature representations pose several challenges for their users, including scale misperception. The second study (chapter 4) then investigated empirically to provide grounded evidence of mismatch tolerances under visual–proprioceptive scale mismatches. Building on these, chapter 5 formulated a structured multidimensional design framework, analyzing existing interaction techniques and how they address such mismatches. Chapter 6) identified challenges within tangible-WiM-VR interaction for a spatial design task, instantiating an interaction technique that operationalizes how mismatch is redistributed by transforming WiM and interaction embodiment.

This dissertation research arc investigates and integrates three core concepts, WiM, VR, and TUI, through the context of spatial design tools. While combinations of these concepts have been introduced in prior work, this dissertation contributes knowledge about how they can be understood together through scale perception, tangible interaction, and system-level design. The findings from the studies, the design framework, and the instantiated interaction technique demonstrate how this combined approach establishes a coherent paradigm.

This dissertation has demonstrated Laminar, an instantiated interaction technique derived from a generative design framework. By examining scale perception and tangible proxy with haptic retargeting in multiple-view virtual environments, this dissertation has established a tangible-WiM-VR paradigm that supports spatial design tasks. Ultimately, this dissertation demonstrates that physical-digital discrepancies do not constitute an interaction barrier, but rather represent manageable design elements, providing a robust foundation and clear architectural design knowledge for future spatial interaction and interfaces to balance spatial cognition with real-world assets.

References

Abruzzo, E., Ellingsen, E., & Solomon, J. D. (Eds.). (2007). *Models: 306090 11.* New York, NY: Princeton Architectural Press.

Abtahi, P., Gonzalez-Franco, M., Ofek, E., & Steed, A. (2019, May). I'm a giant: Walking in large virtual environments at high speed gains. In *CHI 2019 - Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery. Retrieved 2026-02-25, from https://collaborate.princeton.edu/en/publications/im-a-giant-walking-in-large-virtual-environments-at-high-speed-ga/ doi: 10.1145/3290605.3300752

Ahmed, S., Wallace, K. M., & Blessing, L. T. (2003, February). Understanding the differences between how novice and experienced designers approach design tasks. *Research in Engineering Design*, *14*(1), 1–11. Retrieved 2026-02-26, from https://doi.org/10.1007/s00163-002-0023-z doi: 10.1007/s00163-002-0023-z

Aliakseyeu, D., Martens, J.-B., & Rauterberg, M. (2006, July). A computer support tool for the early stages of architectural design. *Interacting with Computers*, *18*(4), 528–555. Retrieved 2026-02-26, from https://www.sciencedirect.com/science/article/pii/S0953543805001128 doi: 10.1016/j.intcom.2005.11.010

Andrzejczak, J., Kozłowicz, W., Szrajber, R., & Wojciechowski, A. (2021). Factors Affecting the Sense of Scale in Immersive, Realistic Virtual Reality Space. In M. Paszynski, D. Kranzlmüller, V. V. Krzhizhanovskaya, J. J. Dongarra, & P. M. Sloot (Eds.), *Computational Science – ICCS 2021* (pp. 3–16). Cham: Springer International Publishing. doi: 10.1007/978-3-030-77977-1_1

Arckit 100 sqm. Architectural Model Building Kit. (2026). Retrieved 2026-05-29, from https://www.arckit.com/products/a100

Arkio. (n.d.). Retrieved 2026-06-25, from https://arkio.is

Armbrüster, C., Wolter, M., Kuhlen, T., Spijkers, W., & Fimm, B. (2008, February). Depth Perception in Virtual Reality: Distance Estimations in Peri- and Extrapersonal Space. *CyberPsychology & Behavior*, *11*(1), 9–15. Retrieved 2026-02-

26, from https://journals.sagepub.com/action/showAbstract doi: 10.1089/cpb.2007.9935

Arnowitz, E. M. (2017). vSpline: Physical Design and the Perception of Scale in Virtual Reality. In *ACADIA 2017: DISCIPLINES & DISRUPTION [Proceedings of the 37th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA) ISBN 978-0-692-96506-1] Cambridge, MA 2-4 November, 2017), pp. 110-117.* CUMINCAD. Retrieved 2026-02-25, from https://papers.cumincad.org/cgi-bin/works/Show?acadia17_110

Autodesk revit software. (n.d.). Retrieved 2026-06-25, from https://www.autodesk.com/ca-en/products/revit/overview

Azmandian, M., Hancock, M., Benko, H., Ofek, E., & Wilson, A. D. (2016, May). Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1968–1979). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/2858036.2858226 doi: 10.1145/2858036.2858226

Ballendat, T., Marquardt, N., & Greenberg, S. (2010, November). Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces* (pp. 121–130). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/1936652.1936676 doi: 10.1145/1936652.1936676

Ban, R., Matsumoto, K., Narumi, T., & Kuzuoka, H. (2022, October). Wormholes in VR: Teleporting Hands for Flexible Passive Haptics. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 748–757). Piscataway, NJ: IEEE. Retrieved 2026-08-15, from https://ieeexplore.ieee.org/abstract/document/9995335 (ISSN: 1554-7868) doi: 10.1109/ISMAR55827.2022.00093

Beall, A. C., Loomis, J. M., Philbeck, J. W., & Fikes, T. G. (1995, April). Absolute motion parallax weakly determines visual scale in real and virtual environments. In *Human Vision, Visual Processing, and Digital Display VI* (Vol. 2411, pp. 288–297). SPIE. Retrieved 2026-02-25, from https://www.spiedigitallibrary.org/conference-proceedings-of-spie/2411/0000/Absolute-motion-parallax-weakly-determines-visual-scale-in-real-and/10.1117/12.207547.full doi: 10.1117/12.207547

Beaudouin-Lafon, M. (2000, April). Instrumental interaction: an interaction model for designing post-WIMP user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 446–453). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/332040.332473 doi: 10.1145/332040.332473

Benedek, J., & Miner, T. (2002). Measuring Desirability: New methods for evaluating desirability in a usability lab setting. *Proceedings of Usability Professionals Association*, *2003*(8-12), 57.

Booth, J. W., Taborda, E. A., Ramani, K., & Reid, T. (2016, March). Interventions for teaching sketching skills and reducing inhibition for novice engineering designers. *Design Studies*, *43*, 1–23. Retrieved 2026-02-26, from https://www.sciencedirect.com/science/article/pii/S0142694X15000952 doi: 10.1016/j.destud.2015.11.002

Bourdeau, S., Lesage, A., Couturier Caron, B., & Léger, P.-M. (2020, April). When Design Novices and LEGO® Meet: Stimulating Creative Thinking for Interface Design. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/3313831.3376495 doi: 10.1145/3313831.3376495

Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2004). *3D User Interfaces: Theory and Practice*. USA: Addison Wesley Longman Publishing Co., Inc.

Bowman, D. A., North, C., Chen, J., Polys, N. F., Pyla, P. S., & Yilmaz, U. (2003, October). Information-rich virtual environments: theory, tools, and research agenda. In *Proceedings of the ACM symposium on Virtual reality software and technology* (pp. 81–90). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/1008653.1008669 doi: 10.1145/1008653.1008669

Bozgeyikli, E., & Bozgeyikli, L. L. (2021, March). Evaluating Object Manipulation Interaction Techniques in Mixed Reality: Tangible User Interfaces and Gesture. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)* (pp. 778–787). Retrieved 2026-02-25, from https://ieeexplore.ieee.org/document/9417663 (ISSN: 2642-5254) doi: 10.1109/VR50410.2021.00105

Bozgeyikli, E., Raij, A., Katkoori, S., & Dubey, R. (2016, October). Point & Teleport Locomotion Technique for Virtual Reality. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (pp. 205–216). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/2967934.2968105 doi: 10.1145/2967934.2968105

Braun, V., & Clarke, V. (2012). Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). Washington, DC, US: American Psychological Association. doi: 10.1037/13620-004

Braun, V., & Clarke, V. (2019, August). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597.

Retrieved 2026-02-26, from https://doi.org/10.1080/2159676X.2019.1628806 (_eprint: https://doi.org/10.1080/2159676X.2019.1628806) doi: 10.1080/2159676X.2019.1628806

Brereton, M., & McGarry, B. (2000, April). An observational study of how objects support engineering design thinking and communication: implications for the design of tangible media. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 217–224). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/332040.332434 doi: 10.1145/332040.332434

Burns, E., Razzaque, S., Panter, A., Whitton, M., McCallus, M., & Brooks, F. (2005, March). The hand is slower than the eye: a quantitative exploration of visual dominance over proprioception. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* (pp. 3–10). Retrieved 2026-02-25, from https://ieeexplore.ieee.org/document/1492747 (ISSN: 2375-5334) doi: 10.1109/VR.2005.1492747

Cheng, L.-P., Ofek, E., Holz, C., Benko, H., & Wilson, A. D. (2017, May). Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3718–3728). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/3025453.3025753 doi: 10.1145/3025453.3025753

Choi, I., Culbertson, H., Miller, M. R., Olwal, A., & Follmer, S. (2017, October). Grabity: A Wearable Haptic Interface for Simulating Weight and Grasping in Virtual Reality. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (pp. 119–130). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/3126594.3126599 doi: 10.1145/3126594.3126599

Clarence, A., https://orcid.org/0000-0001-9704-228X, View Profile, Knibbe, J., https://orcid.org/0000-0002-8844-8576, View Profile, ... View Profile (2024, May). Stacked Retargeting: Combining Redirected Walking and Hand Redirection to Expand Haptic Retargeting's Coverage. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). New York, NY: Association for Computing Machinery. Retrieved 2026-08-14, from https://dl.acm.org/doi/full/10.1145/3613904.3642228 doi: 10.1145/3613904.3642228

Cockburn, A., Karlson, A., & Bederson, B. B. (2009, January). A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv.*, *41*(1), 2:1–2:31. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/1456650.1456652 doi: 10.1145/1456650.1456652

Cohen, D. J., & Bennett, S. (1997). Why can't most people draw what they see? *Journal*

of Experimental Psychology: Human Perception and Performance, *23*(3), 609–621. doi: 10.1037/0096-1523.23.3.609

Combe, E., Posselt, J., & Kemeny, A. (2008). 1:1 scale perception in virtual and augmented reality. In *18th international conference on artificial reality and telexistence.* Retrieved 2026-02-25, from https://www.semanticscholar.org/paper/1%3A1-Scale-Perception-in-Virtual-and-Augmented-Combe-Posselt/ed20a7581c04899bb719fe3c62023f4a443cb43f

Creem-Regehr, S. H., Stefanucci, J. K., & Bodenheimer, B. (2022, December). Perceiving distance in virtual reality: theoretical insights from contemporary technologies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1869), 20210456. Retrieved 2026-02-26, from https://doi.org/10.1098/rstb.2021.0456 doi: 10.1098/rstb.2021.0456

Cuijpers, R. H., Brenner, E., & Smeets, J. B. J. (2008, December). Consistent haptic feedback is required but it is not enough for natural reaching to virtual cylinders. *Human Movement Science*, *27*(6), 857–872. Retrieved 2026-02-26, from https://www.sciencedirect.com/science/article/pii/S0167945708000626 doi: 10.1016/j.humov.2008.07.003

Cummings, J. J., & Bailenson, J. N. (2016, April). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology*, *19*(2), 272–309. Retrieved 2026-03-10, from https://doi.org/10.1080/15213269.2015.1015740 (_eprint: https://doi.org/10.1080/15213269.2015.1015740) doi: 10.1080/15213269.2015.1015740

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of space and motion* (pp. 69–117). San Diego, CA, US: Academic Press. doi: 10.1016/B978-012240530-3/50005-5

Danyluk, K., Ens, B., Jenny, B., & Willett, W. (2021, May). A Design Space Exploration of Worlds in Miniature. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/3411764.3445098 doi: 10.1145/3411764.3445098

Doma, O. O., & Sener, S. M. (2022, June). Dreamscape Bricks VR: An Experimental Virtual Reality Tool for Architectural Design. *Interaction Design and Architecture(s)*(52), 234–258. Retrieved from https://ixdea.org/52_13/ doi: 10.55612/s-5002-052-013

Duchowski, A. T. (2002, November). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 455–470. Retrieved 2026-02-26, from https://doi.org/10.3758/BF03195475 doi: 10.3758/

BF03195475

Ens, B., Hincapié-Ramos, J. D., & Irani, P. (2014, October). Ethereal planes: a design framework for 2D information space in 3D mixed reality environments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction* (pp. 2–12). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-05-14, from https://dl.acm.org/doi/10.1145/2659766.2659769 doi: 10.1145/2659766.2659769

Ernst, M. O., & Banks, M. S. (2002, January). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. Retrieved 2026-02-26, from https://www.nature.com/articles/415429a doi: 10.1038/415429a

Fang, C. M., & Harrison, C. (2021, October). Retargeted Self-Haptics for Increased Immersion in VR without Instrumentation. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (pp. 1109–1121). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3472749.3474810

Fechner, G. T. (1860). *Elemente der psychophysik* (Vol. 1). Leipzig: Breitkopf und Härtel.

Feick, M., Bateman, S., Tang, A., Miede, A., & Marquardt, N. (2020, November). Tangi: Tangible Proxies For Embodied Object Exploration And Manipulation In Virtual Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 195–206). IEEE Computer Society. Retrieved 2026-02-25, from https://www.computer.org/csdl/proceedings-article/ismar/2020/850800a195/1pystVP5LFK doi: 10.1109/ISMAR50242.2020.00042

Fitzmaurice, G. W., Ishii, H., & Buxton, W. A. S. (1995, May). Bricks: laying the foundations for graspable user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 442–449). USA: ACM Press/Addison-Wesley Publishing Co. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/223904.223964 doi: 10.1145/223904.223964

Gagnier, K. M., Atit, K., Ormand, C. J., & Shipley, T. F. (2017). Comprehending 3D Diagrams: Sketching to Support Spatial Reasoning. *Topics in Cognitive Science*, *9*(4), 883–901. Retrieved 2026-02-26, from https://onlinelibrary.wiley.com/doi/abs/10.1111/tops.12233 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/tops.12233) doi: 10.1111/tops.12233

Gogel, W. C., & Da Silva, J. A. (1987, July). Familiar size and the theory of off-sized perceptions. *Perception & Psychophysics*, *41*(4), 318–328. Retrieved 2026-02-26, from https://doi.org/10.3758/BF03208233 doi: 10.3758/BF03208233

Gravity Sketch | 3D sketching and design software. (n.d.). Retrieved 2023-10-25, from https://www.gravitysketch.com/

Grechkin, T. Y., Chihak, B. J., Cremer, J. F., Kearney, J. K., & Plumert, J. M. (2013). Perceiving and acting on complex affordances: How children and adults bicycle across

two lanes of opposing traffic. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 23–36. doi: 10.1037/a0029716

Han, D., Kim, D., & Cho, I. (2022, November). PORTAL: Portal Widget for Remote Target Acquisition and Control in Immersive Virtual Environments. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology* (pp. 1–11). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-08-14, from https://dl.acm.org/doi/10.1145/3562939.3565639 doi: 10.1145/3562939.3565639

Harley, D., Tarun, A. P., Germinario, D., & Mazalek, A. (2017, June). Tangible VR: Diegetic Tangible Objects for Virtual Reality Narratives. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (pp. 1253–1263). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/3064663.3064680 doi: 10.1145/3064663.3064680

Harrison, B. L., Ishii, H., Vicente, K. J., & Buxton, W. A. S. (1995, May). Transparent layered user interfaces: an evaluation of a display design to enhance focused and divided attention. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 317–324). USA: ACM Press/Addison-Wesley Publishing Co. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/223904.223945 doi: 10.1145/223904.223945

Hettiarachchi, A., & Wigdor, D. (2016, May). Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1957–1967). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/2858036.2858134 doi: 10.1145/2858036.2858134

Horiuchi, Y., Inoue, T., & Okada, K.-i. (2012, February). Virtual stage linked with a physical miniature stage to support multiple users in planning theatrical productions. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces* (pp. 109–118). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/2166966.2166989 doi: 10.1145/2166966.2166989

Hornsey, R. L., Hibbard, P. B., & Scarfe, P. (2020, August). Size and shape constancy in consumer virtual reality. *Behavior Research Methods*, *52*(4), 1587–1598. Retrieved 2026-02-26, from https://doi.org/10.3758/s13428-019-01336-9 doi: 10.3758/s13428-019-01336-9

Hosokawa, T., Takeda, Y., Shioiri, N., Hirano, M., & Tanaka, K. (2008, February). Tangible design support system using RFID technology. In *Proceedings of the 2nd international conference on Tangible and embedded interaction* (pp. 75–78). New York, NY, USA:

Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/1347390.1347408 doi: 10.1145/1347390.1347408

Hou, Y., Cui, H., Chen, R., & Zeng, W. (2024, May). Understanding the Impact of Referent Design on Scale Perception in Immersive Data Visualization. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1–7). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-27, from https://dl.acm.org/doi/10.1145/3613905.3650783 doi: 10.1145/3613905.3650783

Hsu, T.-W., Tsai, M.-H., Babu, S. V., Hsu, P.-H., Chang, H.-M., Lin, W.-C., & Chuang, J.-H. (2020, March). Design and Initial Evaluation of a VR based Immersive and Interactive Architectural Design Discussion System. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 363–371). IEEE. Retrieved 2026-02-27, from https://ieeexplore.ieee.org/document/9089521 (ISSN: 2642-5254) doi: 10.1109/VR46266.2020.00056

Huk, T. (2006). Who benefits from learning with 3D models? the case of spatial ability. *Journal of Computer Assisted Learning*, *22*(6), 392–404. Retrieved 2026-02-26, from https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2729.2006.00180.x (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2729.2006.00180.x) doi: 10.1111/j.1365-2729.2006.00180.x

Ishii, H. (2008a, February). Tangible bits: beyond pixels. In *Proceedings of the 2nd international conference on Tangible and embedded interaction* (pp. xv–xxv). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/1347390.1347392 doi: 10.1145/1347390.1347392

Ishii, H. (2008b, June). The tangible user interface and its evolution. *Commun. ACM*, *51*(6), 32–36. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/1349026.1349034 doi: 10.1145/1349026.1349034

Ishii, H., & Ullmer, B. (1997, March). Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (pp. 234–241). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/258549.258715 doi: 10.1145/258549.258715

Jackson, S. A., & Marsh, H. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of Sport & Exercise Psychology*, *18*(1), 17–35. doi: 10.1123/jsep.18.1.17

Jacob, R. J., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008, April). Reality-based interaction: a framework for post-WIMP interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 201–210). New York, NY, USA: Association for Computing Machinery.

Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/1357054.1357089 doi: 10.1145/1357054.1357089

Jennings, N., & Matviienko, A. (2025, December). Reimagining Go-Go: Effects of Signifiers in VR Selection Techniques. In *Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology* (pp. 1–10). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-08-14, from https://dl.acm.org/doi/10.1145/3756884.3766014 doi: 10.1145/3756884.3766014

John, N. W. (2002). Using stereoscopy for medical virtual reality. *Studies in Health Technology and Informatics*, *85*, 214–220.

Ju, W., & Leifer, L. (2008). The Design of Implicit Interactions: Making Interactive Systems Less Obnoxious. *Design Issues*, *24*(3), 72–84. Retrieved 2026-02-26, from https://www.jstor.org/stable/25224184

Kakehi, Y., Yamaoka, J., Akatsuka, D., & Naemura, T. (2009, August). Tablescape animation: a support system for making animations using tabletop physical objects. In *SIGGRAPH 2009: Talks* (p. 1). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/1597990.1598060 doi: 10.1145/1597990.1598060

Keehner, M., Hegarty, M., Cohen, C., Khooshabeh, P., & Montello, D. R. (2008). Spatial Reasoning With External Visualizations: What Matters Is What You See, Not Whether You Interact. *Cognitive Science*, *32*(7), 1099–1132. Retrieved 2026-02-26, from https://onlinelibrary.wiley.com/doi/abs/10.1080/03640210801898177 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1080/03640210801898177) doi: 10.1080/03640210801898177

Kelly, J. W., Cherep, L. A., & Siegel, Z. D. (2017, July). Perceived Space in the HTC Vive. *ACM Trans. Appl. Percept.*, *15*(1), 2:1–2:16. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3106155 doi: 10.1145/3106155

Kelly, J. W., Donaldson, L. S., Sjolund, L. A., & Freiberg, J. B. (2013, October). More than just perception–action recalibration: Walking through a virtual environment causes rescaling of perceived space. *Attention, Perception, & Psychophysics*, *75*(7), 1473–1485. Retrieved 2026-02-26, from https://doi.org/10.3758/s13414-013-0503-4 doi: 10.3758/s13414-013-0503-4

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993, July). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220. Retrieved 2026-02-26, from https://doi.org/10.1207/s15327108ijap0303_3 (_eprint: https://doi.org/10.1207/s15327108ijap0303_3) doi: 10.1207/s15327108ijap0303_3

Kim, H.-J., Kim, J.-W., & Nam, T.-J. (2016, May). miniStudio: Designers' Tool for Prototyping Ubicomp Space with Interactive Miniature. In *Proceedings of the 2016 CHI*

Conference on Human Factors in Computing Systems (pp. 213–224). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-25, from https://dl.acm.org/doi/10.1145/2858036.2858180 doi: 10.1145/2858036.2858180

Kim, J., & Hancock, M. (2025, May). Investigating Digital, Tangible, and Paper-Based Room Design at a Small Scale. In *Graphics Interface 2025 (GI '25)* (Vol. 1). Canadian Human-Computer Communications Society (CHCCS). Retrieved from https://doi.org/10.1145/3769872.3769879 doi: 10.1145/3769872.3769879

Kiyokawa, K., Takemura, H., & Yokoya, N. (2000). SeamlessDesign for 3D Object Creation. *IEEE MultiMedia*, *7*(1), 22–33. Retrieved 2026-02-26, from https://doi.org/10.1109/93.839308 doi: 10.1109/93.839308

Korakakis, G., Pavlatou, E. A., Palyvos, J. A., & Spyrellis, N. (2009, February). 3D visualization types in multimedia applications for science learning: A case study for 8th grade students in Greece. *Computers & Education*, *52*(2), 390–401. Retrieved 2026-02-26, from https://www.sciencedirect.com/science/article/pii/S0360131508001449 doi: 10.1016/j.compedu.2008.09.011

Kudrowitz, B., Te, P., & Wallace, D. (2012, August). The influence of sketch quality on perception of product-idea creativity. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, *26*(3), 267–279. Retrieved 2026-02-26, from https://doi.org/10.1017/S0890060412000145 doi: 10.1017/S0890060412000145

LaViola, J. J., Feliz, D. A., Keefe, D. F., & Zeleznik, R. C. (2001, March). Hands-free multi-scale navigation in virtual environments. In *Proceedings of the 2001 symposium on Interactive 3D graphics* (pp. 9–15). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-27, from https://dl.acm.org/doi/10.1145/364338.364339 doi: 10.1145/364338.364339

Lee-Cultura, S., & Giannakos, M. (2020, July). Embodied Interaction and Spatial Skills: A Systematic Review of Empirical Studies. *Interacting with Computers*, *32*(4), 331–366. Retrieved 2026-02-26, from https://doi.org/10.1093/iwcomp/iwaa023 doi: 10.1093/iwcomp/iwaa023

Liu, X., xl5040@nyu.edu, https://orcid.org/0009-0005-1294-7108, View Profile, Jia, D., bobby003@umn.edu, ... View Profile (2025, September). Reality Proxy: Fluid Interactions with Real-World Objects in MR via Abstract Representations. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology* (pp. 1–16). New York, NY: Association for Computing Machinery. Retrieved 2026-08-14, from https://dl.acm.org/doi/full/10.1145/3746059.3747709 doi: 10.1145/3746059.3747709

Liu, Y., yangliu.hci@gmail.com, https://orcid.org/0009-0005-9839-0864, View Profile, Mikkelsen, T., thormik@cs.au.dk, ... View Profile (2025, September). At a Glance to Your Fingertips: Enabling Direct Manipulation of Distant Objects Through

SightWarp. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology* (pp. 1–12). New York, NY: Association for Computing Machinery. Retrieved 2026-08-14, from https://dl.acm.org/doi/full/10.1145/3746059.3747653 doi: 10.1145/3746059.3747653

Liu, Z., Wu, J., Wang, L., Li, X., & Im, S. K. (2025, July). Proxy Importance Based Haptic Retargeting With Multiple Props in VR. *IEEE Transactions on Visualization and Computer Graphics*, *31*(7), 3987–4002. Retrieved 2026-02-26, from https://ieeexplore-ieee-org.proxy.lib.uwaterloo.ca/document/10507021 doi: 10.1109/TVCG.2024.3392743

London 21034 | Architecture | Buy online at the Official LEGO® Shop CA. (2026). Retrieved 2026-05-29, from https://www.lego.com/en-ca/product/london-21034

Loomis, J., & Knapp, J. (2003). Visual perception of egocentric distance in real and virtual environments. In *Virtual and adaptive environments: Applications, implications, and human performance issues* (pp. 21–46). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. doi: 10.1201/9781410608888.pt1

Luo, E.-X., Tang, K. Y., Xu, S.-Z., Tong, Q., & Zhang, S.-H. (2024). Walking Telescope: Exploring the Zooming Effect in Expanding Detection Threshold Range for Translation Gain. In F.-L. Zhang & A. Sharf (Eds.), *Computational Visual Media* (pp. 252–273). Singapore: Springer Nature. doi: 10.1007/978-981-97-2095-8_14

Maher, M. L., & Kim, M. J. (2005). Do Tangible User Interfaces Impact Spatial Cognition in Collaborative Design? In Y. Luo (Ed.), *Cooperative Design, Visualization, and Engineering* (pp. 30–41). Berlin, Heidelberg: Springer. doi: 10.1007/11555223_4

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016, November). Sample Size in Qualitative Interview Studies: Guided by Information Power. *Qualitative Health Research*, *26*(13), 1753–1760. Retrieved 2026-02-26, from https://doi.org/10.1177/1049732315617444 doi: 10.1177/1049732315617444

Marino, B. F. M., Stucchi, N., Nava, E., Haggard, P., & Maravita, A. (2010, April). Distorting the visual size of the hand affects hand pre-shaping during grasping. *Experimental Brain Research*, *202*(2), 499–505. Retrieved 2026-02-26, from https://doi.org/10.1007/s00221-009-2143-4 doi: 10.1007/s00221-009-2143-4

Marougkas, A., Troussas, C., Krouska, A., & Sgouropoulou, C. (2024, February). How personalized and effective is immersive virtual reality in education? A systematic literature review for the last decade. *Multimedia Tools and Applications*, *83*(6), 18185–18233. Retrieved 2026-02-27, from https://doi.org/10.1007/s11042-023-15986-7 doi: 10.1007/s11042-023-15986-7

Matthews, B. J., Reichherzer, C., & Smith, R. T. (2023, April). Remapped Interfaces: Building Contextually Adaptive User Interfaces with Haptic Retargeting. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (pp.

1–4). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3544549.3583912 doi: 10.1145/3544549.3583912

Matthews, B. J., Reichherzer, C., Thomas, B. H., & Smith, R. T. (2023, October). MaskWarp: Visuo-Haptic Illusions in Mixed Reality using Real-time Video Inpainting. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (pp. 767–768). Piscataway, NJ: IEEE. Retrieved 2026-08-15, from https://ieeexplore.ieee.org/abstract/document/10322261 (ISSN: 2771-1110) doi: 10.1109/ISMAR-Adjunct60411.2023.00167

Matthews, B. J., Thomas, B. H., Von Itzstein, G. S., & Smith, R. T. (2022, March). Shape Aware Haptic Retargeting for Accurate Hand Interactions. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 625–634). Piscataway, NJ: IEEE. Retrieved 2026-08-18, from https://ieeexplore.ieee.org/document/9756806 (ISSN: 2642-5254) doi: 10.1109/VR51125.2022.00083

McGee, D. R., Cohen, P. R., Wesson, R. M., & Horman, S. (2002, April). Comparing paper and tangible, multimodal tools. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 407–414). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/503376.503449 doi: 10.1145/503376.503449

Medeiros, D., Wilson, G., Mcgill, M., & Brewster, S. A. (2023, April). The Benefits of Passive Haptics and Perceptual Manipulation for Extended Reality Interactions in Constrained Passenger Spaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–19). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3544548.3581079 doi: 10.1145/3544548.3581079

Medium by adobe. (n.d.). Retrieved 2023-10-25, from https://www.adobe.com/ca/products/medium.html

Microsoft, & Royal Institute of British Architects. (2018). *Digital Transformation in Architecture* (Tech. Rep.). London, UK: Microsoft UK / Royal Institute of British Architects.

Milgram, P., & Kishino, F. (1994, December). A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Information Systems*, *vol. E77-D, no. 12*, 1321–1329.

Mine, M. R., Brooks, F. P., & Sequin, C. H. (1997, August). Moving objects in space: exploiting proprioception in virtual-environment interaction. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (pp. 19–26). USA: ACM Press/Addison-Wesley Publishing Co. Retrieved 2026-06-19, from https://dl.acm.org/doi/10.1145/258734.258747 doi: 10.1145/258734.258747

Morris, M. (2006). *Models: architecture and the miniature.* Chichester, West Sus-

sex : Wiley-Academy. Retrieved 2026-02-25, from http://archive.org/details/modelsarchitectu0000morr

Muender, T., Reinschluessel, A. V., Drewes, S., Wenig, D., Döring, T., & Malaka, R. (2019, May). Does It Feel Real? Using Tangibles with Different Fidelities to Build and Explore Scenes in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3290605.3300903 doi: 10.1145/3290605.3300903

Nakanishi, Y. (2012, June). Virtual prototyping using miniature model and visualization for interactive public displays. In *Proceedings of the Designing Interactive Systems Conference* (pp. 458–467). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/2317956.2318024 doi: 10.1145/2317956.2318024

Noghabaei, M., Heydarian, A., Balali, V., & Han, K. (2020, March). Trend Analysis on Adoption of Virtual and Augmented Reality in the Architecture, Engineering, and Construction Industry. *Data*, *5*(1), 26. Retrieved 2026-02-26, from https://www.mdpi.com/2306-5729/5/1/26 doi: 10.3390/data5010026

Norman, J. F., Norman, H. F., Clayton, A. M., Lianekhammy, J., & Zielke, G. (2004, February). The visual and haptic perception of natural object shape. *Perception & Psychophysics*, *66*(2), 342–351. Retrieved 2026-02-26, from https://doi.org/10.3758/BF03194883 doi: 10.3758/BF03194883

Ogawa, N., Narumi, T., & Hirose, M. (2018, March). Object Size Perception in Immersive Virtual Reality: Avatar Realism Affects the Way We Perceive. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 647–648). Retrieved 2026-02-26, from https://ieeexplore.ieee.org/document/8446318 doi: 10.1109/VR.2018.8446318

Oyelere, S. S., Bouali, N., Kaliisa, R., Obaido, G., Yunusa, A. A., & Jimoh, E. R. (2020, October). Exploring the trends of educational virtual reality games: a systematic review of empirical studies. *Smart Learning Environments*, *7*(1), 31. Retrieved 2026-02-27, from https://doi.org/10.1186/s40561-020-00142-7 doi: 10.1186/s40561-020-00142-7

Pausch, R., Burnette, T., Brockway, D., & Weiblen, M. E. (1995, September). Navigation and locomotion in virtual worlds via flight into hand-held miniatures. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques* (pp. 399–400). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-27, from https://dl.acm.org/doi/10.1145/218380.218495 doi: 10.1145/218380.218495

Perlin, K., & Fox, D. (1993, September). Pad: an alternative approach to the computer

interface. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques* (pp. 57–64). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/166117.166125 doi: 10.1145/166117.166125

Piumsomboon, T., Lee, G. A., & Billinghurst, M. (2018, April). Snow Dome: A Multi-Scale Interaction in Mixed Reality Remote Collaboration. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1–4). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3170427.3186495 doi: 10.1145/3170427.3186495

Pivovar, J., DeGuzman, J., & Rosenberg, E. S. (2022, March). Virtual Reality on a SWIM: Scalable World in Miniature. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 912–913). Piscataway, NJ: IEEE. Retrieved 2026-08-14, from https://ieeexplore.ieee.org/abstract/document/9757678 doi: 10.1109/VRW55335.2022.00307

Plumert, J. M., & Kearney, J. K. (2014). How Do Children Perceive and Act on Dynamic Affordances in Crossing Traffic-Filled Roads? *Child Development Perspectives*, *8*(4), 207–212. Retrieved 2026-02-26, from https://onlinelibrary.wiley.com/doi/abs/10.1111/cdep.12089 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cdep.12089) doi: 10.1111/cdep.12089

Pohl, H., Lilija, K., McIntosh, J., & Hornbæk, K. (2021, May). Poros: Configurable Proxies for Distant Interactions in VR. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3411764.3445685 doi: 10.1145/3411764.3445685

Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996, November). The go-go interaction technique: non-linear mapping for direct manipulation in VR. In *Proceedings of the 9th annual ACM symposium on User interface software and technology* (pp. 79–80). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/237091.237102 doi: 10.1145/237091.237102

Racz, A., & Zilizi, G. (2018, June). VR Aided Architecture and Interior Design. In *2018 International Conference on Advances in Computing and Communication Engineering (ICACCE)* (pp. 11–16). Retrieved 2026-02-26, from https://ieeexplore.ieee.org/document/8441714 doi: 10.1109/ICACCE.2018.8441714

Rauterberg, M., Fjeld, M., Krueger, H., Bichsel, M., Leonhardt, U., & Meier, M. (1998, April). BUILD-IT: a planning tool for construction and design. In *CHI 98 Conference Summary on Human Factors in Computing Systems* (pp. 177–178). New York, NY,

USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/286498.286657 doi: 10.1145/286498.286657

Razzaque, S., Swapp, D., Slater, M., Whitton, M. C., & Steed, A. (2002, May). Redirected walking in place. In *Proceedings of the workshop on Virtual environments 2002* (pp. 123–130). Goslar, DEU: Eurographics Association.

