

Synergistic Embodiment: A Framework for Shared Control and Multi-Limb Interaction in Virtual Reality

HONGYU ZHOU



THE UNIVERSITY OF
SYDNEY

Supervisor: Assoc. Prof Anusha Withana
Associate Supervisor: Lec. Zhanna Sarsenbayeva
Associate Supervisor: Prof. Judy Kay

A thesis submitted in fulfilment of
the requirements for the degree of
Doctor of Philosophy

School of Computer Science
Faculty of Engineering
The University of Sydney
Australia

11 June 2026

Abstract

Collaboration is moving beyond shared screens and virtual spaces toward the shared execution of bodily action. As VR, robotics, and AI rapidly converge, the augmented body is no longer just a personal avatar, but is becoming a shared, programmable medium for coordination. Whether bodies are co-controlled by multiple users, extended through supernumerary limbs, or partially governed by semi-autonomous agents, bodily action is increasingly distributed across people and algorithms. These systems unsettle a implicit assumption in HCI: that embodiment relies on a single user, a single physical boundary, and a unified perception-action loop. When this loop fractures across distributed agents, traditional models of agency, control, and responsibility break down. Together, these shifts define an emerging condition that I term **collaborative embodiment**.

This thesis investigates collaborative embodiment as a unified problem rather than as a collection of isolated interaction techniques. I first conducted a systematic survey of 137 studies to map the methodological landscape of collaborative VR and identify key gaps in how shared control, visual perspective, and limb augmentation have been studied. Building on these insights, I then designed and evaluated three VR systems: *CoplayingVR*, which examines shared hand control in gaming; *One Body, Two Minds*, which investigates dynamic perspective switching for remote collaboration with virtual supernumerary limbs; and *Juggling Extra Limbs*, which explores how users coordinate with semi-autonomous virtual arms.

Across these systems, I employed a mixed-methods approach that combined performance measures, workload and embodiment questionnaires, and qualitative observations of user behaviour. The findings reveal three recurring patterns. First, shared control can improve novice performance and engagement, but only when coordination costs and role ambiguity are carefully managed. Second, flexible perspective strategies are essential for balancing situational awareness, comfort, and embodied experience in dynamic collaborative tasks. Third, as limb autonomy increases, users shift from direct manipulation toward delegation and

supervisory forms of control, treating supernumerary limbs less as fully embodied appendages than as collaborative extensions.

Based on these contributions, this thesis advances **Synergistic Embodiment** as a conceptual lens for understanding collaborative embodiment in VR and introduces the *Interaction Elements Matrix (IEM)* as a first-order design framework for reasoning about control mapping, perspective strategy, feedback design, and autonomy. Together, these contributions provide empirical evidence, methodological guidance, and structured design support for collaborative VR. More broadly, they reposition collaborative embodiment as a central HCI problem for future immersive systems and clarify how shared virtual bodies can be designed, negotiated, and experienced.

CHAPTER 1

Acknowledgements

I am deeply grateful to the many mentors, collaborators, friends, and family who have supported me throughout this journey.

Above all, I owe my deepest gratitude to my primary supervisor, Anusha Withana. Whenever I hit a wall, his calm encouragement always kept me going. I especially want to thank him for everything he did during the COVID pandemic. Being stuck away from Australia for two years, but Anusha made sure my PhD never stopped. From sorting out remote lab access to simply checking in on how I was doing, he really did everything he could to help me get through that time. I know that my path through this PhD was not always straightforward, and his patience throughout that process meant more to me than I can fully express. I would like to express my sincere gratitude to my supervisor for his continuous support and insightful guidance, which helped me navigate this field and reach the finish line of my PhD journey.

I am equally indebted to my co-advisor, Zhanna Sarsenbayeva, whose exacting standards and patient guidance transformed both my writing and my grasp of HCI research. Zhanna took the time to coach me on crafting clear, compelling prose, walking me through drafts and helping me find my academic voice. She also introduced me to the methodological heart of HCI: user centered evaluation, experimental design, and theory driven analysis, which ensure I built a solid foundation for all my future studies. Her combination of intellectual rigor and genuine empathy taught me how to question assumptions, anticipate alternative interpretations, and design experiments that truly illuminate user experience. I am profoundly grateful for Zhanna's mentorship, which has shaped my approach to research and set me on a path of thoughtful, rigorous inquiry.

I also extend special gratitude to Judy Kay, my third advisor, for her generous guidance and thought provoking questions, which helped me to deepen my theoretical framing and

strengthen my experimental designs. Judy's perspective always brought new dimensions to our discussions and made this thesis far richer.

To my committee members, thank you for your valuable feedback and encouragement. The time and care each of you devoted to reading early drafts and offering constructive critiques truly shaped this final manuscript.

I owe a massive thank you to Andrea Bianchi, my brilliant mentor, collaborator and good friend. Crossing paths with Andrea has been one of the greatest gifts of my PhD, far beyond anything I ever expected. Working with him taught me so much, as he always addressed my seemingly naive questions with immense patience, empathy, and a level of seriousness I deeply appreciated given his seniority. Observing his work firsthand, I was inspired by his uncompromising research standards and meticulous attention to detail. Beyond just guiding me through specific projects, his unwavering support empowered me to grow into an independent researcher. I have also been moved by his selfless commitment to academic service, which has shown me what it means to be a responsible and contributing member of the HCI community. Through his example, I have come to see that the best researchers are not only intellectually rigorous but also generous mentors who help others grow with care, patience, and high standards. Even now, during moments of burnout, reflecting on his work ethic instinctively reminds me to hold myself to a higher standard. I didn't fully realize the impact of all this until I started mentoring my own students. Because I was lucky enough to experience his warmth and sincerity firsthand, I constantly catch myself trying to treat my students the same way, hoping to pass on that same feeling of being supported. For all this, thank you, Andrea. I feel so lucky and proud to be in the same research community as you.

I am profoundly thankful for Masahiko Inami's invitation to visit the University of Tokyo, which allowed me to experience a truly inspiring lab culture. Instead of forcing his students into a fixed mold, he empowers people from incredibly diverse backgrounds to lean into their unique strengths and pursue the exact research they love. He nurtures everyone according to their true nature, patiently polishing their individual talents and dreams. He doesn't push a fish to climb a tree, but instead helps everyone find their own ocean. Beyond his academic vision, Inami-sensei's wit and warmth create a naturally encouraging environment that feels like a breath of fresh air for any researcher. In his lab, the freedom to explore ideas openly allowed

my imagination to flourish, knowing that every spark of inspiration would be met with sincere support and belief. The projects I completed there are particularly special to me as they grew out of that unique, inclusive culture, and I remain grateful to have experienced his unwavering dedication to research firsthand. His pure, unwavering dedication to research is something I admire, and I feel so fortunate to have been part of such an inspiring environment.

I would also like to extend my deep appreciation to Eduardo Velloso. His high standards for methodological rigor have deeply refined the way I approach study design. But honestly, I am grateful for everything he did for the whole group. When he first got to Sydney, he went completely out of his way to put together that Bayesian statistics course and all those experience sharing workshops for us. I got so much out of them. Also, I am especially grateful for the times I went to him for guidance. He was always so open and sincere about sharing what he's learned over the years, and he guided me through every challenge with incredible patience. I am also deeply grateful for how Eduardo recognizes every small milestone in a student's progress. His consistent encouragement and praise provided me with the confidence and courage I needed to keep moving forward. I deeply appreciate his wonderful energy and how he always has everyone's back. Even though I only met Eduardo toward the last year of my PhD, he has still had a very meaningful impact on me, and I am so happy that I had the chance to know him before I graduated.

My heartfelt thanks go to my dear friends and colleagues. Shakyani, thank you for being such a wonderful friend. Working with you has been great, but I am even more grateful for the warmth and comfort you brought to my everyday life. I will always remember us organizing lab events together and how much you were there for me emotionally over the years. It is incredibly rare to find someone who is both a brilliant researcher and such a genuinely good-hearted person, which is exactly why I look up to you so much. Wendi, crossing paths with you has been one of the best parts of my PhD, and I cannot thank you enough for always encouraging me. Beyond research, whether you realized it or not, just your spirit and the way you carry yourself have taught me so much about life. I am honestly just so happy that someone as truly admirable as you came into my life, and I cannot imagine how different this whole experience would have been without you. Yihao, I really lucked out getting to work with you every day. It is rare to find someone who is both a brilliant collaborator and

just a genuinely great friend to hang out with. I am really going to miss those moments where we would just chat and suddenly come up with a cool new idea. Thank you for always sharing your thoughts so freely and for being someone I could always count on. Praneeth, watching your pure dedication to research has been incredibly inspiring. It has been a real privilege to work alongside you and learn from your focus. I feel very lucky to have shared this journey with someone as driven as you. Adele, you have been there since before my PhD even officially started. Thank you for always taking the time to help me and for giving me such great advice whenever I needed it. Marvin, thank you for always being ready with a thoughtful perspective, whether we were talking about work or just life in general. I have always admired your patience and your incredibly calm approach to everything. Pamuditha, having you there right at the beginning of my Sydney journey made that time so beautiful, and it is a memory I will always cherish. I really hope your next chapter brings you everything you are hoping for. Finally, a massive thank you to everyone in the AID Lab. Thank you all for the fun, the shared moments, and for making this entire journey so memorable.

I also want to say a massive thank you to everyone at Inami-sensei's lab. Every single day of my time at UTokyo is a memory I truly cherish. I still miss so many of the little things, like our grocery runs to make hotpot together at night, heading over to Shimokitazawa for dinner, and even those bean sprouts you managed to grow right inside the lab. I'm also just so grateful for how much you supported my research. It meant the world that everyone was so willing to jump in for my user studies and give me such thoughtful and high-quality feedback. I still find myself thinking about life there, whether it was the simple joy of having a raw egg and miso soup with my cafeteria lunch or waking up excited for those amazing salmon onigiri. I love the vibe of the space itself, whether we were all hanging out around the big round table in the middle of the lab or just sinking into that round lounge chair right next to the Jizai arms. Even now, my dreams are often filled with the moonlight I looked up at by the railway tracks on our way to Shimokitazawa, and the memory of all of us just walking along together. It's been a while since I left, but my memories of those days are still so incredibly vivid. Meeting you all was honestly one of the best parts of my PhD, and it was an absolute privilege to spend time with such a creative and passionate group of people.

And to my amazing co-authors, way more than the papers we published, what I really care about is just the time we spent together. I still completely believe that research is really just about “pursuing what you love, alongside the people who just get you,” and I feel incredibly lucky to have crossed paths with you on this journey. I will always remember our boundless conversations, and those sudden calls we made to each other the moment inspiration struck, no matter what we were doing. To me, every single one of you is an utterly unique and fascinating soul in this universe. Even though those projects are done, I really miss the days we spent just laughing, chatting, and being there for each other, and I am so thankful that research brought us together in this huge world to become such great friends.

To my friends in Sydney, thank you for making this city feel like home. Even though we didn't meet because of research, your presence alone made Sydney so much more beautiful to me. I will always remember our pilates sessions, learning to dive, going surfing, and exploring every corner of Australia together. I especially cherish that 2 AM night dive in Wollongong where we caught sea urchins among the reefs. Honestly, those were the sweetest sea urchins I've ever tasted (even if that was just in my head :-). I even have a scar on my knee from that night that still hasn't faded, but it's ok, I love it. Every time I see it, I am taken back to the feeling of that freezing cold water in the middle of the night. Moments like these made my life as a researcher so colorful and happy, and I am just so glad I chose to spend these years in such a beautiful city.

I am endlessly thankful to my parents, who have always supported me with quiet strength and unconditional love. Their resilience and kindness have shaped the person I am today. To my sister, thank you for being my lifelong companion and cheerleader, no matter the distance.

I am also deeply grateful to the friends who have accompanied me through different stages of life, those from my hometown, university, lab and community. Your humor, patience, and encouragement sustained me in moments of doubt and brought joy to this journey. I feel fortunate to have been surrounded by such generous and inspiring people.

Authorship Attribution Statement

The research presented in Chapters 3, 4, 5, and 6 of this thesis is derived from publications and works in progress for which I am the lead author and have made the primary contributions. This work was made possible through the generous support of several funding bodies. My supervisor, Anusha Withana, was supported by an Australian Research Council Discovery Early Career Award from the Australian Government (DE200100479). We also gratefully acknowledge financial support from the University of Sydney’s Digital Sciences Initiative under the DSI Research Pilot Project grant scheme and from the Neurodisability Assist Trust, Australia. My doctoral research was supported by the Faculty of Engineering Research Scholarship. All studies described in these chapters were conducted with ethics approval under project number 2019/553.

Chapter 3: Hongyu Zhou, Yihao Dong, Masahiko Inami, Zhanna Sarsenbayeva, Anusha Withana. 2025. A Survey on Methodological Approaches to Collaborative Embodiment in Virtual Reality.

The paper is under review in ACM Computing Surveys (CSUR). My contributions included formulating the survey scope and research questions, designing the literature search and screening protocol, collecting and analyzing the reviewed studies, synthesizing the methodological taxonomy, and writing the manuscript drafts.

Chapter 4: Hongyu Zhou, Treshan Ayesh, Chenyu Fan, Zhanna Sarsenbayeva, Anusha Withana. 2024. CoplayingVR: Understanding User Experience in Shared Control in Virtual Reality The Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT).

This work was published in IMWUT journal and presented at the UbiComp 2024 conference. As the primary contributor, I was responsible for conceptualizing the project,

leading the system design, and overseeing the user study and data analysis. I also took the lead in manuscript preparation and the subsequent revision process.

Chapter 5: Hongyu Zhou, Xincheng Huang, Winston Wijaya, Yi Fei Cheng, David Lindlbauer, Eduardo Velloso, Andrea Bianchi, Zhanna Sarsenbayeva, Anusha Withana. 2026. One Body, Two Minds: Alternating VR Perspective During Remote Teleoperation of Supernumerary Limbs

The paper is published in the Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI). My contributions included proposing the concept, contributing to the system design, conducting the user study and data analysis, and writing the drafts.

Chapter 6: Hongyu Zhou, Tom Kip, Yihao Dong, Andrea Bianchi, Zhanna Sarsenbayeva, Anusha Withana. 2025. Juggling Extra Limbs: Identifying Control Strategies for Supernumerary Multi-Arms in Virtual Reality.

The paper was published in the Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI) and received a best paper honorable mention award. My contributions included proposing the concept, contributing to the system design, conducting the user study and data analysis, and writing the drafts.

Hongyu Zhou

Assoc. Prof. Anusha Withana

11th June 2026

Statement of Originality

I, Hongyu Zhou, hereby declare that this thesis titled, “Synergistic Embodiment: A Framework for Shared Control and Multi-Limb Interaction in Virtual Reality”, is submitted to the University of Sydney in fulfillment of the requirements for the Degree of Doctor of Philosophy.

This work is original and represents my own research, except where due acknowledgement and citation have been made. I further declare that this thesis has not been previously submitted, either in its entirety or in part, to this or any other academic institution for any degree, diploma, or other qualification.

Hongyu Zhou

11th June 2026

Statement of AI Acknowledgement

Generative artificial intelligence tools, such as ChatGPT, have been used solely for language refinement purposes, including grammar and spelling corrections, in accordance with the University of Sydney's policies. The substantive content, analysis, and interpretation are entirely my own. I take full responsibility for the submitted work and confirm its compliance with the University's guidelines and academic integrity requirements.

Hongyu Zhou

11th June 2026

Contents

Abstract	ii
Chapter 1 Acknowledgements	iv
Authorship Attribution Statement	ix
Statement of Originality	xi
Statement of AI Acknowledgement	xii
Contents	xiii
List of Figures	xxiv
Chapter 1 Introduction	1
1.1 Motivation	1
1.2 Challenges and Knowledge Gaps	3
1.3 Thesis Statement	3
Terminology.....	4
1.4 Research Questions	4
1.5 Research Contributions.....	5
Conceptual Contribution.....	5
Empirical Contribution.	5
Methodological Contribution.....	5
Design Contribution.....	5
1.6 Broader Impact	6
1.7 Thesis Outline.....	7
Chapter 2 Background and Related Work	10
2.1 Embodiment in Virtual Reality	10

2.1.1	Ownership and Agency in VR.....	11
2.1.2	Single-User VR Embodiment Practices	12
2.1.3	Shared Control & Perspective Switching	12
2.1.3.1	Collaboration Control Paradigms	12
2.1.3.2	Perspective Strategies in Collaborative VR.....	13
2.1.4	Supernumerary Limb Interaction	15
2.1.4.1	Multi-arm Mapping and Control Strategies	15
2.1.4.2	Challenges and Countermeasures in Multi-limb Interaction	17
2.1.5	Methodological Considerations	18
2.1.5.1	Key Evaluation Metrics	18
	Quantitative Measures	18
	Qualitative and Subjective Measures.....	19
2.1.5.2	Task and Scenario Design.....	20
2.1.5.3	Participant Recruitment and Experimental Validity	21
2.2	Why Existing Literatures Are Insufficient for Collaborative Embodiment.....	22

Chapter 3	A Survey on Methodological Approaches to Collaborative Embodiment in Virtual Reality	25
3.1	INTRODUCTION	25
3.2	BACKGROUND	27
3.2.1	Design Approaches for Embodiment in VR	28
3.2.1.1	Self-Representation and Avatar Embodiment.....	28
3.2.1.2	Shared Control and Multi-User Interactions.....	29
3.2.1.3	Haptic Feedback and Sensory Integration.....	30
3.2.1.4	Immersive Interaction Design	30
3.2.1.5	Social Presence and Co-Embodiment.....	31
3.2.2	Methodological Concerns in HCI and Embodiment Research	32
3.2.2.1	Participants.....	32
3.2.2.2	Study Environment.....	32
3.2.2.3	Metrics	33
3.2.2.4	Technological Limitations	33

3.3	METHODOLOGY	34
3.3.1	Search Strategy	34
3.3.2	Study Selection	35
3.4	RESULTS	36
3.4.1	Measures and Metrics	37
3.4.1.1	Subjective Measures	37
3.4.1.2	Objective Measures	39
3.4.2	Study Tasks	40
3.4.3	Collaboration Methods	43
3.4.4	Study Parameters	44
3.4.4.1	Participants	44
3.4.4.2	Technological Tools and Systems	46
3.4.4.3	Study Settings	47
3.5	DISCUSSION	48
3.5.1	Implications for Collaborative Embodiment	48
3.5.1.1	Insights into Collaborative Dynamics	48
3.5.1.2	Practical Challenges and Solutions	49
3.5.2	Methodological Reflections on Collaborative Embodiment	50
3.5.2.1	Study Design and Metrics	50
3.5.2.2	Diversity in Experimental Contexts	51
3.5.3	Technological and Methodological Gaps Specific to Collaborative Embodiment	52
3.5.3.1	Technological Advances and Applications	52
3.5.3.2	Limitations in Participant Representation	53
3.5.3.3	Inconsistencies in Methodological Standards	54
3.5.4	Future Directions for Collaborative Embodiment Research	55
3.5.4.1	Expanding Experimental Scope	55
3.5.4.2	Enhancing Inclusivity and Scalability	56
3.6	LIMITATIONS	56
3.7	CONCLUSION	57

3.8	How the Survey Motivates the thesis's Conceptual Architecture	58
Chapter 4	CoplayingVR	61
4.1	Abstract	62
4.2	Introduction	63
4.3	RELATED WORK	65
4.3.1	Collaborative and Multiplayer Games	65
4.3.2	Shared Control in Video Games	66
4.3.3	Shared Body Experiences	67
4.3.4	Measuring Experience	67
4.4	DESIGN and implementation of 'CoplayingVR'	68
4.4.1	Input Control Mechanism	69
4.4.2	Rendering Visual Output	70
4.4.3	Implementation Approach	70
4.5	STUDY DESIGN	71
4.5.1	VR Game Tasks	71
4.5.1.1	Control Task	71
4.5.1.2	Balancing Task	72
4.5.1.3	Shooting Task	73
4.5.2	Experimental Setup	74
4.5.3	Experimental Data	74
4.5.3.1	Player Experience	74
4.5.3.2	Performance	74
4.5.3.3	Agency and body ownership	75
4.5.4	Experimental Conditions and Procedure	76
4.5.4.1	Participants	76
4.5.4.2	Experimental Conditions	76
4.5.4.3	Procedure	76
4.6	RESULTS	78
4.6.1	Effects of Playing Mode on Game Experience	78
4.6.1.1	Control Task (Figures 4.4 and 4.3):	79

4.6.1.2	Balancing Task (Figures 4.4 and 4.3):	79
4.6.1.3	Shooting Task (Figures 4.4 and 4.3):	80
4.6.1.4	Overall Gaming Experience	81
4.6.2	User Performance – Prior Experience and Task Completion Time.....	82
4.6.2.1	Control Task – Task completion time for mixed pair group:.....	83
4.6.2.2	Control Task – Task completion time for experienced players and novice players groups:	83
4.6.2.3	Balancing Task – Task completion time for mixed pair group:	84
4.6.2.4	Balancing Task – Task completion time for experienced players and novice players groups:	84
4.6.2.5	Shooting Task – Error Rate for mixed pair group:	84
4.6.2.6	Shooting Task – Error Rate for experienced players and novice players groups:	85
4.6.2.7	Behavioral Observations:	85
4.6.3	Agency & Body Ownership	86
4.6.4	Qualitative Results.....	88
4.6.5	Summary of Results	91
4.7	Discussion	91
4.7.1	Task-Specific Effects of Shared Control.....	92
4.7.1.1	The Control Task	92
4.7.1.2	The Balancing Task	92
4.7.1.3	The Shooting Task	93
4.7.2	Shared Control as Collaborative Mechanism and Scaffolding Strategy ...	94
4.7.3	Broader Implications Beyond Gaming	95
4.7.4	Design Implications	95
4.7.5	Alternative Configurations and Future Directions in Shared Control.....	96
4.8	Limitations and Future Work.....	97
4.9	Conclusion.....	98
4.9.1	Contribution to the Thesis	99

5.1	Abstract	101
5.2	Introduction	102
5.3	RELATED WORK	105
5.3.1	Perspective Control	105
5.3.2	Embodiment in Collaboration	106
5.3.3	Supernumerary Robotic Limbs	107
5.4	FORMATIVE STUDY	108
5.4.1	Method	109
5.4.2	Findings	110
5.4.2.1	Guest Control Drops Without Perspective Flexibility (5/5 dyads). . .	110
5.4.2.2	Hosts Stop Moving Due to Role Confusion (3/5 dyads).	111
5.4.2.3	Prolonged Shared View Causes Cybersickness (5/5 dyads).	111
5.4.2.4	Camera Shifts Disrupt Guest's VSL Control (4/5 dyads).	112
5.4.2.5	Summary of Findings.	112
5.5	DESIGN AND IMPLEMENTATION	112
5.5.1	Embedded Anchored View (C1)	113
5.5.2	Out-of-body View (C2)	114
5.5.3	Perspective Switching Mechanism	115
5.5.4	VSLs Independent Control and Stabilization	115
5.5.5	Implementation Details	117
5.6	EXPERIMENTAL METHOD	117
5.6.1	Experimental Conditions and Procedure	118
5.6.1.1	Experimental Conditions	118
5.6.1.2	Participants	118
5.6.2	Experimental Tasks	118
5.6.2.1	Transportation Task	118
5.6.2.2	Factory Task	119
5.6.3	Experimental Data Collection	120
5.6.3.1	Qualitative Data	120
5.6.3.2	Performance Measures	121

5.6.3.3	Subjective Feedback	121
5.6.3.4	Physiological Data - Heart Rate Monitoring	122
5.6.3.5	Procedure	122
5.6.4	Data Analysis	123
5.7	RESULTS	124
5.7.1	Physiological Responses Measures (RQ1)	125
5.7.2	Workload and Fatigue Analysis	126
5.7.3	Embodiment Questionnaire Analysis (RQ2)	128
5.7.4	Task Performance Analysis (RQ3)	129
5.7.5	Perspective Switching Behaviors (RQ4)	131
5.7.6	Qualitative Findings	132
5.7.6.1	Role-asymmetric Embodiment Trade-offs	132
5.7.6.2	Importance of Gaming Experience over VR Experience	133
5.7.6.3	Strategic Minimization of Perspective Switching	134
5.7.6.4	Emergence of Natural Gesture-Based Communication Strategies	135
5.8	DISCUSSION	135
5.8.0.1	Reducing Workload and Physiological Strain through Perspective Design (RQ1)	136
5.8.0.2	Balancing Embodiment and Spatial Awareness: Role-specific Considerations (RQ2)	136
5.8.0.3	Optimizing Collaborative Performance through Adaptive Perspective Switching (RQ3, RQ4)	138
5.8.1	Implications for the Design of Collaborative Teleoperation Systems	138
5.8.2	Limitations and Future Work	139
5.9	CONCLUSION	141
5.9.1	Contribution to the Thesis	141
Chapter 6	Juggling Extra Limbs	143
6.1	Abstract	144
6.2	INTRODUCTION	144
6.3	RELATED WORK	147

6.3.1	Virtual Reality and Robotic Body	147
6.3.2	Supernumerary Robotic Limbs	148
6.3.3	Mapping Control Strategies in SRL	149
6.3.4	Wizard-of-Oz Methodology in Supernumerary Robotic Limbs Research	150
6.4	DESIGN AND IMPLEMENTATION	150
6.4.1	Design Considerations	151
6.4.2	Implementation	152
6.5	EXPERIMENTAL METHOD	153
6.5.1	Experimental Conditions and Procedure	153
6.5.1.1	Participants	153
6.5.1.2	Experimental Conditions	153
6.5.1.3	Procedure	155
6.5.2	Experimental Tasks	156
6.5.2.1	Basic Control Task	156
6.5.2.2	Factory Task	157
6.5.3	Experimental Setup	158
6.5.4	Experimental Data	159
6.5.4.1	Qualitative data	159
6.5.4.2	Performance	159
6.5.4.3	Embodiment Experience	159
6.6	RESULTS	159
6.6.1	Quantitative Findings	160
6.6.2	Qualitative Findings	162
6.6.2.1	“It felt weird, like they were someone else’s arms” — Body Ownership, Agency, and Control Embodiment	163
6.6.2.2	“I felt like I was juggling!” — Control Strategies and Their Perceived Performance	165
6.6.2.3	Summary	167
6.7	DISCUSSION	168

6.7.1	Implications for Control strategies and VSLs Control.....	169
6.7.2	Optimizing Control and Performance in Multiple VSLs Systems.....	171
6.7.3	Challenges and Opportunities of Non-Fixed Mapping	173
6.8	Limitations and Future Work.....	174
6.9	CONCLUSION	175
Chapter 7	Discussion	178
7.1	Limits of Existing Approaches	178
7.2	Design Tensions Across the Empirical Studies	180
7.3	Interaction Elements Matrix	182
7.3.1	Primary Design Dimensions	183
7.3.2	Cross-Cutting Moderators	184
7.3.3	Interaction Elements as Configurations	185
7.4	The Explanatory Value of the IEM	186
7.5	Research Propositions Emerging from the IEM	188
7.6	Addressing the Research Questions.....	189
7.6.1	RQ1: How do different visual perspectives and control mechanisms influence user performance, experience (including comfort and workload), and the sense of embodiment in collaborative VR tasks involving shared avatars and supernumerary limbs?.....	190
7.6.2	RQ2: What interaction strategies do users develop and prefer when co-controlling a single avatar, especially with varying levels of limb autonomy or when distinct user roles are implied by the control scheme? .	191
7.6.3	RQ3: How can insights from these user experiences be synthesized into a design framework to guide the creation of more effective and intuitive collaborative VR systems?.....	192
7.7	Methodological Discussion	193
7.7.1	Mixed Methods as a Requirement Rather Than an Option.....	193
7.7.2	Progressive Complexity as an Analytic Device.....	194
7.7.3	Prototypes as Instruments of Conceptual Discovery	195
7.7.4	Methodological Implications for Future Research.....	196

7.8	Boundary Conditions of the Framework	196
7.9	Limitations and Future Directions	197
7.9.1	Limitations	198
	Methodological reporting and published-chapter constraints.....	198
	Scope of the Survey.....	198
	Experimental Contexts.....	199
	Participant Diversity.....	199
	Wizard-of-Oz Autonomy.....	199
7.9.2	Future Directions	199
	Real-World and Longitudinal Studies.....	199
	Adaptive and Intelligent Systems.....	200
	Multi-Modal Embodiment.....	200
	Scalability to Larger Groups.....	200
	Cultural and Demographic Variation.....	200
	Physical Supernumerary Limbs and Hybrid Systems.....	200
	Framework Generalization and Refinement.....	201
7.10	Conclusion.....	201
Chapter 8	Conclusion	203
Appendix A	Proofs in Chapter 5	206
1	Physiological Data Processing Details	206
Appendix A	Proofs in Chapter 6	208
1	Autonomy level details.....	208
1.1	Low-autonomy level.....	208
1.2	High-autonomy level	209
2	Operator Action Guidelines	209
2.1	Command Instructions.....	210
2.2	Demonstration Instructions.....	210
2.3	Delegation Instructions	210
2.4	Labeling Instructions	210

Appendix Bibliography	212
Bibliography	212
3 Appendix A.....	260
Appendix A Proofs in Chapter 5	261
1 Physiological Data Processing Details	261
Appendix A Proofs in Chapter 6	263
1 Autonomy level details.....	263
1.1 Low-autonomy level.....	263
1.2 High-autonomy level	264
2 Operator Action Guidelines	264
2.1 Command Instructions	265
2.2 Demonstration Instructions.....	265
2.3 Delegation Instructions	265
2.4 Labeling Instructions	265

List of Figures

1.1	Conceptual overview of multi-user, multi-limb embodiment in collaborative virtual environments. Subfigure (a) illustrates a scenario in which multiple users collectively control a single avatar, coordinating several virtual arms toward a shared objective. Subfigure (b) depicts parallel collaboration, where each user independently operates their own avatar and set of virtual limbs while working toward a common goal. Subfigure (c) shows a mixed-control configuration in which a user supervises both directly controlled and semi-autonomous supernumerary limbs, highlighting the continuum between manual control and delegated agency. Together, these scenarios illustrate the design space and research challenges addressed in this thesis, including coordination, agency, embodiment, and system feedback.	2
3.1	The PRISMA flow diagram illustrating the stages of our systematic search, screening, and selection process.	36
3.2	Objective Evaluation Matrix for Methods of Collaboration	39
3.3	Frequency Of Task Types In Collaborative VR Research	44
3.4	Participant Demographics Across Study Contexts	46
4.1	Shared control concept and implementation. (a) Two users are able to independently control different limbs of the same avatar (b) by sharing the local input/output control through a networked server.	62
4.2	The three gaming tasks used in the user study. Each hand is controlled by a separate player. (a) Control task: Two hands are used to move a peg into a hole. (b) Balancing task: Two hands are used to tilt the maze to drive the ball to the correct path. (c) Shooting task: One player controls the bow to aim, the other shoots the arrow. The hand models used in this study were downloaded from the Oculus Developer website. The bow and arrow models were sourced from Quaternius.	72

- 4.3 The averages of Likert ratings in shared control and single-player modes on the first three game experience measures. The three conditions in each graph are Control Task, Balancing Task and Shooting Task. * denotes $p < .05$. 78
- 4.4 The averages for Likert scale ratings for game experience measures in shared control and single-player modes on the three conditions Control Task, Balancing Task, and Shooting Task. * indicates $p < .05$ and ** indicates $p < .01$. 80
- 4.5 The figure shows the averages for: (a) the task completion time and (b) the averages for the error rates in the experienced players group; (c) the task completion time and (d) the error rate in the novice players group; and, (e) the task completion time and (f) the error rate in the mixed pair group. * indicates $p < .05$ and ** is $p < .01$. 86
- 5.1 Illustration of three perspectives for remote collaborative teleoperation using virtual supernumerary limbs (VSLs): (a) *Host* user physically walking with VSLs attached to their back; (b) SHARED EMBODIED VIEW: the *guest* shares the *host*'s position with independent head rotation; (c) EMBEDDED ANCHORED VIEW: the *guest* views from the *host*'s head location through a stabilized virtual window; (d) OUT-OF-BODY VIEW: the *guest* independently navigates a drone-like external viewpoint using a handheld controller (the drone serves as a visual metaphor for the decoupled virtual camera). 100
- 5.2 Formative study configuration showing (a) the stable first-person perspective from the *host*, (b) the *guest*'s co-located but independently rotatable view illustrating visual instability, and (c) the avatar embodiment with VSLs used for near-body manipulation. 109
- 5.3 Illustration of the EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW. (a) The *guest* user's viewpoint is spatially aligned with the *host*'s viewpoint but rendered as a video-like window. (b) The *guest* user independently controls a virtual drone camera with 6 degrees-of-freedom (6-DoF), enabling flexible positional and rotational adjustments independent of the *host*'s movements (the drone serves as a visual metaphor for the decoupled virtual camera). 114
- 5.4 Guest perspectives organized as an *Anchor–View* matrix relative to the host. (a) **Locked Egocentric:** *same anchor, same view*; guest camera is rigidly locked to

- the host's head pose and view. (Shown for comparison; not studied, as it grants the guest no agency and renders them a passive observer under full host control). (b) **SHARED EMBODIED VIEW**: *same anchor, different view*; guest shares the host's position but rotates the camera independently. (c) **EMBEDDED ANCHORED VIEW**: *different anchor, same view*; guest remains independently positioned but views a stabilized window of the host's egocentric feed. (d) **OUT-OF-BODY VIEW**: *different anchor, different view*; guest drives a free-floating 6-DoF drone camera, decoupled from the host. 116
- 5.5 Task overview: Transportation and Factory. Panels (a–b) depict the *Transportation* task: (a) colour-coded boxes placed on a central table; (b) matching colour-coded bins distributed throughout the VR environment. Panels (c–d) depict the *Factory* task: (c) geometrically distinct holes on a wall panel as precise insertion targets; (d) collaborative insertion of objects (e.g., cubes) into the corresponding holes. 119
- 5.6 Physiological responses for *hosts* and *guests* under EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW across two different tasks. (a) HRV (RMSSD), a time-domain index of heart-rate variability. (b) Peak heart rate (bpm) by role and task. (c) Event-triggered heart rate aligned to perspective switches from the SHARED EMBODIED VIEW to the target view (SHARED EMBODIED VIEW → EMBEDDED ANCHORED VIEW or SHARED EMBODIED VIEW → OUT-OF-BODY VIEW). * indicates $p < .05$ and ** indicates $p < .01$. 125
- 5.7 Subjective measures (VAS-F and NASA-TLX) for *hosts* and *guests* across tasks: (a) *Host* scores for Factory Task, (b) *Guest* scores for Factory Task, (c) *Host* scores for Transportation Task, and (d) *Guest* scores for Transportation Task. Scores are compared between EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW. * indicates $p < .05$, ** indicates $p < .01$ and *** indicates $p < .001$. 128
- 5.8 Mean embodiment scores ($\pm$ SE) across three perspective conditions (EMBEDDED ANCHORED VIEW, OUT-OF-BODY VIEW, and SHARED EMBODIED VIEW), separated by user role (Host and Guest). Subplots represent AEQ factors: (a) Agency, (b) Body Ownership, (c) Presence/Identification, and (d) Sensory

Integration/Self-location. * indicates $p < .05$, ** indicates $p < .01$ and *** indicates $p < .001$.