Renner, R. S., Velichkovsky, B. M., & Helmert, J. R. (2013, December). The perception of egocentric distances in virtual environments - A review. *ACM Comput. Surv.*, *46*(2), 23:1–23:40. Retrieved 2026-03-10, from https://dl.acm.org/doi/10.1145/2543581.2543590 doi: 10.1145/2543581.2543590

Rosahl, S. K., Gharabaghi, A., Hubbe, U., Shahidi, R., & Samii, M. (2006, May). Virtual Reality Augmentation in Skull Base Surgery. *Skull Base*, *16*(2), 59–66. Retrieved 2026-02-26, from https://pmc.ncbi.nlm.nih.gov/articles/PMC1502039/ doi: 10.1055/s-2006-931620

Salim, F., Jaworski, P., Kaftan, M., Friedrich, E., Urquiza, R., Oh, S., ... Østergaard, J. (2011). Informing architecture and urban modeling with real-world data on 3D tangible interfaces and augmented displays. In *Integration Through Computation - Proceedings of the 31st Annual Conference of the Association for Computer Aided Design in Architecture* (pp. 226–233).

Schjerlund, J., Hornbæk, K., & Bergström, J. (2022, April). OVRlap: Perceiving Multiple Locations Simultaneously to Improve Interaction in VR. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3491102.3501873 doi: 10.1145/3491102.3501873

Schubert, G., Riedel, S., & Petzold, F. (2013). Seamfully Connected: Real Working Models as Tangible Interfaces for Architectural Design. In J. Zhang & C. Sun (Eds.), *Global Design and Local Materialization* (pp. 210–221). Berlin, Heidelberg: Springer. doi: 10.1007/978-3-642-38974-0_20

Shi, H., hshi22@m.fudan.edu.cn, https://orcid.org/0009-0007-6876-5922, View Profile, Luo, H., luohz23@mails.tsinghua.edu.cn, ... View Profile (2025, April). ReachPad: Interacting with Multiple Virtual Screens using a Single Physical Pad through Haptic Retargeting. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–17). New York, NY: Association for Computing Machinery. Retrieved 2026-08-14, from https://dl.acm.org/doi/full/10.1145/3706598.3713629 doi: 10.1145/3706598.3713629

Shi, R., Zhu, N., Liang, H.-N., & Zhao, S. (2021, October). Exploring Head-based Mode-Switching in Virtual Reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 118–127). Retrieved 2026-02-26,

from https://ieeexplore.ieee.org/document/9583741 (ISSN: 1554-7868) doi: 10.1109/ISMAR52148.2021.00026

Shin, Y., https://orcid.org/0000-0002-5099-2779, View Profile, Kim, G. J., https://orcid.org/0000-0001-9880-8021, & View Profile. (2022, November). Leveraging VR Techniques for Efficient Exploration and Interaction in Large and Complex AR Space with Clipped and Small FOV AR Display. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology* (pp. 1–10). New York, NY: Association for Computing Machinery. Retrieved 2026-08-14, from https://dl.acm.org/doi/abs/10.1145/3562939.3565613 doi: 10.1145/3562939.3565613

SketchUp VR viewer. (n.d.). Retrieved 2023-10-25, from https://help.sketchup.com/en/sketchup-viewer/sketchup-viewer-vr

Slater, M. (2009, December). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557. Retrieved 2026-03-10, from https://pmc.ncbi.nlm.nih.gov/articles/PMC2781884/ doi: 10.1098/rstb.2009.0138

Someren, M., Barnard, Y., & Sandberg, J. (1994). The think aloud method: a practical approach to modelling cognitive processes. *Knowledge Based Systems*. Retrieved 2026-02-26, from https://www.semanticscholar.org/paper/The-think-aloud-method%3A-a-practical-approach-to-Someren-Barnard/8b2d1abb3bbd735d616e480c102e793a3d6ab965

Steinicke, F., Bruder, G., Jerald, J., Frenz, H., & Lappe, M. (2010, January). Estimation of Detection Thresholds for Redirected Walking Techniques. *IEEE Transactions on Visualization and Computer Graphics*, *16*(1), 17–27. Retrieved 2026-05-26, from https://ieeexplore.ieee.org/abstract/document/5072212 doi: 10.1109/TVCG.2009.62

Stoakley, R., Conway, M. J., & Pausch, R. (1995, May). Virtual reality on a WIM: interactive worlds in miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 265–272). USA: ACM Press/Addison-Wesley Publishing Co. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/223904.223938 doi: 10.1145/223904.223938

Suwa, M., & Tversky, B. (1997, October). What do architects and students perceive in their design sketches? A protocol analysis. *Design Studies*, *18*(4), 385–403. Retrieved 2026-02-26, from https://www.sciencedirect.com/science/article/pii/S0142694X97000082 doi: 10.1016/S0142-694X(97)00008-2

Swapp, D., Pawar, V., & Loscos, C. (2006, May). Interaction with co-located haptic feedback in virtual reality. *Virtual Reality*, *10*(1), 24–30. Retrieved 2026-02-26, from https://doi.org/10.1007/s10055-006-0027-5 doi: 10.1007/s10055-006-0027-5

Tanenbaum, T. J., Williams, A. M., Desjardins, A., & Tanenbaum, K. (2013, April). Democratizing technology: pleasure, utility and expressiveness in DIY and maker practice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2603–2612). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/2470654.2481360 doi: 10.1145/2470654.2481360

Teather, R. J., & Stuerzlinger, W. (2011, March). Pointing at 3D targets in a stereo head-tracked virtual environment. In *2011 IEEE Symposium on 3D User Interfaces (3DUI)* (pp. 87–94). Retrieved 2026-02-26, from https://ieeexplore-ieee-org.proxy.lib.uwaterloo.ca/document/5759222 doi: 10.1109/3DUI.2011.5759222

Thomas, B. H. (2020, January). Examining User Perception of the Size of Multiple Objects in Virtual Reality. *Applied Sciences*, *10*(11), 4049. Retrieved 2026-02-26, from https://www.mdpi.com/2076-3417/10/11/4049 doi: 10.3390/app10114049

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004, October). Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments? *Presence*, *13*(5), 560–571. Retrieved 2026-02-26, from https://ieeexplore.ieee.org/document/6788739 doi: 10.1162/1054746042545292

Ullmer, B., & Ishii, H. (2000). Emerging frameworks for tangible user interfaces. *IBM Systems Journal*, *39*(3.4), 915–931. Retrieved 2026-02-27, from https://ieeexplore.ieee.org/document/5387042 doi: 10.1147/sj.393.0915

Underkoffler, J., & Ishii, H. (1999, May). Urp: a luminous-tangible workbench for urban planning and design. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 386–393). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/302979.303114 doi: 10.1145/302979.303114

Vallerand, O. (2018). Architecture for the Masses: the CCA's The University Is Now On Air. *Canadian Architect*, 46–47.

Vardouli, T., & Buechley, L. (2014). Open Source Architecture: An Exploration of Source Code and Access in Architectural Design. *Leonardo*, *47*(1), 51–55. Retrieved 2026-02-26, from https://www.jstor.org/stable/43834115

Vogel, D., & Balakrishnan, R. (2004, October). Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (pp. 137–146). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/1029632.1029656 doi: 10.1145/1029632.1029656

Waller, D., & Richardson, A. R. (2008, March). Correcting Distance Estimates by In-

teracting With Immersive Virtual Environments: Effects of Task and Available Sensory Information. *Journal of Experimental Psychology. Applied*, *14*(1), 61–72. doi: 10.1037/1076-898X.14.1.61

Wang Baldonado, M. Q., Woodruff, A., & Kuchinsky, A. (2000, May). Guidelines for using multiple views in information visualization. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 110–119). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/345513.345271 doi: 10.1145/345513.345271

Ware, C., & Lewis, M. (1995, May). The DragMag image magnifier. In *Conference Companion on Human Factors in Computing Systems* (pp. 407–408). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/223355.223749 doi: 10.1145/223355.223749

The wild integration | autodesk construction cloud. (n.d.). Retrieved from https://construction.autodesk.com/workflows/construction-software-integrations/the-wild/

Wingrave, C., Haciahmetoglu, Y., & Bowman, D. (2006, March). Overcoming World in Miniature Limitations by a Scaled and Scrolling WIM. In *3D User Interfaces (3DUI'06)* (pp. 11–16). Retrieved 2026-02-26, from https://ieeexplore.ieee.org/document/1647500 doi: 10.1109/VR.2006.106

Witmer, B. G., & Kline, P. B. (1998, April). Judging Perceived and Traversed Distance in Virtual Environments. *Presence*, *7*(2), 144–167. Retrieved 2026-02-26, from https://ieeexplore.ieee.org/document/6787840 doi: 10.1162/105474698565640

Yang, M. C., & Cham, J. G. (2006, May). An Analysis of Sketching Skill and Its Role in Early Stage Engineering Design. *Journal of Mechanical Design*, *129*(5), 476–482. Retrieved 2026-02-26, from https://doi.org/10.1115/1.2712214 doi: 10.1115/1.2712214

Yang, X., https://orcid.org/0000-0001-6234-4587, View Profile, Zhang, Y., https://orcid.org/0000-0002-7549-4725, View Profile, ... View Profile (2023, October). Redirected Placement: Evaluating the Redirection of Passive Props during Reach-to-Place in Virtual Reality. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology* (pp. 1–11). New York, NY: Association for Computing Machinery. Retrieved 2026-08-14, from https://dl.acm.org/doi/abs/10.1145/3611659.3615688 doi: 10.1145/3611659.3615688

Yeh, Y.-Y., & Silverstein, L. D. (1992, October). Spatial Judgments with Monoscopic and Stereoscopic Presentation of Perspective Displays. *Human Factors*, *34*(5), 583–600. Retrieved 2026-02-26, from https://doi.org/10.1177/001872089203400506 doi:

10.1177/001872089203400506

Zenner, A., & Krüger, A. (2019, May). Drag:on: A Virtual Reality Controller Providing Haptic Feedback Based on Drag and Weight Shift. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3290605.3300441 doi: 10.1145/3290605.3300441

Zhai, S., & Milgram, P. (1998, January). Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 320–327). USA: ACM Press/Addison-Wesley Publishing Co. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/274644.274689 doi: 10.1145/274644.274689

Zheng, W., Xiong, D., Li, J., Jiang, J., Weng, C., Zhou, J., & Fan, M. (2024, December). Investigating Size Congruency Between the Visual Perception of a VR Object and the Haptic Perception of Its Physical World Agent. In *Proceedings of the 17th International Symposium on Visual Information Communication and Interaction* (pp. 1–9). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/3678698.3678706 doi: 10.1145/3678698.3678706

Zhu, F., Sidenmark, L., Sousa, M., & Grossman, T. (2023, October). PinchLens: Applying Spatial Magnification and Adaptive Control-Display Gain for Precise Selection in Virtual Reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (pp. 1221–1230). Piscataway, NJ: IEEE. Retrieved 2026-08-15, from https://ieeexplore.ieee.org/abstract/document/10316437 (ISSN: 2473-0726) doi: 10.1109/ISMAR59233.2023.00139

Zimmerman, J., Forlizzi, J., & Evenson, S. (2007, April). Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 493–502). New York, NY, USA: Association for Computing Machinery. Retrieved 2026-02-26, from https://dl.acm.org/doi/10.1145/1240624.1240704 doi: 10.1145/1240624.1240704

Šafhalter, A., Vukman, K. B., & Glodež, S. (2016, June). The Effect of 3D-Modeling Training on Students' Spatial Reasoning Relative to Gender and Grade. *Journal of Educational Computing Research*, *54*(3), 395–406. Retrieved 2026-02-26, from https://doi.org/10.1177/0735633115620430 doi: 10.1177/0735633115620430

APPENDICES

Appendix A

Technical Implementation Details

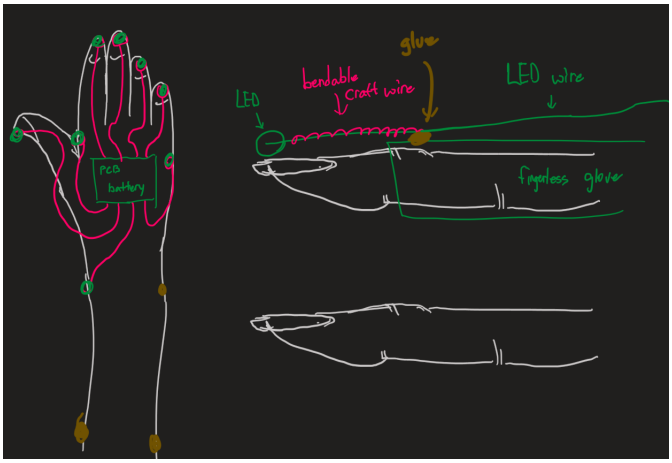
A.1 Developing a Finger-Tracking Glove

I needed to track the precise finger movements for studies in which participants interacted with a tangible proxy. The primary purpose of the tracking was to align the virtual hand representation with real hand movement in real time. While proprietary solutions for this purpose exist, I opted to make an in-house solution for cost-effectiveness and maintainability across studies.

A.1.1 Using the OptiTrack Algorithm

An early solution considered using OptiTrack's hand-tracking algorithm directly, leveraging asymmetry in passive markers across the finger, hand, and arm. However, frequent marker and orientation flipping due to passive tracking led to tracking instability, making this approach unreliable. Instead, I developed a simplified finger-tracking system using active markers, which are distinguishable from one another, enabling real-time tracking with submillimeter precision.

The early sketch (Figure A.1A) shows possible placements of active markers for OptiTrack's internal finger/hand tracking algorithm, and the quick test prototype (Figure A.1) shows how I tested potential wire routing for this purpose.



(A) Early sketch



(B) Prototyping look-and-feel

Figure A.1: Early sketch and prototype exploring marker placement and wire routing for the finger-tracking glove.

A.1.2 Developing a Simplified Solution

To accommodate scaling of hand representation in VR, control over the hand representation algorithm was needed. Rather than using the OptiTrack algorithm, I developed a simplified version using active markers in Unity3D.

Active markers, a controller printed circuit board, and a battery needed to be housed on the glove for the system to operate. A work glove was used to attach the PCB on the back and active markers near the fingertip of the glove (Figure A.2). While this approach allowed robust tracking performance, it was necessary to

1. accommodate different-sized hands across participants;
2. provide direct tactile feedback on the fingerpad.

Using half-fingered or fingerless gloves could address these issues, but attaching the active markers securely was problematic. To attach markers to the fingertip, nail rings were used. This approach minimized coverage of the fingerpad for tactile feedback and allowed small objects to be attached to the artificial nails.

In the later iteration of development, I used a half-fingered glove that attached a PCB and a battery directly on the back of the hand, with five nail rings to attach active markers

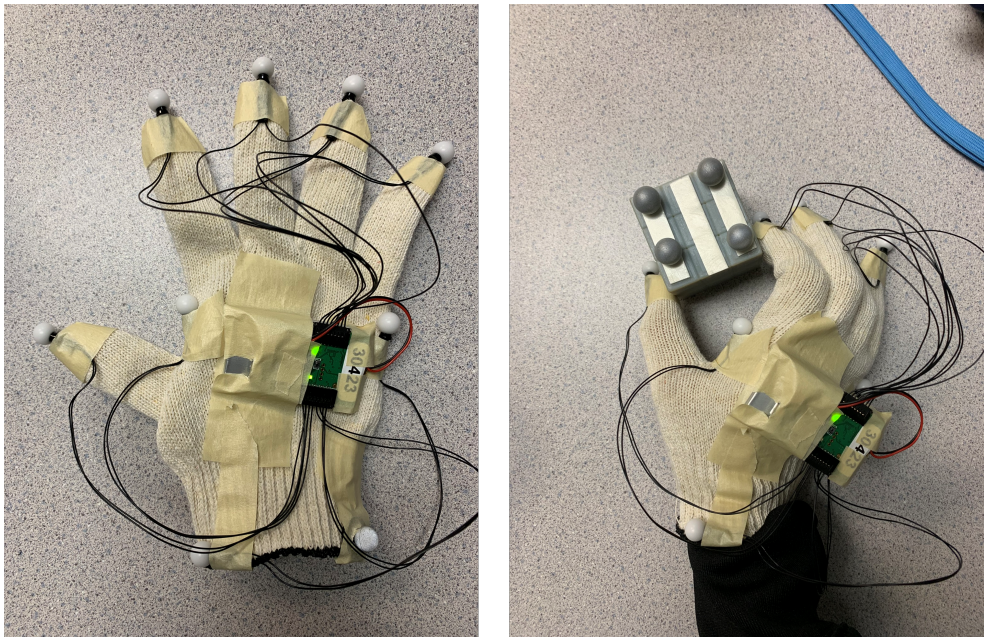


Figure A.2: Early working prototype glove for finger/hand tracking

(Figure A.3). The resulting setup provided tracked fingertip positions for the thumb and fingers used in the interaction tasks. These marker positions were streamed into Unity3D and mapped to the virtual hand representation used in the study. This allowed the virtual hand representation to be adjusted independently from the tangible proxy, which was necessary for testing scale and mismatch conditions.

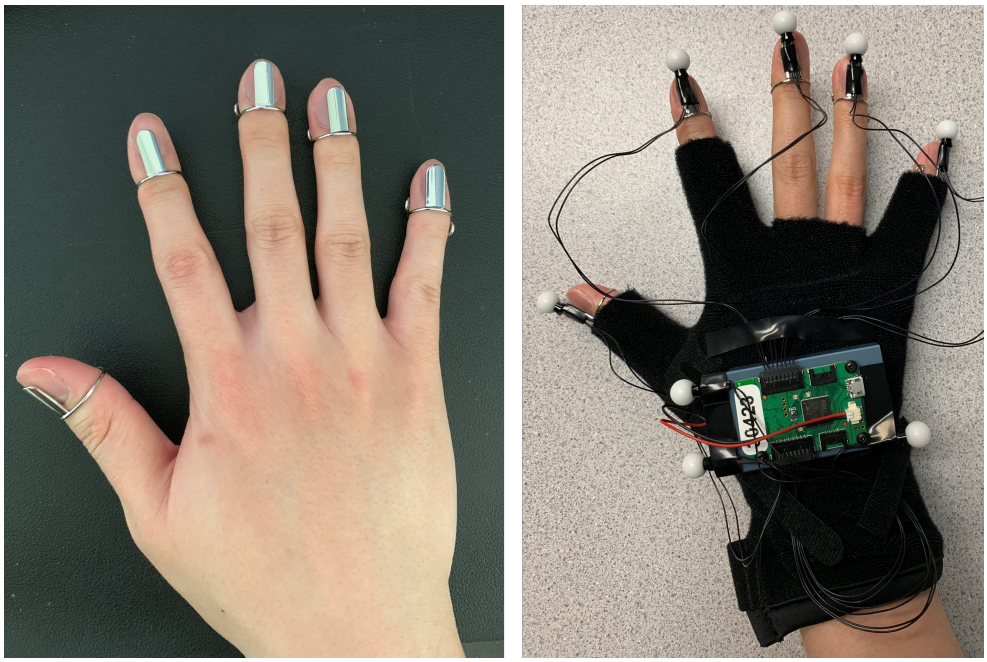


Figure A.3: Nail rings and prototypes using them.

Appendix B

Study Materials and Data

This chapter provides the study materials and raw data used acquired across the studies in this dissertation.

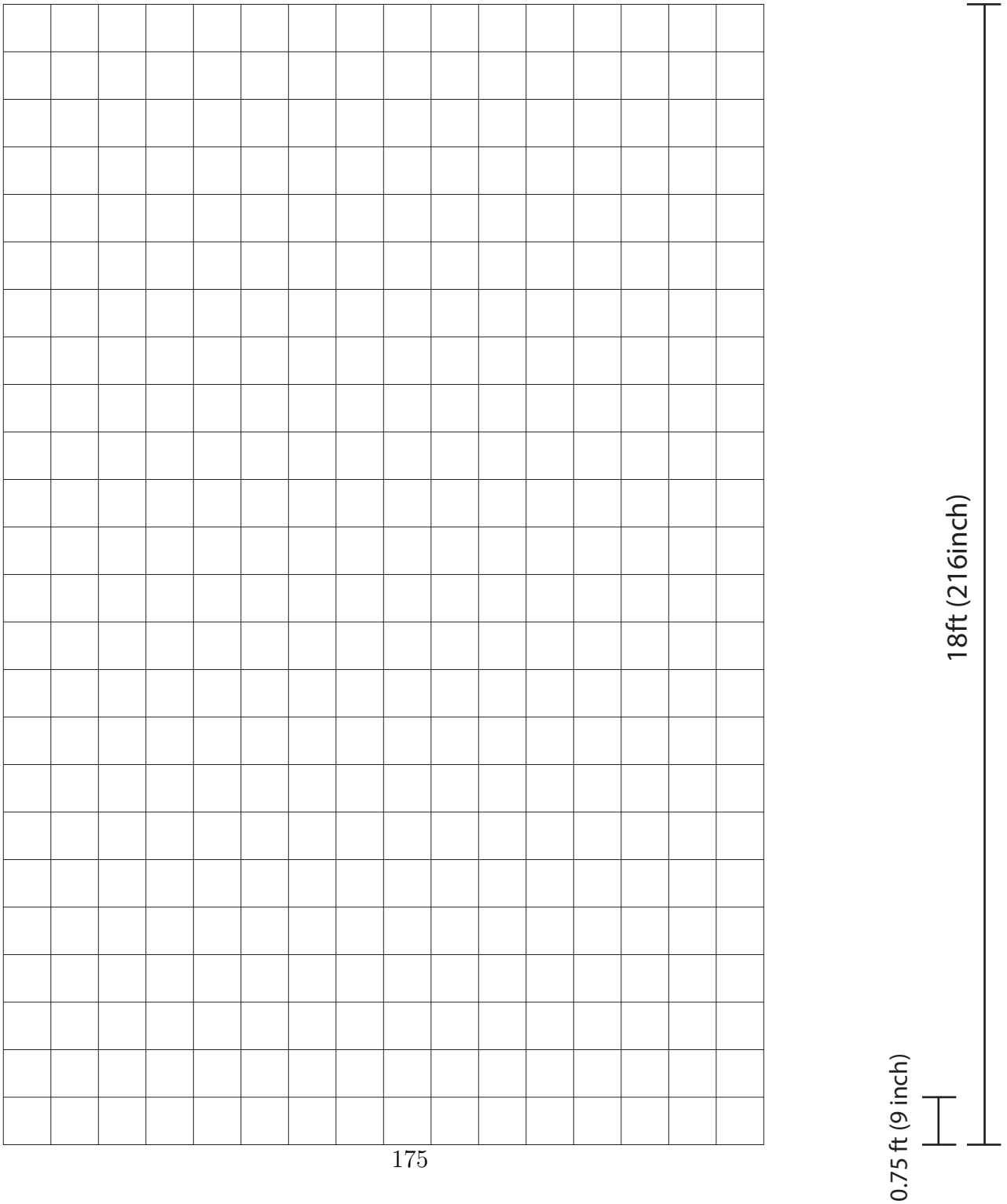
B.1 Materials and Data for Chapter 3

B.1.1 Print-outs Grid Papers

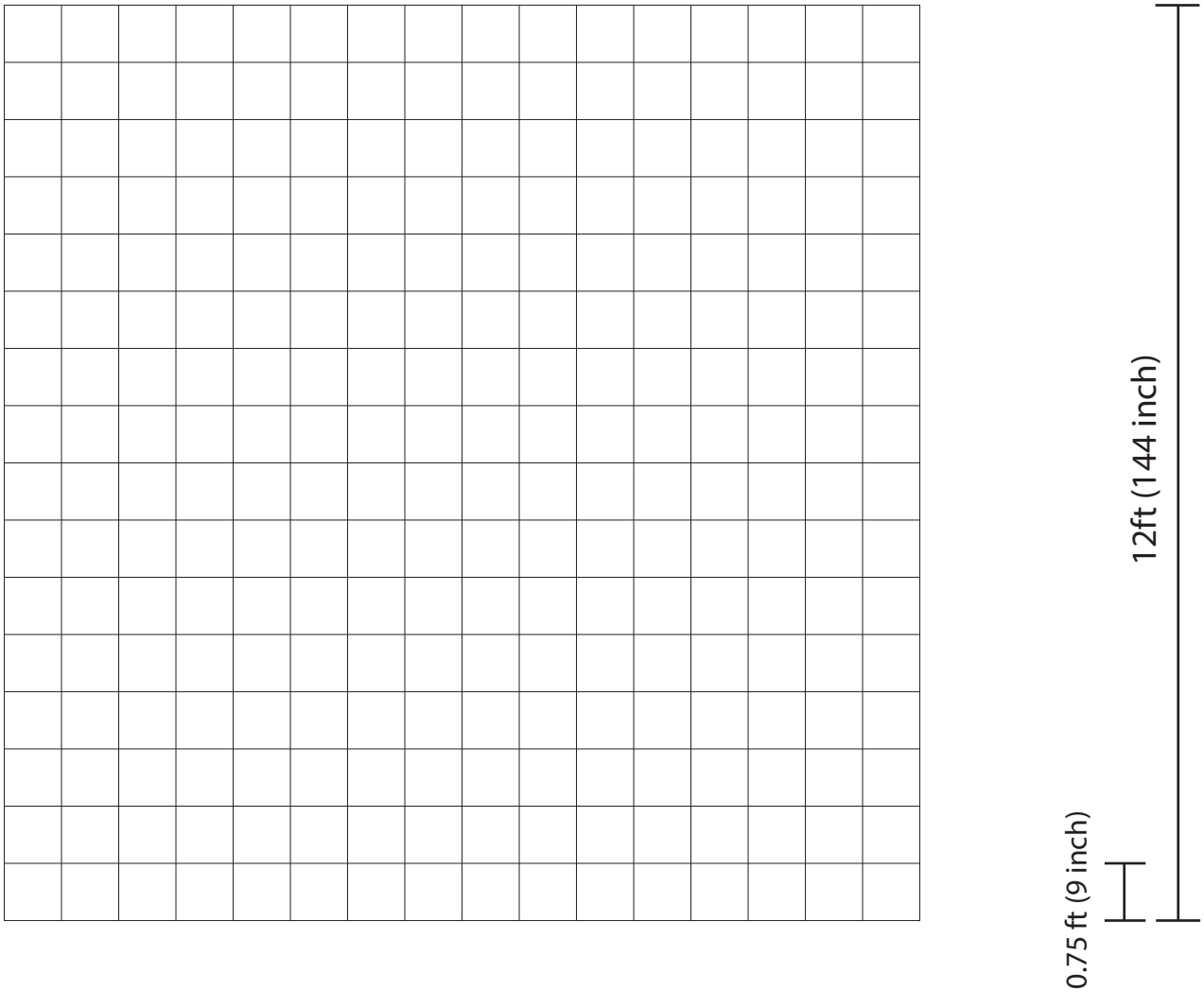
The following grid papers were provided to Paper and LEGO conditions participants to print out before coming to their study sessions.

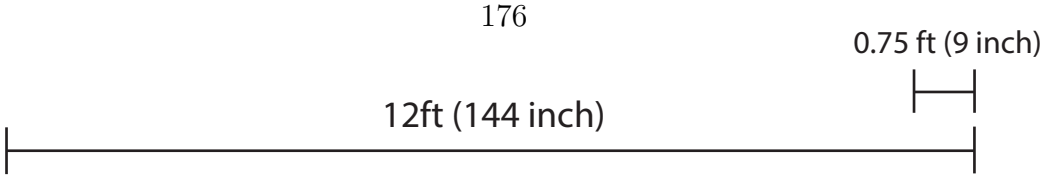
living room

(1/2)



(2/2)





B.1.2 Instruction and Interview Guide

The following guide was used for a semi-structured interview in chapter 3.

Interview guide

1. words chosen

- I’ve noticed that you picked ‘word’, can you tell me more about it?
    - what are the similar experience or activity that you might feel the same?

(pick three words that I thought the participant would have chosen on the first selection)

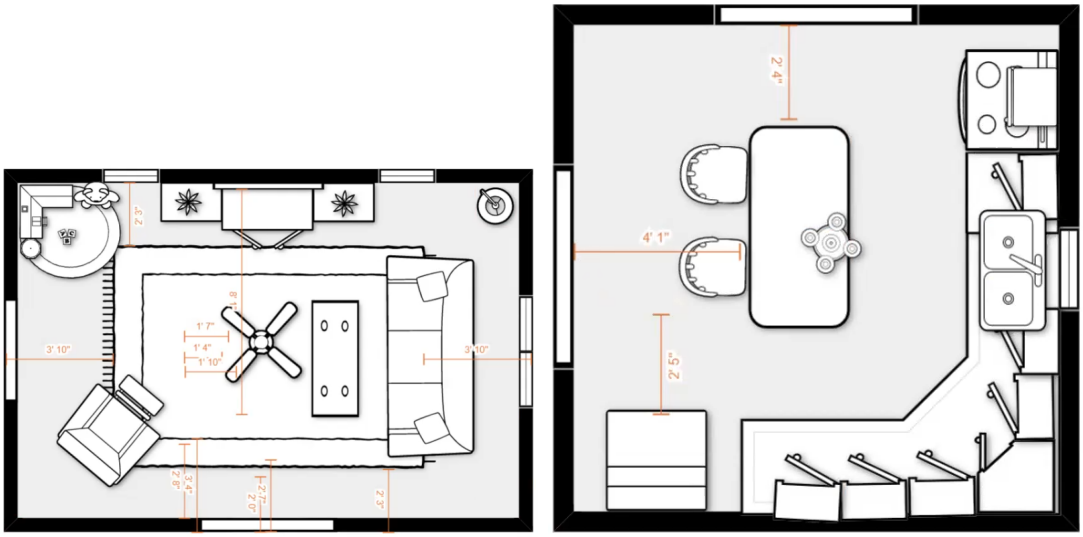
- I’ve noticed you didn’t pick *word* can you tell me more about that?
    - what are the similar experience or activity that you might feel the same?

2. general experience

- [ ] how was your experience of this study using your ‘**tool**’?
- [ ] did you plan the overall design in ur head? or make it out on the fly as you go?
- [ ] how did you like this **tool** for interior design?
- [ ] did you have any struggle or difficulty to achieve what you were doing?
    - [ ] ‘yes’:
        - [ ] what was it?
        - [ ] how did you overcome?
    - [ ] ‘no’:
- [ ] based on your experience in this study, what do you think are the strength (and weakness) of this tool?
- [ ] would you use this tool for similar activities?
    - [ ] ‘yes’: why?
        - [ ] how would you use this tool? about the same? or do you have other ways using this tool?
    - [ ] ‘no’: why?

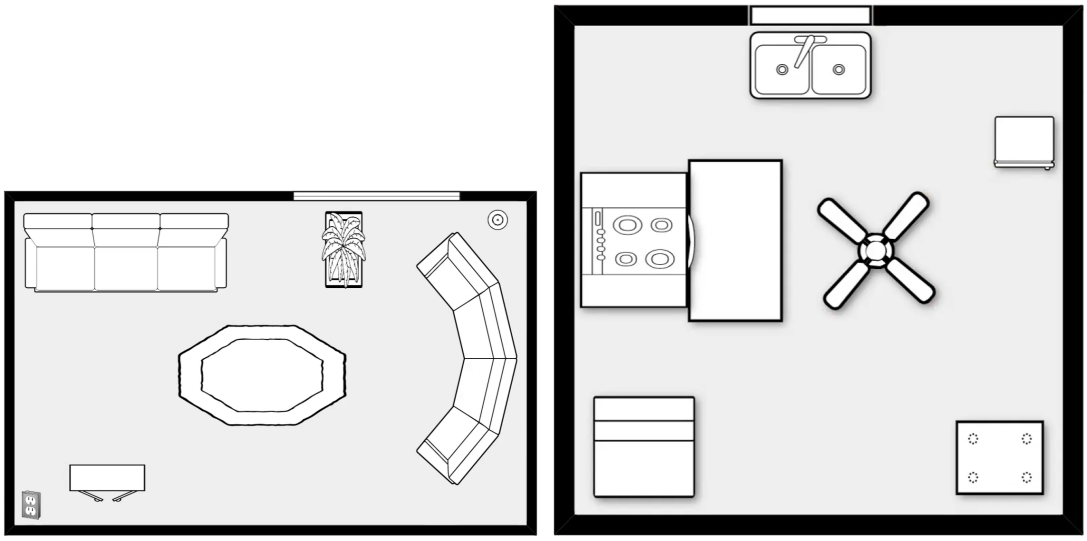
- [ ] do you have any other recommendations? for example, perhaps adding some other tools that can accelerate the process or that can give you more quality of life while you are designing?