128

- 5.9 Effects of perspective mode and task phase. (a) Completion time: OUT-OF-BODY VIEW led to significantly faster task completion in the Transportation task. (b) Error count: OUT-OF-BODY VIEW resulted in significantly fewer errors in the Factory task. (c) Switch frequency by task phase. Switching was infrequent in the baseline SHARED EMBODIED VIEW condition (median ≤ 1 across phases). In comparison, participants switched more frequently to EMBEDDED ANCHORED VIEW during precision subtasks, and to OUT-OF-BODY VIEW during navigation and remote-response phases. * indicates $p < .05$, ** $p < .01$, and *** $p < .001$. 130
- 6.1 Wizard-of-Oz approach and experimental tasks. (a) The participants' VSLs were controlled by a human operator behind a partition to simulate semi-autonomous supernumerary limbs. (b) Basic Control Task: In this task, participants used both their own arms and the VSLs to press buttons. (c) Factory Task: Participants used their own arms and the VSLs to collaboratively insert found objects into the appropriate holes. 145
- 6.2 Two tasks used in the user study: the two avatar's limbs were controlled by the participants, while the VSLs were controlled by a human operator. (a) Basic Control Task: Participants used the avatar's limbs and VSLs to simultaneously actuate illuminated buttons in a sequence. (b) Factory Task: One hand grasps a falling target, while the other rotates a Shape Sorter to match and insert the target into the correct hole based on shape and color. 157
- 6.3 The figure shows the averages for *low-autonomy* level and *high-autonomy* levels in (a) the task completion time in the basic control task and factory task; (b) the error rates in the basic control task and factory task; (c) the embodiment questionnaire. * indicates $p < .05$ and ** indicates $p < .01$. 160
- 6.4 The figure illustrates the participant-issued instructions, including **Command Instructions**, **Demonstration Instructions**, **Delegation Instructions** and **Labeling Instructions**, along with the WoZ operator's reaction times to these four instruction types. A single colored segment may represent multiple instructions of the same

	type. It also compares the reaction times across two distinct autonomy levels for both the Basic Control Task and the Factory Task.	162
6.5	This figure shows the flow of the follow and repeat strategies performed by participants.	170
6.6	This figure shows the organizing strategies performed by participants.	171
7.1	Interaction Elements Matrix (IEM): a first-order design space for collaborative embodiment in VR, organized around control mapping, perspective strategy, feedback design, and autonomy level. The matrix summarizes the principal dimensions surfaced across Chapters 4–6 and relates the studied prototypes to adjacent design possibilities.	186

CHAPTER 1

Introduction

1.1 Motivation

Most HCI and VR systems are built around a familiar assumption: one user controls one body through a unified loop of perception and action. Emerging collaborative VR systems increasingly challenge this assumption. Users may co-control a shared avatar, operate supernumerary limbs, or collaborate through distributed viewpoints and semi-autonomous embodied interfaces. In such settings, control, perception, and agency are no longer contained within a single user-body pair, but are distributed across multiple users, virtual limbs, and system processes. As a result, collaborative embodiment is not simply a novel application scenario in VR. It represents a distinct design problem that requires HCI to reconsider how embodied interaction is structured, negotiated, and evaluated.

This problem has become increasingly urgent because advances in full-body tracking, rendering, and low-latency networking now make such systems technically viable in practice. Full-body tracking systems such as HybridTrak support real-time pose reconstruction with lightweight hardware, lowering barriers to embodied interaction [1]. Rendering breakthroughs, including real-time path tracing methods such as ReSTIR, enable more visually convincing and immersive environments [2]. At the same time, progress in low-latency networking makes synchronous multi-user VR increasingly feasible, with recent work showing how latency directly shapes collaboration quality and user experience [3].

These developments have made possible new forms of interaction in which multiple users can jointly inhabit and operate shared virtual bodies, collaboratively manipulate objects, and

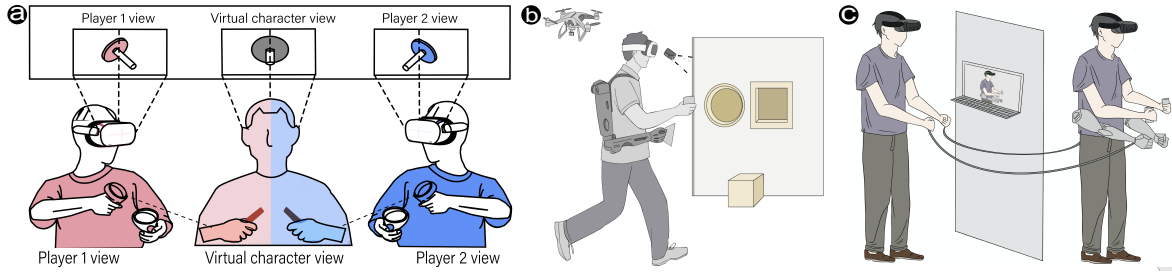


FIGURE 1.1. Conceptual overview of multi-user, multi-limb embodiment in collaborative virtual environments. Subfigure (a) illustrates a scenario in which multiple users collectively control a single avatar, coordinating several virtual arms toward a shared objective. Subfigure (b) depicts parallel collaboration, where each user independently operates their own avatar and set of virtual limbs while working toward a common goal. Subfigure (c) shows a mixed-control configuration in which a user supervises both directly controlled and semi-autonomous supernumerary limbs, highlighting the continuum between manual control and delegated agency. Together, these scenarios illustrate the design space and research challenges addressed in this thesis, including coordination, agency, embodiment, and system feedback.

coordinate actions in real time [4, 5]. Prior work has shown that shared augmented and virtual environments can support distributed collaboration by improving situational awareness and joint task performance [6, 7]. However, these capabilities also expose unresolved tensions among embodiment, teamwork, comfort, autonomy, and role allocation. Systems that improve coordination may weaken agency. Perspectives that strengthen immersion may reduce awareness. Semi-autonomous limbs may increase efficiency while destabilizing users' sense of control.

These tensions are especially visible in domains such as industrial training, healthcare, remote collaboration, and gaming. In such settings, users must coordinate under asymmetric roles, changing task demands, and partial control over shared embodied resources. What matters is not only whether users can act in VR, but how control should be distributed, when viewpoints should shift, how autonomy should be calibrated, and under what conditions shared virtual bodies can remain both usable and experientially coherent. These questions motivate the central concern of this thesis: how collaborative embodiment can be designed as a principled, rather than ad hoc, form of interaction.

1.2 Challenges and Knowledge Gaps

Despite this promise, realizing effective multi-user, multi-limb collaboration in VR remains challenging. Existing systems often rely on rigid and fixed control mappings, which limit users' ability to negotiate roles, redistribute effort, or share virtual limbs dynamically during demanding tasks [8, 9]. Such rigidity can increase cognitive load, produce input conflicts, and make collaboration brittle when task demands shift unexpectedly.

A second challenge concerns perspective. Many collaborative VR systems provide only limited or static viewpoint configurations, constraining users' ability to monitor joint action, maintain situational awareness, or balance local embodiment with global coordination [10, 11]. Poorly designed viewpoint strategies can reduce both collaboration quality and user comfort, especially in mobile or role-asymmetric scenarios.

A third challenge lies in the integration of supernumerary limbs and semi-autonomous control. While additional virtual limbs expand users' action space, they also introduce new problems of delegation, transparency, and embodiment. Users may struggle to determine what is under direct control, what is being handled autonomously, and how these different forms of control should be coordinated in real time.

Taken together, these issues point to a broader gap in the literature. Although prior work has studied shared control, viewpoint design, and limb augmentation separately, there remains limited systematic understanding of how these factors interact in collaborative VR. In particular, the field still lacks an integrated account of how control mapping, perspective, feedback, and autonomy should be configured to support both effective teamwork and meaningful co-embodiment in realistic, task-driven settings.

1.3 Thesis Statement

This thesis argues that **collaborative embodiment** should be understood as a unified Human–Computer Interaction problem rather than as a collection of isolated interaction techniques.

In collaborative VR, control, perception, and agency are no longer confined to a single user-body loop. Instead, they are distributed across multiple users, viewpoints, and virtual limbs, requiring systems to support the ongoing negotiation of shared action, role allocation, and embodied coordination. To address this challenge, this thesis proposes **Synergistic Embodiment** as a conceptual lens for understanding and designing collaborative embodiment, framing it as a dynamic configuration of shared control, perspective, feedback, and autonomy that can be studied through survey research, system design, and empirical evaluation.

Terminology. In this thesis, *collaborative embodiment* refers to the broader problem space of multi-user embodied interaction in VR, where control, perception, and agency may be distributed across multiple users, viewpoints, and virtual limbs. *Synergistic Embodiment* refers to the conceptual lens proposed in this dissertation for understanding this problem space. The *Interaction Elements Matrix* (IEM) is the design-space representation through which this lens is operationalized. Where used, the *Synergistic Embodiment Framework* (SEF) refers to the set of design principles derived from the empirical findings and organized through this design space.

1.4 Research Questions

Based on the gaps identified above, this thesis investigates the following research questions:

- **RQ1:** How do different visual perspectives and control mechanisms influence user performance, user experience, including comfort and workload, and the sense of embodiment in collaborative VR tasks involving shared avatars and supernumerary limbs?
- **RQ2:** What interaction strategies do users develop and prefer when co-controlling a single avatar, especially with varying levels of limb autonomy or when distinct user roles are implied by the control scheme?

- **RQ3:** How can insights from these user experiences be synthesized into a design framework to guide the creation of more effective and intuitive collaborative VR systems?

1.5 Research Contributions

This thesis makes four primary contributions.

Conceptual Contribution. This thesis advances **Synergistic Embodiment** as a conceptual lens for understanding collaborative embodiment in VR. This lens reframes collaborative VR not simply as a problem of multi-user access or augmented control, but as a design problem in which agency, control, and perception are jointly negotiated across users and virtual limbs.

Empirical Contribution. This thesis presents three empirical investigations that examine complementary dimensions of collaborative embodiment. *CoplayingVR* studies shared control and role allocation in a gaming context. *One Body, Two Minds* investigates perspective switching and role-asymmetric collaboration with virtual supernumerary limbs. *Juggling Extra Limbs* explores how users coordinate with semi-autonomous virtual arms. Together, these systems reveal recurring trade-offs among coordination cost, embodiment, situational awareness, and autonomy.

Methodological Contribution. This thesis demonstrates the value of combining systematic review, system prototyping, controlled user studies, and mixed-method evaluation in the study of collaborative embodiment. Across the thesis, quantitative performance measures, subjective embodiment and workload scales, behavioural observations, and qualitative strategy analyses are used to capture the multi-dimensional nature of embodied collaboration.

Design Contribution. This thesis proposes the *Interaction Elements Matrix (IEM)* as a first-order design framework for reasoning about collaborative embodiment in VR. The IEM synthesizes the principal design dimensions surfaced across the empirical studies and provides

a structured way to think about control mapping, perspective strategy, feedback design, and autonomy in multi-user, multi-limb systems.

1.6 Broader Impact

The studies reported in this thesis show that collaborative VR systems are most effective when they support flexible and intelligible coordination among users. Across the three empirical systems, several recurring findings emerged. First, shared control can improve novice performance and engagement, but only when role ambiguity and coordination overhead are carefully managed. Second, perspective flexibility is essential for balancing immersion, awareness, and comfort in dynamic collaborative tasks. Third, supernumerary limbs become more usable when autonomy is legible and users can delegate tasks strategically rather than relying solely on direct manipulation.

These findings have direct implications for the design of future collaborative VR systems. Designers should support dynamic and role-sensitive control allocation, rather than assuming that fixed mappings will remain appropriate across task phases. Systems should provide perspective configurations that allow users to move between local embodied action and broader collaborative awareness without incurring excessive disorientation. They should also make autonomous or semi-autonomous behaviour understandable and predictable, so that users can coordinate with additional limbs as collaborative extensions rather than opaque system actions.

More broadly, this thesis contributes to HCI by framing collaborative VR as a problem of embodied coordination, shared agency, and role negotiation. Its findings are relevant not only to social VR and gaming, but also to teleoperation, remote training, assistive and rehabilitation technologies, and human–robot collaboration. By showing how shared virtual bodies can be designed, negotiated, and experienced, this thesis contributes to a broader understanding of how future interactive systems may support distributed action, shared agency, and embodied cooperation.

1.7 Thesis Outline

This thesis is structured into three main parts, tracing a progression from theoretical grounding, through empirical investigation, to conceptual synthesis.

Chapter 2 reviews relevant literature and theoretical foundations. It defines key concepts such as embodiment, agency, and ownership in virtual environments, and situates collaborative embodiment within the broader bodies of VR and HCI research. It also reviews prior work on shared control, perspective modulation, and supernumerary limb interaction.

Chapter 3 presents a systematic survey of collaborative embodiment research using the PRISMA framework. This chapter details the literature search process, inclusion criteria, and coding scheme, and provides an overview of tasks, control schemes, visual perspectives, and evaluation methods used across prior work. The survey identifies methodological and conceptual gaps that motivate the empirical investigations in later chapters.

Chapter 4 introduces *CoplayingVR*, a system in which two users jointly control a shared avatar. Through gaming tasks and mixed-method evaluation, this chapter examines how shared control affects performance, game experience, agency, and body ownership, and how coordination costs and scaffolding benefits emerge under shared mapping.

Chapter 5 presents *One Body, Two Minds*, a perspective-switching system for remote collaboration using virtual supernumerary limbs. This chapter compares first-person, embedded-anchored, and out-of-body perspectives, showing how viewpoint strategy shapes cognitive load, spatial awareness, comfort, and collaborative coordination.

Chapter 6 examines *Juggling Extra Limbs*, a Wizard-of-Oz exploration of semi-autonomous virtual arms in a factory-style task scenario. It investigates how users issue commands, delegate tasks, and coordinate with additional limbs under different levels of autonomy, revealing the emergence of supervisory control strategies.

Chapter 7 synthesizes findings from the survey and empirical studies into the *Interaction Elements Matrix* and the broader perspective of **Synergistic Embodiment**. It discusses how control mapping, perspective, feedback, and autonomy can be organized as a first-order design space for collaborative embodiment in VR, and reflects on the methodological implications of this work.

Chapter 8 concludes the thesis by revisiting its primary contributions and outlining future directions for collaborative embodiment research, including adaptive systems, longitudinal studies, and the integration of physical robotic limbs with virtual collaborative embodiment.

Taken together, these chapters present a coherent account of collaborative embodiment as a distinct and consequential design problem in immersive systems. By moving from literature synthesis to system design, empirical study, and conceptual abstraction, this thesis offers both foundational understanding and practical guidance for the design of future multi-user, multi-limb VR experiences.

PART I

Background and Theoretical Grounding

CHAPTER 2

Background and Related Work

This chapter reviews the intellectual foundations of this thesis with a more specific aim than broad coverage alone. Rather than treating embodiment, shared control, perspective switching, supernumerary limbs, and methodology as separate topics, it argues that they jointly define the design problem of collaborative embodiment in VR. Existing work has made important progress on each of these dimensions, but has often examined them in isolation. This chapter therefore not only summarizes the relevant literature, but also identifies the conceptual gap that motivates this thesis: the lack of an integrated account of how control allocation, perspective configuration, feedback legibility, and autonomy calibration interact in multi-user, multi-limb VR.

2.1 Embodiment in Virtual Reality

Embodiment is one of the central concepts in VR research because it explains how users come to experience a virtual body as their own and how this experience shapes interaction, performance, and presence [12, 13, 14, 15, 16, 17]. Early VR embodiment research focused primarily on single-user settings, where self-location, body ownership, and agency were studied as properties of the relationship between one user and one virtual body [18, 19]. This work established the perceptual and technical foundations for embodied interaction, showing how sensorimotor contingencies, visual realism, and temporal congruence influence users' connection to virtual bodies.

Collaborative embodiment extends these concerns into a more complex interaction setting. In multi-user VR, embodiment is no longer only about whether a user feels that a virtual body is

“mine,” but also about how control, perception, and agency may be distributed, negotiated, or shared across multiple users, viewpoints, and virtual limbs. This shift makes collaborative embodiment more than a simple extension of single-user VR. It becomes a distinct design problem that requires attention not only to ownership and agency, but also to coordination, role allocation, distributed perception, and autonomy.

2.1.1 Ownership and Agency in VR

The sense of *ownership* refers to the subjective experience that a virtual limb or body is part of one’s own physical self, a phenomenon first systematically studied in the Rubber Hand Illusion paradigm [20]. In VR, ownership is commonly assessed via self report scales adapted from embodiment questionnaires (e.g., “I felt as if the virtual hand were my own”) [21, 22]. This experience relies on tight sensorimotor contingencies provided by head mounted displays (HMDs) and hand tracking systems, which synchronize visual, tactile, and proprioceptive feedback to produce the illusion of a unified body.

The sense of *agency* denotes the feeling of control over the actions of that virtual limb or body and is measured by items such as “I felt I was controlling the virtual hand” and “Movements of the virtual hand were caused by my own movements” [23]. Agency emerges from the temporal and spatial congruence between a user’s motor commands, captured via hand controllers or trackers, and the resulting virtual movements, and is critical for fluent interaction in both single user and collaborative VR contexts.

While ownership and agency are foundational constructs for embodiment research, they were largely developed in single-user settings. These concepts explain how users relate to a virtual body, but they do not by themselves explain how agency becomes shared, how ownership is negotiated across collaborators, or how embodiment changes when control is distributed across multiple users or semi-autonomous limbs. These limitations make them necessary but insufficient for understanding collaborative embodiment.

2.1.2 Single-User VR Embodiment Practices

Single user VR embodiment systems typically employ full body tracking via optical motion capture or inertial sensors to render the user’s movements onto a virtual avatar in real time. *Body substitution* techniques replace real limbs with virtual counterparts, creating consistent visuomotor contingencies that dramatically enhance presence and ownership, studies report up to a 25% reduction in task completion time and significant increases in reported embodiment when visual, proprioceptive, and tactile feedback are tightly synchronized [20, 22]. *Mirror mapping* approaches extend these effects by reflecting the user’s motions through a virtual mirror or mirrored avatar, aligning visual and proprioceptive cues and improving spatial awareness and motor precision in tasks such as object placement and locomotion [21].

Implementations based on high fidelity tracking (sub 20 ms end to end latency) minimize temporal discrepancies between intention and action, yielding both faster objective performance and higher subjective immersion ratings [22]. Moreover, adaptive smoothing and prediction algorithms further reduce jitter, ensuring that small hand tremors or tracking noise do not break the illusion of ownership.

These single user insights directly inform our approach to collaborative embodiment. In multi user scenarios, such as the shared body mappings explored in Chapter 4, we must preserve the same strict sensorimotor alignment across multiple input channels and devices. Similarly, maintaining low latency, congruent feedback is critical when users co-control supernumerary limbs (Chapter 6) or switch perspectives dynamically (Chapter 5). Thus, the techniques and performance benchmarks established in single user VR serve as the technical and experiential baseline from which we build our shared control systems.

2.1.3 Shared Control & Perspective Switching

2.1.3.1 Collaboration Control Paradigms

In multi user VR, two dominant shared control paradigms have emerged. In the *parallel collaboration* paradigm, each participant independently controls their own virtual limbs and

coordinates actions via shared workspace cues or verbal communication. This approach maintains individual agency and clear control boundaries, but can introduce significant synchronization overhead when tasks demand tight timing or precision, as demonstrated in collaborative object passing tasks where misaligned hand trajectories led to increased error rates and verbal coordination delays [24, 25].

By contrast, the *co embodiment* paradigm merges inputs from multiple users into a single virtual body. For example, two users may each control two of four hands, collectively operating one avatar. This shared body mapping can foster a stronger sense of “we” and streamline division of labor, participants often report heightened engagement and a unified action experience, but it also raises challenges of input arbitration, conflict resolution, and role clarity [26].

These paradigms inform our CoplayingVR prototype (Chapter 4). There, we implement co embodiment by assigning User A to the avatar’s left hands and User B to the right, then compare performance and subjective experience against a parallel configuration where each user independently controls two hands. Our findings reveal that while parallel setups minimize input conflicts, they increase the verbal coordination burden; co embodiment reduces the need for verbal cues but requires robust arbitration mechanisms to resolve simultaneous grasp attempts and preserve users’ sense of agency.

This literature establishes the importance of control allocation and conflict management in multi-user VR. However, it remains largely control-centric: it focuses on who controls what, but gives less attention to how viewpoint, feedback, and autonomy reshape the experience of shared agency and collaboration. As a result, shared control is often treated as a coordination problem alone, rather than as part of a broader problem of collaborative embodiment.

2.1.3.2 Perspective Strategies in Collaborative VR

Visual perspective is a central design factor in collaborative VR because it shapes how users perceive their own virtual body, their collaborators, and the shared task space. In collaborative

embodiment, perspective is not merely a display parameter. It directly influences situational awareness, workload, coordination, and the quality of embodied experience.

Prior work commonly distinguishes three perspective configurations. In a *head-coupled first-person view*, the virtual camera follows the user's head orientation, providing strong sensorimotor alignment and a high sense of immersion, but often limiting visibility of collaborators' actions and the broader workspace context [27, 28]. An *embedded-anchored view* augments the egocentric display with a secondary inset or picture-in-picture window that presents an allocentric overview or a partner's viewpoint, thereby balancing immersion with improved situational awareness and coordination support [29, 30]. An *out-of-body* or *third-person view* positions the camera at a distance from the avatar, offering a broader overview of spatial relationships and shared goals, but often at the cost of reduced first-person presence and ownership [31, 32].

Empirical studies show that perspective choice strongly affects both task performance and user experience. First-person views preserve tight coupling between movement and visual feedback, which helps sustain embodiment, but can increase the need for verbal coordination when collaborators or task-relevant objects fall outside the user's field of view [21]. Embedded-anchored views can improve task accuracy and anticipation in joint manipulation tasks by allowing users to monitor both local action and global context without fully breaking immersion [33]. Third-person perspectives can support global coordination in complex collaborative tasks by revealing occluded objects, partner actions, and shared spatial structure, although they may reduce subjective presence and ownership [34].

The literature further suggests that no single perspective is optimal across all collaborative scenarios. Instead, the most effective configuration often depends on task demands. In fine-grained assembly, navigation, and remote manipulation tasks, third-person or scene-focused views can improve visibility of partner actions and reduce coordination errors [35]. Conversely, tasks requiring shared attention or close embodied alignment, such as collaborative reading or joint design, may benefit from synchronized first-person views that preserve immersion and local focus [36]. In gaming and other dynamic scenarios, decoupling head movement from avatar control or switching between perspectives can also reduce discomfort while

maintaining engagement [37, 38]. Some studies further indicate that perspective switching can preserve task performance and embodiment without substantially diminishing ownership over the virtual body [39].

Taken together, these findings show that perspective in collaborative VR is not simply a matter of camera placement, but a mechanism for distributing perception across users and across phases of a task. However, existing work remains largely camera-centric. It typically compares viewpoints as isolated visual configurations, without fully examining how perspective interacts with shared control mappings, role asymmetry, or supernumerary limb coordination. Likewise, dynamic viewpoint switching has rarely been studied together with questions of delegation, autonomy, or multi-user body sharing. As a result, the literature still lacks an integrated account of how perspective strategies should be configured when users share bodies, negotiate asymmetric roles, or coordinate with semi-autonomous limbs. This is not simply a question of which camera view is best, but of how distributed perception should be designed within collaborative embodiment. These gaps motivate the perspective-centered investigations later in this thesis, particularly the empirical studies in Chapter 5 and the synthesis presented in Chapter 7.

2.1.4 Supernumerary Limb Interaction

2.1.4.1 Multi-arm Mapping and Control Strategies

The effective control of supernumerary limbs is a central research theme in both robotics and virtual reality. A primary goal is to develop intuitive mechanisms that allow users to manage these additional limbs efficiently alongside their natural ones [40]. Research on wearable robotics has explored various forms of augmenting the human body, such as robotic legs, fingers, and arms, with a focus on designing appropriate control mappings [41, 42]. Prior research has predominantly employed direct and fixed control mappings, where specific commands are explicitly assigned to the movement of particular body parts [8]. Common examples include remapping existing body movements, such as utilizing foot or shoulder gestures, or using proprioceptive interfaces like foot pedals to operate the extra limbs [8].

Brain-Computer Interfaces (BCI) have also been explored for direct neural control [43]. While these methods have demonstrated effectiveness in controlled environments, they often depend heavily on fixed control mappings, which can restrict adaptability in dynamic and varying task conditions [44, 45].

To overcome the limitations of fixed mappings, researchers have begun to explore more flexible and dynamic approaches. One such strategy involves dynamic control switching, which allows users to manage and seamlessly alternate control among multiple supernumerary limbs in real-time [17]. For example, systems can be designed to use voice commands to dynamically switch focus and control between different virtual robotic arms, enabling a more fluid and context-aware interaction model [46]. This approach moves beyond static one-to-one mapping and allows for a more adaptive control scheme.

Furthermore, to reduce the cognitive load associated with continuously managing multiple limbs, the concept of semi-autonomous mapping has emerged. In this paradigm, supernumerary limbs are imbued with a degree of autonomy, allowing them to learn, adapt, and act semi-independently to respond to changing task requirements [47]. This shifts the user's role from a direct operator to a supervisor who provides high level instructions [48]. For instance, a user might demonstrate a simple action and instruct the system to repeat it autonomously, freeing up the user to focus on more complex tasks [49, 50, 51]. This strategy allows users to delegate repetitive or lower-priority actions to the virtual robotic arms and adjust their control strategy by switching between manual control and automation depending on task complexity [52]. However, this approach introduces new challenges related to trust, predictability, and maintaining a sense of agency when control is shared with an autonomous system [53].

This literature provides a strong foundation for understanding how extra limbs can be mapped, switched, and delegated. However, it remains largely augmentation-centric. Most prior work focuses on how a single user controls additional limbs, rather than on how augmented control is coordinated across multiple users, viewpoints, and task roles. As a result, supernumerary limb research has not yet been fully integrated with broader questions of collaborative embodiment.

2.1.4.2 Challenges and Countermeasures in Multi-limb Interaction

While supernumerary limb systems offer significant potential for augmenting human capabilities, they also introduce considerable usability challenges. A primary issue is the increased cognitive load and mental strain placed upon the user. Controlling multiple limbs simultaneously requires constant attention shifting and complex coordination, which can be mentally exhausting and lead to fatigue [54, 55, 56]. As noted in a recent survey, supernumerary limbs “require significant cognitive adaptability” from the user [57]. This cognitive burden is exacerbated in low-autonomy settings, where frequent manual switching and step-by-step guidance for each limb increase mental strain and can reduce task efficiency [58, 59].

Another challenge lies in managing feedback and learning. While feedback is essential for control, multiple active limbs can generate a redundant or overwhelming stream of visual and auditory information, making it difficult for the user to maintain a clear sense of the overall task state. Furthermore, there is an inherent learning cost as users must adapt their motor skills and develop new coordination strategies. Participants in studies often describe an initial sense of disconnect, remarking that the virtual arms “felt weird, like they were someone else’s arms” before they gradually felt more natural with practice [60, 61, 62].

To understand and address these challenges, researchers have explored user developed strategies, sometimes using a Wizard-of-Oz methodology to simulate advanced autonomous capabilities that are not yet fully realized in current technology [63, 64]. Such studies provide valuable insights into how users naturally cope with the complexity of multi-limb control. For example, a key countermeasure that users employ is task delegation. One study found that participants preferred to delegate repetitive or lower priority actions to the semi-autonomous arms, allowing them to reserve their cognitive resources and their own hands for more complex or high priority tasks [52]. This strategic division of labor allows for more efficient multitasking [65].

To further streamline interactions, users have also been observed developing organizing strategies, such as labeling or categorizing objects by their properties (e.g., color or function) to simplify commands and improve coordination with the supernumerary arms [66]. These

findings underscore the need for adaptive control mechanisms and intelligent system design that can support users in managing cognitive load, balancing manual and autonomous control, and developing effective strategies for complex multi-limb interaction [67].

These studies make clear that multi-limb interaction is not only a technical control problem, but also a problem of cognitive organization, feedback interpretation, and strategic delegation. Yet the literature still lacks a unified way of relating these issues to shared agency, role asymmetry, and distributed perception in collaborative VR. This thesis builds on these insights by examining supernumerary limb interaction as part of a broader collaborative embodiment problem.

2.1.5 Methodological Considerations

The study of collaborative embodiment in virtual reality necessitates a multifaceted methodological approach to capture the complexities of user performance, subjective experience, and interaction dynamics [68, 69]. Research in this domain typically employs a combination of quantitative and qualitative methods to provide a comprehensive evaluation of novel systems and interactions [70].

2.1.5.1 Key Evaluation Metrics

Evaluating collaborative embodiment involves a blend of objective performance indicators and subjective user reported data. The choice of metrics often depends on whether the study's focus is on task efficiency, interaction quality, or the phenomenological experience of the user [69].

Quantitative Measures. Objective metrics are crucial for providing quantifiable data on user performance and system effectiveness [71]. Among the most common metrics are those related to task performance. Task completion time is a straightforward indicator of efficiency [72], while task accuracy and error rates are used to evaluate the precision of task outcomes

[73]. These metrics are fundamental in assessing how well users can control and interact within the virtual environment [70].

To gain deeper insights into the collaborative process itself, researchers also analyze interaction quality. This includes metrics such as gaze and gesture analysis to track non-verbal cues and evaluate user attention [73], and the measurement of communication frequency to capture the dynamics of teamwork [74]. Furthermore, physiological synchrony, measured via signals like heart rate and electrodermal activity, is an emerging metric used to assess shared engagement and emotional alignment between users [72, 73].