B.1.3 Final Design Snapshots



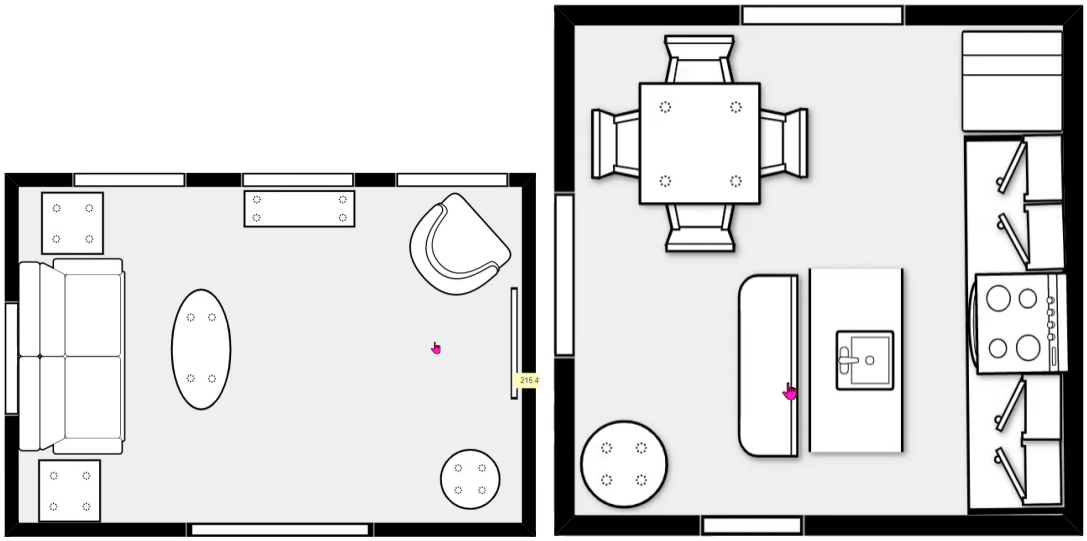
(A) P1: living room (B) P1: kitchen

Figure B.1: P1 Final design



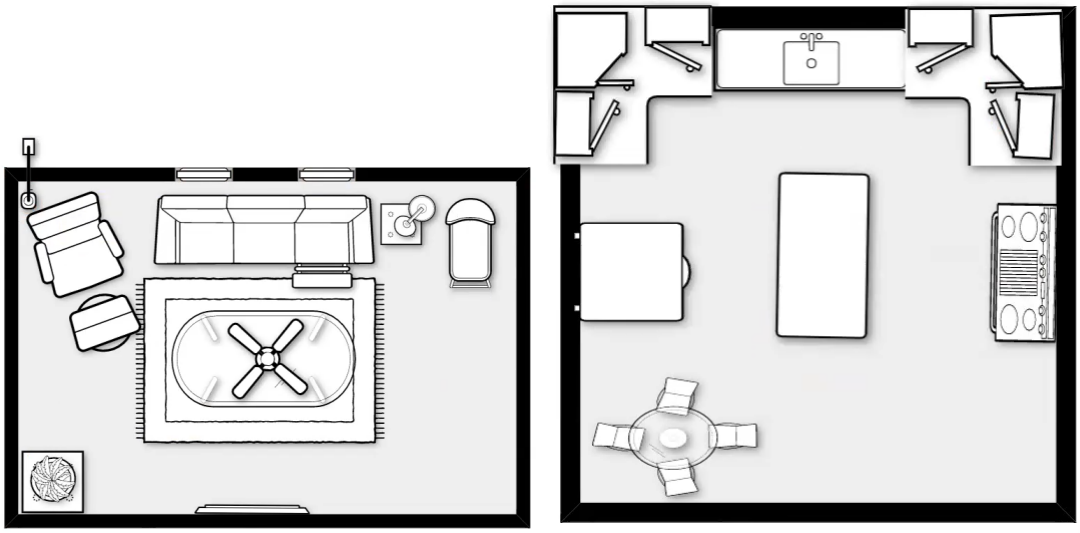
(A) P2: living room (B) P2: kitchen

Figure B.2: P2 Final design



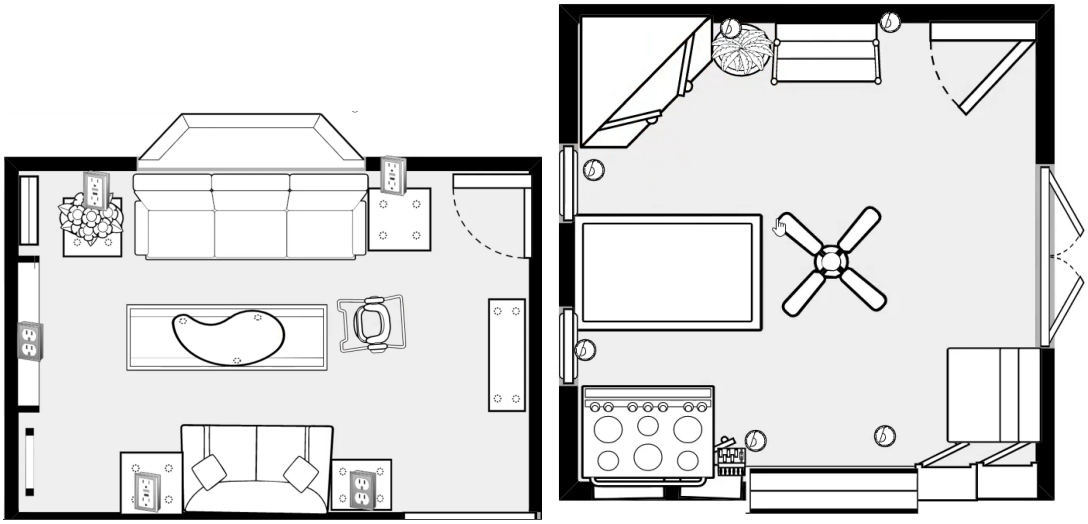
(A) P3: living room (B) P3: kitchen

Figure B.3: P3 Final design



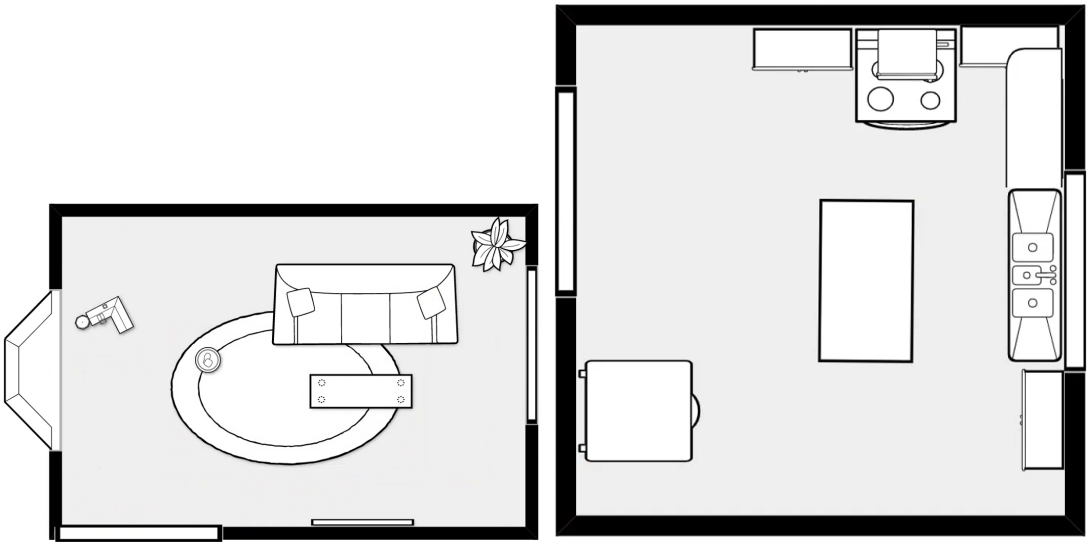
(A) P4: living room (B) P4: kitchen

Figure B.4: P4 Final design



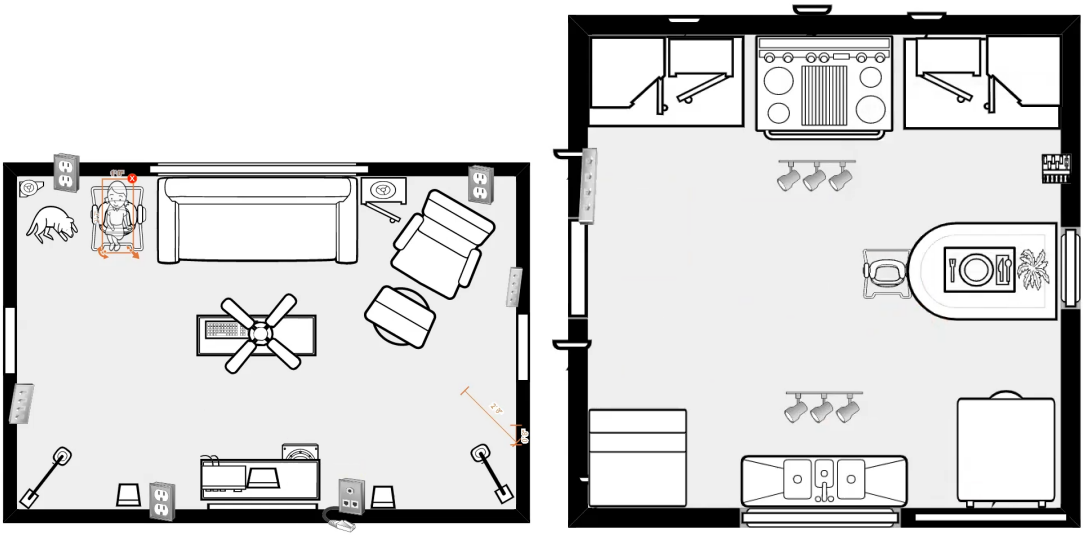
(A) P5: living room (B) P5: kitchen

Figure B.5: P5 Final design



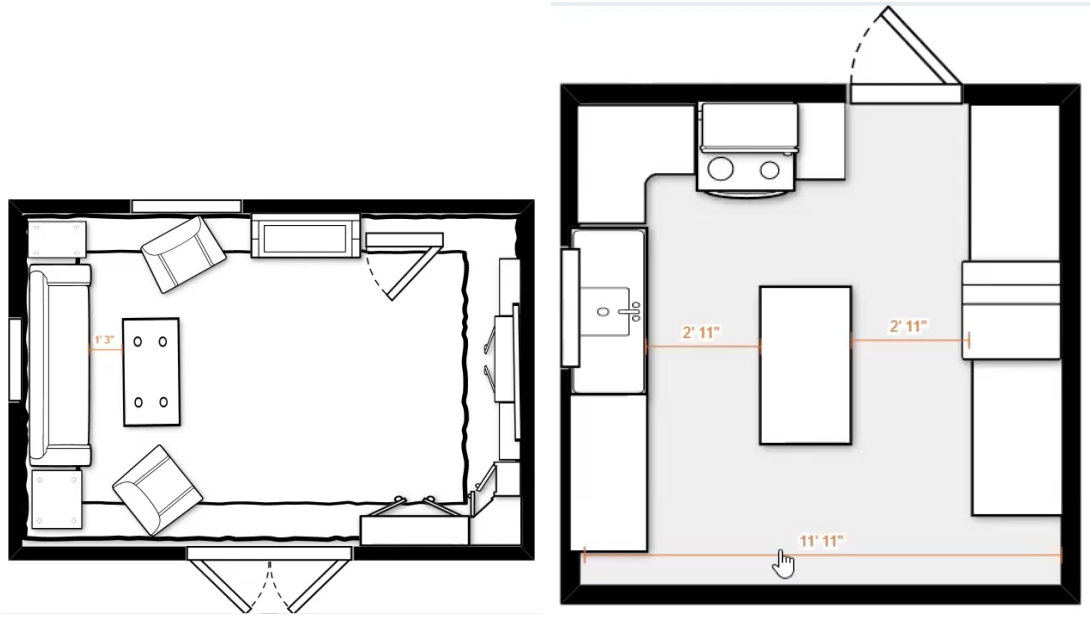
(A) P6: living room (B) P6: kitchen

Figure B.6: P6 Final design



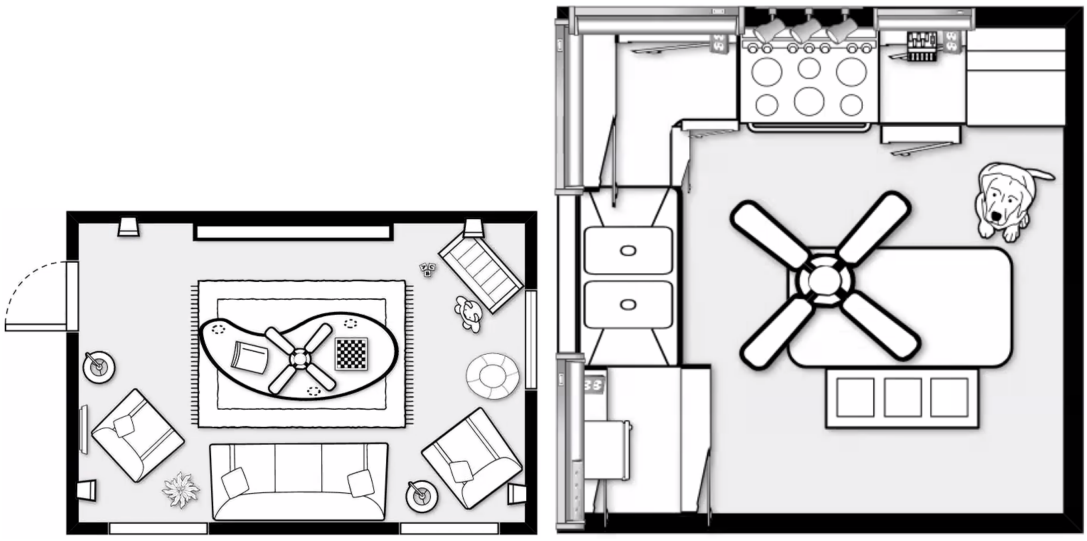
(A) P7: living room (B) P7: kitchen

Figure B.7: P7 Final design



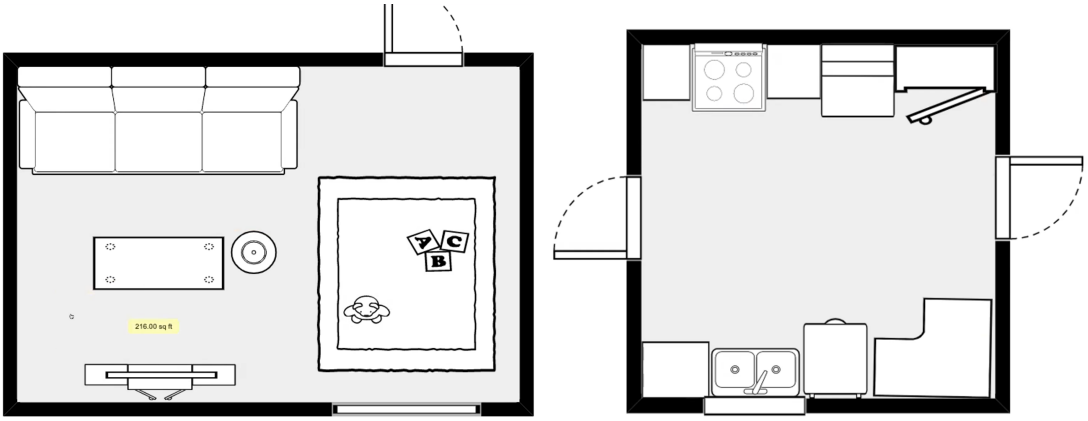
(A) P8: living room (B) P8: kitchen

Figure B.8: P8 Final design



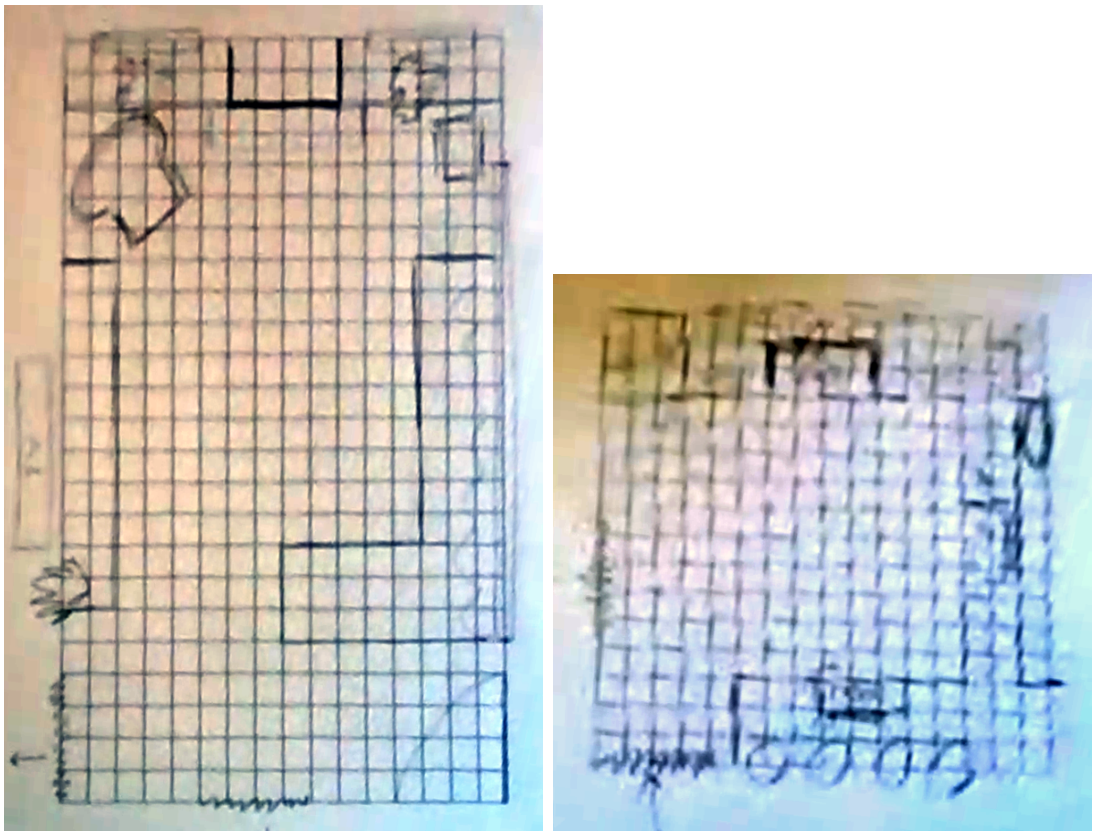
(A) P9: living room (B) P9: kitchen

Figure B.9: P9 Final design



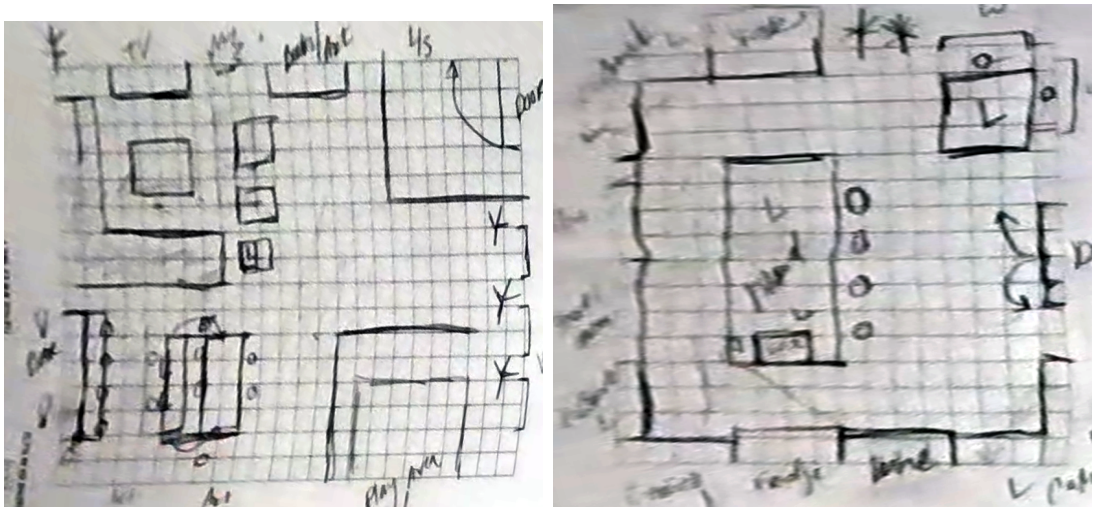
(A) P10: living room (B) P10: kitchen

Figure B.10: P10 Final design



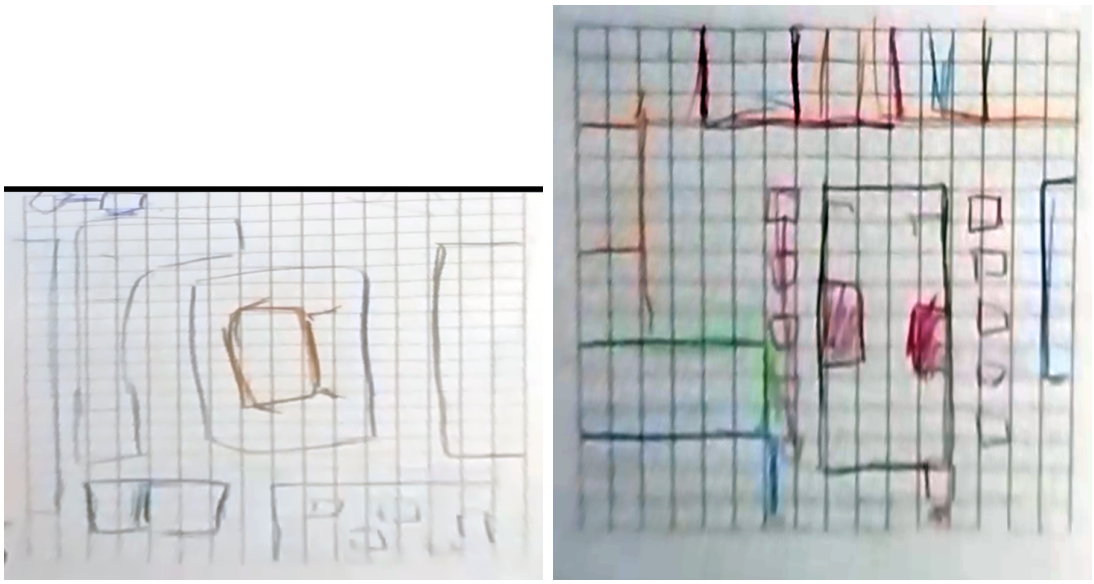
(A) P11: living room (B) P11: kitchen

Figure B.11: P11 Final design



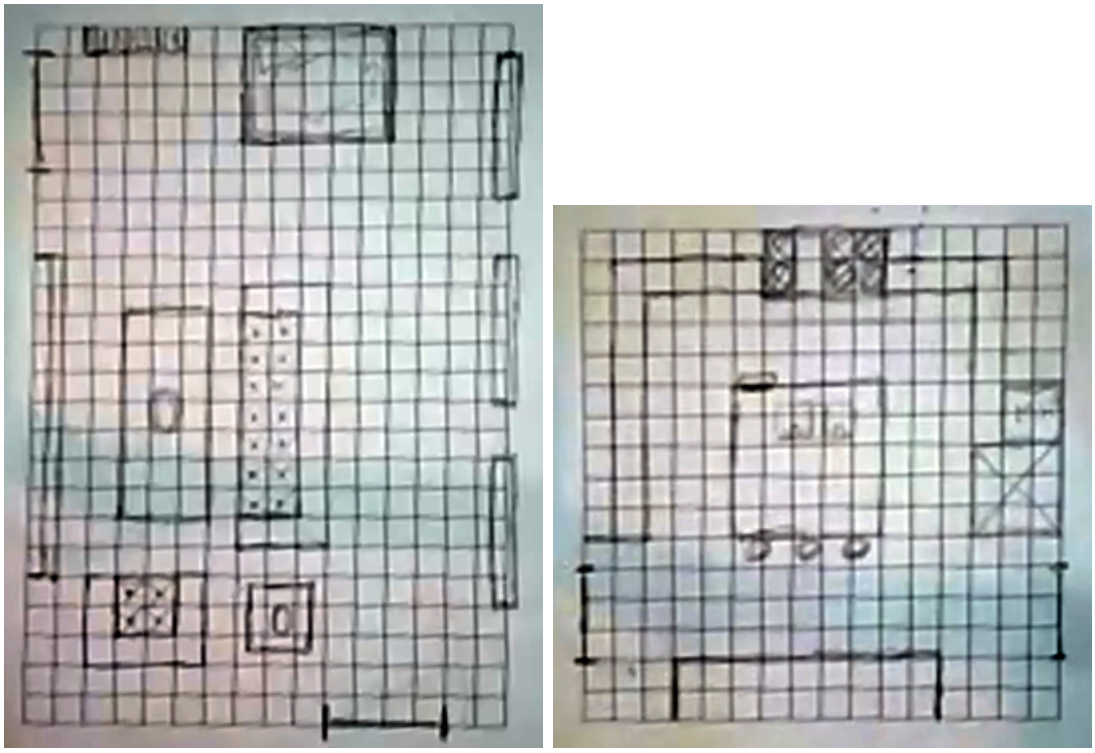
(A) P12: living room (B) P12: kitchen

Figure B.12: P12 Final design



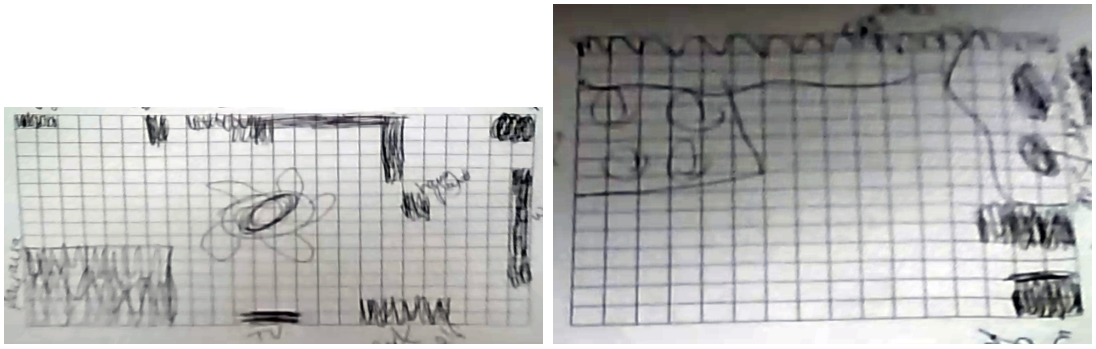
(A) P13: living room (B) P13: kitchen

Figure B.13: P13 Final design



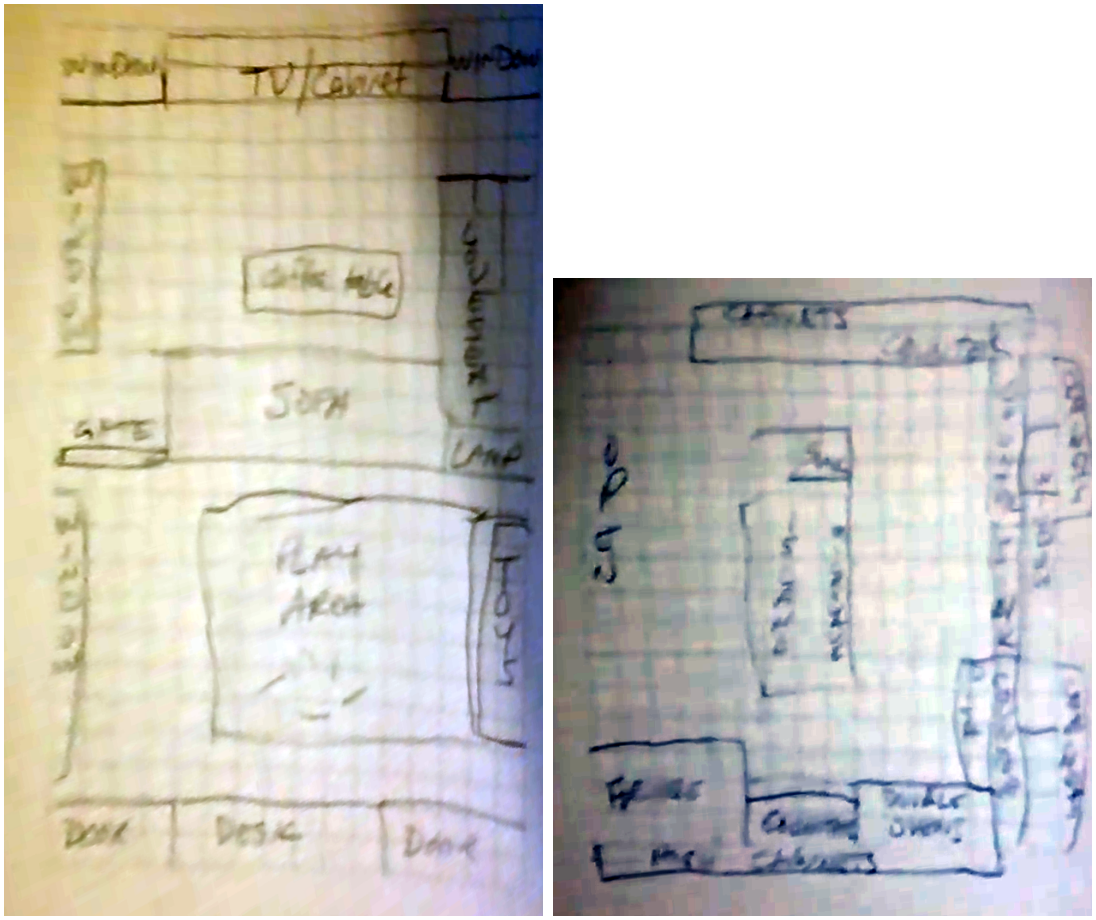
(A) P14: living room (B) P14: kitchen

Figure B.14: P14 Final design



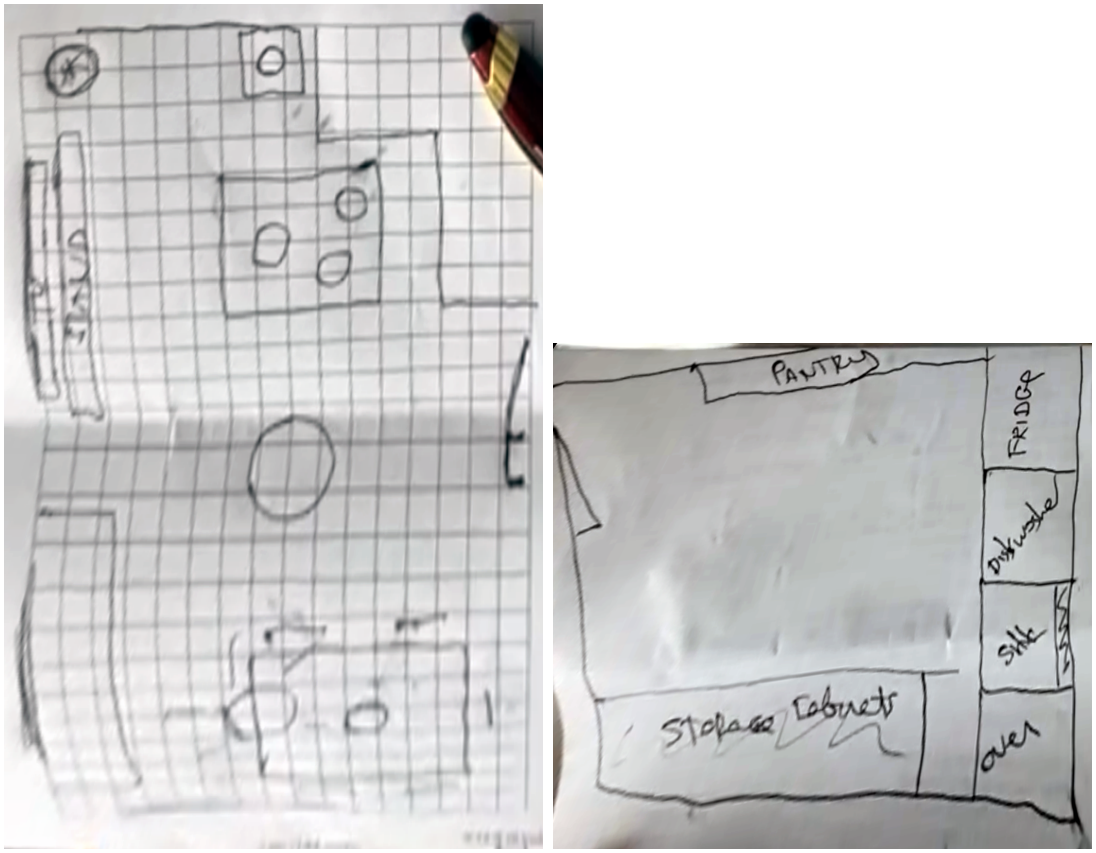
(A) P15: living room (B) P15: kitchen

Figure B.15: P15 Final design



(A) P16: living room (B) P16: kitchen

Figure B.16: P16 Final design



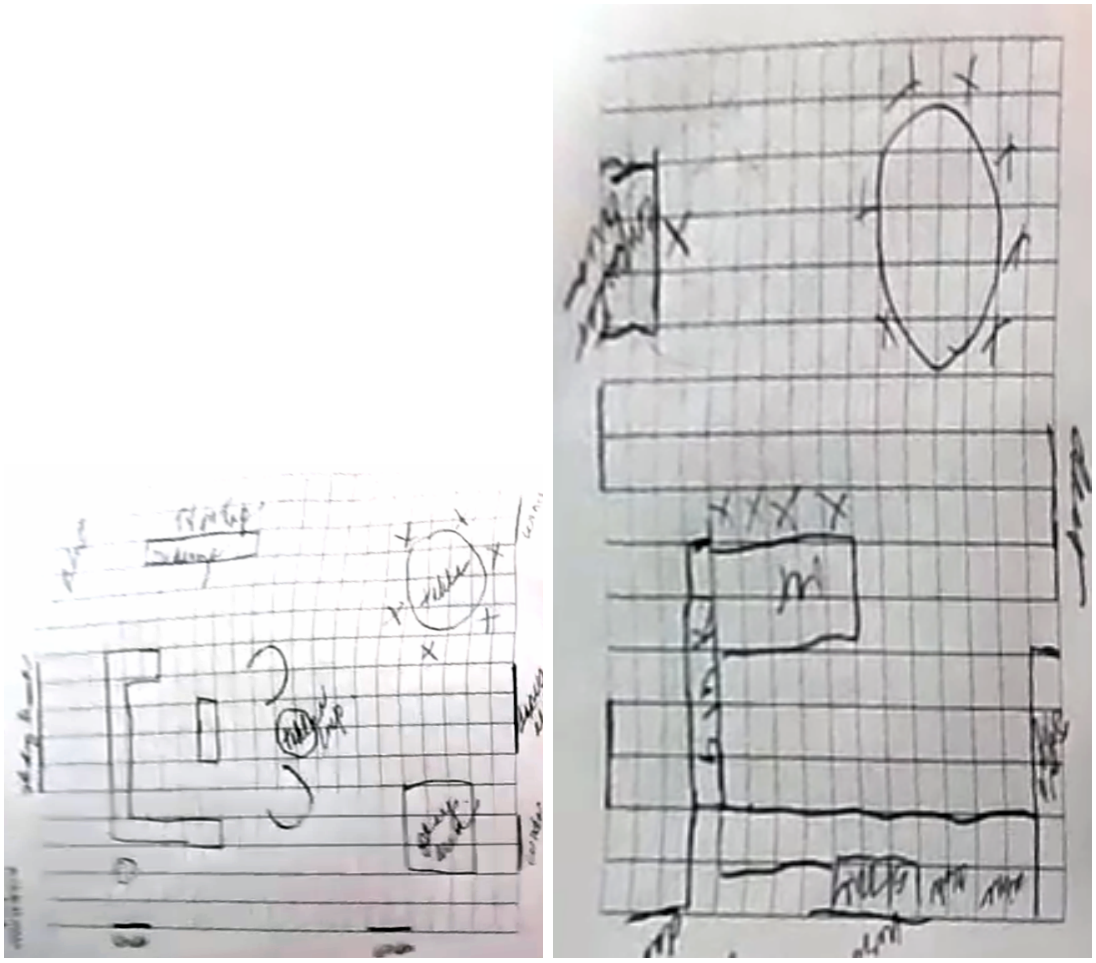
(A) P17: living room (B) P17: kitchen

Figure B.17: P17 Final design



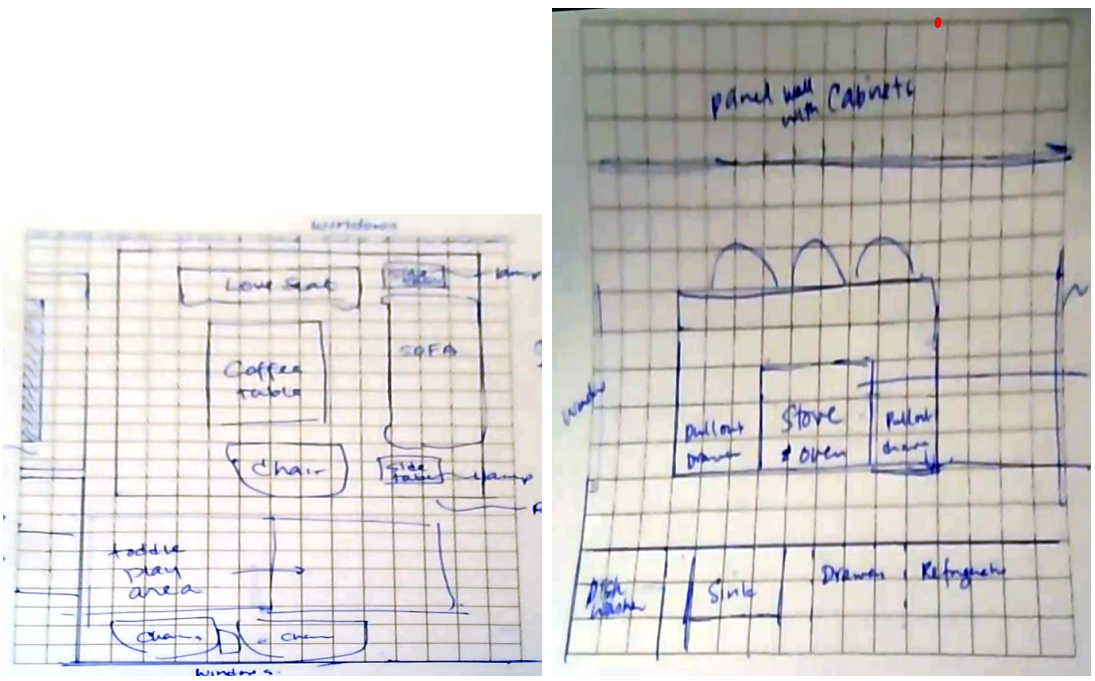
(A) P18: living room (B) P18: kitchen

Figure B.18: P18 Final design



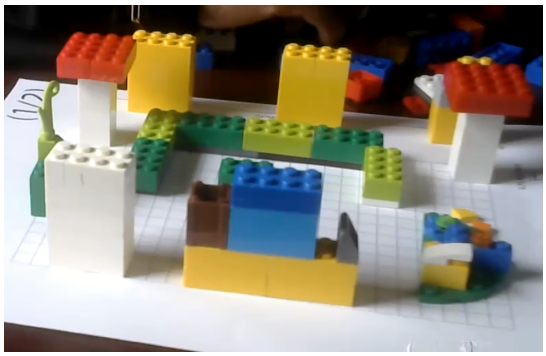
(A) P19: living room (B) P19: kitchen

Figure B.19: P19 Final design

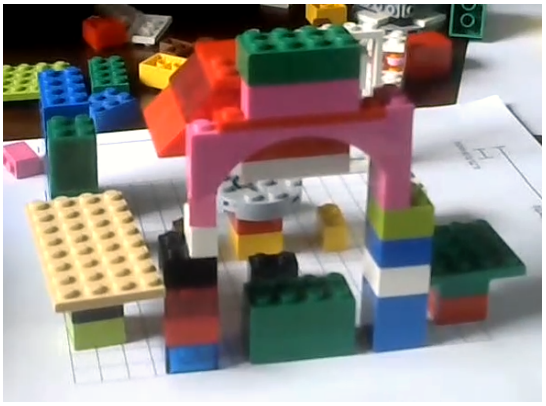


(A) P20: living room (B) P20: kitchen

Figure B.20: P20 Final design

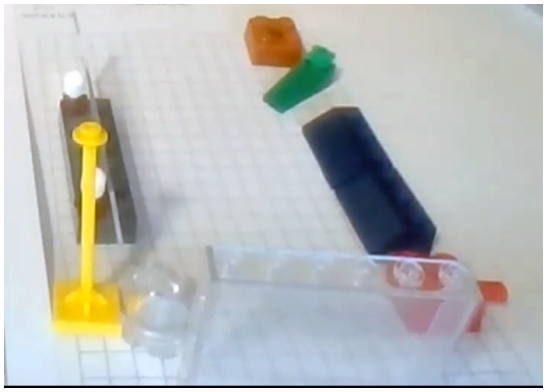


(A) P21: living room

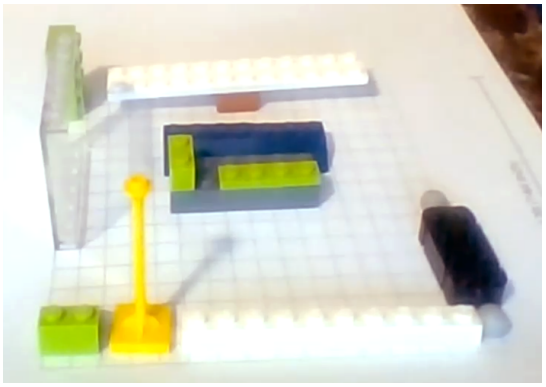


(B) P21: kitchen

Figure B.21: P21 Final design

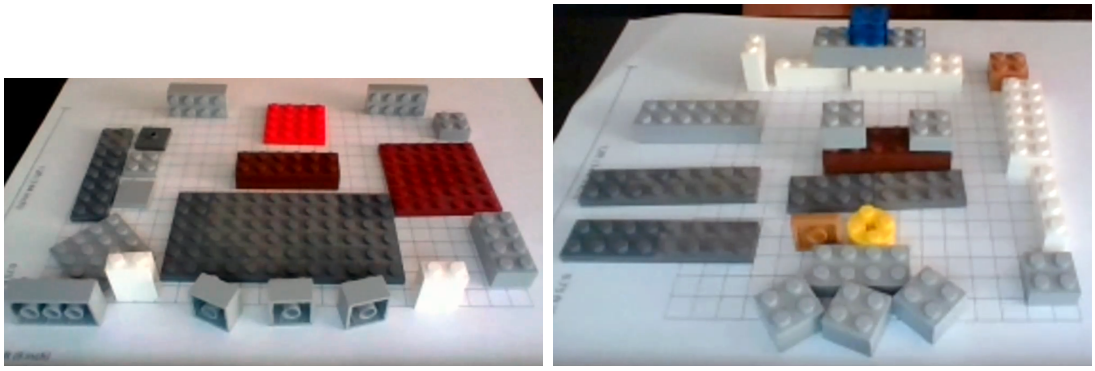


(A) P22: living room



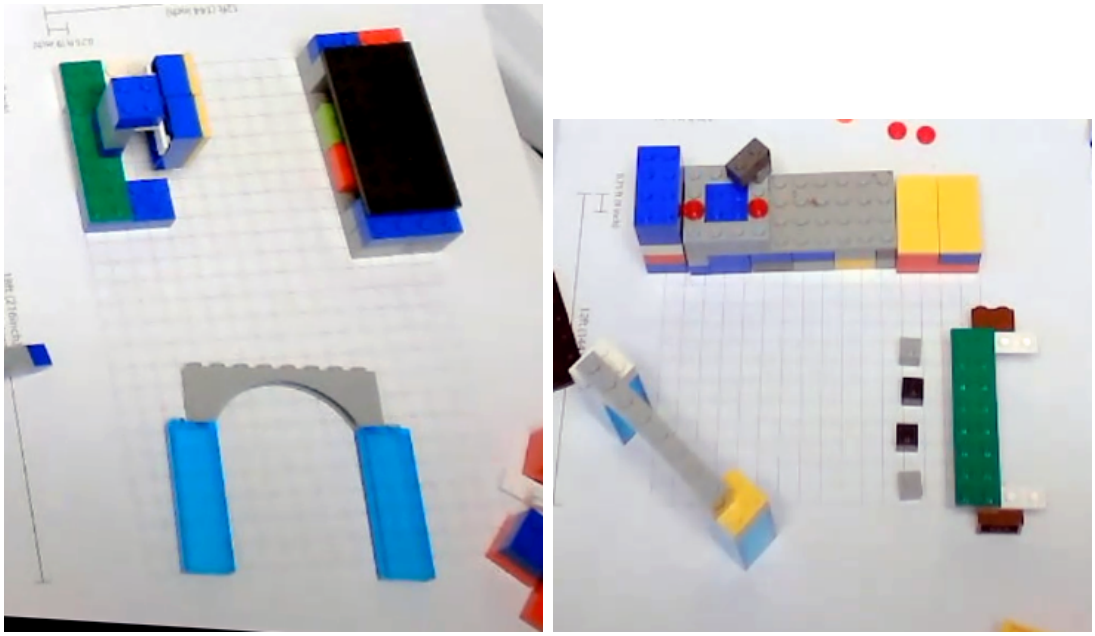
(B) P22: kitchen

Figure B.22: P22 Final design



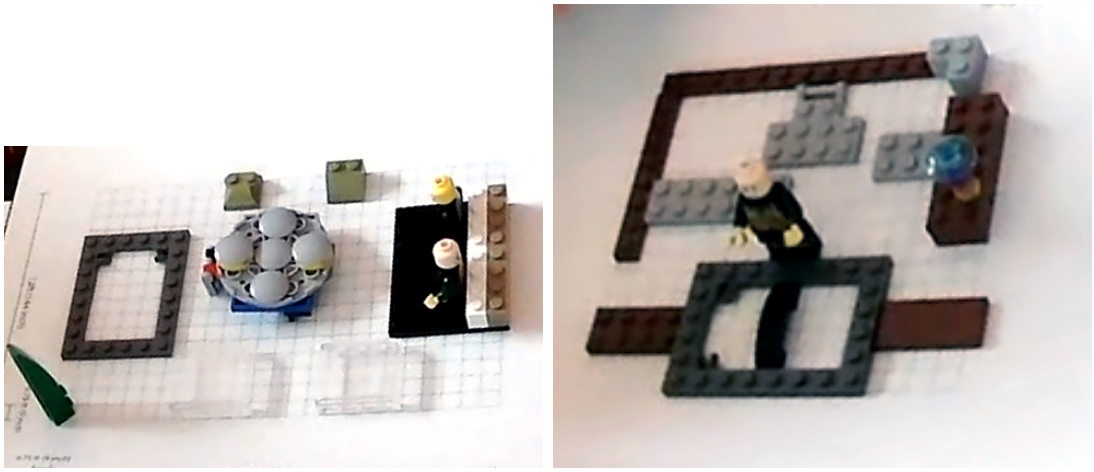
(A) P23: living room (B) P23: kitchen

Figure B.23: P23 Final design



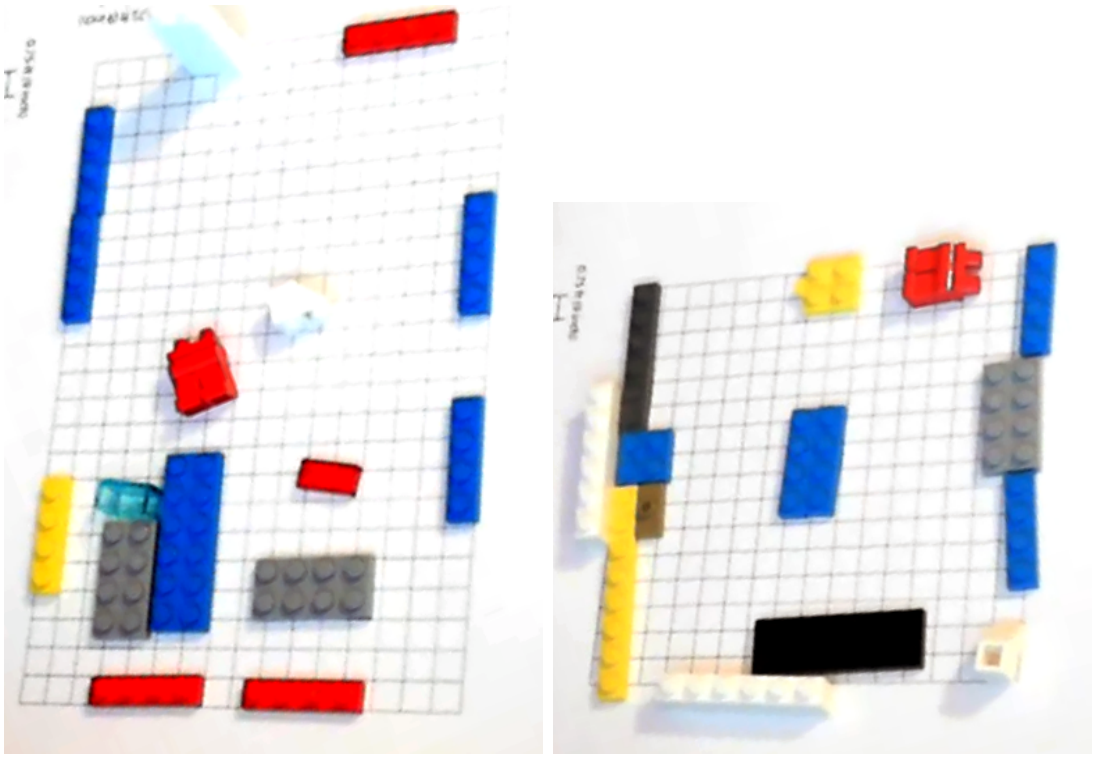
(A) P24: living room (B) P24: kitchen

Figure B.24: P24 Final design



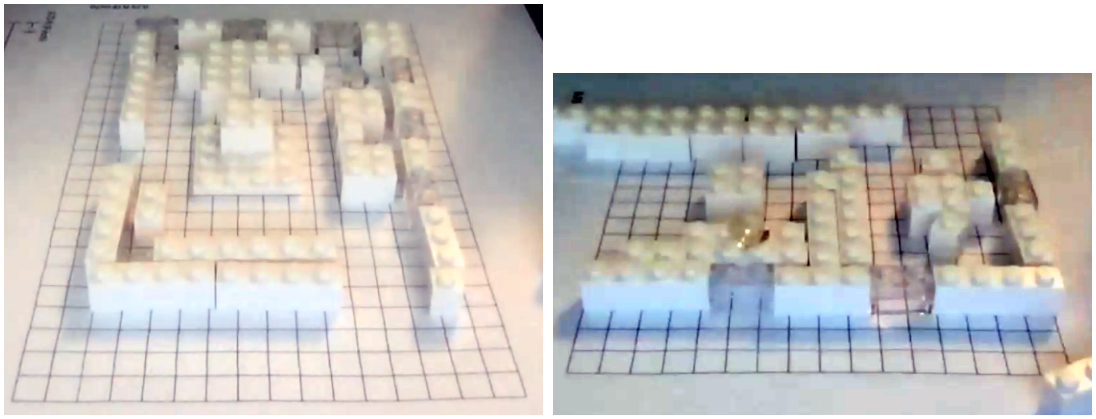
(A) P25: living room (B) P25: kitchen

Figure B.25: P25 Final design



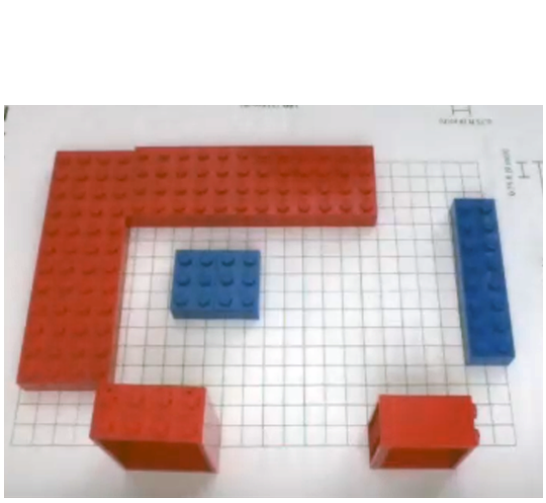
(A) P26: living room (B) P26: kitchen

Figure B.26: P26 Final design

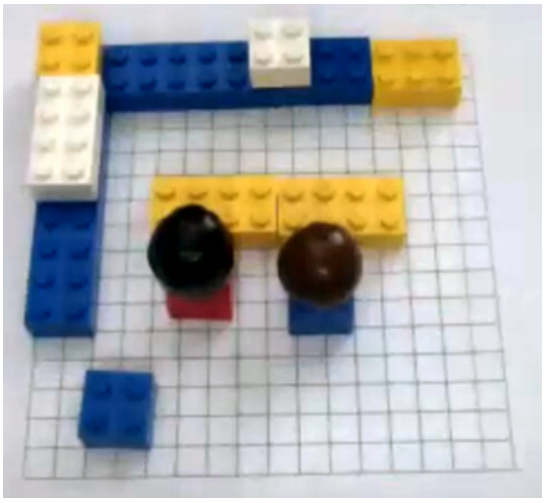


(A) P27: living room (B) P27: kitchen

Figure B.27: P27 Final design



(A) P28: living room

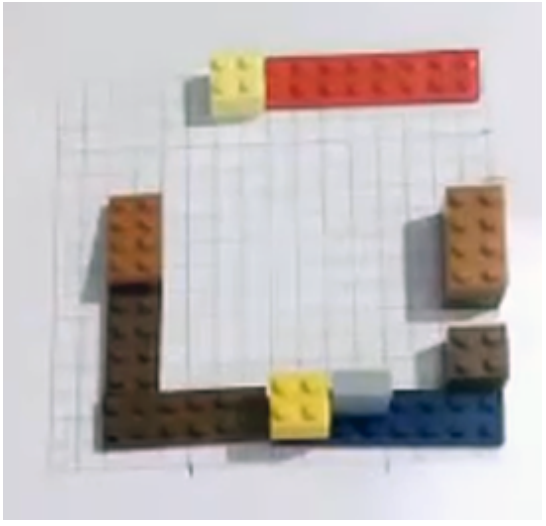


(B) P28: kitchen

Figure B.28: P28 Final design

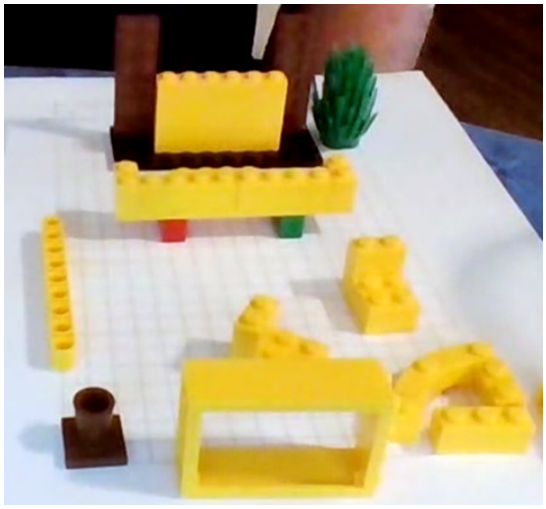


(A) P29: living room

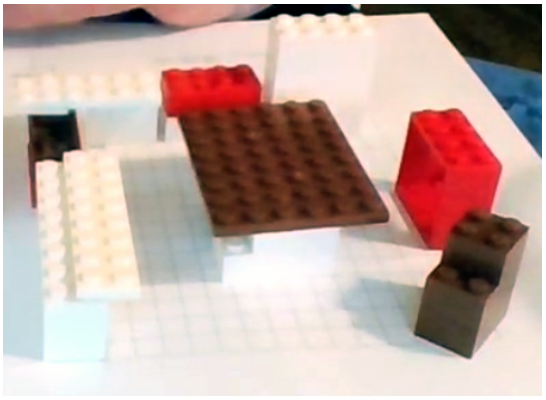


(B) P29: kitchen

Figure B.29: P29 Final design



(A) P30: living room



(B) P30: kitchen

Figure B.30: P30 Final design

B.2 Materials and Data for Chapter 4

B.2.1 Observation Notes

Below shows notes taken observing participants. These raw notes were typed and refined after each participant.