Qualitative and Subjective Measures. To understand the user's personal experience, which is central to embodiment research, subjective and qualitative measures are indispensable [75]. A primary method for gathering rich, in-depth data is the use of semi-structured interviews and think-aloud protocols, where the resulting transcripts are often analyzed using methods like thematic analysis to identify emergent strategies and perceptions [76, 77].

In addition, a wide array of standardized questionnaires are frequently employed to assess specific facets of the user experience.

- **Presence and Immersion:** Questionnaires such as the Igroup Presence Questionnaire (IPQ) and the Slater-Usoh-Steed (SUS) questionnaire are widely used to assess the user's sense of "being there" in the virtual environment [78]. The Game Experience Questionnaire (GEQ) can also be used to measure dimensions like Sensory and Imaginative Immersion (SII) in gaming contexts [79].
- **Embodiment:** To measure the user's connection to their virtual avatar, researchers utilize tools like the Body Ownership Questionnaire (BOQ) [80] and the Avatar Embodiment Questionnaire (AEQ) [81], which assess critical components such as the sense of body ownership and agency.
- **Workload and Comfort:** The NASA Task Load Index (NASA-TLX) is a multi-dimensional tool frequently used for assessing cognitive and physical workload [82]. Given the potential for motion sickness in VR, the Simulator Sickness Questionnaire (SSQ) is also a staple for ensuring and measuring user comfort [83].

The combination of these objective and subjective measures allows researchers to build a holistic understanding of how new forms of collaborative embodiment, such as those involving supernumerary limbs, impact the user from multiple perspectives.

2.1.5.2 Task and Scenario Design

The design of experimental tasks and scenarios is fundamental to evaluating the performance and experience of collaborative embodiment in VR [24]. The choice of task directly shapes the nature of interaction and the cognitive demands placed on users. A review of the literature reveals a wide spectrum of tasks, ranging from abstract, controlled experimental setups to more ecologically valid, domain specific activities [84]. These tasks can be broadly grouped into categories such as object manipulation and assembly, collaborative problem solving, and simulation and testing, among others [25].

A key principle in designing these tasks is the management of difficulty and complexity. Many studies adopt a strategy of progressive complexity, starting with simple, foundational tasks before moving to more demanding scenarios. For example, research investigating supernumerary limb control often begins with a basic “Reaching” task, which involves the precise action of reaching for a target. Such a task is chosen for its “straightforward objective, singular focus, and absence of distracting gameplay elements,” providing a controlled environment to examine fundamental multi-limb coordination [24]. This foundational stage allows participants to familiarize themselves with the system and provides a baseline for performance measurement [85, 86].

Building on this, researchers then introduce more complex tasks that reflect real-world applications. These tasks often require advanced skills such as asymmetrical bilateral coordination, where each limb performs a distinct action, like one hand grasping an object while the other positions it [87]. An example is a “Factory Task” designed to simulate high-pressure, multitasking environments found in industrial settings, which requires not only manipulation but also precision, timing, and coordination under pressure [88]. Similarly, collaborative tasks can involve jointly assembling virtual objects or navigating through a shared space [89]. This progression from simple, everyday actions like grasping and pointing to complex,

simulated industrial workflows allows researchers to systematically investigate how users adapt their strategies and how system variables, such as perspective or autonomy, affect performance and user experience across different levels of task demand [85]. This balance between experimental control and ecological relevance is critical for generating both reliable and applicable findings in the field of collaborative embodiment [16].

2.1.5.3 Participant Recruitment and Experimental Validity

The methodological rigor and generalizability of collaborative embodiment research are heavily dependent on participant recruitment strategies and the choice of experimental design [68]. The selection of participants, in particular, can significantly influence the outcomes of a study. The literature reveals considerable variation in participants' prior VR experience, with studies recruiting a mix of novice users with limited exposure (38%), intermediate users (42%), and expert users with significant prior experience (20%) [90]. This choice presents a trade-off: while including novices allows for the study of learning curves and initial impressions, it risks introducing bias in performance data if not properly managed. Consequently, some researchers choose to recruit only experienced users to “avoid distractions caused by learning VR controls, ensuring focus on interacting with supernumerary arms” [91]. To mitigate experience based differences, many studies implement onboarding or familiarization sessions before the main tasks begin [92].

In terms of experimental design, controlled experiments are common, often employing a within-subjects design where every participant experiences all conditions [93]. This approach is effective for minimizing the impact of individual differences on the results [94]. To reduce order effects, researchers frequently counterbalance the sequence of conditions presented to participants [95]. The alternative is a between-subjects design, which uses distinct control and experimental groups to compare outcomes across different populations, though this often requires a larger sample size [96].

Finally, the validity of findings is shaped by the study's context. A majority of research in this area (67%) is conducted in laboratory-based environments, which offer high precision in controlling variables and measuring metrics like task efficiency [97]. However, this high

degree of control can come at the cost of ecological validity, as it may not fully represent real-world interaction dynamics [98]. In contrast, field studies (17%) conducted in naturalistic settings like museums or workplaces offer more authentic insights but face challenges in maintaining control and often involve smaller sample sizes [99]. This tension between experimental control and real-world relevance remains a central methodological consideration in the field, with a growing recognition of the need for hybrid designs that can bridge this gap.

Taken together, these methodological considerations suggest that collaborative embodiment cannot be adequately studied through a single metric, task type, or experimental setting. Because the phenomenon spans performance, workload, agency, ownership, coordination, and strategy, it requires mixed methods and carefully staged task design. This methodological complexity mirrors the conceptual complexity identified throughout this chapter: collaborative embodiment emerges not from any one variable in isolation, but from the interaction of control, perspective, feedback, autonomy, and task context.

2.2 Why Existing Literatures Are Insufficient for Collaborative Embodiment

The literature reviewed above establishes several important building blocks for this thesis. Research on embodiment explains how ownership and agency arise under tight sensorimotor coupling. Shared control literature shows how multiple users can coordinate within or around a shared task space. Perspective research demonstrates that viewpoint choice shapes awareness, comfort, and coordination. Supernumerary limb research explores how users adapt to additional effectors under fixed, dynamic, or semi-autonomous mappings. However, these literatures remain only partially connected.

First, shared control work is largely **control-centric**. It tends to focus on who controls what and how conflicts can be reduced, but gives less attention to how viewpoint, feedback, and autonomy reshape the experience of shared agency and collaborative embodiment.

Second, perspective work is often **camera-centric**. It compares first-person, third-person, and hybrid views, but rarely studies how these perspectives interact with shared mappings, asymmetric roles, or additional limbs within the same collaborative system.

Third, supernumerary limb work is largely **augmentation-centric**. It emphasizes mapping techniques, command schemes, and usability of extra limbs, but less often addresses how augmented control is coordinated across multiple users or how autonomy changes the social and embodied character of interaction.

As a result, the field still lacks a unified account of collaborative embodiment. What remains under-theorized is not any one of these dimensions in isolation, but how they must be configured together when users share bodies, distribute perception, and negotiate control in real time. In particular, existing literatures do not yet offer a framework that can jointly explain **role asymmetry**, **delegation**, **shared agency**, and **distributed perception** across multi-user, multi-limb systems.

This thesis addresses this gap by treating collaborative embodiment as a design problem organized around a set of interdependent dimensions:

- **Control allocation:** how control is assigned, shared, or reassigned across users and limbs.
- **Perspective configuration:** how users access egocentric, exocentric, or hybrid views of the body, partner, and task space.
- **Feedback legibility:** how the system communicates control state, role state, stability, and autonomy through visual, auditory, or haptic cues.
- **Autonomy calibration:** how manual control, delegation, and semi-autonomous behaviour are balanced.
- **Moderators:** how task structure, role asymmetry, and temporal transitions shape which combinations are effective.

Seen in this way, the problem space of collaborative embodiment is not captured by any one prior literature alone. Rather, it emerges at the intersection of these design dimensions. This

observation motivates the survey in Chapter 3, the empirical systems in Chapters 4–6, and the synthesis framework developed later in this thesis.

A Survey on Methodological Approaches to Collaborative Embodiment in Virtual Reality

This chapter serves a broader role in the thesis than documenting the state of the art. It presents a systematic survey of collaborative embodiment methods in virtual reality, examining how prior research has addressed control, perspective, feedback, and autonomy across multi-user VR applications. As collaborative VR becomes increasingly prevalent in domains such as education, healthcare, gaming, and remote teamwork, there is a growing need to understand how embodiment constructs, including agency, body ownership, and perspective, can be extended from individual to collaborative contexts. By reviewing and synthesizing existing literature, this chapter identifies common interaction paradigms, design strategies, and evaluation methodologies, while also revealing the fragmented structure of the field and the conceptual gaps that motivate the empirical and conceptual contributions of this thesis. In this sense, the survey provides a field-level foundation for the thesis’s later architecture, rather than functioning only as background context.

The remainder of this chapter is based on my paper, “A Survey on Methodological Approaches to Collaborative Embodiment in Virtual Reality” (currently under review), and is kept unchanged to comply with university thesis requirements.

3.1 INTRODUCTION

Virtual reality (VR) is increasingly used in education [12], healthcare [13], gaming [14], and remote collaboration [15], where immersive environments support new forms of interaction and user experience [16, 17]. Within this broader landscape, collaborative embodiment

concerns how users experience self-location, agency, body ownership, and virtual bodies in multi-user settings [18, 19]. Rather than treating embodiment as an individual experience alone, collaborative embodiment examines how users coordinate with others through shared avatars, distributed control, co-presence, and embodied interaction. In contexts such as education [100], gaming [16], and remote collaboration [101, 102], these issues shape how users engage with one another in shared virtual spaces. As VR systems move toward richer multi-user applications, there is a need to understand how collaborative embodiment has been studied, evaluated, and methodologically supported.

Although substantial progress has been made in understanding individual embodiment in VR, collaborative embodiment remains less clearly consolidated. Multi-user VR introduces methodological challenges that are not fully captured by single-user embodiment research, including synchronization between users, distribution of control, communication through embodied cues, and the maintenance of presence during joint tasks [103]. As VR applications expand across education [104], gaming [16], and remote collaboration [15], these challenges become increasingly relevant to both research and system design. Existing studies have proposed a range of techniques for supporting shared control, co-presence, haptic interaction, social communication, and immersive collaboration. However, the methodological approaches used to study these techniques have not yet been systematically reviewed. This survey therefore examines how collaborative embodiment has been operationalized in VR research, with attention to synchronization, shared control, immersive interaction, task design, and evaluation practice.

This survey focuses on three methodological dimensions of collaborative embodiment research: the measures and metrics used, the task designs adopted, and the collaboration methods examined. These dimensions are central to understanding how prior studies define and evaluate collaborative embodied experience. By analyzing them together, we aim to identify recurring practices, methodological limitations, and areas where comparison across studies remains difficult. The goal is not only to catalogue existing work, but also to clarify how evaluation standards, task choices, and collaboration mechanisms shape the interpretation of collaborative embodiment in VR.

Our findings show that research on collaborative embodiment in VR is methodologically diverse. Laboratory studies remain common because they allow researchers to control variables and measure outcomes such as task completion time, accuracy, and synchronization. Field, remote, and hybrid studies provide complementary evidence about collaboration in more varied contexts, but remain less common. Across the reviewed corpus, many studies rely on performance-oriented measures such as task accuracy and user engagement, while measures such as physiological synchrony, gaze behavior, and gesture analysis are increasingly used to examine coordination and social interaction. We also identify limitations in participant reporting, task standardization, and metric selection, which complicate comparison across studies. Based on these findings, we discuss how future work can develop more transparent, comparable, and context-sensitive approaches to studying collaborative embodiment in VR.

3.2 BACKGROUND

The concept of embodiment in virtual reality (VR) builds on earlier theories of presence and telepresence, which framed VR not merely as a set of display technologies but as a mediated experience of being in an environment [105]. Slater and Wilbur later distinguished immersion as an objective property of a system’s sensory and interactive fidelity from presence as the user’s subjective response, arguing that immersive systems can support a sense of “being there” in virtual environments [106]. Over time, embodiment research expanded beyond immersion to incorporate more specific aspects, including self-location, agency, and body ownership, which are essential to understanding how users relate to their virtual bodies [22]. These components were originally developed to improve the user’s interaction with and control over the virtual environment in single-user VR setups, enhancing the sense of personal presence and natural engagement within the space.

In the field of Human-Computer Interaction (HCI), research on embodiment has been crucial for designing more intuitive and immersive VR systems. HCI frameworks like user-centered design have played a significant role in shaping how embodiment is understood, especially through the study of interaction design and how users perceive their avatars in VR. Studies

have shown that the representation of the body in VR, including first-person perspectives and avatar synchronization, significantly impacts users' sense of control and immersion [107]. These design advancements have been crucial for improving how users experience and interact with virtual environments, making them feel more present and engaged, even in multi-user or collaborative contexts.

3.2.1 Design Approaches for Embodiment in VR

3.2.1.1 Self-Representation and Avatar Embodiment

Self-representation in virtual environments has been studied through a range of avatar design approaches, from self-similar bodies to more stylized, idealized, or dissimilar representations [108, 109]. A key element in avatar design is customization, through which users can adjust appearance features such as body shape, skin tone, hairstyle, clothing, and facial features to express aspects of their identity and preferred self-presentation [110, 111]. Empirical work in social VR shows that many users want avatars to feel personal and, in some cases, visually similar to their physical selves, which can support self-presentation, attachment, and identification with the avatar [109, 110]. At the same time, avatar-mediated environments also allow users to explore idealized, alternative, or multiple identities, making dissimilar or creative avatars important for flexible self-expression [112, 113].

Avatar embodiment is commonly understood in terms of body ownership, agency, and self-location [22, 18]. In immersive VR, first-person perspective over a virtual body, especially when combined with visuomotor synchrony, can support the illusion that the virtual body is one's own [114, 115]. Virtual mirrors can further support embodiment by allowing users to observe the avatar's body and its movements from an external viewpoint. Prior work shows that synchronous mirror reflections of avatar movements can increase virtual body ownership [116]. More recent work also suggests that avatar customization, particularly when performed from a first-person perspective, can influence embodiment and body-related perception [117]. Together, these design elements shape how users perceive, control, identify with, and inhabit virtual bodies.

3.2.1.2 Shared Control and Multi-User Interactions

Shared control in VR has been explored most directly through virtual co-embodiment, where two users, agents, or other entities share control of a single avatar. A common implementation is a weighted-average control scheme, in which the motion of the shared avatar is computed from the users' movements according to assigned control weights [118]. This allows researchers to study how control distribution affects agency, co-presence, body ownership, and coordination during joint action [119, 120]. Recent work further shows that adjusting control weights during training can support motor-skill learning and retention, suggesting that co-embodiment can be used not only for collaboration, but also for teacher–learner motor guidance [121, 119].

Related ideas also appear in shared-control robotics and crowd-control systems. Multi-operator single-robot systems examine how multiple operators can jointly control one robot, including through visual or haptic coupling between operators [122]. Outside VR, systems such as Legion show how multiple remote crowd workers can coordinate real-time control of existing graphical interfaces, with the system mediating their inputs into a single controlled action stream [123]. These examples are not equivalent to avatar co-embodiment, but they help situate shared control as a broader design problem involving input aggregation, role division, and user agency.

Multi-user virtual environments also depend on communication mechanisms that help users coordinate action and interpret one another's intentions. In social VR, nonverbal communication is often carried through avatar body movement, gesture, posture, gaze, and facial expression, rather than through speech alone [124, 125]. Gaze behavior is especially important in shared immersive environments, where inferred or realistic avatar gaze can affect the perceived quality of communication [126]. More expressive avatar systems, including those that represent body movement, hand gestures, and facial expressions, have been shown to improve social presence, interpersonal attraction, and task performance in collaborative VR settings [127]. Facial expressions can also affect social presence in collaborative VR, particularly when combined with other communicative cues [128]. This work suggests that

collaborative embodiment depends not only on how control is distributed, but also on how users perceive, interpret, and respond to each other's embodied signals.

3.2.1.3 Haptic Feedback and Sensory Integration

Haptic feedback has long been studied as a way to add touch, force, and bodily sensation to virtual environments. In VR, common haptic approaches include vibrotactile cues, force feedback, passive haptics, and wearable or encountered-type devices, each providing different levels of physicality and actuation [129, 130, 131]. In shared virtual environments, force feedback has been shown to support collaborative manipulation and can improve task performance and perceived presence during joint interaction [132, 133]. More recent work on virtual co-embodiment has examined how haptic feedback and control distribution affect sense of agency, co-presence, body ownership, and motion synchrony when two users share a single avatar [120]. Related co-embodiment studies also show that sensory and intention-related feedback can shape how users experience limbs controlled by another person. Hapuarachchi and Kitazaki found that knowing a partner's movement intention can increase ownership and agency toward a partner-controlled limb in a joint avatar [19]. Follow-up work further showed that connection-induced upper-body movements can contribute to embodiment toward a limb controlled by another during virtual co-embodiment [134]. This work suggests that haptic and multisensory feedback should be treated as part of the coordination infrastructure for shared control, rather than as an optional realism feature.

3.2.1.4 Immersive Interaction Design

Embodied VR interaction also depends on how users act on and receive feedback from the virtual environment. Three-dimensional user interface research commonly organizes immersive interaction around tasks such as selection, manipulation, travel, and system control [135]. Within this design space, hand tracking and gesture-based input have been explored as ways to support direct manipulation and mid-air interaction. Bachmann et al. review three-dimensional human-computer interaction with a focus on the Leap Motion Controller, covering hand tracking, gesture recognition, and their use in 3D interaction tasks [136]. These

interaction techniques are relevant to embodiment because they can align the user's physical actions with visible changes in the virtual world. However, they should not be treated as sufficient for embodiment on their own. Rather, they provide interaction channels through which users can select, manipulate, navigate, and coordinate within virtual environments. When combined with appropriate visual, proprioceptive, and haptic feedback, these channels can support more coherent sensorimotor loops between the user's body, the avatar, and the surrounding virtual environment [22, 137].

3.2.1.5 Social Presence and Co-Embodiment

Social presence in multi-user virtual reality (VR) concerns the extent to which users experience other virtual humans or avatars as socially present, rather than merely co-located in the same virtual space [138]. Surveys of social interaction with agents and avatars in immersive VR identify avatar behaviour, visual representation, gaze, gesture, and other social cues as important factors shaping how users perceive and respond to others in virtual environments [139]. In location-based multi-user VR, Wienrich et al. showed that social interdependence can increase social presence and cooperation during group VR experiences [140]. Rather than simply replicating face-to-face behaviour, Roth et al. proposed visual augmentations for eye contact, joint attention, and grouping, showing that such augmentations can increase social presence and influence participants' social behaviour in a virtual museum setting [141]. This work suggests that social presence and co-embodied collaboration depend not only on avatar realism, but also on how avatars make attention, orientation, and social signals visible during interaction.

Existing work shows that VR embodiment is shaped by visual representation, sensorimotor alignment, haptic feedback, and social cues. Yet these mechanisms are often studied separately, leaving unclear how they combine in collaborative settings and how their joint effects are evaluated. This raises a methodological question: how do prior studies connect design choices to collaborative embodied experience through tasks, measures, and evaluation practices?

3.2.2 Methodological Concerns in HCI and Embodiment Research

Methodological trade-offs play a crucial role in the field of Human-Computer Interaction (HCI) and embodiment research. While a comprehensive discussion of all methodological concerns is beyond the scope of this paper, we focus on five critical aspects that are particularly relevant to studying collaborative embodiment: sampling and representation of users, study environments, evaluation metrics, technological limitations, and ethical considerations.

3.2.2.1 Participants

Participant selection shapes how collaborative embodiment findings can be interpreted. Studies should report participant characteristics that may affect shared control and multi-user VR experience, such as prior VR experience, familiarity between collaborators, motor ability, and demographic background. In virtual co-embodiment, for example, Fribourg et al. examined how users' level of control over a shared avatar and prior knowledge of movement affected sense of agency [118]. Similarly, ShareYourReality studied co-embodied participant pairs under different avatar control distributions and haptic feedback conditions [120]. Rather than treating sample size or diversity as fixed requirements, collaborative embodiment studies should justify participant recruitment in relation to the study design, expected effects, and type of analysis [142]. When comparison groups are used, their purpose should also be made explicit, for example whether they test differences in experience, expertise, familiarity, or interaction condition.

3.2.2.2 Study Environment

Study environments also shape what can be concluded from collaborative embodiment research. Controlled laboratory settings allow researchers to manipulate factors such as control distribution, movement synchrony, haptic feedback, and visual feedback with high precision. This is useful in studies such as virtual co-embodiment and ShareYourReality, where the goal is to isolate how specific interaction factors affect agency, co-presence, body ownership, or coordination [118, 120]. However, controlled settings may not capture

the interruptions, environmental constraints, social context, and task variability present in everyday collaborative VR use. The level of immersion also matters: prior work distinguishes technological immersion from the user's subjective experience of presence [106], and meta-analytic evidence suggests that immersive system features can affect perceived presence [143]. Study environment should therefore be treated as part of the methodological design, not merely as a reporting detail.

3.2.2.3 Metrics

Embodiment studies in VR commonly combine behavioural measures with subjective reports. Behavioural and task-level measures, such as task completion time, accuracy, error rate, movement similarity, and motion synchrony, describe how users act through or coordinate with a virtual body. In virtual co-embodiment, for example, prior work has used control distribution, motor behaviour, and motion synchrony to examine how shared control affects users' actions and perceived agency [118, 120]. Subjective measures capture the experiential side of embodiment, including body ownership, agency, self-location, co-presence, and user preference [22, 18, 144]. Using both types of measures is useful because task performance alone does not indicate whether users feel located in, in control of, or connected to a virtual body.

3.2.2.4 Technological Limitations

Technological constraints can affect both presence and embodiment in VR. Tracking quality, latency, field of view, stereoscopy, and sensory feedback influence the degree of immersion provided by the system and can shape the user's subjective sense of presence [106, 143]. For avatar embodiment, mismatches between the user's movement and the avatar's visual or haptic feedback can weaken agency and body ownership, especially when visuomotor synchrony is disrupted [22]. In multi-user and co-embodiment settings, additional challenges arise from synchronizing shared avatar motion, distributing control between users, and providing feedback that helps participants understand who controls which action [118, 120].

These limitations should therefore be treated as methodological factors when interpreting collaborative embodiment studies, rather than as implementation details alone.

Collaborative embodiment is shaped not only by interaction design, but also by methodological choices such as technical fidelity, participant characteristics, study setting, and evaluation measures. These factors affect how findings about shared agency, body ownership, and co-presence should be interpreted. However, they are often reported separately rather than synthesized as part of the field's evaluation practice. This leaves open the question of how collaborative embodiment has been operationalized and measured across VR research.

3.3 METHODOLOGY

This review adheres to the four-phase guidelines outlined by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [145]. Initially, a search strategy was implemented to retrieve study records from publication databases. Subsequently, titles and abstracts were screened to exclude studies that were not relevant to the review's scope. In the third phase, the full-text articles were evaluated according to predefined eligibility criteria. Ultimately, studies that met these criteria were selected for inclusion in the analysis.

3.3.1 Search Strategy

The initial search was performed based on the topic of this review to include the most relevant studies. Considering the two main parts of the topic, collaborative embodiment and virtual reality, the query statement was defined below:

```
(embodiment OR "body ownership" OR agency OR
presence OR
  "co-embodiment")
AND
(collaborative OR collaboration OR "multi-user" OR
shared OR
  "co-presence" OR "shared perspective")
AND
("virtual reality" OR VR OR "immersive
environment*" OR
  "virtual environment*")
```

The asterisk symbol (*) in the query represents arbitrary non-space characters. This query was crafted to capture studies discussing the general concept of embodiment as well as those focusing specifically on collaborative aspects within virtual reality environments. The search targeted titles and abstracts across four major databases: ACM Digital Library, IEEE Xplore, Springer Link, and Elsevier Scopus (including Science Direct). To ensure relevance, searches in the Springer and Elsevier databases were restricted to conference papers or journal articles in Human-Computer Interaction and Virtual Reality. The publication range was set from 2014 to 2024 to include recent studies. Notably, since Springer Link lacks an option to limit searches to titles or abstracts, we adopted a modified approach: first, performing a full-text search and then filtering results using a Structured Query Language (SQL) statement, identical to the queries applied in ACM, IEEE, and Elsevier databases.

3.3.2 Study Selection

A total of 1,633 query results were retrieved from the databases (381 from ACM, 194 from IEEE, 862 from Springer, and 196 from Elsevier). After removing duplicates by title, 949 studies were initially included for screening and eligibility assessment, as shown in Figure 3.1. The selection criteria for these studies were as follows:

Criterion 1 — Publication category: The study must be published as a peer-reviewed technical paper. Publications such as progress reports, opinion pieces, or dissertations are excluded.

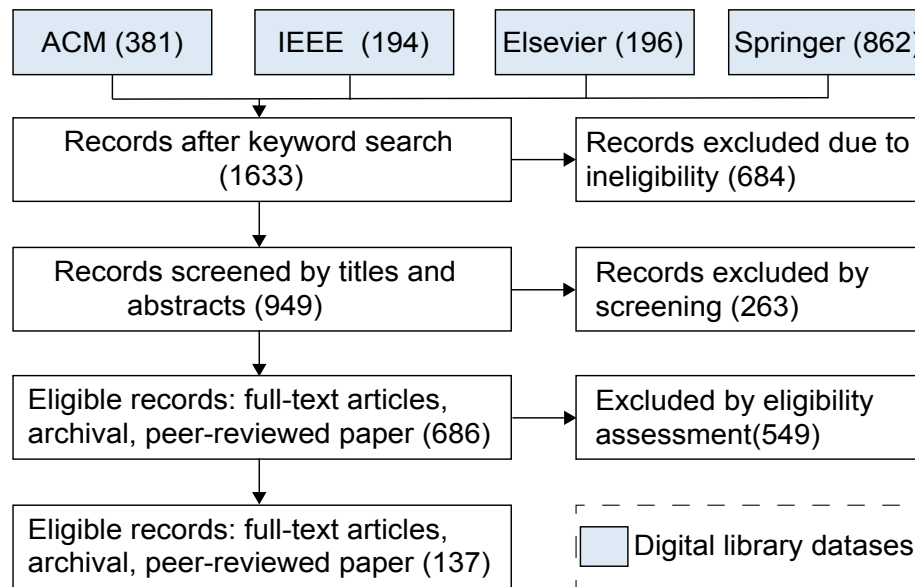


FIGURE 3.1. The PRISMA flow diagram illustrating the stages of our systematic search, screening, and selection process.

Criterion 2 — Involvement of virtual reality: The study must involve virtual reality. Studies using only other technologies, such as augmented reality (AR), extended reality (XR), or enhanced reality (ER), are excluded.

Criterion 3 — Collaborative embodiment: The study must involve collaborative embodiment. Studies focusing solely on single-user embodiment, such as experiences involving a single user's virtual body, are excluded.

Overall, 263 studies were excluded during title and abstract screening, and 549 studies were excluded following the full-text eligibility assessment. In total, 137 studies were included in this review for analysis.

3.4 RESULTS

In this section, we present the findings from our systematic review of collaborative embodiment in Virtual Reality (VR). Our analysis identified a total of 137 papers published between 2014 and 2024, encompassing 150 studies. The majority of papers contained a single study (127), while 7 papers included 2 studies, and 3 papers featured 3 or more studies, consistent

with previously reported local HCI standards [142]. To provide deeper insights into their specific methodological and conceptual contributions, we evaluate and present each study individually, rather than grouping them by paper. The analysis focuses on the measures and metrics, study tasks, collaboration methods, and study parameters employed in the reviewed studies. By examining these elements, we aim to identify common practices, methodological strengths, and gaps in the current research, which will inform future studies and the development of more effective and inclusive collaborative VR systems.

3.4.1 Measures and Metrics

3.4.1.1 Subjective Measures

TABLE 3.1. Overview of Questionnaires and Key Metrics Measured

Questionnaire	Percentage of Papers	Key Metrics Measured
Igroup Presence Questionnaire (IPQ)	35%	Spatial Presence, Involvement, Experienced Realism
Slater-Usoh-Steed Questionnaire (SUS)	25%	Sense of “being there,” immersion relative to reality
NASA Task Load Index (NASA-TLX)	30%	Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, Frustration
System Usability Scale (SUS)	20%	Usability of VR interfaces
User Experience Questionnaire (UEQ)	15%	Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, Novelty
Networked Minds Social Presence Inventory	18%	Co-presence, Psychological Involvement, Behavioral Engagement
Social Presence Questionnaire (SPQ)	22%	Co-presence, Perceived Interaction Quality, Emotional Connectedness
Simulator Sickness Questionnaire (SSQ)	40%	Nausea, Oculomotor Symptoms, Disorientation
Presence Questionnaire (PQ)	28%	Sensory Fidelity, Control Factors, Distraction
Body Ownership Questionnaire (BOQ)	12%	Agency, Self-Location, Ownership

This section summarizes subjective measures commonly used to evaluate collaborative embodiment in VR. Figure 3.2 shows how different collaboration methods are assessed across multiple criteria. Across the reviewed studies, presence is often measured with instruments such as the IPQ [146] and PQ [147], which assess users' subjective sense of being in a virtual environment. Usability is usually evaluated separately, most commonly through the SUS [148], which captures perceived ease of use rather than presence.

Social presence is another recurring evaluation target in collaborative VR. Studies use measures such as the Networked Minds Social Presence Inventory [149] and social-presence questionnaires adapted from virtual-environment research [150] to examine co-presence, mutual awareness, and perceived interaction with others. These measures are relevant when collaboration depends on users' ability to perceive and respond to partners' actions in real time.

Workload and usability measures are also common. NASA-TLX [151] is used to assess perceived mental, physical, and temporal demand, while SUS [148] provides a compact measure of system usability. These instruments help distinguish whether a collaborative VR system is difficult because of the task, the interface, or the interaction technique.

Embodiment is measured less consistently. Some studies use body-ownership questionnaires adapted from body-illusion research [20, 152], while others use standardized avatar embodiment measures that separate ownership, agency, and self-location [18]. These measures are especially relevant when the study involves shared avatars, co-embodiment, or distributed control.

Simulator sickness is commonly assessed with the SSQ [83]. This is important because discomfort can affect task performance, subjective experience, and willingness to continue collaborative VR interaction. Overall, the reviewed studies draw on questionnaires for presence, social presence, workload, usability, embodiment, and physical comfort, but the choice of instruments varies considerably. This variation makes direct comparison across studies difficult and points to the need for clearer reporting of what each measure is intended to capture.

3.4.1.2 Objective Measures

We began by examining the objective measures used to evaluate collaborative virtual reality (VR) systems. These measures were typically used to describe task performance, interaction behaviour, communication patterns, and system-level responsiveness. Across the 113 studies, we identified several recurring metric categories. Technical quality and QoE-related system measures [153, 154] were the most frequent, with 20 studies (15%) reporting indicators such as system responsiveness, latency, frame rate, or interaction stability. Gaze and gesture analysis [126, 125] was reported in 15 studies (11%), often to examine attention, nonverbal signalling, coordination, or communication behaviour. Task accuracy [135] and communication frequency [155] were each used in 15 studies (11%). Task accuracy captured the correctness of task outcomes, while communication frequency described verbal or nonverbal exchanges during collaboration.