1. People tend to undershoot for docking (Z-axis). People struggle to dock (making it green) for Z-axis(forward). Definitely related to depth-perception in VR → WiM didn't help, maybe because it was placed far? If it provided overlooking overview, this should have been corrected? → another human factor problem related to desk height, human height when seated etc.

2. grab action. The study setup is designed for index+thumb, but some people really prefer to use middle finger instead of index finger. I had to enforce and ask them to use index+thumb for correct detection. People don't necessarily use fingertip (end of finger) for physical interaction. This could be more inner, closer to where the last joint(I don't know the name) is

3. For conditions (mostly mismatch conditions), I saw cases where their fingers collide with the tangible proxy. I think it is because the mismatch. e.g., bigger proxy (than visual) will collide with their finger most often because their visual expectation and their hand coordination [finger "undershooting"? For the tangible proxy]

B.2.2 Semi-structured Interview Questions

1. Which part of the study was hardest?
2. What helped you complete the task? (sensory reliance)
3. Did you have any strategy, or cheat developed?
4. Did you notice the mismatch conditions easily? Did it bother you? Do you think these conditions affected your performance?

B.2.3 Participants' Response

This section shows participants' responses to the questions. This was typed during/after the interview, and responses were interpreted thematically. These show raw notes taken individually for each participant, corresponding to numbered questions.

p1

1. Visual estimation was hardest
2. refreshing the feeling of the miniature (physical models), as reference, and middle one was hardest to tell. and small and big were ok
3. having tangible object against visual, basically comparison. and inter-trials (Q2 first, Q1 later)
4. didn't bother the performance

p2

1. 1˜5 naming confusing about the numbering of the options: Visual and tangible both were hard: 3% 5%:Physical: largest easy, 2% really hard, between 3˜5% most confusing.: Visual: visually hard compared to physical

2. Field of view: how big the visual, how much it filled (my view), proportion of miniature room
3. Visually the proportion; Physically, proprioception, hand(finger) how wide finger
4. Didn't bother the performance

p3

1. visual estimation hard
2. compared size first, answered second question in mind and derive Q1 from that.
3. (same answer above)
4. noticed mismatch, but didn't affect performance (manipulation, the docking task)

p4

1. Visual estimation was hard, because "3D volume"
2. How wide between finger; Compare to my hand
3. Sometimes, miniature becomes very small, so I know it is minimum size (it can be in this study); Feels too big, have to create mapping between what I saw
4. Bigger projection, Not really,

p5

1. First few questions confusing, instruction was little unclear? Didn't know what I am supposed to do; too easy? a. Option picking. b. Distracting about the surrounding, but overall helpful to estimate the visual scale c. Easy manipulation
2. Measuring with fingers. And comparing proprioceptive measure vs. visual measure a. Try to move finger first (probably needed some time to adapt their move/perception), try to see how fast, sensitive of the finger movement (this participant caught the concept of C/D ratio)
3. Same as above. Refreshing time (break time for refreshing reference=physical model) was helpful
4. A lot of mismatch; a. Felt "i was right(correct)" b. It bothered me a lot of the mismatch. Manipulation was quite easy. Needed some time to adapt (sensory calibration) less than a second. c. Bigger haptic was easier than the other way.