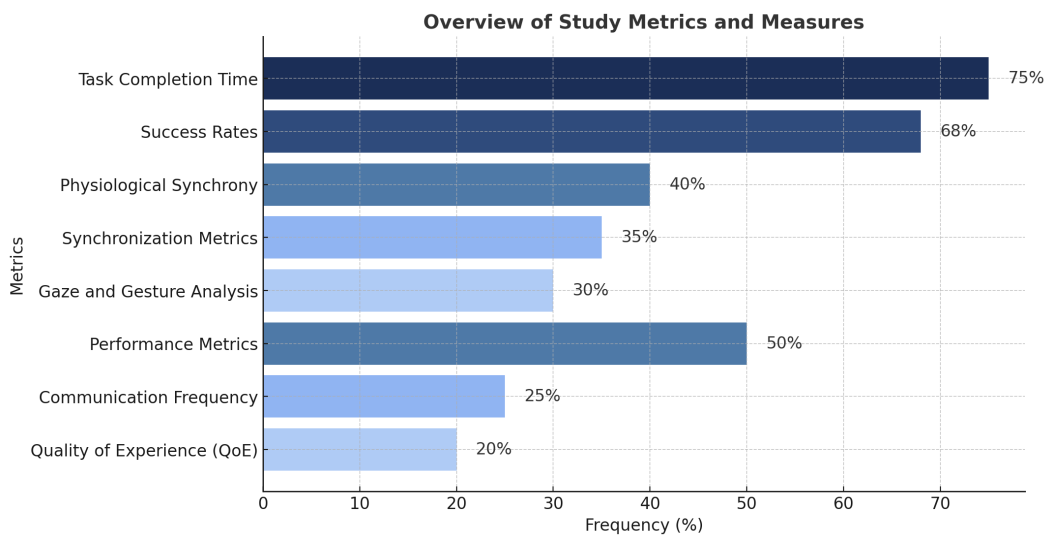


FIGURE 3.2. Objective Evaluation Matrix for Methods of Collaboration

Furthermore, *physiological synchrony* [156] was used in 13 studies (10%), typically through measures such as heart rate, electrodermal activity, or related physiological signals to examine temporal alignment between participants. *Task completion time* [157, 158], a common indicator of efficiency, was reported in 12 studies (9%) to evaluate how quickly participants

completed collaborative tasks. Similarly, *behavioural synchronization* [159, 120] was measured in 12 studies (9%), particularly in tasks involving coordinated timing, shared movement, or distributed control. Lastly, *success or error rates* [157, 158] were included in 11 studies (8%) to capture task effectiveness and identify points where collaborative interaction broke down. This distribution of metrics shows that collaborative VR studies draw on a mixture of system-level, behavioural, physiological, and task-performance measures. However, these measures are not interchangeable. Clearer reporting of what each metric is intended to capture would make it easier to compare findings across studies and to distinguish task efficiency, coordination quality, embodiment, and social experience.

3.4.2 Study Tasks

TABLE 3.2. Overview of Task Categories, Specific Tasks, Measure Variables, and Study Distribution

Task Category	Example Tasks	Typical Measures	Representative References
Object Selection and Activation	Selecting objects, pointing, clicking, highlighting, activating interface elements	Task completion time, accuracy, error rate	[135]
Interface Interaction	Menu navigation, system configuration, file manipulation, tool usage	Usability, efficiency, user satisfaction, error rate	[160, 161]

Object Manipulation and Assembly	Dragging and dropping, object assembly, fine motor tasks, force feedback interaction	Dexterity, task completion time, accuracy, force precision	[133, 132, 162]
Physical Activity-Based Tasks	Virtual sports, synchronized movement games, target shooting	Movement precision, engagement, task success rate	[159, 163]
Game-Based Tasks	Collaborative games, multiplayer strategy games, competitive gaming	Collaboration quality, engagement, task success rate	[163, 16]
Collaborative Problem Solving	Resource management, joint construction, shared control tasks, decision-making activities	Time to solution, teamwork measures, task success rate	[162, 161, 16]
Creative and Expressive Tasks	Annotation, sketching, collaborative authoring, creating virtual objects	Output quality, accuracy, task completion time	[164, 160, 161]
Daily Virtual Activities	Meetings, shared productivity tasks, social interaction, routine virtual activities	Usability, user satisfaction, completion time	[165, 160]

Exploration and Navigation	Collaborative navigation, exploration games, coordinated navigation	Navigation efficiency, coordination, task success rate	[166, 163]
Simulation and Testing	Input device testing, simulation setup, system evaluation	Accuracy, device responsiveness, setup time	[120, 167]

We analyzed the types of tasks utilized in the studies to evaluate the performance of participants in collaborative virtual reality (VR) environments in Figure 3.3. Among the 150 studies reviewed, most employed specific tasks to quantify user performance, interaction quality, and coordination, while a smaller subset relied on qualitative methods, such as interviews and surveys, without quantifiable tasks. Due to the diversity in task naming and definitions across the studies, we grouped tasks into broader categories based on their similarities. This process was conducted independently by two authors, achieving strong interrater reliability Cohen's Kappa ($\kappa = 0.91$, $p < 0.01$). The resulting categorization revealed ten primary task groups commonly used in the studies. Tasks such as selecting objects, pointing, clicking, and activating interface elements were grouped under "Object Selection and Activation." Activities that involved interacting with or navigating virtual interfaces, including menu navigation, system configuration, and virtual file manipulation, were categorized as "Interface Interaction." Tasks requiring manipulation or assembly of virtual objects, often involving fine motor skills, were grouped into "Object Manipulation and Assembly." Tasks that emphasized physical movements, such as virtual sports or synchronized movement games, were classified as "Physical Activity-Based Tasks."

Several studies incorporated gameplay to evaluate coordination, decision-making, and strategy, which we categorized under "Game-Based Tasks." Similarly, problem-solving tasks requiring teamwork, resource management, or shared control were grouped into "Collaborative Problem

Solving.” Tasks encouraging creativity or expression, such as annotation, sketching, or creating virtual objects, were categorized as “Creative and Expressive Tasks.” Routine tasks adapted for VR environments, such as sending emails, checking weather updates, and scheduling, were grouped under “Daily Virtual Activities.” Tasks involving navigation or exploration in virtual spaces were categorized as “Exploration and Navigation,” while those focusing on testing or simulation scenarios, such as input device testing, were placed under “Simulation and Testing.”

While most studies reported the use of at least one task to quantify user performance, approximately 23% of the studies did not employ specific tasks, instead relying on qualitative insights from user feedback or observational data. This categorization highlights the diversity of tasks used in collaborative VR research, ranging from precise motor skill evaluations to routine and exploratory activities. A detailed overview of these tasks and their corresponding measure metrics is provided in Table 3.2, offering a foundation for standardizing task design in future research.

Understanding task design choices is crucial for future research on collaboration, as tasks shape the nature of interaction and the cognitive demands placed on users. Clear task frameworks will help researchers develop studies that more effectively capture the complexities of collaborative behavior in embodied virtual environments.

3.4.3 Collaboration Methods

In Table 3.5, we identified 23 collaboration methods across the reviewed studies. Rather than treating these methods as a single group, we use them to distinguish different ways in which collaborative VR systems organize control, roles, communication, and shared activity. Shared-control and co-embodiment methods allow multiple users or entities to influence a shared avatar or virtual body, making control distribution central to agency and coordination [118, 120, 16]. Related work on supernumerary limbs extends this problem by adding extra virtual effectors that users must coordinate with their own actions [17]. Other methods structure collaboration through role assignment, perspective management, and remote

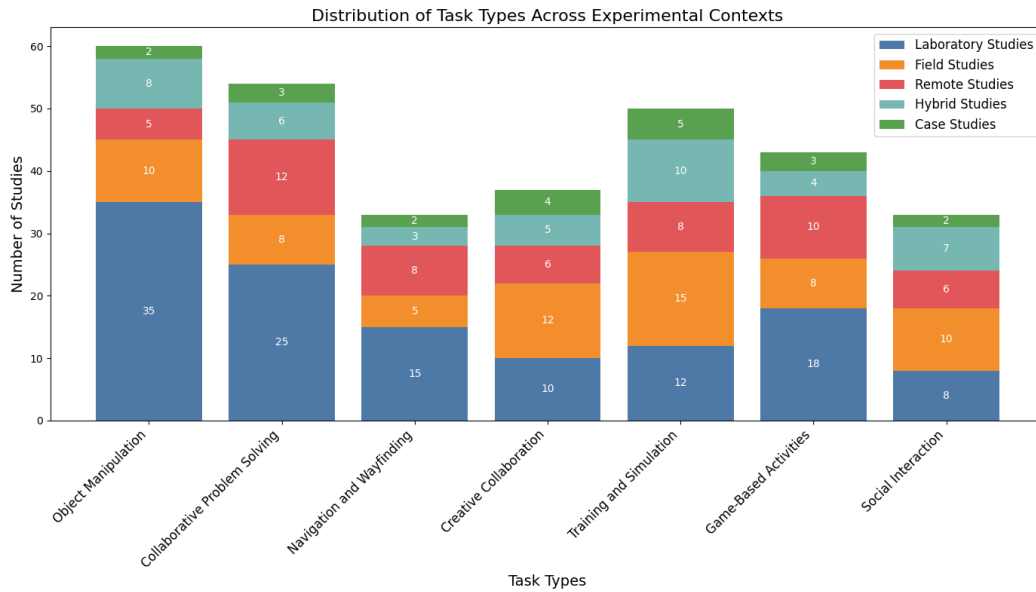


FIGURE 3.3. Frequency Of Task Types In Collaborative VR Research

coordination. Role-based or asymmetric systems can divide responsibilities between users, while collaborative authoring and editing systems support shared work through tools, timelines, or spatial workspaces [160, 161]. Haptic and multimodal collaboration methods add tactile or sensory channels to support joint manipulation and co-embodied coordination [133, 132, 120]. Game-based, navigation, problem-solving, and co-creation tasks provide recurring contexts for studying how users coordinate goals, divide labour, and negotiate shared attention [163, 162, 16]. Finally, communication-oriented methods use gaze, gesture, facial expression, and other avatar signals to support social presence and interpersonal coordination in multi-user VR [126, 125, 124, 141]. Overall, these methods show that collaboration in VR is shaped by more than the number of users in a shared space. It also depends on how control is distributed, how roles and viewpoints are organized, what communication channels are available, and how tasks require users to coordinate action. This motivates analysing collaboration methods as design mechanisms rather than as a flat list of system features.

3.4.4 Study Parameters

3.4.4.1 Participants

The reviewed 137 studies demonstrated notable variability in participant recruitment (Figure 3.4), with sample sizes ranging from 4 to 300 participants per study ($M = 22$, $Med = 18$). These sample sizes align with general practices in Human-Computer Interaction (HCI) research, though variations emerged based on study design and purpose. For instance, smaller exploratory studies involving qualitative assessments frequently recruited fewer participants

Collaboration Category	Primary Goal	Nature of Collaboration	Typical Setting
Embodied Shared Control	Task execution via shared embodiment and viewpoint alignment	Co-control of a shared avatar/system; synchronized viewpoint strategies	Engineering / robotics / VR avatar operations; telepresence co-design
Supernumerary; Haptic-Assisted Collaboration	Capability enhancement with extra limbs and tactile channels	Human-in-the-loop assist; shared tactile feedback; safety-aware coordination	Virtual multitasking environments; assistive/SRL-enhanced workflows
Distributed; Mixed-Reality Collaboration	Remote co-work in shared hybrid spaces	Remote collaboration across locations/devices	Telepresence, MR co-design, team-based simulations
Role-Structured; Tool-Mediated Work	Role-differentiated problem solving and flexible division of labor	Leader–follower shifts, joint tool use, on-the-fly role reassignment	Training scenarios, creative/engineering studios, team tasks
Multi-Modal; AI-Mediated Orchestration	Fuse modalities and AI support to improve fluency	Multi-sensory coordination; AI facilitation, guidance, and mediation	Immersive VR experiences; social VR platforms; collaborative editing
Training, Learning; Decision-Centric Collaboration	Skill development, consensus building, and joint planning/execution	Team training, collective deliberation, shared simulations and resources	Education/training; strategy and design reviews; engineering virtual labs

TABLE 3.5. Collaboration Methods Overview

included onboarding or familiarization sessions to mitigate performance differences due to varied expertise levels.

The role distribution within collaborative tasks also differed across studies. Leader-follower roles were assigned in 30% of the studies, where one participant assumed a guiding role (e.g., navigator or instructor). In contrast, shared control roles, where all participants had equal control over the task, were implemented in 62% of the studies. Notably, 8% of the studies reported dynamic role switching, where participants transitioned between roles based on task

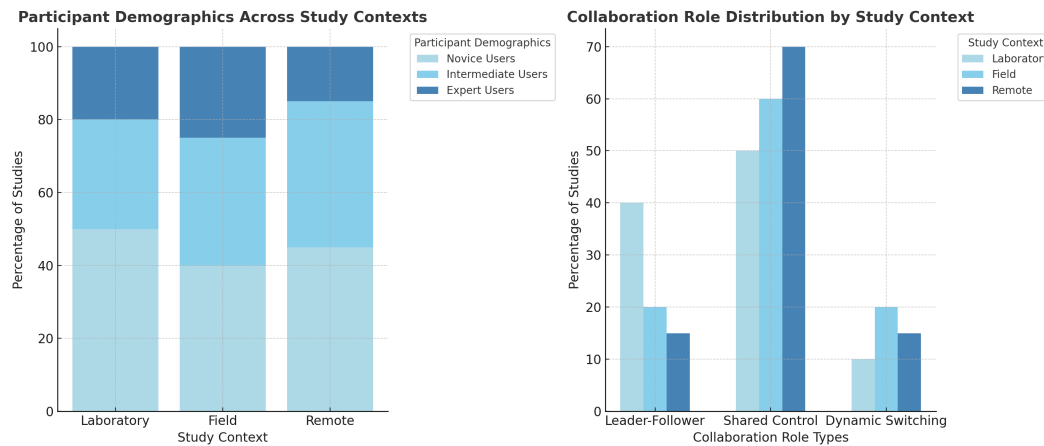


FIGURE 3.4. Participant Demographics Across Study Contexts

progression or emergent team strategies. This highlights the importance of role design in understanding collaborative dynamics and task performance.

Examining study parameters, including participant demographics and technological tools, helps future researchers understand the practical constraints and design decisions that influence collaborative outcomes. This perspective is critical for creating more inclusive, scalable, and realistic collaborative VR studies.

3.4.4.2 Technological Tools and Systems

Hardware Platforms: The reviewed studies used a range of VR hardware to support collaborative embodiment. Head-mounted displays (HMDs) were the most frequently reported devices, appearing in 106 of 137 studies (76%), reflecting their role as the main display platform for immersive VR research [135]. Standalone HMDs, including Meta Quest-class devices, appeared in 39 studies (28%), often in studies that required easier setup or room-scale deployment. Haptic devices were reported in 47 studies (34%) and included gloves, handheld devices, vibrotactile actuators, force-feedback devices, and wearable haptic systems. These devices differ in the type of feedback they provide, including vibration, force, contact, and motion cues [129, 130, 131]. In practice, lightweight vibrotactile actuators appeared more often than high-fidelity force-feedback devices, since they are easier to deploy in collaborative studies but provide a more limited form of tactile feedback [137]. Beyond standard VR

peripherals, 27 studies (19%) used non-traditional interfaces. Examples include robotic or haptic proxies for remote presence and tangible interaction [168, 169], multi-user redirected walking systems [170, 171], and custom tangible or authoring interfaces for collaborative 3D work [164, 161]. These systems show that collaborative embodiment is shaped not only by display type, but also by how users move, touch, share space, and manipulate virtual objects.

Software Platforms: VR software platforms varied across the reviewed studies, ranging from commercial engines to custom-built research frameworks. Real-time engines and networking frameworks, such as Unity with Photon or Unreal Engine, appeared in 83 studies (59%), often supporting shared scenes, synchronized avatars, and multi-user interaction. Representative examples include CoplayingVR, which used a networked Unity-based setup for shared-control VR interaction [16], and Ubiq, a Unity-based toolkit for building flexible social VR experiences [172]. Specialized toolkits appeared in 29 studies (21%) and were usually developed for particular collaborative tasks or domains. For example, Hyve-3D supported immersive collaborative 3D sketching [164], Spacetime supported individual and collaborative editing in VR [161], and REVERIE integrated autonomous agents and avatars in a web-based collaborative virtual environment [173]. Web-based or browser-oriented solutions appeared in 20 studies (14%). These systems often prioritized remote access, streaming, or lightweight participation over maximum graphical fidelity. Examples include VRComm, an end-to-end web system for real-time photorealistic social VR communication [174], VR2Gather, an open-source Unity package for adaptive multi-party real-time communication [175], and Space Connection, a 3D tele-immersion platform for web-based gesture-collaborative games and services [176].

3.4.4.3 Study Settings

The reviewed studies showed diverse methodologies and settings (Figure 3.4). Laboratory-based studies dominated (67%), leveraging controlled environments for tasks like comparing VR features or exploring embodiment effects. Field-based studies (17%) occurred in naturalistic settings, including museums and collaborative workspaces, focusing on real-world applicability.

Incentives were reported in 34% of studies, often involving monetary rewards or VR-related items, particularly in studies requiring extended or specialized participation. Methodologically, 39% were controlled experiments comparing VR techniques, while 17% were case studies describing unique VR applications.

Additionally, 10% employed longitudinal designs to observe long-term trends, and 10% focused on remote collaboration or VR meetings, reflecting a growing interest in distributed VR systems. This range highlights the field's focus on balancing experimental rigor with real-world relevance.

3.5 DISCUSSION

3.5.1 Implications for Collaborative Embodiment

The study of collaborative embodiment goes beyond technical feasibility, revealing how shared virtual interactions reshape teamwork. This section synthesizes theoretical insights and practical challenges in multi-user VR, bridging empirical findings with real-world applications. While synchronization metrics and nonverbal cues inform teamwork dynamics, issues like cognitive overload, accessibility, and engagement highlight the need for adaptive design. By analyzing behavioral drivers of collaboration alongside systemic barriers, we propose strategies to translate embodiment theories into scalable solutions. The following subsections explore these implications, advocating for a holistic approach to advancing collaborative VR systems.

3.5.1.1 Insights into Collaborative Dynamics

We first examine measures that capture collaborative dynamics in multi-user VR, including synchronization, communication, and task-sharing. Task completion time and success rates are commonly used to describe task efficiency and outcome quality, but they provide limited insight into how collaboration unfolds. Measures such as behavioural synchronization, gaze coordination, and physiological synchrony can offer additional evidence about shared

attention, arousal coupling, or coordinated action, although they should not be interpreted as direct measures of collaboration quality on their own [177, 159]. Current collaborative VR systems rarely use these measures for real-time adaptation. Related work in physiological computing shows that physiological signals can support adaptive interfaces, but applying this logic to collaborative embodiment remains underdeveloped [178]. For example, future systems could adapt communication cues, task pacing, or support tools based on behavioural or physiological indicators, but such adaptations would need careful validation to avoid over-interpreting noisy signals. Existing work on groupware and workspace awareness shows that effective collaboration depends on making others' actions, intentions, and task progress visible [179, 180]. In collaborative VR, similar principles could be supported through shared task indicators, role displays, progress visualizations, or cues that clarify who is acting and what has changed in the shared environment. These features are especially relevant for collaborative embodiment, where coordination depends not only on verbal communication, but also on embodied signals, shared control, and mutual awareness. Future work should examine how collaborative dynamics change over repeated interactions and more complex tasks. Longitudinal studies could show whether synchronization, communication strategies, and role coordination stabilize over time, and whether the same measures remain meaningful across education, healthcare, remote work, and other collaborative VR contexts.

3.5.1.2 Practical Challenges and Solutions

Collaborative embodiment in VR introduces practical challenges around workload, accessibility, and sustained engagement. Tasks that require real-time coordination, spatial reasoning, or role negotiation can increase cognitive workload, especially when users must divide attention between their own actions and their collaborators' actions [181]. Accessibility is also a concern, since many VR systems rely on physical movement, hand-held controllers, visual feedback, or headset-based interaction patterns that may exclude users with different motor, sensory, or cognitive abilities [182, 183]. Longer or repeated collaborative sessions may also introduce fatigue and uneven participation, which can affect how collaboration quality and embodiment are interpreted. Adaptive interfaces offer one way to address these issues, but they require careful validation. Physiological computing research shows that biosignals

can be used to adapt interfaces, although inferring user state from such signals is complex and context-dependent [178]. In VR, recent work has explored adaptive systems that adjust visual complexity or task demands based on cognitive-load indicators [184]. Similar ideas could be applied to collaborative embodiment by adapting task pacing, communication cues, or guidance when users show signs of overload or repeated errors, but such adaptations should be treated as design hypotheses rather than established solutions. Engagement can also be supported through feedback and gameful elements, provided they are aligned with the task rather than added as decoration [185, 186]. Accessibility can be improved through options such as voice input, alternative controls, seated modes, adjustable interaction ranges, and customizable feedback, particularly for users who cannot rely on the default movement or controller assumptions of VR systems [182, 183]. Future work should examine how these strategies affect collaborative embodiment over repeated use, especially in education, healthcare, and remote work contexts.

3.5.2 Methodological Reflections on Collaborative Embodiment

This section specifically reflects on methodological aspects unique to collaborative embodiment research, examining how experimental design, contextual diversity, and measurement strategies collectively influence our understanding of multi-user VR scenarios. By comparing controlled laboratory studies with field experiments, we highlight the tensions between internal validity and ecological relevance. Additionally, we discuss evolving metrics and participant dynamics that challenge traditional paradigms, emphasizing the need for adaptive methodological frameworks that balance precision with inclusivity. The following subsections offer detailed insights to guide future innovations in collaborative embodiment methodologies.

3.5.2.1 Study Design and Metrics

One methodological limitation in collaborative embodiment research is the limited use of longitudinal designs. Repeated-session studies can show how participants adapt to VR collaboration over time, including changes in workload, communication strategies, presence, and role coordination [187, 188]. This is especially relevant for collaborative embodiment, where

coordination, shared agency, and familiarity with partners or avatars may develop across repeated interactions rather than within a single session. A second issue is the variability of evaluation measures. Our review shows that studies assess collaboration quality and embodiment using different combinations of performance metrics, questionnaires, behavioural measures, and physiological signals. Standardized embodiment questionnaires have been proposed to improve comparability across avatar embodiment studies [18, 189], but collaborative embodiment also requires clearer reporting of what each measure is intended to capture, such as individual embodiment, dyadic coordination, social presence, workload, or task outcome. Task design also requires closer methodological attention. Simplified tasks can isolate specific variables, but they may not capture the coordination demands, role negotiation, and environmental complexity of real collaborative VR use. Prior work on VR methodology argues that virtual environments can support ecological validity when tasks preserve meaningful features of real-world activity while retaining experimental control [190]. Future studies should therefore report not only task outcomes, but also why a task is appropriate for the collaborative embodiment construct being evaluated.

3.5.2.2 Diversity in Experimental Contexts

Laboratory-based studies accounted for 67% of the reviewed corpus. This setting supports control over variables such as task structure, communication channels, avatar configuration, and system latency, which is useful for measuring task efficiency and coordination in collaborative VR [191]. However, controlled settings can limit how far findings transfer to workplace, public, or long-term collaborative use. Prior work has discussed VR as a way to balance experimental control with ecological validity, while also noting that realistic social interaction remains difficult to reproduce in laboratory settings [192]. Field, remote, and hybrid studies can provide complementary evidence about collaboration under more varied technical, social, and environmental conditions, but they also introduce challenges around recruitment, device heterogeneity, data quality, and experimental control [193, 194].

Incentives were reported in 34% of studies. These incentives can support recruitment and retention, especially in longer or more demanding studies, but they should be reported because

they may affect response rates, sample composition, response quality, and study cost [195]. Future work should make the rationale for study setting and incentive structure more explicit, especially when claims are made about real-world collaboration, long-term use, or broad user populations.

3.5.3 Technological and Methodological Gaps Specific to Collaborative Embodiment

Collaborative embodiment research has advanced through technological innovations and interdisciplinary insights, yet challenges remain in balancing inclusivity, scalability, and methodological rigor. While haptic feedback, adaptive avatars, and cross-platform integration showcase VR's potential, issues like user diversity, inconsistent task design, and technical bottlenecks limit broader applicability. By contrasting achievements, such as improved co-presence metrics and AI-mediated collaboration, with gaps in participant representation and methodological standardization, we highlight tensions between technological ambition and practical implementation. The following subsections examine these strengths and limitations, offering strategies to align innovation with inclusivity and enhance research accessibility.

3.5.3.1 Technological Advances and Applications

Recent collaborative VR systems increasingly combine full-body tracking, haptic feedback, adaptive avatars, and networked multi-user platforms. These technologies make it possible to study richer forms of shared embodiment, including role transitions, synchronized movement, and distributed collaboration. At the same time, they introduce technical constraints that affect how collaborative embodiment can be interpreted. For example, *Whisper* explores multisensory cues for private conversations in social VR, showing how audio and interaction design can support more situated communication practices [196]. Input modality also matters: studies comparing hand tracking, controllers, and audio feedback show that device choices can affect collaborative VR experience [87]. Similarly, work on avatar visual representation shows that how a co-actor is represented can influence task co-representation in VR [197]. Networked collaboration adds further challenges around latency, synchronization, and shared

state consistency. Serious-game and co-embodiment studies illustrate how different collaborative motion modes can affect usability and perception during shared virtual-body control [198]. Open-source and real-time frameworks such as Ubiq and VR2Gather provide useful infrastructure for building social and collaborative VR systems, but they also show that multi-user VR requires careful handling of networking, synchronization, and device heterogeneity [172, 175]. Rendering approaches such as foveated, cloud, and low-power rendering may reduce computational load, but their relevance to collaborative embodiment depends on how they affect latency, visual fidelity, and interaction timing [199]. Haptic feedback remains important but technically constrained. Existing haptic devices vary in feedback type, resolution, force range, and wearability, which affects what kinds of touch, contact, or force cues can be studied [200, 130, 131]. Cross-platform and mixed-device collaboration introduces additional mismatches in visual fidelity, tracking accuracy, input precision, and interaction affordances. Adaptive avatar systems such as Mini-Me show one way to support asymmetric MR collaboration by adjusting avatar scale, position, and orientation for the local user [201]. Future work should therefore treat sensor accuracy, latency, haptic fidelity, accessibility, and open infrastructure as methodological conditions that shape collaborative embodiment, rather than as implementation details alone [182, 183].

3.5.3.2 Limitations in Participant Representation

Participant representation affects how findings in collaborative embodiment studies should be interpreted. In the reviewed corpus, many studies report basic demographics and prior VR experience, but fewer provide detailed information about collaborator familiarity, domain expertise, accessibility needs, or cultural background. This limits how confidently findings can be generalized across user groups and use contexts. Prior HCI work has argued for more careful reporting of demographic and diversity dimensions in participant research [202, 203], while VR accessibility studies show that immersive systems can exclude users with different motor, sensory, or cognitive abilities if accessibility is not considered during study design [182, 183]. Onboarding is another methodological concern. Many collaborative VR studies involve participants with limited prior VR experience, but do not always describe how users were introduced to devices, interaction techniques, roles, or collaboration tasks. This matters

because unfamiliarity with VR can affect task performance, workload, and interpretation of shared-control behaviour. Prior work on VR onboarding highlights the need to design tutorials and introductory procedures that help users understand interaction possibilities before evaluation begins [204]. Role dynamics also require more explicit analysis. Collaborative embodiment studies often assign asymmetric roles, shared-control responsibilities, or leader-follower structures, but role transitions and role imbalance are not always analyzed as part of the results. In remote collaborative VR, task distribution has been shown to affect spatial processing and mental workload [181]. Future studies should therefore report how roles are assigned, whether participants can switch roles, and how role structure affects agency, coordination, workload, and satisfaction.

3.5.3.3 Inconsistencies in Methodological Standards

Collaborative embodiment studies vary substantially in task design, metrics, and role structure. This diversity is useful for exploring a young research area, but it makes comparison difficult when studies do not clearly explain how tasks, roles, and measures were chosen. Task designs range from abstract laboratory tasks to domain-specific activities, while metrics such as task completion time, success rate, workload, social presence, and physiological synchrony are applied with different assumptions. Role structures also vary, including fixed leader-follower roles, shared control, asymmetric collaboration, and role switching, but their effects are not always analyzed explicitly. These inconsistencies create challenges for replication and interpretation. Prior work on VR research methods has argued that reproducibility depends on clearer reporting of experimental procedures, hardware and software setup, participant-facing protocols, and sources of human error [205, 206]. In our corpus, similar issues arise when task parameters, onboarding procedures, calibration steps, role assignments, or analysis choices are only briefly described. VR onboarding research further shows that familiarization and tutorial design can affect how users learn to interact with unfamiliar systems [204]. Related fields offer useful models for improving transparency. Accessibility research, for example, has used systematic coding to examine participant groups, methodological choices, and gaps in representation [207]. Collaborative XR research has also begun to develop task taxonomies that clarify how collaboration can be organized by task properties and interaction

methods [208]. For collaborative embodiment, similar reporting practices could make it easier to compare studies, identify when findings depend on a specific task or system configuration, and build cumulative knowledge across VR research.

3.5.4 Future Directions for Collaborative Embodiment Research

The future of collaborative embodiment research calls for methodological refinement and societal relevance. While progress has been made, key challenges remain in long-term adaptation, interdisciplinary integration, and equitable design. This section explores expanding experimental paradigms, such as longitudinal studies and inclusive system design, while leveraging neuroscience, AI, and human-robot interaction to develop novel metrics and adaptive frameworks. Ensuring scalable architectures and diverse user cohorts will better align VR systems with global collaboration needs. The following subsections propose strategies to bridge these gaps, balancing innovation with ethical responsibility and accessibility.

3.5.4.1 Expanding Experimental Scope

Experimental work on collaborative embodiment in VR has often used short-term, task-specific studies. Such studies are useful for isolating variables, but they provide limited insight into how users learn to coordinate, negotiate roles, and adapt to shared control over repeated interactions. Longitudinal studies of collaborative virtual environments show that workload, presence, emotional recognition, and group interaction can change over time [187, 188]. For collaborative embodiment, this suggests a need to examine how synchronization, communication strategies, and shared agency develop across repeated sessions rather than only within single-session studies. Expanding the experimental scope also requires drawing on methods from related areas such as psychology, education, physiological computing, robotics, and AI. These areas can contribute measures and study designs for examining cognitive load, learning, shared attention, role negotiation, and adaptive support [178, 181]. However, such methods should be used carefully, since physiological or behavioural signals do not directly explain collaboration quality without contextual interpretation. Future work should therefore

combine controlled experiments with longitudinal, applied, and mixed-methods studies in contexts such as education, healthcare, remote work, and distributed teamwork.

3.5.4.2 Enhancing Inclusivity and Scalability

Collaborative embodiment studies would benefit from clearer reporting of participant diversity, including cultural background, accessibility needs, professional expertise, prior VR experience, and familiarity between collaborators. Prior HCI work has shown that demographic reporting and diversity considerations shape how participant-based findings should be interpreted [202, 203]. VR accessibility research also shows that systems relying on visual feedback, hand-held controllers, physical movement, or headset-based interaction can exclude users with different motor, sensory, or cognitive abilities [182, 183]. These concerns are especially relevant when collaborative embodiment findings are used to inform systems for broader user populations. Scalability is another open issue. Many collaborative embodiment studies involve pairs or small groups, whereas education, remote work, and training scenarios may require larger or distributed teams. Open-source and networked social VR frameworks. Gather, provide infrastructure for building and evaluating multi-user VR systems, but they also highlight practical challenges around synchronization, networking, device heterogeneity, and shared-state management [172, 175]. Future work should therefore examine how inclusive interaction options, adaptive role structures, and scalable system architectures affect collaboration quality across different user groups and deployment contexts.

3.6 LIMITATIONS

Our survey has several limitations. First, it is strictly limited to the results of our search within the domain of collaborative embodiment in virtual reality. Although we aimed to conduct an extensive review, some relevant papers might have been excluded due to the use of alternative keywords or their publication in interdisciplinary venues outside computer science. Furthermore, we focused on peer-reviewed academic literature, which may have resulted in the exclusion of valuable insights from industry reports or non-peer-reviewed studies.

While we systematically classified existing methodological approaches, this classification might oversimplify hybrid or unconventional frameworks, potentially misrepresenting studies that do not fit neatly into predefined categories. Moreover, the rapid pace of technological development in virtual reality may limit the long-term applicability of some findings discussed in this survey. Finally, our survey focused on collaborative contexts, leaving out individual embodiment techniques or their integration with collaborative mechanisms, which could provide a more comprehensive understanding of embodiment in VR. These limitations highlight opportunities for future research to expand and refine the insights presented in this survey.

3.7 CONCLUSION

This paper presents a comprehensive survey of methodological approaches to collaborative embodiment in Virtual Reality (VR). Our review analyzed a diverse set of studies focusing on how collaborative embodiment is operationalized and measured within multi-user VR environments, covering metrics, participant interaction methods, and technological tools designed to enable collaborative experiences. Our findings reveal key challenges in standardizing methods for evaluating embodiment, including the lack of consistent metrics, the limited exploration of diverse collaboration contexts, and the underrepresentation of longitudinal studies in the field. We systematically categorized existing methodological approaches and provided a structured overview to guide researchers in designing and evaluating future collaborative embodiment studies.