p6

1. Comparison haptic vs. visual scale, when visual was small hard to dock as well
2. CAD modelling education helped a lot. Modelling and 3D printing etc. a. Having a reference(physical model), and 100% scale in virtual model. b. Manipulation was, it was seamless, natural interaction with the objects. (because using hands directly)
3. Made decision before completing docking a. Visual first → haptic second
4. Both extreme was easier to notice. a. Visual was smaller, b. The task was easy overall. c. The colored corner was 'accessible'

p7

1. This participant - Red/green color blindness; a. Small one(visually). Text, color, box difficult to focus b. (Focus different between canvas and scene - Unity stuff)
2. Visually focused. visual main. Tangible is for manipulation task
3. Orientation first, specially. Smaller ones! Orientation first and position. Dragging/sliding on the surface a. Same strategy as above.
4. Huge diff - yes. Slight (bothered). Also depends on the previous trials. a. Text information was a little helpful. b. Mismatching; a bit bothered. If small, used background a bit. i. Mismatch didn't really bother the performance. ii. Task was simple. And it was repetitive task (almost same movement, and direction overall) so didn't really need 'much time' to adjust/adapt

p8

1. Object manipulation, positional thing the most. Remembering (Q1) was hard. too
2. WiM red/blue feedback on the receptacle('ghostimage') to judge
3. Relied on previous trial
4. Quite easy to notice the differences (mismatch condition) a. Second question was first in mind (the comparison) and ... hard to get the first question (Q2 first → Q1 later) b. I don't think I needed some time to adjust and adapt to the mismatch condition. c. Smaller one was harder.

p9

1. Once grasp the physical object, notice the changes; needed some time to adjust (recalibrate sensory). a. Because It is volumetric, just had to guess b. Depending on the size of the model, smaller ones - relied on full scale (the reference model) i. Macro (big enough) → quick WiM ii. Micro (visually small ones) → background information more (the full scale scene) iii. Manipulation - easy, quick 'calibration' time

2. Bounding of box of the 'ghostimage' to align the side table (manipulation)

3. Rounding number (in mind), a. Physical sensation than it looked; perception vs. 'feel' b. Extreme ones are easy, otherwise it is middle c. Visually inspect first and followed the intuition d. Calibrate physically measure e. Expected (visually) → recalibrate physical, motor skill

4. As soon as interact(touch), understand the mismatch. a. Some bothersome, needed to recalibrate/ readapt b. Not really affected the performance c. Middle one was easiest(intuitive), big was too big(occlusion), small was too small (can't see the WiM) d. 'mouse sensitivity' (the concept of C/D ratio)

p10

1. First question (visual estimation); roughly estimate.

2. For Later trials, previous trials helped for subsequent trials. Reference refresh during quick break and practice run helped (Bare eye feeling (having to experience real, physical object) helped me).

3. Small objects, smaller, big objects → bigger.; read the trend of the study conditions.(if visual is small, small physical objects were expected)

4. Depends on. Position. World in miniature; surrounding context. a. Felt easy enough; b. Was hard when mismatch (minor) - the smaller, the harder;

p11

1. Visual estimation was hardest, sometime making it green (docking). "I think I didn't do well on distinguishing"

2. Hand feeling first.