Our survey identifies three critical research gaps: the need for a unified set of metrics to evaluate embodiment across different collaboration contexts, the necessity to explore cross-disciplinary methodologies to enhance the richness of VR studies, and the importance of investigating long-term impacts of collaborative embodiment in VR on user performance and experience. While current methods have provided valuable insights, further development is needed to enable accessible, scalable, and inclusive VR systems for diverse user groups.

This work offers a foundation for future researchers and designers seeking to enhance collaborative embodiment in VR by providing a structured overview of existing practices, challenges, and opportunities. It also highlights the need for interdisciplinary collaborations and iterative methodologies to deepen the understanding of how collaborative embodiment directly impacts teamwork, interaction, and shared control in virtual environments. By addressing the identified gaps, the field can advance toward more effective, scalable, and user-centered solutions for collaborative VR.

3.8 How the Survey Motivates the thesis's Conceptual Architecture

The survey reveals that collaborative VR research has made progress across several important dimensions, including control mapping, perspective strategy, feedback design, and supernumerary limb interaction. However, these dimensions are often investigated separately. Many studies focus on one design variable at a time, such as a specific control mapping or a particular viewpoint configuration, without situating it within a broader account of collaborative embodiment.

In particular, the literature tends to be fragmented along at least four lines. Work on shared control is often control-centric, emphasizing allocation and arbitration. Perspective studies are frequently camera-centric, comparing view types without fully integrating them with control and role structure. Research on supernumerary limbs is often augmentation-centric, focusing on limb usability for individual users rather than collaborative coordination. Feedback, while clearly important, is comparatively under-theorized as a design dimension linking agency, awareness, and autonomy.

These observations directly motivate the empirical structure of this thesis. Chapter 4 focuses on mapping and control allocation under shared-avatar interaction. Chapter 5 examines perspective distribution and dynamic switching in collaborative multi-limb tasks. Chapter 6 investigates autonomy and delegation in semi-autonomous supernumerary limb interaction.

Chapter 7 then synthesizes these findings into the Interaction Elements Matrix, which organizes the design space surfaced by both the survey and the empirical studies.

Seen in this way, the Interaction Elements Matrix is not introduced as an abstract taxonomy imposed after the fact. Rather, it emerges from two sources: the survey's identification of fragmented design dimensions in prior work, and the empirical studies' demonstration that these dimensions interact in practice. The survey therefore provides the field-level rationale for the thesis's later conceptual synthesis.

The survey also informs the thesis's empirical progression. Chapter 4 examines control allocation and shared mapping in a co-embodied gaming context. Chapter 5 investigates perspective distribution and dynamic switching in collaborative multi-limb interaction. Chapter 6 explores autonomy and delegation in semi-autonomous supernumerary limb control. Chapter 7 synthesizes these strands into the Interaction Elements Matrix, which organizes the design space surfaced through both the survey and the empirical studies.

PART II

Systems and Experiments

CHAPTER 4

CoplayingVR

CoplayingVR: Understanding User Experience in Shared Control in Virtual Reality

This chapter examines the first core dimension of collaborative embodiment in this thesis: **control allocation**. Although the empirical setting is a VR gaming environment, the significance of *CoplayingVR* is not limited to shared gameplay. Rather, it provides a controlled setting for examining how shared agency emerges when two users act through one virtual body. The chapter therefore asks not only whether shared control changes player experience, but also what it reveals about control mapping, role structure, and the redistribution of embodied action across collaborators.

To investigate this problem, we designed *CoplayingVR*, a networked VR system in which two users co-control the hands of a shared avatar. Through a mixed-methods study comparing shared-control and single-player conditions across multiple game tasks, this chapter examines performance, workload, body ownership, agency, and user experience. In doing so, it shows that redistributing control can generate both *coordination cost* and *scaffolding benefit*: shared control may increase communication demands and conflict, but it can also support novice participation, improve performance, and sustain engagement through distributed effort.

Within this thesis, this chapter examines collaborative embodiment through the problem of control allocation. Its findings provide a basis for the later analyses of perspective and autonomy by showing that, even in the absence of additional limbs or viewpoint switching, the distribution of control across users already has important consequences for both collaboration and embodied experience.

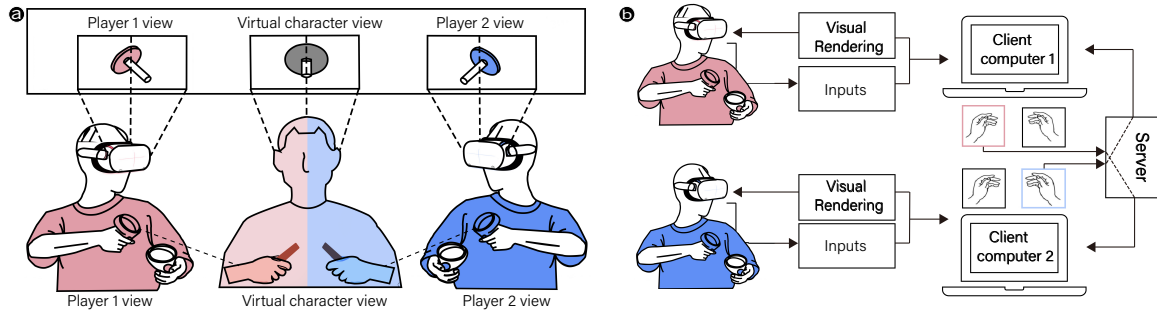


FIGURE 4.1. Shared control concept and implementation. (a) Two users are able to independently control different limbs of the same avatar (b) by sharing the local input/output control through a networked server.

The remainder of this chapter is based on my published paper, “CoplayingVR: Understanding User Experience in Shared Control in Virtual Reality” (Zhou, H., Ayesh, T., Fan, C., Sarsenbayeva, Z., Withana, A. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, <https://doi.org/10.1145/3678508>), and is kept unchanged in accordance with university thesis requirements.

4.1 Abstract

Shared control of avatars in virtual reality is emerging as a promising method to enhance collaborative interactions. While the literature investigates shared control as an input modality to understand the design of functional experiences, the user experience associated with shared control remains underexplored. In this paper, we investigate the user experience of shared control using a networked VR environment, “CoplayingVR”. Through a user study comprised of three gaming tasks in VR with 48 participants, we collect both quantitative and qualitative data on the user experience and performance of shared control. Our findings reveal that the shared control mode significantly improved user experience factors and brought novelty to the play. Furthermore, shared control significantly enhanced the task efficiency for novice users. Finally, we discuss how our findings inform the implications for research and practice of shared control for the growing body of research in VR.

4.2 Introduction

Virtual Reality (VR) has altered the way we interact with digital environments, bringing a new era of immersive experiences in education [209], work [210], skill acquisition [211], gaming [212], and well being [213, 214, 215, 216, 217]. An emerging interaction paradigm in VR is *shared control* – the concept where several individuals or entities share a common virtual avatar or form [19, 118, 119]. The approach draws inspiration from work done in non-VR contexts, for instance using shared control for collaboratively controlled robots [218] and networked tele-operation systems [219] and also more recent work in physical shared control using technologies such as electrical muscle stimulation (EMS) and pneumatics [220, 221, 222]. In VR, shared control allows users to simultaneously experience and control a single avatar from a first-person perspective, facilitating the sharing of skills and knowledge for enhanced collaboration [118].

Previous literature examines shared control as an input modality to design functional experiences [118], and to investigate performance [223] and user immersion [224] in VR. Moreover, the majority of previous research adopted systems where user interactions were restricted and evaluated in a highly controlled manner. For instance, users were only allowed to make specific movements, such as raising their hands [119]. As a result, the user experience (UX) associated with these control mechanisms in a realistic VR setting remains underexplored. Therefore, there is a need to deepen the understanding of the impact of shared control interactions on user engagement, performance, agency, and body ownership in a typical VR scenario.

To address this gap, we designed a system, *CoplayingVR*, focusing on unrestrained shared hand control in VR games. Hands are the primary medium for interactions in VR and the previous works show they can be selected as an implementation for investigating user experience of shared control [225]. *CoplayingVR* lets two users play VR games together. At the beginning, each user chooses one hand (left or right), which they will have control over in a shared avatar. Once selected, the user's hand movements with respect to the avatar are not constrained within the game. We believe that *CoplayingVR* can bring insights from a

scenario closer to a real-world application. Compared to the controlled cases investigated in literature [119, 223], we hypothesise that *CoplayingVR* approach can positively enhance user performance and experience for both experienced and novice users in a typical VR scenario. Through a user study using a series of VR games implemented on the *CoplayingVR* system, this paper investigates the user experience of shared control over virtual avatars. Particularly, we aim to answer three research questions: RQ1: *How does the shared control of hands influence user experience in VR games?* RQ2: *How does the shared experience of player expertise influence user performance in VR games?* and RQ3: *What are the effects of shared control on the sense of agency and body ownership for VR users?*

The paper presents results from a user study comprised of both *shared control* and *single-player* modes in VR with 48 participants. In the study, participants performed three gaming tasks from accuracy incentive to time incentive using the two control modes. We collect both quantitative and qualitative data on the user experience and performance of shared control. Our findings reveal that the shared control mode significantly improved user experience factors such as the challenging nature of gameplay, immersion within the game, negative affect, positive affect and brought novelty to the gaming experience [226]. Furthermore, shared control significantly enhanced the performance of novice users, and we did not observe a statistically significant performance decrease for experienced users. Additionally, there was a comparatively significant increase in the sense of agency and body ownership as the study progressed under the shared control mode. We further discuss how our findings inform the implications for research and practice of shared control for the growing body of research in VR, and we anticipate these will inform the design of future shared control systems.

The key contributions of the paper are:

- We present *CoplayingVR*, a system that allows sharing control of hands between two players in a typical VR game-play setting.
- We show empirical results on the impact of shared control on 1) user experience in VR games, 2) user performance, and 3) agency and body ownership.

- We discuss our findings and provide implications for both research and design. In particular, we provide design guidelines to benefit future research and the development of systems aimed at integrating shared control within VR gaming environments.

4.3 RELATED WORK

Collaborative multiplayer games such as Little Big Planet [227] and Donkey Kong [228] have been popular since the early days of gaming, allowing more than one player to collaborate and work towards a common goal. In comparison, shared control is employed as a collaborative input mechanism, where multiple players simultaneously control the same game element or characters [229]. Particularly, in collaboration, users take cooperative efforts with overlapping roles [230] while shared control focuses on mutual reliance with divided control [225]. In this paper, we aim to analyse the user experience of shared control in a VR setting. To this end, in this section, we discuss the literature on multiplayer and collaborative games, shared control approaches and methods used to evaluate user experience in VR and gaming settings.

4.3.1 Collaborative and Multiplayer Games

Collaboration is defined as the action of individuals or groups working together to achieve a common goal (i.e. a shared objective), for instance, seminal works by Johnson et al. have laid a robust foundation for understanding computer-supported collaboration through environments with predefined tasks [231]. While there are numerous works that focus on computer supported collaboration [232, 233, 234, 235, 236], since our tasks are primarily in games, we focus on literature in collaborative games. There are several aspects of collaborative games studied in previous research which facilitate complex social interactions and collective problem-solving. For instance, El-Nasr et al. present a set of collaborative design patterns and a Cooperative Performance Metrics (CPMs) for analysing and evaluating collaborative games [237]. Mitchell et al. focused on investigating social collaboration through digital bodily play using the HandyFeet platform, which combines physical activity with collaborative gaming, underscoring the potential to boost social interactions and teamwork [238].

Some research focuses on design patterns and game mechanics, such as studies by Rocha, Mascarenhas and Prada [239] and Zagal, Rick and Hsi [240]. These studies offer valuable insights of user experience collaborative gaming environments. Such in-depth evaluations of user experience of shared control in VR is lacking in literature and are essential to examine the unique challenges of the concept. Our study design was inspired by the existing collaborative game design strategies and critically examines the applicability of those within shared control in VR. Particularly, we paid close attention to game mechanics in the multiplayer games and how measures such as CPMs reflect on the complementarity, shared goals, shared puzzles and shared objects between players [237].

4.3.2 Shared Control in Video Games

There are several works that explore shared control in video or 2D games. For instance, El-Nasr et al. devised and executed various conventional input mediators and visual overlays to implement shared control within video games [241]. These approaches include allocating control segments, assignment of specific control actions, and time-based control allocation. Lessel et al. introduced HedgewarsSGC, a test environment to investigate shared game control in situations where multiple players compete against one another while sharing control of various game components [242]. These studies have made significant contributions on 2D and traditional video game platforms, which inspire current research in shared control, however they are done in context very different to VR. Specifically, these approaches often focus on the distribution of decision-making without addressing the heightened sensory and spatial interactions, and the psychological engagement that immersive VR environments necessitate [229]. In contrast, our research extends these concepts into VR, focusing on immersive shared control and its impact on user experience. By transitioning from the straightforward shared decision-making in 2D games to complex shared control of hands in VR, we explore how physical embodiment and psychological factors intertwine to affect collaboration.

4.3.3 Shared Body Experiences

Different aspects of sharing control of virtual bodies have been studied to explore the perceptual aspects of embodiment and performance. Goldberg et al. proposed the idea of a multi-operator single robot (MOSR) system, where multiple players control a single robot [218]. Later, an advanced MOSR was introduced where the player controls a remote actor who responds to input in real time [219]. Such studies primarily focused on the performance of remote operations, and less on experiences. More recent studies expanded these to explore the effects of virtual body autonomy resulting from shared control by mirroring first-person perspective and multisensory information [243], and presented insights into the psychological aspects of shared control in a joint avatar, highlighting influences of a partner's intentions and goals [19]. Furthermore, phenomena such as "enfacement", where simultaneous multisensory stimulation of blended faces, leading participants to perceive morphed images as themselves, creates interesting experiential opportunities [244]. These findings underscore the potential of VR to manipulate and enhance user experiences through shared control.

Sharing control is also used in application towards learning and work. Kodama et al. [119] investigated VR approaches to enhance motor skill learning, comparing it with both solitary learning and sharing the teacher's perspective. Kasahara et al. [245] investigated the efficiency effects of sharing first-person perspectives at work. Similarly, Lasecki et al. [123] introduced Legion, a system that allows a population of remote anonymous workers to control an existing player interface by forwarding the worker's mouse and keyboard input to an input mediator. These studies have a highly controlled experimental conditions, and the main focus of the study is utility and performance. Our aim is to fill this gap by investigating the user experience of shared control in a unrestrained controlled setting using VR games.

4.3.4 Measuring Experience

Since our study aims to measure player experience in shared control, selecting the appropriate method and metrics is critical. Particularly, we used game-like tasks, so we decided to focus primarily on gaming experience metrics. There are tools developed to measure different

aspects of user experience such as gamefulness [246] and players' feelings in games [247]. IJsselsteijn, De Kort and Poels [226] developed the Game Experience Questionnaire (GEQ) as an inclusive and comprehensive tool, which has been widely used to measure player experiences in various settings, including recent applications in virtual reality exergames [248]. The GEQ is composed of several modules, including the Core Questionnaire, Social Presence Module, and Post-Game Module, allowing for a multifaceted evaluation. Some recent criticism of its factor structure issues limits its suitability for all gaming contexts [249]. In this study, we focus on understanding specific aspects of shared control in VR, so the broad dimensions covered by the GEQ such as psychological and emotional aspects of gaming are still relevant and valuable. We specifically applied GEQ to assess dimensions like Challenge, Competence, Flow, Negative Affect, Positive Affect, Sensory and Imaginative Immersion (SII), and Tension, which are crucial for evaluating shared control in VR. We conducted additional evaluation of results to verify the consistency of finding of the GEQ to address the concerns.

Additionally in shared control, participants may feel as if their own body has been replaced by the self-avatar or the other players [18]. Embodiment is complex, as it involves not only ownership of the avatar's body but also agency, co-presence, and external appearance [18]. In terms of agency, a lack of body agency and autonomy in embodiment may have negative effects on social, perceptual, and behavioral performance [250]. In the shared control context in VR, players might perceive the avatar not only as an extension of themselves but also as an intertwined identity with other players, thereby influencing cooperative behaviors and gaming strategies. Thus, we employed an additional questionnaire to provide measurement of agency and ownership [251, 252]. Finally, as performance metrics, we collected completion times, and error rates where applicable.

4.4 DESIGN and implementation of 'CoplayingVR'

Shared control is a concept where a system or a process is controlled by multiple individuals or entities, as opposed to a single entity having complete control. In computer games, shared

control is implemented by distributing the control over the game character or elements among multiple players [225]. The implementation of shared control varies depending on the type of game and the mechanics involved. In this section we discuss design considerations and the implementation details of our 'CoplayingVR' system.

Sykownik, Emmerich and Masuch [225] presented an approach for the classification of the concept of shared control for a game character based on the concept of Locus of Manipulation (LoM). A LoM is defined as the "in-game position of the player's ability to assert control over the game-world" [253]. It denotes a clear in-game event that serves as evidence of a player's intentional manipulation of the virtual gaming environment. In the context of a VR game, we chose to implement the concept of "sharing control of distinct LoM that establish a coherent entity" [225] using hands since they serve as the primary means of interacting with the virtual world. This concept implies that every player has control over a separate LoM, but these LoMs come together to form one entity. Precisely, the control of one of the virtual hands of the virtual gaming character is given to one player so that the two players control the two hands independently (see Figure 6.1 for more details). However, as the two hands form a single gaming character, this prevents players from operating totally independently and involves a high level of player inter-dependency. However, it is important to note that it is not possible to isolate shared control from collaboration completely in games [225].

4.4.1 Input Control Mechanism

For the control mechanism, we mapped a virtual character's two hands to the controllers of two different participants, each of whom is wearing their own VR headset. The participants decided which of their hands (left or right) they would like to be mapped to control the virtual character's corresponding hand. Our study does not use combination of inputs by addition, average, and other methods as in previous work [254], but employs a control alternation between users. However, we prevented movements such as walking since this may lead to sharing control over the whole body, a condition that is out of the scope of the paper.

4.4.2 Rendering Visual Output

As the two participants are sharing the control of two virtual hands of a virtual character, the question of how virtual vision will be shared in a collaborative setting is an important consideration in the design of the shared control system. Collectively, the system creates a first-person experience of shared hand control by tethering virtual hands and head to a single body, creating the visual illusions as if they belong to the player. Although both hands are visually present, each player's control is limited to their selected hand, the other virtual hand reflects the other player. Furthermore, a virtual view that does not correlate to the participant's own head movement could create motion sickness. We mitigated this by letting players control the head rotation creating an independent perspective from the fixed body avatar location. Every gameplay action, triggered by either of the participant, is distributed to both headsets through the network, so they can see the changes happening in the virtual environment in real-time.

4.4.3 Implementation Approach

We used Unity platform to implement our 'CoplayingVR' concept. The system enabled two players to share a single virtual environment and synchronously reflect their real-time interactions on the same avatar. Our setup was based on two Oculus Quest 2 head-mounted displays and controllers. The application ran at a constant frame rate of 90 Hz on Unity 2021. Each pair of devices was connected to a separate computer using a USB cable and both computers were physically connected to the same network to minimize latency.

In the virtual environment, the two users were simultaneously tethered to a single anthropomorphic virtual character. The two players were physically co-located a few meters apart to accommodate user movements. The system allowed two playing modes, 'single-player mode' and 'shared control mode'. In the single-player mode, player's movement relied entirely on their own actions. In the shared control mode involving two players, we combined the hand movement inputs from both players.

We chose the “Photon Unity Networking” (Photon PUN) framework for implementing the shared control mode. We selected Photon since it provides Photon Cloud – a software as a service (SaaS) solution which enabled us to implement our multiplayer version of the game. Furthermore, we used an on-premises server application (Photon Server) that can be installed locally and a multiplayer game can be hosted on-site. This allowed us to reduce the delay between the participants’ actions to 15ms.

4.5 STUDY DESIGN

In this paper, we aim to evaluate the user experience and performance of shared control of virtual characters in VR. For our study, we designed three VR game tasks based on shared control. We based our tasks on the three types of VR games: casual games – *Control Task*, puzzle games – *Balancing Task*, and action games – *Shooting Task* and collected data on both user experience and performance. Our participants performed the game tasks in two playing modes: single-player mode (individually) and shared control mode. Participants were allowed to communicate with each other during the tasks.

4.5.1 VR Game Tasks

4.5.1.1 Control Task

The first task in our study, is a traditional peg-in-hole task (Figure 4.2-a) [255], which is a control task commonly used in manipulation studies [256]. This task required participants to build basic coordination, spatial awareness, and cooperative strategies through manipulating an object in space. Additionally, it needs precise coordination of movements from each hand (controlled by different players). Such skills are imperative as they form the foundation for the subsequent tasks. It was specifically chosen as the first task due to its simplicity to understand, focus on the manipulation, and lack of other gameplay elements. The task required participants to properly align a virtual peg in the game environment and insert it into a hole. To ensure both participants contributed to the control in the shared control mode,

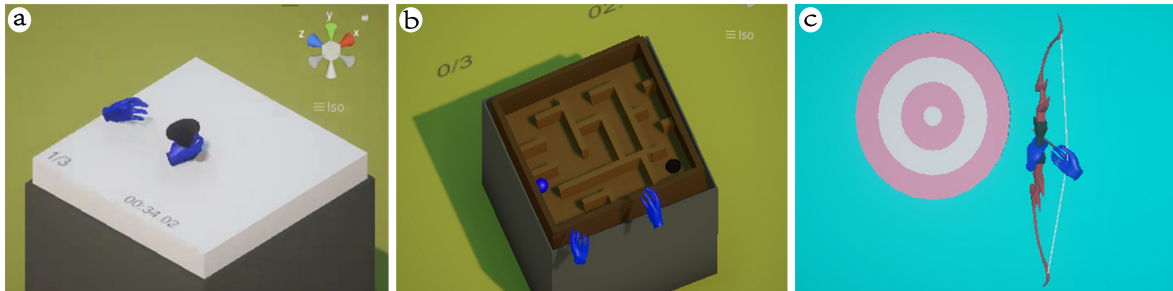


FIGURE 4.2. The three gaming tasks used in the user study. Each hand is controlled by a separate player. (a) Control task: Two hands are used to move a peg into a hole. (b) Balancing task: Two hands are used to tilt the maze to drive the ball to the correct path. (c) Shooting task: One player controls the bow to aim, the other shoots the arrow. The hand models used in this study were downloaded from the Oculus Developer website. The bow and arrow models were sourced from Quaternius.

the positional control is done by one player and the rotation is done by the other. Both players need to grab the peg simultaneously to control the two aspects to minimize turn taking. Whoever initiated the control aspect (rotation or movement) had complete control of that aspect regardless of the second player's input. The other player could control the other aspect simultaneously.

In total, the task required to complete 3 peg-in-hole activities. Participants were required to successfully position each peg into its designated hole, with each peg appearing only after the previous peg had been successfully positioned. For this task, the user controlled the location, the depth, the combination of axes required for rotation. Tasks were marked complete once the peg was correctly inserted into the hole reaching a specific depth. We recorded the task completion time, positions and rotations of hands and heads of the players.

4.5.1.2 Balancing Task

The second task is a balancing board game with puzzle elements, which was inspired by the work of Seif El-Nasr et al. [237] (Figure 4.2-b). This task requires participants to work together to balance a maze board, thereby enhancing coordination and introducing control over two axes, which emphasizes the necessity for synchronized movements and reactions to navigate the maze successfully. The activity involved participants guiding a ball into a hole in

a maze by tilting the board in two axes. They used virtual hands to control how the board is tilted. The task involved completing three levels, each comprising three mazes that became progressively more complex. Aim of the progression was to develop the participants' skills. In the shared control mode, to successfully guide the ball into the hole the participants needed to be able to work together to balance the tilt angle. We recorded the task completion time for each level along with positions and rotations of the hands and heads, the ball, and the table.

4.5.1.3 Shooting Task

The third task was a fast-paced archery game that involved shooting a target, inspired by the work of Loparev et al. [229] (Figure 4.2-c). This task builds on the coordination developed in previous activities. It applies and extends these skills through precise hand-eye coordination, timing and teamwork. It brings a challenging scenario and thus placed third in the order of tasks.

The players are offered two primary control options: "Aiming" and "Shooting". One virtual hand could be employed to grasp the bow and aim at the target, while the other hand could be used to pull the string for shooting arrows. Upon stretching the string of the bow, an arrow is spawned and ready to be released. Within the virtual environment, targets periodically appeared and advances towards the virtual player at a consistent velocity. Over the course of the game, a total of 50 targets were available, with each accurately hit target resulting in the score of one point. Participants were encouraged to maximize their shooting prowess to achieve a higher overall score.

We recorded data pertaining to archery shots, specifically recording the elapsed time between stretching the bowstring and releasing the arrow for each shot, the count of accurate shots made, and the total number of shots taken by participants along with the hand and head positions and rotations.

4.5.2 Experimental Setup

We conducted the experiment in a 2.5 x 3.5 meter laboratory room. User activities were recorded using GoPro cameras for later analysis. Both users wore VR headsets and could communicate with each other.

4.5.3 Experimental Data

4.5.3.1 Player Experience

In our study, we employed the Game Experience Questionnaire (GEQ)[226] to evaluate the gaming experience of participants since it has been used in similar studies [257, 248, 258]. We tested the GEQ's factor stability and verified that the reliability is in acceptable to good range using Cronbach's alpha analysis [249].

The specific questions rated the Challenge, Competence, Flow, Negative Affect, Positive Affect, Sensory and Imaginative Immersion (SII) and Tension of game experience. Challenge measured person's perception of how difficult the game was [259]. Competence included a sense of achievement, excitement and pride during the game [260]. Flow measured the gamer's perception of the "total absorption in the game and the non-self-conscious enjoyment of it" [261]. Negative Affect summarized feelings related to a bad mood and boredom that came from the gaming experience [260]. Positive Affect included feelings of happiness and enjoyment that came from the gaming experience [260]. Sensory and Imaginative Immersion (SII) described mental involvement of the user in the game [257]. Tension refers to feelings of annoyance, frustration, and pressure perceived during the game [260].

4.5.3.2 Performance

We measured the performance of participants using three variables: the task completion time, the error rate and the behavioral measurements (how participants engaged their hands in the game). Specifically, the task completion time for the *Control Task* is calculated from when the peg appeared until it is successfully inserted into the hole. Similarly, in the *Balancing*

TABLE 4.1. Questionnaire used in the sense of body ownership and agency test

Variable	Question
Body Ownership	Q1: “I felt as if the two virtual hands were my hands”
	Q2: “It felt as if the virtual hands I saw was someone else”
	Q3: “It seemed as if I might have more than one hand”
Agency	Q1: “It felt like I could control the virtual two hands as if they were my own hands”
	Q2: “The movements of the virtual hands were caused by my movements”
	Q3: “I felt as if the movements of the virtual hands were influencing my own movements”
	Q7: “I felt as if the virtual hand was moving by itself”

Task, completion time is measured from the moment the ball appears in the game until it is successfully inserted into the hole. For the *Shooting Task*, since it has a fixed play time, we quantified the error rate as the performance metric by tallying the number of accurately hit targets out of a total of 50 presented during the game. We did not let participants see the performance outcomes during the experiment to avoid comparisons and competition.

4.5.3.3 Agency and body ownership

As a measure of the degree of experience with the shared virtual body, we assessed agency and body ownership by using Avatar Embodiment Questionnaire (AEQ) [18]. We selected the earlier version compared to a more recent questionnaire [189] because of its strong focus on sense of agency. We removed some questions that were not directly related to our experiment. Our aim was to make it more focused and effective for our specific research needs. We administered the questionnaire (Table 4.1) before and after the experiment. Specifically, since the training session initially introduced participants to the virtual environment and avatar, mirroring the first exposure phase in [18], we decided to have participants complete the questionnaire after the training session and then again after the third task. They completed it in different tablets individually.

4.5.4 Experimental Conditions and Procedure

4.5.4.1 Participants

We recruited 48 participants (28 female, 20 male) for the study from a university student population representing different backgrounds. They were recruited individually, and not as pairs. Background information such as gender identity, VR gaming experience were collected and used to create participant pairs. We also did not consider handedness as a factor since we used both hands in the experiment, and let the participants choose which hand they like to use in the game. However, a majority of participants let us randomly assign them to each hand. The participants' age ranged from 21 to 34 years old ($M=24$, $SD=3.05$).

4.5.4.2 Experimental Conditions

We employed a within-subjects experimental design. All participants completed the experiment in both conditions – shared control and a single-player mode. To minimize any potential fatigue or learning effects, we counterbalanced the order of conditions presented to the participants with a break between each task and conditions. This ensured that any fatigue did not disproportionately affect one condition over the other, thereby maintaining the integrity of our results.

4.5.4.3 Procedure

First, we briefed participants on the experiment's purpose and procedure. Then, the participants gave the written consent. During this stage, we collected participants' background information (e.g., age, gender, previous VR gaming experience (novice or experienced players), gaming preferences).

Next, we assisted the participants to wear VR headsets and handed them the controllers. Prior to starting the experiment, our participants went through a training session to familiarize themselves with the equipment and tasks. The training session lasted approximately 5 minutes. After the training was completed, participants began the experiment. The participants always

followed the same task sequence: Control Task → Balancing Task → Shooting Task. In our study's setup, participants were physically located a few meters apart from each other and could communicate with each other verbally. Participants in the same group completed the experiment in the same order to avoid the potential impact on the experience gap.

When allocating the participants into pairs, we accounted for their previous gaming experience. In total we had 3 groups of participants:

- (1) The first group was comprised of participants (N=16) who had prior experience in playing VR games, i.e., both of the participants within one gaming session had previous experience of playing VR games.
- (2) The second group was comprised of participants (N=16) both with and without previous gaming experience in VR, i.e., within one gaming session we had one participant who had prior experience playing VR games, whereas the second participant did not have prior VR experience.
- (3) Finally, the third group had participants (N=16) with no prior experience of playing games in VR, i.e., both participants within one gaming session have not played VR games before.

After completing each task, participants took off their VR headsets, and completed the questionnaire. We ensured they did not discuss the answers to the questionnaire during this stage.

At the end of the experiment, we conducted brief semi-structured interviews first individually and then in pairs. Individual interviews allowed each participant to express their personal experiences and views. This was followed by a joint interview in pairs to explore their collective experience and management of shared control in the VR environment. Questions focused on their perceptions of performance under different conditions, the effects of shared control on their experience, and any challenges faced during the game.

The experiment lasted for approximately 40 minutes including the briefing, training, data collection, and final interviews. We did not provide any compensation for our participants

for their participation. The study received ethics clearance from the Human Research Ethics Committee (HREC) of the University of Sydney (2019/553).

4.6 RESULTS

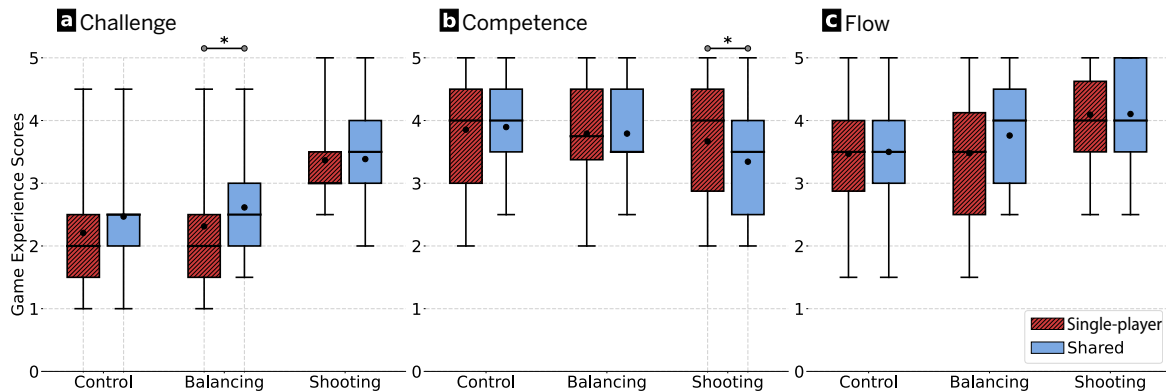


FIGURE 4.3. The averages of Likert ratings in shared control and single-player modes on the first three game experience measures. The three conditions in each graph are Control Task, Balancing Task and Shooting Task. * denotes $p < .05$.

4.6.1 Effects of Playing Mode on Game Experience

In this section, we analyse and present the results from the game experience questionnaire. Our aim is to understand how the game experience is affected by shared control and single-player modes.