3. Holding, measuring. second question first → deduct that from for the first questions

4. Sometimes - extreme case was easy. a. Little bit bothered. Mismatch conditions was kind of affected my performance. (little bit) "because it was less intuitive". Needed some time to adjust (recalibrate) but it was really quick calibration

p12

1. Converting the visual to physical. Visual estimation was hardest. a. Mostly visual stuff

2. Little both, tactile thing mostly though, moving. Grasp/pickup the model, inspect it (3D,volume feel). Chain mapping: physical mass model at full scale → virtual model at full scale → virtual miniature model

3. Middle fingers, measure. Rotating first (to align the orientation for docking task) then moved to the receptacle(ghostimage for docking); middle finger was supportive role for (a) handling delicate manipulation (rotation etc), (b) once pick up the model, sometimes used middle finger to support the weight, Balance.

4. Extreme case was easy to notice the discrepancy. Smaller ones kind of hard to notice. a. Not really bother; didn't affect performance; for detailed task, in competitive gaming session, might 'prefer' to have matching conditions. b. Needed some time for visual feel (visual inspection). Try to calculate (Q1) first. c. Sometimes, wanted to tell bigger number because 'felt' bigger(physically in hand), but needed to focus on the visual only. [this tells interference to the visual dominance when paired with haptic discrepancy?]

B.3 Materials and Data for Chapter 6

B.3.1 Subjective Measures Surveys

The following shows NASA-TLX and a custom 7-point Likert for 6 statements used for subjective measures.

NASA-TLX

Please use the sliders to answer each question below

Mental Demand: How mentally demanding was the task?	
Physical Demand: How physically demanding was the task?	
Temporal Demand: How hurried or rushed was the pace of the task?	
Performance: How successful were you in accomplishing what you were asked to do?	
Effort: How hard did you have to work to accomplish your level of performance?	
Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?	

Custom Likert

Please rate your response according to the statement below

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
I could easily predict how the system would respond to my movement	○	○	○	○	○	○	○
This interaction style allows me to stay focused on the task without interruption	○	○	○	○	○	○	○
My sense of my own hand and body felt stable during interaction	○	○	○	○	○	○	○
The moment of touching the physical object felt visually consistent with the VR interaction	○	○	○	○	○	○	○
I had to consciously think about how to control the system	○	○	○	○	○	○	○

The interaction felt coherent as a whole, rather than a set of separate mechanisms	○	○	○	○	○	○	○

B.3.2 Interview Guidelines

Main Evaluation for the two strategies

The following questions were asked during the main evaluation in chapter 6.

- Did the two interaction styles feel different to you? If so, how?
- Where did you feel most constrained in each version, your body, or the space?
- Did either version make fine adjustments feel harder or easier?
- Did you trust one version more than the other to behave as you expected?
- What made you stop adjusting the layout when you did?
- If you had to use one of these long terms, which would you choose, and why?

Hybrid Compensation

The following questions were asked after the main evaluation for the hybrid scale compensation strategy.

- Does this version address what you experienced from either strategy?
- What feels improved, and what still feels unresolved?
- Does this introduce new issues?