During this study, 288 game experience questionnaires were administered, and our participants answered a total of 4032 questions for the game experience test in both single-player and shared control modes. In both modes, we calculated the average Likert ratings for seven different aspects of the game experience for each of the games and for the overall gaming experience. These aspects are Challenge, Competence, Flow, Negative Affect, Positive Affect, Sensory and Imaginative Immersion (SII) and Tension [226]. The data was not normally distributed, therefore, we conducted the Wilcoxon Signed-Rank Test to investigate the differences.

In light of concerns regarding the GEQ's factor stability, we conducted a Cronbach's alpha analysis to verify the internal consistency of its scales within our study. The Cronbach's alpha and 95% confidence interval respectively for each factor are: Challenge - 0.849 (0.731, 0.916), Competence - 0.741 (0.538, 0.855), SII - 0.805 (0.652, 0.891), Flow - 0.748 (0.550, 0.859), Tension - 0.740 (0.536, 0.854), Negative Affect - 0.777 (0.602, 0.875), and Positive Affect - 0.711 (0.484, 0.838). Findings suggest that our results are within acceptable to good range in terms of consistency.

4.6.1.1 Control Task (Figures 4.4 and 4.3):

First, we analysed the gaming experience in the *Control Task*. For the overall game experience, we found statistically significant differences for measures of "Negative Affect", "Positive Affect", and "Sensory and Imaginative Immersion (SII)". Precisely, "Negative Affect" in the shared control mode ($M = 1.20$, $SD = 0.41$) was significantly lower ($Z = -4.1084$, $p < 0.01$) compared with the single-player mode ($M = 1.79$, $SD = 0.78$). Furthermore, "Positive Affect" in the shared control mode ($M = 3.90$, $SD = 0.63$) was significantly higher ($Z = -3.6935$, $p < 0.05$) than in the single-player mode ($M = 3.37$, $SD = 0.74$). Finally, "Sensory and Imaginative Immersion (SII)" in the shared control mode ($M = 3.85$, $SD = 0.80$) was significantly higher ($Z = -4.6456$, $p < 0.01$) than in the single-player mode ($M = 3.00$, $SD = 0.85$).

For measures of "Challenge" ($Z = -1.704$, $p = 0.089$), "Competence" ($Z = -0.2656$, $p = 0.787$), "Flow" ($Z = -0.0089$, $p = 0.992$), and "Tension" ($Z = -1.7623$, $p = 0.078$), we did not find statistically significant differences in averages between the two modes.

4.6.1.2 Balancing Task (Figures 4.4 and 4.3):

For the *Balancing Task*, we found statistically significant differences in measures of "Challenge", "Negative Affect", "Positive Affect", and "Sensory and Imaginative Immersion". In particular, "Challenge" in the shared control mode ($M = 2.61$, $SD = 0.83$) was significantly higher ($Z = -2.3251$, $p < 0.05$) than the single-player mode ($M = 2.31$, $SD = 0.85$). "Negative Affect" in the shared control mode ($M = 1.23$, $SD = 0.47$) was significantly lower

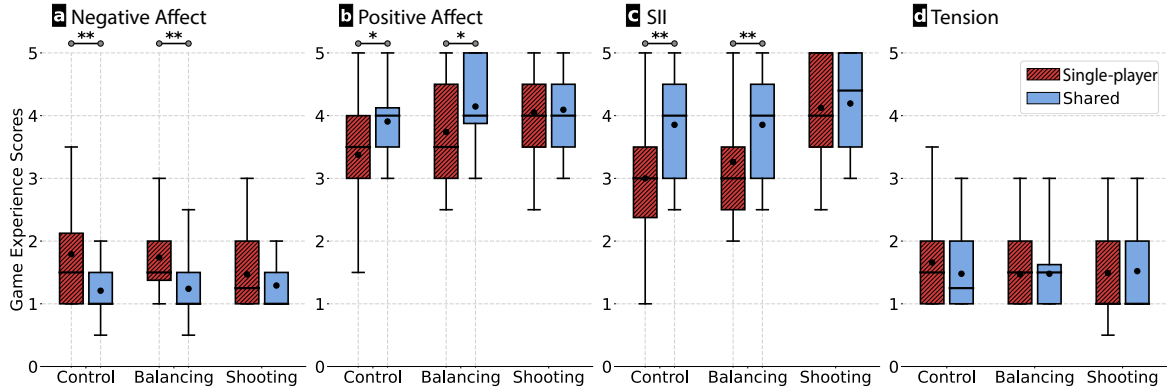


FIGURE 4.4. The averages for Likert scale ratings for game experience measures in shared control and single-player modes on the three conditions Control Task, Balancing Task, and Shooting Task. * indicates $p < .05$ and ** indicates $p < .01$.

($Z = -3.3306$, $p < 0.01$) compared to that in the single-player mode ($M = 1.73$, $SD = 0.61$). Furthermore, “Positive Affect” in the shared control mode ($M = 4.14$, $SD = 0.69$) was significantly higher ($Z = -2.6472$, $p < 0.05$) compared to that in the single-player mode ($M = 3.73$, $SD = 0.74$). “Sensory and Imaginative Immersion” in the shared control mode ($M = 3.85$, $SD = 0.75$) was significantly higher ($Z = -3.8375$, $p < 0.01$) than that in the single-player mode ($M = 3.26$, $SD = 0.86$). However, for measures of “Competence” ($Z = -0.2057$, $p = 0.833$), “Flow” ($Z = -1.82$, $p = 0.068$), and “Tension” ($Z = -0.1429$, $p = 0.888$), we did not find their averages in the shared control mode and their averages and in the single-player mode are statistically significant different.

4.6.1.3 Shooting Task (Figures 4.4 and 4.3):

For this task, we only found statistically significant differences for measures of “Competence”. Particularly, “Competence” in the shared control mode ($M = 3.20$, $SD = 0.89$) was significantly lower ($Z = -2.4515$, $p < 0.05$) compared to the single-player mode ($M = 3.66$, $SD = 0.91$). However, we did not find statistically significant differences between averages for “Challenge” ($Z = -0.1282$, $p = 0.896$), “Flow” ($Z = -0.0841$, $p = 0.936$), “Negative Affect” ($Z = -1.6337$, $p = 0.103$), “Positive Affect” ($Z = -0.1666$, $p = 0.865$), “Sensory and Imaginative Immersion” ($Z = -0.6273$, $p = 0.528$), and “Tension” ($Z = -0.1429$, $p = 0.888$).

4.6.1.4 Overall Gaming Experience

For the overall game experience, we found statistically significant differences for measures of “Challenge”, “Negative Affect”, “Positive Affect”, and “Sensory and Imaginative Immersion”. For instance, “Challenge” in the shared control mode ($M = 2.82$, $SD = 0.82$) was significantly higher ($Z = -2.4277$, $p < 0.05$) than that of the single-player mode ($M = 2.62$, $SD = 0.94$). “Negative Affect” in the shared control mode ($M = 1.24$, $SD = 0.41$) was significantly lower ($Z = -5.4028$, $p < 0.01$) compared to the single-player mode ($M = 1.66$, $SD = 0.67$). In alignment with this, “Positive Affect” in the shared control mode ($M = 4.04$, $SD = 0.65$) was significantly higher ($Z = -4.0242$, $p < 0.01$) than in the single-player mode ($M = 3.72$, $SD = 0.77$). Furthermore, “Sensory and Imaginative Immersion” in the shared control mode ($M = 3.96$, $SD = 0.76$) was significantly higher ($Z = -5.6477$, $p < 0.01$) compared to the single-player mode ($M = 3.46$, $SD = 0.93$). For measures of “Competence” ($Z = -0.7437$, $p = 0.459$), “Flow” ($Z = -1.2781$, $p = 0.200$), and “Tension” ($Z = -0.9241$, $p = 0.357$), we did not find statistically significant differences between averages for the two modes.

To summarise, our findings in section 5.1 answer RQ1: “*How does the shared control of hands influence user experience in VR games?*”. Precisely, our results demonstrate a significant effect of the shared control mode on user experience in the *Control Task* in terms of “Negative Affect”, “Positive Affect”, and “Sensory and Imaginative Immersion”. We further found that the shared control mode in the *Balancing Task* had a significant effect on “Challenge”, “Negative Affect”, “Positive Affect”, and “Sensory and Imaginative Immersion”. We have also found that the shared control mode had a significant effect on “Competence” in the *Shooting Task*. For the overall user experience in VR games, our findings show that the shared control mode had a statistically significant effect on “Challenge”, “Negative Affect”, “Positive Affect”, and “Sensory and Imaginative Immersion”.

4.6.2 User Performance – Prior Experience and Task Completion Time

In this section, we aim to answer RQ2: *How does the shared experience of player expertise influence user performance in VR games?* By analyzing task completion time, error rate, and behavioral observations using the recorded video, we aim to uncover how sharing VR gaming experience influences participants' performance in the shared control and the single-player modes.

We compared the task completion time between different modes in the *Control Task* and the *Balancing Task*. Task completion time in the *Shooting Task* was the same for all participants since the mechanics of the game relied on the number of accurately shot targets within a specified time. Nevertheless, for the *Shooting Task*, we measured the error rates in two playing modes.

To analyse the data we grouped the pairs of our participants into three groups. The groups were as follows: one comprised of participants (N=16) both with and without prior VR gaming experience, referred to as the *mixed pair* group; the second group (N=16) had participants who had prior VR gaming experience, referred to as *experienced players*; and the final group (N=16) was comprised of participants without any prior VR gaming experience, referred to as *novice players*. Sample sizes of the groups align with the standards used in HCI research identified by Caine et al. [142].

For the *mixed pair* group, we had 8 *novice players* and 8 *experienced players* to complete the tasks in the single-player mode; however, for the shared control mode, these participants were grouped together to complete the tasks. To analyse the performance of the *mixed pair* group we looked at the differences in performance between the *experienced players* in the single-player mode, *novice players* in the single-player mode, and mixed-experience in the shared control mode. For the *experienced players* and *novice players* groups, we looked at the differences in performance between the shared control mode and the single-player mode for each of the groups. In our study, we used Aligned Rank Transformation (ART) to analyse the effects of two different modes on participant performance in three tasks. The ART test allows us to effectively handle non-parametric data and complex interactions, suitable for our

study's groups and varying distributions. The test revealed statistically significant differences in performance data between the three groups.

4.6.2.1 Control Task – Task completion time for mixed pair group:

For the *Control Task* (Figure 4.5(e)), the ART analysis revealed that the task completion time in the shared control mode ($M = 0.77$, $SD = 0.03$) for *mixed pairs* group was significantly shorter ($F(1, 15) = 5.62$, $p < 0.05$) than the task completion time for *novice players* in the single-player mode ($M = 1.69$, $SD = 0.36$). Furthermore, the task completion time in the *Control Task* for the *experienced players* (Figure 4.5(e)) in the single-player mode ($M = 0.74$, $SD = 0.15$) was significantly shorter ($F(1, 15) = 6.45$, $p < 0.05$) than the task completion time for the *novice players* in the single-player mode ($M = 1.69$, $SD = 0.36$).

However, our analysis did not reveal any statistically significant difference ($F(1, 15) = 0.85$, $p = 0.400$) when comparing the task completion time of the *Control Task* for *experienced players* in the single-player mode ($M = 0.74$, $SD = 0.15$) versus *mixed pairs* group in the shared control mode ($M = 0.77$, $SD = 0.03$). The ART analysis did not reveal a statistically significant interaction effect ($F(1, 15) = 2.07$, $p = 0.17$).

4.6.2.2 Control Task – Task completion time for experienced players and novice players groups:

Additionally, the ART analysis revealed that the *novice players* (Figure 4.5(c)) in the shared control mode exhibited significantly shorter task completion times ($M = 1.50$, $SD = 0.25$, $F(1, 15) = 4.32$, $p < 0.05$) than in the single-player mode ($M = 1.72$, $SD = 0.38$). For the *experienced players*, our analysis revealed no significant difference between the *experienced players'* performance in the shared control mode ($M = 0.72$, $SD = 0.18$) versus their single-player mode ($M = 0.74$, $SD = 0.15$).

4.6.2.3 Balancing Task – Task completion time for mixed pair group:

In the *Balancing Task*, our analysis showed that there is a statistically significant difference ($F(1, 15) = 11.38, p < 0.05$) in task completion time between *experienced players* ($M = 0.73, SD = 0.10$) and *novice players* ($M = 1.99, SD = 0.87$) in the single-player mode. Similarly, no significant difference ($F(1, 15) = 0.70, p = 0.660$) was found when comparing the task completion times for *experienced players* in the single-player mode ($M = 0.73, SD = 0.10$) with *mixed pair* group in the shared control mode ($M = 0.77, SD = 0.05$).

4.6.2.4 Balancing Task – Task completion time for experienced players and novice players groups:

For the *novice players* (Figure 4.5(c)) in the *Balancing Task*, ART analysis showed significantly shorter completion times in the shared control mode ($M = 1.50, SD = 0.30, F(1, 15) = 6.57, p < 0.05$) compared to their single-player mode ($M = 1.99, SD = 0.87$). Furthermore, our analysis revealed no significant differences ($F(1, 15) = 0.42, p = 0.68$) in completion times for *experienced players'* performance between shared control mode ($M = 0.77, SD = 0.05$) and single-player mode ($M = 0.73, SD = 0.10$). The ART analysis showed a significant interaction effect ($F(1, 15) = 4.22, p < 0.05$) between player experience and game mode on task completion times in the Balancing Task.

4.6.2.5 Shooting Task – Error Rate for mixed pair group:

We then analysed the performance data of players in the *Shooting Task*. ART analysis revealed that the error rate for *mixed pairs* (Figure 4.5(f)) group in the shared control mode ($M = 27.25, SD = 2.65$) was significantly higher ($F(1, 15) = 5.65, p < 0.05$) than the error rate for *novice players* in the single-player mode ($M = 15.75, SD = 1.66$). Furthermore, there was a significant difference ($F(1, 15) = 6.78, p < 0.05$) in error rates when comparing the performance of *experienced players* in the single-player mode ($M = 12.25, SD = 1.38$) to *mixed pairs* group in the shared control mode ($M = 27.25, SD = 2.65$). Additionally, our analysis indicated that the error rate for the *experienced players* in the single-player mode

($M = 12.25$, $SD = 1.38$) was significantly lower ($F(1, 15) = 5.42$, $p < 0.05$) than for the *novice players* ($M = 15.75$, $SD = 1.66$) in the *Shooting Task* for the *mixed pairs* group.

4.6.2.6 Shooting Task – Error Rate for experienced players and novice players groups:

The ART analysis shows that the *novice players* (Figure 4.5(d)) had a significantly lower error rate in the shared control mode ($M = 15.00$, $SD = 1.50$, $F(1, 15) = 8.67$, $p < 0.01$) compared to their performance in the single-player mode ($M = 15.75$, $SD = 1.66$). However, no significant difference was found ($F(1, 15) = 0.58$, $p = 0.457$) when comparing the error rate of the *experienced players* in the shared control mode ($M = 12.50$, $SD = 1.40$) with their single-player mode ($M = 12.25$, $SD = 1.38$). The analysis indicated a significant interaction effect ($F(1, 15) = 4.95$, $p < 0.05$) between the player experience and game mode.

To summarise, our findings in section 5.2 answer RQ2: “*How does the shared experience of player expertise influence user performance in VR games?*”. Particularly, our results demonstrate that *mixed pairs* group, where novice users are put together with the experienced users, performed better in terms of task completion time compared to their performance in the single-player mode for the *Control Task*. Furthermore, our results show that when *novice players* work together in the shared control mode, their performance is significantly better compared to when they complete the tasks in the single-player mode. This statement holds true for task completion time for the *Control* and the *Balancing Tasks* and error rate in the *Shooting Task*.

4.6.2.7 Behavioral Observations:

In our behavioural observations, we looked at the movements of the participants’ real hands that were not engaged in performing the tasks in the shared control mode. We found that during the *Control Task*, the participants actively used their hands unintentionally, even though they were not using them in VR. This observation indicates that the participants perhaps tried to cooperate with their non-engaged hand subconsciously during the shared control mode. However, surprisingly, this behaviour of our participants changed in the *Shooting Task*, which

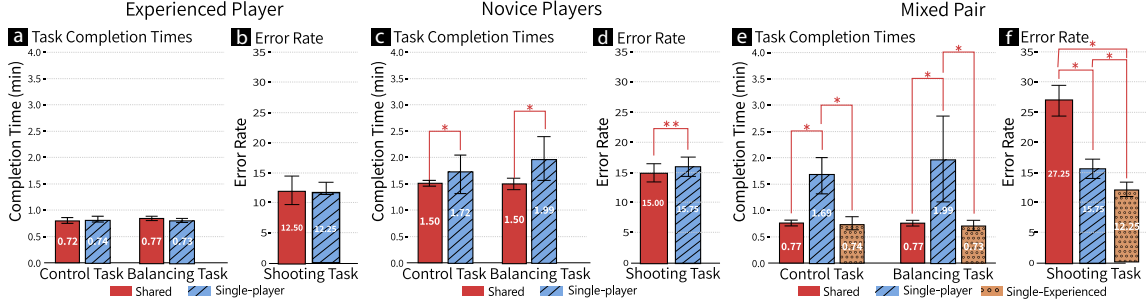


FIGURE 4.5. The figure shows the averages for: (a) the task completion time and (b) the averages for the error rates in the experienced players group; (c) the task completion time and (d) the error rate in the novice players group; and, (e) the task completion time and (f) the error rate in the mixed pair group. * indicates $p < .05$ and ** is $p < .01$.

was the last task. It appears with experience, the participants exhibited minimal movement with their non-engaged hands, engaging more with the shared control mode.

4.6.3 Agency & Body Ownership

We analysed how the sense of agency and body ownership changed throughout the experiment. To measure the sense of “agency” and “body ownership” we used the embodiment standardized questionnaire identified by Gonzalez-Franco and Peck [18]. For this, we followed the approach described in [118], and performed two measurements of agency and body ownership throughout the study:

- *First Phase* – before the gaming experience (i.e., before the Control Task)
- *Last Phase* – after the gaming experience (i.e., after the Shooting Task)

The data met the test assumptions since it was normally distributed. Therefore, we used two-way repeated measures ANOVA to analyse the effect of the phases on the sense of “agency” and “body ownership”.

In the “sense of agency”, two-way repeated measures ANOVA showed that the phases of the task had a statistically significant effect ($F(1, 47) = 4, p < 0.01$). The post hoc comparisons using Tukey’s HSD indicated that the sense of agency in the shared control mode

($M = 10.50$, $SD = 3.88$) was significantly lower ($F(1, 47) = 33.18$, $p < 0.01$) than in the single-player ($M = 15.21$, $SD = 2.13$) in the *first phase*.

After the *last phase* of the study, we observed a significant interaction effect between task phases and game modes concerning “agency”. The “agency” of the shared control mode after the *last phase* ($M = 12.49$, $SD = 3.31$) showed a statistically significant increase compared to the same before the *first phase* ($F(1, 47) = 16.52$, $p < 0.05$). This increase indicates a substantial improvement in the sense of agency for participants in the shared control mode as they progress through the experiment. No statistically significant difference ($p > 0.05$) was observed between the first and last phases for “agency” in the single-player mode. Similarly, there was no significant difference between the sense of “agency” after the *last phase* in the shared control mode compared with the single-player. This trend highlights the potential of shared control environments to foster a sense of agency in participants, approaching the levels for a sense of “agency” in the single-player modes.

Similarly, we employed a two-way repeated measures ANOVA to see if the sense of “body ownership” changed throughout the experiment. Our findings indicate that “body ownership” increased towards the end of the experiment ($F(1, 47) = 4.06$, $p < 0.05$). Tukey’s HSD post hoc test revealed that the “body ownership” was significantly lower ($F(1, 47) = 58.81$, $p < 0.01$) in the shared control mode ($M = 2.13$, $SD = 1.20$) compared to the single-player mode ($M = 4.56$, $SD = 1.42$) before the *first phase*. After the *last phase*, the “body ownership” in the shared control mode ($M = 2.92$, $SD = 1.62$) significantly increased ($F(1, 47) = 4.27$, $p < 0.05$) compared to before the *first phase* ($M = 2.13$, $SD = 1.20$). When comparing the “body ownership” within different modes, our results reported that after the *last phase*, in the shared control mode ($M = 2.92$, $SD = 1.62$), the sense of “body ownership” was significantly lower ($F(1, 47) = 33.04$, $p < 0.01$) than in the single-player mode ($M = 4.77$, $SD = 1.51$). Furthermore, there was no statistically significant effect of the phase ($p > 0.05$) in the single-player mode on “body ownership”. We additionally evaluated the interaction effect between task phases and the game modes in our analysis. The data indicated no statistically significant difference ($p > 0.05$) between task phases and game modes concerning “body ownership”.

To summarise, our results in section 5.3 answer RQ3: *What are the effects of shared control on the sense of agency and body ownership for VR users?* Precisely, our findings demonstrate that the sense of agency and body ownership increases throughout the gaming experience in the shared control mode.

4.6.4 Qualitative Results

We used thematic analysis to analyse the qualitative data collected from our participants during the semi-structured interviews. First, we read through each of the responses. Then the two authors of the paper independently completed the initial coding of the data. Next, the two authors compared their initial codes and finalised the codes based on their similarity. To validate the consistency of the thematic analysis, the data was independently coded by two researchers. The inter-rater reliability, measured using Cohen's Kappa, was found to be $\kappa=0.63$, signifying the reliability and precision of the coding process employed in our study. We then independently coded participants' responses according to the finalised codes. Lastly, we identified the themes by reviewing our coding and described below.

Theme 1 – Innovation and Novelty in Gameplay through Shared Control.

Many participants (N=13) stated that the shared control brings new features to the game. For example, one participant mentioned that the shared control mode brings innovation to the game experience. "I like both [modes], but the shared control mode is more fresh for me" (P15). Another participant highlighted, "Shared control changes how we play the game together — it's like we're teaming up in a whole new way" (P22). Furthermore, several participants (N=3) mentioned that shared control mode brought more fun to the gaming experience "Shared control one is more fun" (P28). Additionally, another participant emphasized the distinct nature of the game experience within the shared control mode, particularly in contrast to traditional multiplayer games. "I [used to] avoid multiplayer games for their lack of real teamwork. But in [this] VR games, like moving the ball together. It's new and fun, making us work together in a way I've never experienced before" (P23). "It's like you're seeing a whole new side to an old game", said another participant (P19). Also, one participant added, "With

shared control, playing feels more like a dance, where every move is smoothly linked to the next” (P34).

This highlights the unique challenges and experiences of shared control instead of general collaboration interaction. Finally, one participant mentioned that playing games in the shared control mode was very interesting. “I found it interesting that two people fused into one body and shared each other’s experiences” (P08). All the new features were derived from feedback provided by the participants.

Theme 2 – Collaborative Complexity – Discovering Challenges of Shared Control.

The majority of our participants (N=22) found the shared control game mode to be more challenging in the playing process because it was hard to communicate with each other during the game. “Hard to chat with others, but I like multiplayer games” (P18). “I think the hardest part is we [could not] to talk with each other” (P24). Another participant highlighted the impact of this communication barrier, “The need for quick decisions made it difficult to coordinate with my partner, adding an extra layer of challenge to the game” (P35). Moreover, some participants (N=7) also reported that the challenge of the shared control mode depended on the pace of the game: “I failed too many times in the last game. I found it hard to keep up with the pace” (P13). “When the game picks up speed, it’s not only about how fast you react but also about how quickly you can get on the same page with your teammate” (P26). Several participants (N=10) mentioned that the type of the task (game) also affected the difficulty of the shared mode: “For the shared control in third one [game], I should always find the position to control my arrow, because another player should shoot for it” (P23). Furthermore, according to some participants (N=8), fast-paced games were the hardest to play in the shared control mode: “The shared mode is difficult when it’s fast-paced, but I feel I have more of a challenge to do it.” In addition, one participant mentioned that they found it challenging to manipulate their hands in the shared control mode: “I think it’s hard for controlling hand movement” (P03). Furthermore, several of our participants (N=3) mentioned that they were subconsciously willing to engage their non-rendered hand in the study. For example, P14 mentioned: “In the first two tasks, I felt the urge to use my left [non-rendered] hand, like an instinctive reaction, but not in the last task, he [another participant] and I have become

better at working together”. We also found the shared control mode highlighted the challenge related to the immersive nature of VR (N=3); for example participants said, “Each level is a fresh adventure where we need to sync up our moves and even our thoughts” (P12). “It’s like you’re merging minds with your teammates to tackle the game” (P21). Moreover, P29 noted, “The times we had to communicate and synchronize our actions were the most intense and satisfying moments”. Furthermore, from our observations, we notice that in the first two tasks (*Control Task*, *Balancing Task*), participants use their non-engaged hand more in the experiment as compared to the last task (*Shooting Task*). We argue this is because their sense of body ownership has increased with time, as our quantitative results demonstrate.

Theme 3 – Empowering Choice – Personal Preference for Shared Control.

We also found that the fondness of a certain mode depended on participants’ personal preferences. Many participants (N=15) said that they preferred the shared control mode because they did not have experienced it before: “I liked the shared control mode because I did not have much experience [in using it]” (P11). This sentiment was echoed by another, who highlighted the immersive quality of shared control, saying, “The shared control felt more engaging; it was like being part of the game rather than just playing it” (P24). Some participants (N=10) also reported that the mode preference depended on the type of the game. “I thought the second game had a good shared control mode, but the third game was faster for me to play on my own than in pairs” (P8). This perspective was further supported by another participant “For quick-paced games, I prefer going solo because it lets me react faster without needing to sync with someone else” (P17). Whereas few participants (N=4) mentioned that their preference also depended on the pace of the game: “I liked the single-player version of the third game. Because the two-player mode was so fast-paced that it made it difficult for us to communicate in time” (P8). “In the third game the [*Shooting Task*], I found that the single-player mode would be easier. For the second game the [*Balancing Task*], the shared mode is more interesting because the second game the ([*Balancing Task*]) has a clear job for each person” (P7). Also, one participant reported that the shared control mode would enhance the game experience: “I think this could add more to the gaming experience” (P02). Furthermore, many participants noted a deeper emotional engagement with the game even

though the preferences varied widely initially (N=2). P26 talked about getting used to it: “It was tough to adapt at first, but once we clicked, it quickly became my favourite mode to play”. Also, one participant reflected on the learning curve, “Mastering this mode feels really rewarding. It’s like going on a shared adventure” (P39).

4.6.5 Summary of Results

To summarise, our results demonstrated a statistically significant effect of the shared control mode on user experience during certain game tasks. Furthermore, our results demonstrate that in the shared control mode, the performance of novice players improves. Moreover, our results also showed that the sense of agency and ownership of the shared hands increases throughout the gaming experience. Finally, our qualitative data revealed that the shared control mode introduced new game features and new challenges to the gaming experience. Many people reported that the shared control mode brought innovation to the game, whereas other participants found the shared control mode to be more difficult. In addition, our results showed that the participants enjoyed or preferred the shared control mode because of their personal preference.

4.7 Discussion

This chapter examined shared control in VR through a gaming context, but its significance extends beyond gameplay itself. The three tasks functioned as controlled probes that made coordination burden, input conflict, role assignment, and shared agency visible under different temporal and interaction demands. Viewed from the perspective of this thesis, the central contribution of *CoplayingVR* is therefore not only that shared control can change player experience, but that it exposes a more general collaborative embodiment problem: how control can be distributed across multiple users without collapsing usability, engagement, or the sense of embodied action.

Across the study, the findings suggest that shared control introduces both opportunities and constraints. On the one hand, it can enrich engagement, increase perceived challenge, support

positive affect, and improve novice performance. On the other hand, it can also amplify coordination costs, especially in fast-paced tasks that require tightly synchronized action. This tension between *scaffolding benefit* and *coordination cost* becomes the central theme of the discussion below.

4.7.1 Task-Specific Effects of Shared Control

4.7.1.1 The Control Task

In the *Control Task*, shared control improved task efficiency, particularly for novice players. Performance results suggest that the skillfulness of experienced players may have scaffolded the performance of their novice counterparts, echoing previous work on collaborative support and co-performance [262]. This result is especially important because it shows that shared control is not merely a coordination mechanism. It can also function as a form of in-situ support, redistributing task burden without fully removing the novice user from embodied participation.

Qualitative findings further show that participants were enthusiastic about the novelty of the shared control mode, suggesting receptivity to this form of interaction [229]. This novelty was not superficial. It was accompanied by a significant decrease in “Negative Affect” and an increase in “Positive Affect,” indicating that users not only tolerated shared control, but often found it enjoyable in this slower and more structured task. These observations are consistent with prior work suggesting that the introduction of advanced collaborative features can enhance enjoyment in game contexts [263].

4.7.1.2 The Balancing Task

Among the three tasks, the *Balancing Task* appeared to be the most naturally suited to shared control. Participants frequently described it as the most appropriate game for co-play, likely because its pacing gave users enough time to anticipate each other’s actions and recover from minor misalignments. This observation is important because it suggests that the success

of shared control depends not only on the control scheme itself, but also on the temporal structure of the task.

As in the *Control Task*, participants reported a higher sense of “Challenge” in the shared control mode, consistent with prior findings [225]. At the same time, the shared mode increased “Positive Affect,” reduced “Negative Affect,” and improved “Sensory and Imaginative Immersion.” These results suggest that, in moderately paced collaborative tasks, shared control may deepen engagement rather than merely introduce friction. The enhanced immersion may stem from a stronger sense of partnership and mutual dependence, which in turn made the gameplay more socially and experientially meaningful [264, 229].

4.7.1.3 The Shooting Task

The *Shooting Task* produced a different pattern. As the fastest-paced task, it placed the greatest demand on precise timing and real-time coordination. In this setting, shared control reduced perceived “Competence” relative to the single-player mode. Participants reported feeling less capable, less synchronized, and less effective, often attributing this to misalignment, asynchronicity, and difficulty coordinating with their partner. This suggests that the costs of distributed control become particularly visible when tasks demand fast, tightly coupled, high-frequency action [265, 266].

Performance data reinforce this interpretation. In the shared control mode, participant pairs produced significantly higher error rates than novice users in the single-player mode. One explanation is that experienced users, despite their skill, were unable to fully capitalize on it because they had to constantly align strategies and timing with another player [267]. By contrast, the higher error rates of novice users in the single-player mode relative to experienced users align with prior work on skill and visual familiarity [268]. Together, these findings suggest that shared control is not uniformly beneficial. Its value depends strongly on whether the interaction pace allows users to coordinate, anticipate, and negotiate control effectively.

4.7.2 Shared Control as Collaborative Mechanism and Scaffolding Strategy

Taken together, the findings reveal a central trade-off in collaborative embodiment: redistributing control can produce both coordination cost and scaffolding benefit. Shared control introduces additional dependencies between users. It can increase communication demands, timing sensitivity, and conflict, especially in fast-paced contexts. At the same time, it can lower the effective burden of complex tasks for less experienced users, allowing them to perform alongside more experienced partners without being excluded from embodied action.

This latter point is particularly important. The benefits observed for novice users suggest that shared control should not be understood only as a collaborative mechanism. It can also function as a pedagogical and assistive strategy. By allowing a novice player to act through a shared body with a more capable partner, the system redistributes task difficulty and provides real-time support without fully removing the novice from participation. In this sense, shared control offers a promising design strategy for training, coaching, and support-oriented applications, where learning and participation matter as much as independent task execution [269, 270, 271].

This interpretation is reinforced by behavioural observations. Participants often attempted to move their non-engaged hand together with the engaged one, revealing an initial tendency toward synchronized bilateral motor coordination. Over time, however, this behaviour diminished, particularly in later stages of the study. This adaptation indicates that users can gradually accept asymmetric distributions of control and incorporate partner-controlled limbs into their interaction model [119]. The increase in agency and body ownership toward the end of the gaming experience further supports this point: prolonged co-control may allow users to transition from initial conflict toward a more stable experience of shared agency.

4.7.3 Broader Implications Beyond Gaming

The contribution of this chapter is not limited to game design. The gaming tasks serve as an experimental lens for a broader HCI problem: how embodied action should be distributed across users in a shared virtual body. In such systems, the central question is not only whether people can collaborate, but how control should be partitioned, how conflicts should be resolved, and under what conditions shared agency can be sustained.

This broader perspective connects the chapter to existing work on non-immersive shared control and teleoperation. Shared control has long been explored in robotics, rehabilitation, and collaborative interfaces [218, 219, 272]. Compared to prior 2D shared control models [225], VR adds distinctive challenges and opportunities. Because users act through an embodied first-person interface, shared control in VR affects not only task execution, but also immersion, ownership, agency, and emotional engagement. This extension is important: the same control mechanism that supports coordination may also change how users experience the body through which they act.

These observations are especially relevant for serious games, rehabilitation, and assisted participation. Prior work has shown that shared control can support learning [273, 274], rehabilitation [275, 276], and serious gaming [272]. In such contexts, shared control could be used not only to improve task performance, but also to scaffold learning, maintain motivation, and create emotionally supportive interaction. For example, experienced users could act as coaches or trainers for novice users, helping them participate in training or rehabilitation tasks while preserving a sense of contribution and engagement [277, 278, 279, 280]. The increases in “Positive Affect” and reductions in “Negative Affect” observed in this study suggest that shared control may support such contexts not only instrumentally, but affectively as well.

4.7.4 Design Implications

The findings from this chapter suggest three design implications for shared control in VR.

- *Balance coordination cost against scaffolding benefit.* Shared control can improve novice performance and sustain engagement, but only when the coordination demands of the task remain manageable. Designers should therefore evaluate shared control not only in terms of efficiency, but also in terms of how much additional synchronization burden it introduces.
- *Favour tasks that reward cooperation without requiring constant micro-synchronization.* Tasks such as balancing and controlled manipulation appeared better suited to shared control than highly time-critical shooting tasks. When designing shared control experiences, task pacing, temporal tolerance, and opportunities for anticipation should be considered core design variables.
- *Support sustained co-play and adaptation.* Participants gradually adapted to asymmetric control and reported increased agency and body ownership over time. This suggests that repeated cooperation can help users internalize the shared control scheme. Systems should therefore support ongoing co-play, clear role structure, and progressive familiarization rather than assuming immediate fluency.

4.7.5 Alternative Configurations and Future Directions in Shared Control

This study focused on a specific control distribution in which each user controlled one arm of the avatar. Future work could expand this paradigm to other configurations, such as sharing the same limb, splitting forearm and upper-arm control, or extending shared control to the torso and lower limbs. Depending on the context, different fusion strategies such as averaged motion [134], divisional control [281], or time-multiplexed switching [282] may prove more effective. Exploring these alternatives is important not only for improving gameplay, but also for understanding the broader design space of shared control in collaborative embodiment.

4.8 Limitations and Future Work

Several limitations should be considered when interpreting the findings of this chapter. First, the set of game tasks was deliberately constrained. This was necessary to isolate the effect of shared control and to reduce confounding variables, but it also limits the range of interaction dynamics captured by the study. In real-world gaming and collaborative VR applications, user preferences, task structures, and difficulty levels are often more varied. Future work should therefore examine how shared control performs under a wider range of task demands and how game difficulty can be tuned to maintain a sense of fairness and meaningful contribution across partners.

Second, the participant pool limits the generalizability of the findings. Most participants were in a relatively narrow age range, and partner familiarity was not systematically controlled. Shared control may evolve differently when collaborators know one another well, have pre-existing strategies, or develop stable coordination patterns over repeated use [283]. Future work should investigate how social familiarity, repeated exposure, and long-term co-play affect the development of shared agency, body ownership, and collaborative strategy.

Third, the laboratory setting improved experimental control but reduced ecological validity. Although task counterbalancing helped mitigate order effects, laboratory-based task sequences may still shape learning, fatigue, and adaptation differently than naturalistic settings [284]. Future studies would benefit from more longitudinal or ecologically grounded designs that capture how shared control is used over time and in more realistic contexts.

A fourth limitation concerns measurement. While the Game Experience Questionnaire (GEQ) captured several useful dimensions of player experience, it does not fully address immersion and embodiment in ways that are specific to VR [249]. Future work should consider incorporating more VR-focused instruments, such as the Immersion Experience Questionnaire (IEQ) [285], alongside embodiment-specific measures to better understand how shared control shapes presence, ownership, and agency over time.

Finally, this work points to several promising application areas. Shared control could be especially valuable in accessibility, rehabilitation, and assistive contexts, where one user may benefit from the embodied support of another [286, 236, 287]. More broadly, future research should move beyond two-hand game interaction to examine full-body shared control, richer control fusion strategies, and task contexts such as coaching, therapy, and training. Such work would help clarify how the findings reported here extend from playful co-control to more consequential forms of collaborative embodiment.

4.9 Conclusion

In this chapter, we investigated how shared control influences user experience, performance, agency, and body ownership through three VR game tasks: the *Control Task*, the *Balancing Task*, and the *Shooting Task*. The findings show that shared control can improve several dimensions of user experience, including perceived challenge, sensory and imaginative immersion, and positive affect, while reducing negative affect in suitable task settings. Importantly, shared control also significantly improved the performance of novice users when they played together with experienced partners.

At the same time, the results show that shared control is not uniformly beneficial. Its success depends strongly on the temporal and coordination demands of the task. While slower and more cooperative tasks benefited from distributed action, the fast-paced *Shooting Task* exposed the costs of misalignment, synchronization burden, and reduced competence. This establishes a key insight for the broader thesis: redistributing control can generate both scaffolding benefit and coordination cost.

Finally, the chapter shows that shared agency can develop over time. As participants continued to act through a shared avatar, their reported sense of agency and body ownership increased, suggesting that users can gradually adapt to asymmetric and partner-mediated control. Taken together, these findings provide foundational evidence for understanding collaborative embodiment as a problem of control allocation and negotiated agency, and they motivate the broader design questions addressed in the following chapters.

4.9.1 Contribution to the Thesis

This chapter makes three thesis-level contributions.

First, it establishes **control mapping** as a central design problem in collaborative embodiment. By showing how two users negotiate action through a shared avatar, it demonstrates that multi-user VR cannot be understood only in terms of individual interaction channels; it must also account for how control is distributed and contested across collaborators.

Second, it reveals a recurring trade-off between **coordination cost** and **scaffolding benefit**. Shared control can increase communication overhead and conflict, but it can also improve novice performance, sustain engagement, and redistribute task difficulty in productive ways.

Third, it provides the first empirical basis for the thesis's later design synthesis. In particular, this chapter motivates the control-allocation dimension of the Interaction Elements Matrix by showing that how control is assigned, shared, and negotiated fundamentally shapes both collaborative performance and embodied experience.

CHAPTER 5

One Body, Two Minds

One Body, Two Minds: Alternating VR Perspective During Remote Teleoperation of Supernumerary Limbs

This chapter examines the second core dimension of collaborative embodiment in this thesis: **perspective configuration**. While prior work often treats viewpoint as a matter of camera selection, this chapter argues that, in collaborative VR, perspective is more fundamentally a problem of **distributed perception**. When two users act through or around a shared embodied system, what matters is not only what each user can control, but also what each user can perceive, monitor, and anticipate.

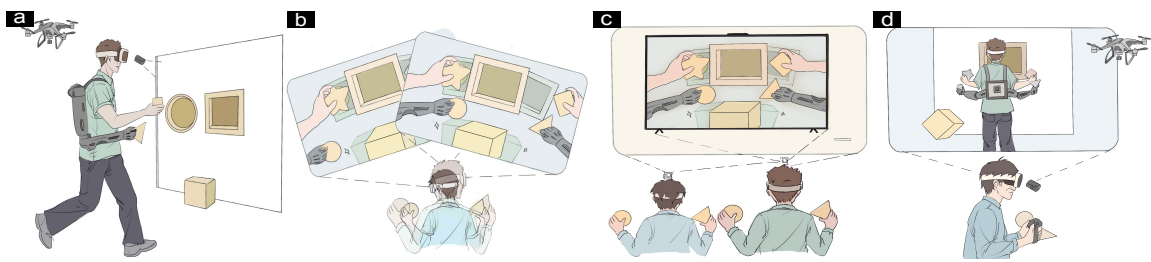


FIGURE 5.1. Illustration of three perspectives for remote collaborative teleoperation using virtual supernumerary limbs (VSLs): (a) *Host* user physically walking with VSLs attached to their back; (b) **SHARED EMBODIED VIEW**: the *guest* shares the *host*'s position with independent head rotation; (c) **EMBEDDED ANCHORED VIEW**: the *guest* views from the *host*'s head location through a stabilized virtual window; (d) **OUT-OF-BODY VIEW**: the *guest* independently navigates a drone-like external viewpoint using a handheld controller (the drone serves as a visual metaphor for the decoupled virtual camera).

To investigate this problem, this chapter studies how dynamic perspective switching supports remote collaboration when two users co-control a shared avatar augmented with virtual supernumerary limbs (VSLs). We begin with a formative study of five dyads, which reveals four recurring difficulties in VSL operation: prolonged visual coupling, constrained viewpoint flexibility, unstable control, and diminished embodiment. Guided by these insights, we introduce three guest-addressable perspective modes, Shared-Embodied View (SHARED EMBODIED VIEW), Embedded-Anchored View (EMBEDDED ANCHORED VIEW), and Out-of-Body View (OUT-OF-BODY VIEW), and implement them in a VR prototype. We then conduct a controlled study with 48 participants to assess how each perspective affects efficiency, comfort, accuracy, and perceived embodiment.

Within this thesis, this chapter contributes to the analysis of perspective as a collaborative design variable rather than merely a viewing condition. The findings suggest that perspective in collaborative VR is closely related to how perceptual access is distributed across collaborators, and that different users may benefit from different forms of visual access depending on their role and the demands of the task. Viewed in this way, collaborative embodiment need not depend on symmetric viewpoints. Instead, stable and task-sensitive asymmetries in perspective may support coordination while preserving usability and comfort.

The rest of the text of this chapter is a published paper: Zhou, H., Fan, C., Dong, Y., Takashita, S., Inami, M., Sarsenbayeva, Z., Withana, A. “One Body, Two Minds: Exploring Perspective Switching for Remote Collaboration with Virtual Supernumerary Limbs” In Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI 2026), and is kept unchanged to reflect the university thesis requirements.

5.1 Abstract

Remote VR teleoperation with supernumerary robotic limbs enables distant users to operate in another’s local space. While a shared first-person view aids hand-eye coordination, locking the guest’s camera to the host’s head can degrade comfort, embodiment, and coordination. Based on a formative study (N=10) using a virtual supernumerary robotic limbs configuration

to stress-test coordination, we propose guest-driven perspective switching from a shared first-person baseline (SHARED EMBODIED VIEW) to two alternatives: (a) a stabilized view with guest-controlled rotation (EMBEDDED ANCHORED VIEW), and (b) a fully decoupled third-person view (OUT-OF-BODY VIEW). We ran a user study with 24 pairs (N=48), who switched between the baseline and proposed views as task demands changed. We measured performance, embodiment, fatigue, physiological arousal, and switching behaviors. Our results reveal role-dependent trade-offs: OUT-OF-BODY VIEW improves navigation efficiency and reduces errors, while EMBEDDED ANCHORED VIEW supports embodiment. We conclude with guidelines: use EMBEDDED ANCHORED VIEW for hand-centric adjustments, OUT-OF-BODY VIEW for navigation and object placement, and ensure smooth transitions.

5.2 Introduction

“When you drift with someone, you feel like there’s nothing to talk about.” This evocative quote from *Pacific Rim* [288] describes the deep mutual understanding achieved through the fictional mental process that allows two users to co-pilot a single machine by sharing bodily sensations, intentions, and thoughts. It captures the promise of shared bodily control and sensorimotor synchrony. Though fictional, such scenarios are becoming increasingly plausible thanks to advances in telepresence and virtual reality (VR), enabling two individuals to act through a shared body, whether physical or virtual, in tightly coordinated ways [289, 16, 290].

Emerging robotic platforms, such as the humanoid Neo [291], suggest a future where remote operators inhabit complex physical bodies. However, controlling locomotion while simultaneously performing precise manipulation can place high cognitive demands on a single operator [292]. Shared control is emerging in these scenarios, for instance, Multiple Operator, Single Robots (MOSR)[293]. A natural emergence of this is co-embodiment, splitting control between a host and a guest, for instance, between navigation and manipulation. This, in turn, raises a central design question: *viewpoint coupling*, that is, how tightly a guest user’s camera is linked to the host’s head and body motion, which directly affects how fluidly the pair can

collaborate. In most VR systems, the remote first-person view is tied to the local user’s head; small head movements therefore propagate to the shared camera, destabilizing near-body control and blurring role boundaries [11, 294]. Simultaneously, advances in AI-based 3D scene reconstruction and world models are making it feasible for teleoperators to work inside navigable virtual replicas of real environments [295, 296, 297, 298, 299]. In such digital twins, anchored or “drone-like” viewpoints can be realised virtually, even while the robot remains grounded in the physical world. Therefore, studying shared viewpoint control in VR offers a way to probe how co-embodied collaborators should manage perspectives for future telepresence systems. In this work, we focus on viewpoint management and shared visualisation, rather than simulating specific robot kinematics or control hardware.

In this paper, we developed a collaborative VR scenario in which two users, a *host* and a *guest*, share control over the same virtual body. To examine viewpoint management under demanding conditions, we augment this body with virtual supernumerary limbs (VSLs). We use this multi-limb configuration not as the sole target application, but as a challenging testbed that maximizes overlap in the near-body workspace, so that insights about viewpoint coupling can generalize to less complex shared-control scenarios. A formative study with five participant pairs identified four empirical patterns: (1) prolonged visual coupling during locomotion co-occurring with *guest* discomfort and blurred self-location; (2) limited perspective flexibility aligning with grasping errors; (3) VSL control instability during the *host*’s head position changes; and (4) role ambiguity leading to *host* pauses. Motivated by these observations, we adopt a task-contingent approach: the *host* remains in a standard HMD first-person view, while the *guest* may switch their viewpoint on demand to align coupling with navigation versus near-body phases.

Building upon these insights, we implemented three perspectives for the *guest*: (i) SHARED EMBODIED VIEW, our design baseline, which co-locates the *guest*’s camera with the *host*’s head position. In this configuration, the *guest*’s head position is locked to the *host*, and moving the *guest*’s head does not change the viewpoint location but only rotates the view direction independently. (ii) EMBEDDED ANCHORED VIEW shows a stereoscopic portal near the *host*,

streaming a steady first-person view for the *guest*; and (iii) OUT-OF-BODY VIEW, a world-anchored, fully decoupled 6-DoF third-person camera for spatial awareness. Operationally, the *host* remains in SHARED EMBODIED VIEW throughout. The *guest* can switch between SHARED EMBODIED VIEW and either EMBEDDED ANCHORED VIEW or OUT-OF-BODY VIEW, and our study investigates two *guest*-driven switching strategies: SHARED EMBODIED VIEW $\leftrightarrow$ EMBEDDED ANCHORED VIEW and SHARED EMBODIED VIEW $\leftrightarrow$ OUT-OF-BODY VIEW.

We then conducted a controlled, within-subjects study with 24 pairs of participants (N=48) to examine how perspective modes affect task performance, embodiment, workload/fatigue, physiology, and guest switching. Results show that OUT-OF-BODY VIEW yielded fewer errors in the Factory task with comparable completion times, while participants switched to EMBEDDED ANCHORED VIEW more frequently for precision-demanding subtasks. Guests reported higher subjective workload and fatigue in OUT-OF-BODY VIEW, although HRV indicated lower physiological stress; participants tended to use EMBEDDED ANCHORED VIEW for precision phases and OUT-OF-BODY VIEW for navigation.

Hence, this research offers the following key contributions:

- **Empirical diagnosis.** A formative study diagnosed two systemic breakdowns, coordination friction and guest disorientation, arising from fixed first-person coupling in co-embodied multi-limb VR control.
- **System & evaluation.** We introduce guest-driven perspective switching and evaluate two regimes (SHARED EMBODIED VIEW $\leftrightarrow$ EMBEDDED ANCHORED VIEW SHARED EMBODIED VIEW $\leftrightarrow$ OUT-OF-BODY VIEW) in a controlled within-subjects study (N=48), measuring performance, workload / fatigue, embodiment, physiology, and switching behaviour.
- **Design implications.** We derive guidelines on when to use SHARED EMBODIED VIEW, EMBEDDED ANCHORED VIEW, and OUT-OF-BODY VIEW to manage view-point coupling and support comfort, coordination, and embodiment in co-embodied multi-limb VR.

5.3 RELATED WORK

In this section, we contextualize our research within existing literature on perspective control, remote embodiment, and supernumerary robotic limbs. We highlight key insights and identify the knowledge gaps our study aims to address.

5.3.1 Perspective Control

Viewpoint selection significantly impacts remote collaboration and coordination, influencing both task performance and user experience. Here, we distinguish coordination from collaboration. Coordination stands for the alignment of actions and timing between collaborators, whereas collaboration refers to the broader process of jointly working toward a shared goal, including planning, decision-making, and mutual understanding [300]. Prior research has explored the effects of different perspectives on remote interactions and their impact on coordination. Galvan et al. [301] demonstrated how alternating between first-person and third-person perspectives can leverage the unique advantages of each viewpoint without significantly harming body ownership of the virtual body. Similarly, Komiyama et al. [302] constructed a system that allows users to freely switch between first-person and third-person views in a remote work setting, enabling precise task execution through first-person images while maintaining situational awareness via third-person viewpoints. Saraiji et al. [289] introduced a wearable system where a teleoperator controls two robotic limbs remotely, using an HMD to observe the environment from a shoulder-mounted perspective, highlighting the importance of viewpoint selection for skill-sharing and collaboration.

However, viewpoint discrepancies between collaborators remain a significant challenge, particularly in object-focused tasks where precise spatial alignment and gesture interpretation are crucial. Nagai et al. [303] explored 360-degree wearable cameras, allowing users to observe and communicate with an entire environment from a remote location, potentially enhancing collaborative spatial awareness. Similarly, Jones et al. [304] highlighted the inherent limitations of handheld perspectives in mobile video chat, as collaborators struggled to perceive shared scenes clearly due to instability and narrow visual fields. Fussell et al. [305]

found that scene-focused camera angles outperformed head-mounted perspectives in physical task coordination, as they provided better visibility of the shared workspace. Further, Tang et al. [306] showed that different perspectives optimize different task demands, with shared perspectives aiding textual reading, while asymmetric perspectives facilitate the creation of shared and personal workspaces.

Beyond egocentric and third-person views, prior work proposed techniques like Worlds-in-Miniature (WIM) [307], which offers a scaled-down, allocentric replica of the environment, later adapted for real-world control [308]. Go-Go [309] enables non-linear arm extension to reach distant objects. These methods expand the design space for remote perspective control and inform our investigation of dynamic viewpoint switching.

In summary, prior work shows that flexible viewpoint control can improve spatial awareness and collaboration in VR and telepresence. However, we still know little about how dynamic, user-driven perspective switching shapes coordination, embodiment, and fatigue when two users co-embody a single avatar and jointly control multiple limbs in a shared virtual environment. Our study introduces adaptive guest-driven switching between SHARED EMBODIED VIEW, EMBEDDED ANCHORED VIEW, and OUT-OF-BODY VIEW, to probe the trade-offs in control fluency, comfort, and visual stability under collaborative embodiment.

5.3.2 Embodiment in Collaboration

In this paper, we distinguish between “self-embodiment”, which describes a user’s ownership/agency over their own avatar (commonly measured with AEQ [18, 310]), and “remote/co-embodiment” or “social presence”, which describe mutual awareness of a partner. These constructs are related but distinct. In our evaluation, we quantify self-embodiment and user experience.

Self-embodiment relates to a user’s ownership, agency, and self-location with respect to their own avatar. Remote embodiment refers to visual proxies that convey a collaborator’s bodily state and support awareness and coordination. Our study quantifies self-embodiment while using lightweight partner cues to aid reference. Research has explored methods where

multiple users share control over a single avatar, using techniques like the weighted-average co-embodiment method, which distributes control percentages between participants [118, 311, 312]. Studies show that increasing control weight improves users' sense of agency, enhancing task coordination [312]. This method has been effective in contexts like VR-based rehabilitation, where even users with less control feel engaged [313, 314]. Further work in multi-operator single robot (MOSR) systems examined how multiple users control a robot, later evolving to include shared control in VR environments with a focus on first-person perspectives and multisensory feedback, affecting psychological aspects like intention alignment [315, 316, 83]. The phenomenon of "enfacement", where users perceive a merged identity, highlights the immersive potential of shared control [146].

Complementary to these control-centric approaches, effective collaboration also depends on how partners are represented and perceived, i.e., embodiment cues that support mutual awareness, reference, and non-verbal coordination. Partner embodiments span a spectrum from minimal surrogates such as the Telepointer [317], through visualisations of partner hands [318, 16, 319, 320, 102], arms [321, 322, 306, 323], and feet [324], to full-body AR/VR telepresence avatars [325, 326, 327]. These representations improve joint attention, reference establishment, and social presence. When gestures alone are insufficient for object reference, raycasting and local scene reconstruction offer stable referents [328, 329].

While remote embodiment is known to support collaboration, its interplay with dynamic perspective control remains underexplored in co-embodied VR scenarios. We examine this relationship in a co-embodied avatar, focusing on how shared-viewpoint schemes shape coordination and embodied experience.

5.3.3 Supernumerary Robotic Limbs

Research on supernumerary robotic limbs (SRL) has investigated various ways of augmenting human bodily capabilities [330, 331, 332, 45, 333], including additional robotic legs [334], fingers [332], and arms [45]. Our study focuses specifically on virtual supernumerary limbs (VSLs), aiming to expand user capabilities within virtual reality environments.

Initially, SRLs emerged from industrial contexts, designed to assist workers in demanding physical tasks such as construction and assembly operations [335, 336, 337]. Over time, the scope of SRL research broadened significantly, encompassing novel designs such as robotic tails to assist balance [338, 339], and extra limbs facilitating complex manipulations [340, 330, 341, 342]. For instance, Maekawa, Kawamura and Inami [338] presented a wearable robotic tail that enhanced stability during physically challenging tasks. Control strategies for SRLs have consistently been a central research theme, with multiple intuitive mechanisms proposed to facilitate efficient limb manipulation. These include methods like remapping existing body movements (e.g., utilizing foot or shoulder gestures)[331, 8, 343], and employing brain-computer interfaces (BCI) for direct neural control[43]. While promising, many of these approaches depend heavily on fixed control mappings, which restrict adaptability to dynamic and varying task conditions. Further, comprehensive user evaluations remain scarce, as most studies are limited to concept validations involving single-user scenarios [331, 44, 45]. In collaborative contexts, sharing body control between two users introduces additional demands, such as division of roles, viewpoint negotiation, and real-time coordination, that fixed mappings do not address.

In summary, SRLs and VSLs demonstrate substantial potential, while most prior evaluations emphasize single-user prototypes with fixed mappings. We complement prior work by studying guest-driven switching from SHARED EMBODIED VIEW to EMBEDDED ANCHORED VIEW or OUT-OF-BODY VIEW, improving coordination, efficiency, and comfort when two users co-control one body.

5.4 FORMATIVE STUDY

To surface early-stage challenges in shared-body collaboration during locomotion, we conducted a diagnostic formative study focused on the limitations of a fixed shared first-person view, which is commonly adopted in prior work [17, 289].

5.4.1 Method

We recruited five dyads ($N = 10$), each composed of a *host* and a *guest*, all with prior VR experience but no exposure to shared-control systems. This ensured that participants were not distracted by the novelty of VR itself or its basic controls, allowing us to better isolate usability challenges specific to shared-body interaction. Both participants were immersed in the same virtual environment via head-mounted displays. The *host* was embodied as a full-body avatar augmented with a pair of Virtual Supernumerary Limbs (VSLs) attached to their back. The *guest* shared the *host*'s first-person view via a camera rigidly anchored to the *host*'s head position. While the *guest* could not change the camera's spatial location, they retained independent control over view orientation through head rotation. This configuration, which we term the SHARED EMBODIED VIEW, served as the baseline design condition (Figure 5.2).

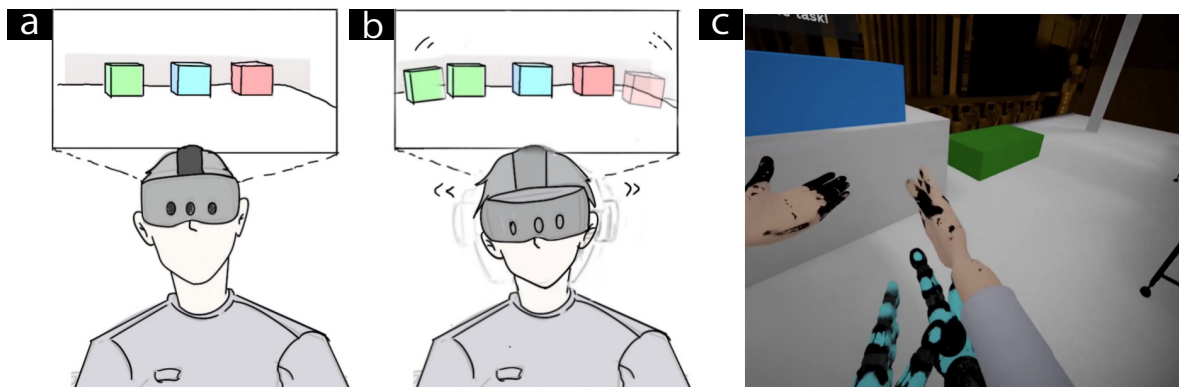


FIGURE 5.2. Formative study configuration showing (a) the stable first-person perspective from the *host*, (b) the *guest*'s co-located but independently rotatable view illustrating visual instability, and (c) the avatar embodiment with VSLs used for near-body manipulation.

Task and Procedure: Participants performed a simplified collaborative task designed to surface visual, operational, and embodiment challenges specific to shared-body teleoperation with VSLs. The task involved the joint transport of small boxes across a 2–3 meter space. The *guest* used VSLs to grasp objects from a table; after a successful grasp, the *host* navigated to a placement zone. Before the main trials, each pair had a short familiarisation period to practise grasping and transporting boxes using the VSLs. They then completed two 5-minute recorded trials. The short duration was intended to foreground early-stage coordination breakdowns

and discomfort, rather than capture long-term adaptation effects. We collected observational notes, head rotation logs, VSL activity traces, and think-aloud comments [344, 345]. Post-trial, participants were briefly interviewed on discomfort, control issues, and coordination experiences [346].

Measures and Analysis. To identify breakdowns in shared-body coordination under fixed first-person coupling, we used three diagnostic indicators: (i) prolonged head rotations by the *host* (lasting >10 s), which are known to induce discomfort in the *guest* [347, 348]; (ii) grasping errors (misses or drops per attempt) to assess VSL precision [349, 350]; and (iii) self-reported loss-of-control from the *guest* and observed host pauses as markers of coordination breakdown [351]. Two authors independently coded observation notes and interview/think-aloud transcripts using these three indicators as deductive anchors. Discrepancies were resolved through discussion, and related codes were grouped into macro-level themes following reflexive thematic analysis [352, 353, 354]. We triangulated qualitative excerpts with synchronized system logs (e.g., head rotation, VSL activity, locomotion events) to support interpretation [355, 356]. Given the small sample size ($N = 10$), we report dyad-level coverage, representative excerpts, and cross-modal log co-occurrence to inform design, not to claim generalizable causality [357].

5.4.2 Findings

5.4.2.1 Guest Control Drops Without Perspective Flexibility (5/5 dyads).

Guests reported losing spatial orientation when the *host* was moving. We define this as disorientation during *host* movement that coincides with grasping errors. Participants remarked, “I kept losing where the arms” (*guest*, P04) and “felt like the hands lagged behind what I wanted to do” (*guest*, P08); *hosts* noticed this too (“When I started walking, he said ‘wait’,” *host*, P03). Grasping errors often occurred shortly after the *host* began moving, according to system logs—suggesting a close relationship, though not a causal one. These findings highlight the need for a viewpoint mode that maintains stable, body-centered alignment during fine motor tasks, allowing *guests* to invoke it as needed for precise control.

View	Camera Anchoring	Orientation	Translation	Render
SHARED EMBODIED VIEW	Host-anchored	Guest head (independent)	No	Full-screen
EMBEDDED ANCHORED VIEW	Decoupled-anchored	Guest head (linked)	No	Stabilized portal
OUT-OF-BODY VIEW	World-anchored (free-floating)	Guest head / controller	Yes (6-DoF)	Full-screen

TABLE 5.1. Differences across views along two axes: position anchoring and view control. Host always stays in SHARED EMBODIED VIEW; the guest switches between SHARED EMBODIED VIEW and one alternative per condition.

5.4.2.2 Hosts Stop Moving Due to Role Confusion (3/5 dyads).

Following prior work on detecting stillness in movement trajectories and HMD logs [358, 359], we identified *host* pauses as periods of minimal movement lasting several seconds during ongoing tasks. *Hosts* often attributed these pauses to uncertainty about whether continued movement would interfere with the *guest*'s manipulation. Participants remarked, "I kept stopping, unsure if moving would disrupt them" (*host*, P03) and "I waited here until he finished" (*host*, P05), with *guests* corroborating this dynamic ("I needed to say 'don't move'" *guest*, P06). These co-occurrences suggest hesitation rooted in unclear role boundaries, rather than deliberate coordination.

5.4.2.3 Prolonged Shared View Causes Cybersickness (5/5 dyads).

When the *guest* shared the host's first-person view, extended walking or turning by the host often coincided with verbal reports of dizziness, disorientation, or difficulty controlling the arms. Participants described feeling dizzy or unsure of their location after several seconds of walking or turning, e.g., "long turns made me dizzy" (*guest*, P02), and "I [kept] losing where the arms [were], and felt nauseous" (*guest*, P06). Discomfort did not arise from sudden turns alone, but more often from sustained motion lasting over 10 seconds, even when the motion was low in intensity. In the system logs, these episodes overlapped with sustained head rotations or locomotion events. Grasping errors were also more likely to occur shortly after such movement sequences. Although we make no causal claims, this repeated co-occurrence points to a possible mismatch between visual motion and control stability during prolonged shared viewing.

5.4.2.4 Camera Shifts Disrupt Guest's VSL Control (4/5 dyads).

When the *guest* shared the host's camera view, sudden or sharp head turns by the *host* often disrupted the *guest*'s ability to control the virtual arms. Participants described losing targeting accuracy or feeling disoriented during these transitions, such as “Hard to keep control when everything shifts” (*guest*, P06) and “every turn jolted my aim” (*guest*, P02). Triangulating with locomotion logs, these breakdowns frequently occurred within short windows after *host* heading-change events, evidence of time-adjacent co-occurrence rather than causality.

5.4.2.5 Summary of Findings.

Our formative study revealed recurring difficulties with a fixed shared first-person view. First, *guests* often lost track of hand position during precise actions, especially when *hosts* moved their head mid-task, leading to grasping errors and hesitation. At the same time, *hosts* frequently paused movement, unsure whether the guest was in control, reflecting ambiguity in coordination roles. Second, shared locomotion episodes where the *host* moved while the *guest* remained visually coupled caused discomfort, disorientation, and disengagement for the guest. These issues clustered into two under-explored challenge areas: (C1) coordination breakdowns during precision tasks, and (C2) discomfort and control loss from shared movement. System logs confirmed that most errors clustered around host motion: about 71% of logged errors (45/63) occurred within 4 s of walking onset or a sharp head turn. These findings highlight the need to restore guest perceptual agency and reduce coordination friction during task transitions.

5.5 DESIGN AND IMPLEMENTATION

To address the issues surfaced in our formative study, we designed two guest-driven perspective modes that directly target the observed challenges. In C1, coordination broke down during precise manipulation due to ambiguous visual control. To address this, we introduce the EMBEDDED ANCHORED VIEW, which grants the guest independent rotational control of the viewpoint. This stabilizes the visual frame during fine-grained tasks and helps clarify

interaction roles. In C2, guests experienced discomfort and disorientation caused by involuntary shared locomotion. To address this, we implemented the OUT-OF-BODY VIEW, a third-person, freely positionable camera that decouples the guest’s perspective from the host’s body, easing spatial confusion and supporting higher-level navigation decisions. The specific distinctions between these perspective modes are summarized in Table 5.1.

5.5.1 Embedded Anchored View (C1)

We implemented EMBEDDED ANCHORED VIEW as a body-anchored yet independently rotatable viewpoint, as shown in Figure 5.4 (a). The system renders a stabilized egocentric feed as a stereoscopic portal, positioned at a fixed offset from the host and oriented toward the host’s head-camera direction. This provides visual stability while preserving directional control. We also apply light IPD tuning and frame smoothing to support accurate manipulation.

Previous research has demonstrated improved depth perception and task accuracy with slight inter-pupillary distance (IPD) adjustments in VR [360, 361]. Thus, we slightly reduced the guest’s effective IPD within the EAV window to enhance binocular depth cues, facilitating precise grasping and alignment tasks.

Moreover, as rapid *host* movements can disrupt the *guest*’s reference and degrade VSL stability [362], we employed Unity SmoothDamp stabilization (exponential smoothing) on EAV camera updates to soften *host* jitter. This preserves a steady body-anchored reference during *host* motion, improving comfort and control precision.

Finally, to reinforce role clarity, we emphasized visual embodiment cues known to reduce hesitation in collaboration [363]. While the shared avatar is visible in all modes, EMBEDDED ANCHORED VIEW shows a steady 3D window that keeps the host’s hands, arms, and VSLs visually locked to the body with correct in-front/behind relationships, making the control locus near the body unambiguous.

5.5.2 Out-of-body View (C2)

OUT-OF-BODY VIEW provides a fully decoupled, world-frame third-person camera with 6-DoF motion (rotation and translation), independent of the *host*'s movement, as shown in Figure 5.4 (b). The *guest* controls the drone camera via joystick inputs. We apply smooth interpolation and bound speed/acceleration, with collision checks to prevent disruptive motion.

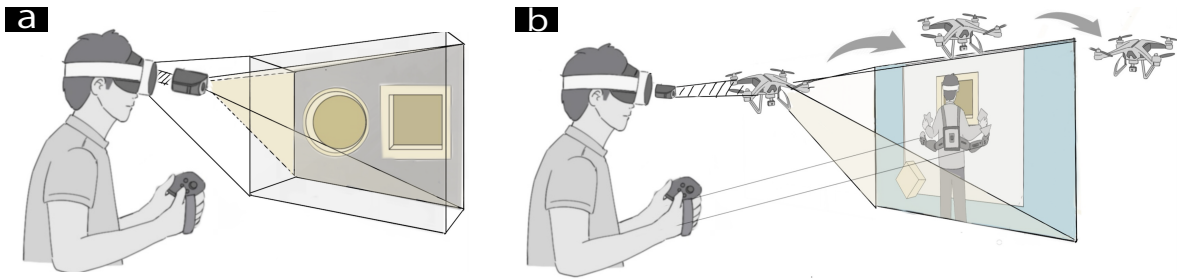


FIGURE 5.3. Illustration of the EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW. (a) The *guest* user's viewpoint is spatially aligned with the *host*'s viewpoint but rendered as a video-like window. (b) The *guest* user independently controls a virtual drone camera with 6 degrees-of-freedom (6-DoF), enabling flexible positional and rotational adjustments independent of the *host*'s movements (the drone serves as a visual metaphor for the decoupled virtual camera).

To improve the *guest*'s spatial awareness and viewpoint flexibility, we implemented a 6-degree-of-freedom (6-DoF) virtual drone camera. The *guest* controlled positional (forward, backward, lateral, vertical) and rotational (pitch, yaw, roll) movements using intuitive VR joystick inputs. Smooth camera transitions were ensured through spherical linear interpolation (Slerp) and linear interpolation (Lerp), enhancing visual comfort and spatial understanding.

Stabilizing VSL control during *host*'s movements required a stable spatial reference frame. Thus, we implemented the drone camera in a fixed world coordinate system, independent of *host* movement [364]. Additionally, real-time collision detection and spatial constraints via Unity's raycasting prevent visual disruptions and maintain stable associations with the shared environment and avatar [365].

5.5.3 Perspective Switching Mechanism

In our system, SHARED EMBODIED VIEW is the default host-centric view and serves as the baseline from which other perspectives are invoked. To overcome limitations of this fixed shared view [366, 367], we implement a *guest-driven* perspective-switching mechanism that lets users fluidly alternate between views based on task demands.

This mechanism allows the *guest* to transition from SHARED EMBODIED VIEW to a condition-specific alternative view using VR controller input, which activates or deactivates dedicated Unity camera components. To preserve comfort and spatial context, each switch is rendered with short cross-fades, input debounce, and on-screen indicators. The design enables the *guest* to time perspective transitions to task needs: invoking EMBEDDED ANCHORED VIEW’s stabilization, body-aligned portal for fine manipulation, or OUT-OF-BODY VIEW’s fully decoupled third-person view for spatial planning. This mechanism directly supports user agency, reduces prolonged visual coupling, and maintains stable VSL control during host movement—addressing the challenges identified in our formative study.

5.5.4 VSLs Independent Control and Stabilization

Formative findings indicated that VSL control became fragile under *host* locomotion and heading changes. We therefore decouple VSL actuation from *host* head/body transforms and drive VSL poses in a world-referenced frame initialized when the *host* is stationary. Using Meta Quest 3 tracking, we monitor head orientation and translation, and flag short-window spikes in angular velocity or position delta as “significant motion.” When flagged, a stabilization mode engages: VSLs are either held in place or eased back to a predefined neutral pose via Unity `SmoothDamp` to avoid abrupt shifts. Guest input immediately disables stabilization and enters active control to preserve precision and responsiveness. This mechanism mitigates motion-induced disruptions without altering the *host*’s baseline view and applies uniformly under SHARED EMBODIED VIEW and EMBEDDED ANCHORED VIEW (the difference is rendering only).

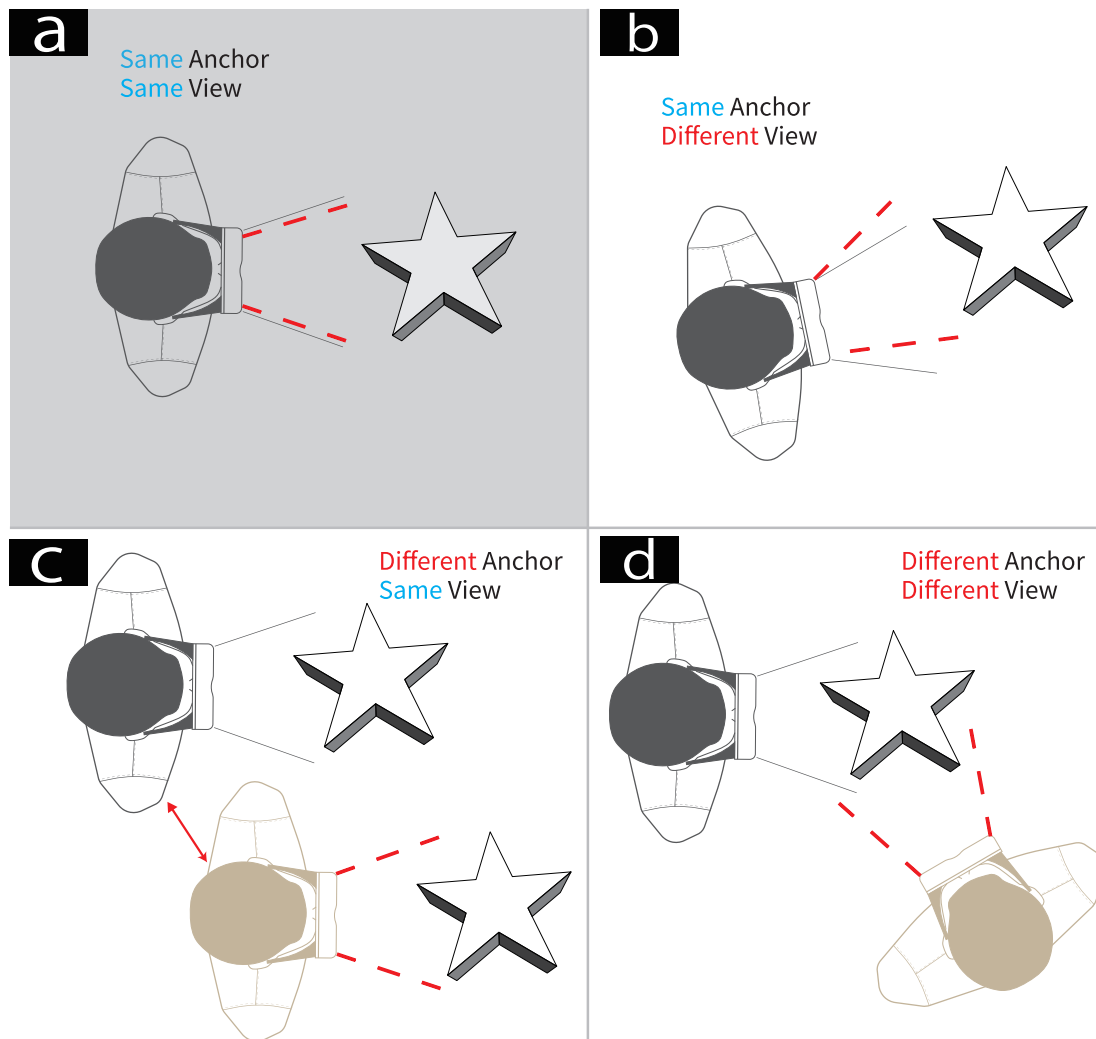


FIGURE 5.4. Guest perspectives organized as an *Anchor-View* matrix relative to the host. (a) **Locked Egocentric:** *same anchor, same view*; guest camera is rigidly locked to the host's head pose and view. (Shown for comparison; not studied, as it grants the guest no agency and renders them a passive observer under full host control). (b) **SHARED EMBODIED VIEW:** *same anchor, different view*; guest shares the host's position but rotates the camera independently. (c) **EMBEDDED ANCHORED VIEW:** *different anchor, same view*; guest remains independently positioned but views a stabilized window of the host's egocentric feed. (d) **OUT-OF-BODY VIEW:** *different anchor, different view*; guest drives a free-floating 6-DoF drone camera, decoupled from the host.

5.5.5 Implementation Details

We developed the system in Unity 2022.3.7f1 and deployed it on two Meta Quest 3 headsets. Both users share one anthropomorphic avatar in VR: the *host* navigates and controls the primary arms, and the *guest* operates the back-mounted VSLs. Each headset was tethered via USB to a dedicated PC on the same local network, achieving latency below 5ms [368]. Multiplayer synchronization used Photon Unity Networking with a locally hosted Photon Server to minimize network hops. The *host* and *guest* were located in separate rooms to preserve a remote collaboration context while enabling natural voice communication via audio only. Perspective control, VSL manipulation, and task logic were integrated into the Unity scene. On-screen indicators show the current mode. A logging module records head kinematics, perspective switches, VSL commands, and grasp events.

5.6 EXPERIMENTAL METHOD

We ran a within-subjects experiment to examine how two guest-controlled view-switching conditions shape user strategies and collaborative behaviors when operating VSLs with a shared-body avatar. Both tasks required tight coordination between the *host* and *guest*, emphasizing precision, timing, and role-dependent experience in dynamic VR scenarios.

To align with our revised research scope, we address the following RQs:

- **RQ1:** How do EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW affect subjective workload, fatigue, and physiological responses during collaboration?
- **RQ2:** How do EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW influence self-embodiment (AEQ) for hosts and guests?
- **RQ3:** How do EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW affect task performance (completion time, errors) in dynamic shared-body collaboration when switching is initiated from the SHARED EMBODIED VIEW?
- **RQ4:** When and why do *guests* switch between SHARED EMBODIED VIEW ↔ EMBEDDED ANCHORED VIEW or SHARED EMBODIED VIEW ↔ OUT-OF-BODY VIEW as task

demands evolve, and how are these switches associated with performance and experience?

5.6.1 Experimental Conditions and Procedure

5.6.1.1 Experimental Conditions

We tested two conditions: SHARED EMBODIED VIEW↔EMBEDDED ANCHORED VIEW and SHARED EMBODIED VIEW↔OUT-OF-BODY VIEW. In both conditions, tasks began in SHARED EMBODIED VIEW; the *guest* could switch to the condition-specific alternative, while the *host* remained in SHARED EMBODIED VIEW throughout. Each session comprised two task runs. Task type and condition order were fully counterbalanced across participant pairs to mitigate learning effects and fatigue, ensuring each pair experienced one run per condition.

5.6.1.2 Participants

We recruited 24 participant pairs (48 participants; 24 female, 24 male; M=22.8 yrs, SD=3.2 yrs; no color vision deficiencies). This sample size aligns with common practice in HCI research involving collaborative VR systems [142]. All participants had prior experience with VR [143]. 16 participants reported regular VR use, while 32 primarily engaged with VR for entertainment. Each participant pair consisted of a *host* and a *guest*.

5.6.2 Experimental Tasks

5.6.2.1 Transportation Task

The first task simulated a mobile manipulation scenario commonly encountered in collaborative work and human-robot collaboration environments [369, 370, 371, 372], as shown in Figure 5.5 (a,b). In each trial, the participant pair (host and guest) started from a designated starting location. Their goal was to collaboratively transport colour-coded boxes placed centrally on a table into corresponding target bins distributed across a virtual room approximately

6×8 meters in size, populated with static obstacles such as tables, walls, and barriers. The *host* navigated the virtual environment freely by physically walking in the tracking space, ensuring balance, obstacle avoidance, and efficient positioning, while the *guest* remotely controlled the VSLs using VR controllers, pressing specific buttons to grasp and release boxes accurately, a role distribution consistent with prior collaborative robotics studies [373, 374]. To enforce joint action, the boxes were designed to be large enough to require two hands to lift and carry, mimicking the size and weight of bulky real-world objects. While two hands stabilized the load, the remaining two were required to interact with the environment by pushing aside obstacles or lifting spring-loaded bin covers, creating situations in which all four hands had to operate simultaneously.

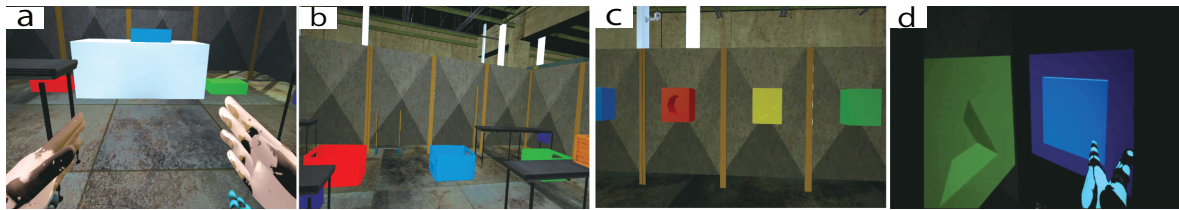


FIGURE 5.5. Task overview: Transportation and Factory. Panels (a–b) depict the *Transportation* task: (a) colour-coded boxes placed on a central table; (b) matching colour-coded bins distributed throughout the VR environment. Panels (c–d) depict the *Factory* task: (c) geometrically distinct holes on a wall panel as precise insertion targets; (d) collaborative insertion of objects (e.g., cubes) into the corresponding holes.

This task required continuous collaboration: the *host* needed to provide optimal positioning for the *guest*'s manipulation tasks, whereas the *guest* had to adapt to continuous perspective shifts caused by the *host*'s movement. The design intentionally introduced continuous demands for spatial navigation and fine-grained manipulations, enabling systematic investigation of the impacts of dynamic viewpoint control, cognitive load, and collaborative efficiency in shared-body remote teleoperation scenarios.

5.6.2.2 Factory Task

The second task required participants to perform precision object manipulation and insertion tasks collaboratively, emphasizing precise spatial coordination and multi-limb synchronization [375], as shown in Figure 5.5 (c,d). In each trial, participants simultaneously controlled

the avatar's four hands (two controlled by the *host* and two by the *guest*) to insert four geometrically distinct objects (e.g., cubes, cones, spheres, and cylinders) into corresponding matching holes on a virtual wall panel. Each insertion required precise alignment and steady positioning to succeed; insertion errors included misalignment, improper orientation, or unsuccessful attempts. The *host* controlled the avatar's locomotion and two primary arms, while the *guest* controlled the two extra arms. Clear visual feedback indicated successful insertion. To encourage multi-limb synchronization, we designed a timed batch-insertion task. All four target holes were always available, and participants were asked to complete all insertions before the trial time limit expired. The pick-up area and insertion panel were spatially separated, encouraging participants to carry multiple objects per trip when possible. Using all four hands in parallel helped reduce walking, minimize idle time between insertions, and improve overall efficiency. While not strictly required, coordinating all four hands became the most effective strategy for completing the task on time. The task specifically aimed to evaluate how dynamic perspective-switching influenced precision tasks requiring fine motor coordination, collaborative synchronization, and role-specific embodiment under varying viewpoints.

5.6.3 Experimental Data Collection

To evaluate task performance and switching strategies, we collected quantitative data covering subjective embodiment, workload and fatigue, performance metrics, physiological responses, and qualitative data through interviews and think-aloud protocols.

5.6.3.1 Qualitative Data

We recorded audio and video during the think-aloud session and post-task semi-structured interviews as qualitative data.

5.6.3.2 Performance Measures

Task performance metrics included completion time, error rate, and perspective-switching behaviors. In the Transportation Task, we recorded completion time (from grasp initiation to successful bin placement) and errors (failed grasps, drops, or wrong-bin placements). In the Factory Task, we measured insertion time per object and recorded insertion errors, which included both translational misalignments (incorrect positional offsets relative to the target hole) and rotational misalignments (incorrect orientation of objects relative to the required target orientation). Participants received no performance feedback during trials.

5.6.3.3 Subjective Feedback

We administered three standardized questionnaires:

The Avatar Embodiment Questionnaire (AEQ) measured self-embodiment, consistent with prior work on VR embodiment [18]. Following prior SRL-in-VR work that administers AEQ in a pre/post-adaptation schedule [23], we followed the same timing logic and extended it to our two perspective-switching strategies by collecting one pre-adaptation AEQ under SHARED EMBODIED VIEW and two post-adaptation AEQ: one after SHARED EMBODIED VIEW → EMBEDDED ANCHORED VIEW and one after SHARED EMBODIED VIEW → OUT-OF-BODY VIEW.

NASA Task Load Index (NASA-TLX) assessed perceived workload (mental, physical, temporal, effort, performance, frustration) [376]. Visual Analog Scale for Fatigue (VAS-F) measured subjective fatigue across 18 dimensions [377]. While established measures of cybersickness such as the Simulator Sickness Questionnaire (SSQ) [83] are available, in this first study we focused on workload (NASA-TLX) and fatigue (VAS-F) as our primary subjective indicators of discomfort in shared viewpoints, to limit questionnaire burden and align the scales with our task demands. We acknowledge the absence of a dedicated cybersickness instrument as a limitation and see it as an important addition for future work.

5.6.3.4 Physiological Data - Heart Rate Monitoring

We continuously recorded heart rate data using a Polar H10 chest strap [378], known for accuracy and minimal interference with VR [379, 380]. Specifically, we analyzed heart rate variability (HRV) and peak heart rate during intensive task phases and critical perspective-switching events to quantify physiological correlates of subjective experience [381, 382]. RR-interval series were preprocessed following psychophysiology recommendations [383]. We removed artefact-contaminated/ectopic intervals, interpolated only short gaps, and computed RMSSD in sliding windows during active task periods (full parameters in Appendix 1). Heart rate variability (HRV) was computed using the time-domain RMSSD index (root mean square of successive differences between adjacent NN intervals). Higher values indicate lower physiological stress.

5.6.3.5 Procedure

We adopted a fully within-subjects, repeated-measures design over two consecutive days. Each dyad experienced both collaborative tasks (Transportation and Factory) and both switching regimes (SHARED EMBODIED VIEW→EMBEDDED ANCHORED VIEW and SHARED EMBODIED VIEW→OUT-OF-BODY VIEW). Tasks were split across days such that each day focused on one task, and task order was balanced across dyads (half started with Transportation on Day 1 and Factory on Day 2, and the remainder received the opposite order). Within each task, the order of the two switching regimes was also balanced across dyads so that each regime appeared equally often in the first and second position. For each task, the number of targets, spatial layout, and time limits were held constant across perspective conditions; only the guest's viewpoint control differed, making task difficulty comparable between EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW.

Participants first completed a demographics form, received a study briefing, and provided informed consent. We then introduced the switching modes, control mappings, and role division: the *host* managed avatar locomotion and native arms and remained in SHARED

EMBODIED VIEW; the *guest* controlled the virtual supernumerary limbs (VSLs) and initiated perspective switches, which occurred only between SHARED EMBODIED VIEW and EMBEDDED ANCHORED VIEW or OUT-OF-BODY VIEW.

Participants completed a 5-minute familiarization session in SHARED EMBODIED VIEW to ensure fluency before entering the main conditions. We then administered the Avatar Embodiment Questionnaire (AEQ) to establish pre-condition embodiment. Following prior SRL-in-VR work [23], we administered AEQ after each condition to obtain post-adaptation scores. This enabled both pre–post contrasts and condition-level comparisons of self-embodiment.

Each daily session began with five minutes of calming music (“Zen mode soundtrack” [384]) to minimize participant fatigue, following a similar approach used in prior research examining sensory conflicts and cognitive fatigue in VR [385]. Short breaks were offered between blocks as needed, and although we did not model learning or fatigue effects explicitly, we did not observe obvious performance degradation over time. Within each daily session, participants performed the assigned task twice, once under each of the two perspective conditions (EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW), with the presentation order of these conditions also counterbalanced. During task execution, the host controlled the avatar’s locomotion and physical arms, while the guest operated the VSLs. Verbal communication was encouraged to allow participants to articulate strategies and decisions.

At the end of both perspective conditions, participants removed their VR headsets and completed the NASA Task Load Index, the Visual Analog Scale for Fatigue, and the AEQ. The study received ethics clearance from the Human Research Ethics Committee (HREC) of the University of Sydney (2019/553).

5.6.4 Data Analysis

We analysed questionnaire, performance, and physiological measures using linear and generalised linear mixed-effects models. For approximately normal continuous outcomes (e.g.,

AEQ factors, NASA–TLX, VAS–F, completion times) we used Gaussian models with identity link; for counts and binary outcomes (e.g., errors, trial success) we used Poisson or binomial models with log/logit links. Unless noted, models included fixed effects of Perspective and, where applicable, User Role and Task, and random intercepts for participant and dyad plus random slopes for within-subject factors when supported by the data. Model diagnostics (residuals, Q–Q plots, overdispersion) and full formulas are reported in Supplementary Materials. Results in the main text are summarised as estimated marginal means (EMMs) with 95% confidence intervals, standardised effect sizes, and p -values for planned contrasts.

To characterise the sensitivity of our design, we conducted a G*Power [386] analysis. For the planned within-subject contrasts between EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW ($\alpha = .05$, power = .80), our sample of $N = 48$ provides sensitivity to paired-samples effects of approximately $d_z \approx 0.41$. For the AEQ analysis with three Perspective levels (SEV, EAV, OOB), a repeated-measures ANOVA approximation (correlation among repeated measures $r = .50$, $\varepsilon = 1.0$) yields a detectable within-subject main effect of $f \approx 0.20$. For between-role comparisons (Host vs. Guest; $n = 24$ per role), an independent-samples approximation indicates 80% power to detect effects of roughly $d \approx 0.80$ at $\alpha = .05$.

5.7 RESULTS

We report quantitative and qualitative findings addressing our four research questions (RQs). RQ1 investigates how perspectives impact subjective workload, fatigue, and physiological responses. RQ2 explores the effects of visual perspectives on embodiment. RQ3 examines how different perspective strategies affect task performance, including completion times and error rates. RQ4 analyzes strategic switching behaviours in response to evolving task demands.

Because of the within-subjects design with repeated measures, we used Generalized Linear Mixed Models (GLMMs) to account for random effects. We refer to *EAV* and *OOB* as the two dynamic switching conditions (SHARED EMBODIED VIEW→EMBEDDED ANCHORED VIEW, SHARED EMBODIED VIEW→OUT-OF-BODY VIEW). For all RQs except RQ2, we

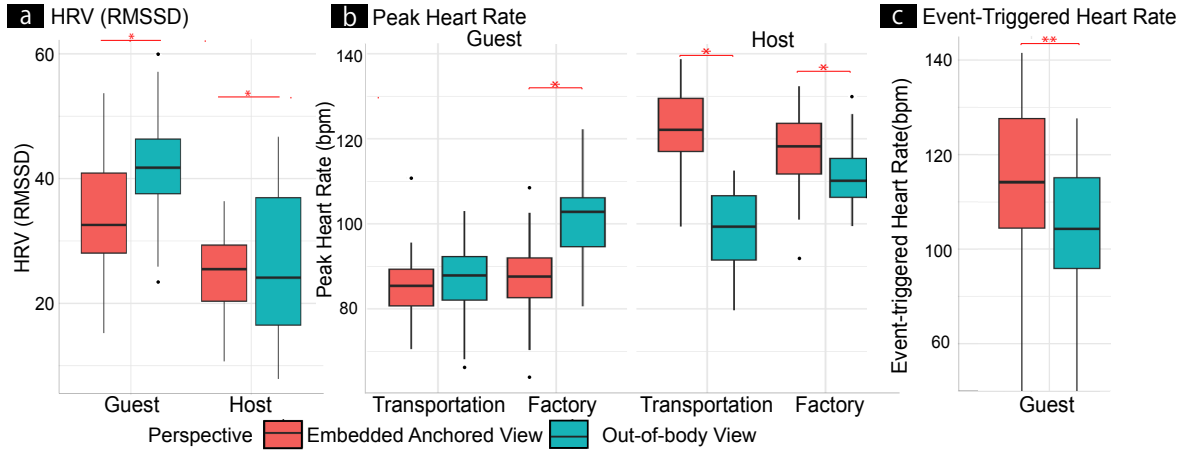


FIGURE 5.6. Physiological responses for *hosts* and *guests* under EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW across two different tasks. (a) HRV (RMSSD), a time-domain index of heart-rate variability. (b) Peak heart rate (bpm) by role and task. (c) Event-triggered heart rate aligned to perspective switches from the SHARED EMBODIED VIEW to the target view (SHARED EMBODIED VIEW → EMBEDDED ANCHORED VIEW or SHARED EMBODIED VIEW → OUT-OF-BODY VIEW). * indicates $p < .05$ and ** indicates $p < .01$.

focus inferential statistics on EAV vs. OOB; SEV is plotted as a baseline but excluded from planned pairwise comparisons unless noted. Results are presented as estimated marginal means (EMMs) with 95% confidence intervals (CIs). Full model details are available in the Supplementary Materials.

5.7.1 Physiological Responses Measures (RQ1)

Complementing our subjective findings (RQ1), we report physiological data (HR, HRV) summarized in Figure 5.6. GLMM analyses revealed significant effects of Perspective, Task, and Role on both peak heart rate and HRV (RMSSD), including Perspective $\times$ Task and Perspective $\times$ Role interactions for peak heart rate (e.g., Perspective main effect for *hosts* in the Transportation baseline: $\beta = -23.0$ bpm, 95% CI $[-25.4, -20.6]$; Perspective $\times$ Task: $\beta = 16.2$ bpm, 95% CI $[13.3, 19.0]$; Perspective $\times$ Role: $\beta = 22.1$ bpm, 95% CI $[19.2, 24.9]$).

Peak Heart Rate. Participants exhibited significantly elevated peak heart rates in OUT-OF-BODY VIEW compared to EMBEDDED ANCHORED VIEW during the precision-oriented Factory task. This effect was more pronounced for *guests* ($M_{OOB} = 102.6$ bpm, 95% CI $[100.2,$

104.9]; $M_{EA} = 87.3$ bpm, 95% CI [85.0, 89.7]; $\Delta M = 15.3$ bpm, $p < .05$). For *hosts*, Factory showed a small but reliable elevation under EMBEDDED ANCHORED VIEW relative to OUT-OF-BODY VIEW ($M_{EA} = 115.4$ bpm, 95% CI [113.0, 117.8]; $M_{OOB} = 108.6$ bpm, 95% CI [106.2, 110.9]; $\Delta M = 6.8$ bpm, $p < .05$). In the navigation-focused Transportation task, *hosts* showed significantly higher peak heart rates in EMBEDDED ANCHORED VIEW ($M_{EA} = 118.4$ bpm, 95% CI [116.1, 120.8]) than OUT-OF-BODY VIEW ($M_{OOB} = 95.4$ bpm, 95% CI [93.1, 97.8]; $\Delta M = 23.0$ bpm, $p < .05$), while *guests* showed no reliable difference ($M_{EA} = 90.4$ bpm, 95% CI [88.0, 92.7] vs. $M_{OOB} = 89.4$ bpm, 95% CI [87.1, 91.8]; $\Delta M = 1.0$ bpm, $p > .05$).

Heart Rate Variability (RMSSD). We observed a significant Perspective $\times$ Role interaction ($\beta = 5.0$ ms, 95% CI [0.4, 9.6]). For *guests*, RMSSD was higher in OUT-OF-BODY VIEW ($M = 42.5$ ms, 95% CI [40.2, 44.8]) than in EMBEDDED ANCHORED VIEW ($M = 35.7$ ms, 95% CI [33.4, 38.0]; $\Delta M = 6.8$ ms, $p < .05$), indicating lower physiological stress in the observer perspective. For *hosts*, RMSSD was also higher in OUT-OF-BODY VIEW than in EMBEDDED ANCHORED VIEW ($M_{OOB} = 27.3$ ms, 95% CI [25.0, 29.6]; $M_{EA} = 25.4$ ms, 95% CI [23.2, 27.7]), although the magnitude of the difference was modest.

Event-triggered heart-rate analyses confirmed a significant immediate increase when switching from EMBEDDED ANCHORED VIEW to OUT-OF-BODY VIEW (+11.1 bpm, 95% CI [9.0, 13.2], $p < .01$), highlighting the anticipatory cognitive demand associated with viewpoint adjustments.

These physiological results corroborate our subjective findings, showing that OUT-OF-BODY VIEW reduced physiological stress and arousal compared to EMBEDDED ANCHORED VIEW, particularly for guests. This supports RQ1 by highlighting how observer perspectives can mitigate workload during collaboration.

5.7.2 Workload and Fatigue Analysis

This analysis addresses RQ1 by comparing subjective workload and fatigue between EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW across task and role conditions, as

shown in Figure 5.7. GLMMs revealed significant Perspective $\times$ Task and Perspective $\times$ Role interactions for workload (NASA-TLX). During the Factory task, both *hosts* and *guests* reported higher workload in OUT-OF-BODY VIEW compared to EMBEDDED ANCHORED VIEW ($p < .05$ for *hosts*, $p < .001$ for *guests*; $\Delta = 12.03$, 95% CI [8.02, 16.04] for *hosts*; $\Delta = 17.35$, 95% CI [13.27, 21.43] for *guests*; *hosts*: OUT-OF-BODY VIEW = 57.60 ± 5.87 , EMBEDDED ANCHORED VIEW = 45.57 ± 4.29 ; *guests*: OUT-OF-BODY VIEW = 54.58 ± 5.06 , EMBEDDED ANCHORED VIEW = 37.23 ± 5.30). For the Transportation task, no reliable workload difference was observed for *guests* ($p > 0.10$; $\Delta = 4.77$, 95% CI [0.95, 8.58]; OUT-OF-BODY VIEW = 48.95 ± 5.66 , EMBEDDED ANCHORED VIEW = 44.18 ± 4.52), whereas *hosts* reported significantly higher workload in OUT-OF-BODY VIEW than EMBEDDED ANCHORED VIEW ($p < .05$; $\Delta = 4.82$, 95% CI [4.82, 4.82]; OUT-OF-BODY VIEW = 53.17 ± 5.16 , EMBEDDED ANCHORED VIEW = 48.35 ± 4.95).

For subjective fatigue (VAS-F), *guests* reported significantly higher fatigue in OUT-OF-BODY VIEW than EMBEDDED ANCHORED VIEW during both Factory ($p < .05$; $\Delta = 3.25$, 95% CI [0.05, 6.45]; OUT-OF-BODY VIEW = 22.65 ± 4.99 , EMBEDDED ANCHORED VIEW = 19.40 ± 2.67) and Transportation ($p < .001$; $\Delta = 2.67$, 95% CI [0.25, 5.09]; OUT-OF-BODY VIEW = 27.22 ± 4.97 , EMBEDDED ANCHORED VIEW = 24.55 ± 5.11). For *hosts*, we observed significantly higher fatigue in EMBEDDED ANCHORED VIEW than OUT-OF-BODY VIEW during Transportation ($p < .05$; $\Delta = -16.46$, 95% CI [-19.84, -13.08]; EMBEDDED ANCHORED VIEW = 42.67 ± 4.92 , OUT-OF-BODY VIEW = 26.21 ± 3.71), and a marginally higher fatigue during Factory, although the effect did not reach significance ($p > 0.05$; $\Delta = -12.96$, 95% CI [-15.29, -10.63]; EMBEDDED ANCHORED VIEW = 37.77 ± 3.35 , OUT-OF-BODY VIEW = 24.81 ± 4.99).

Together, these results support RQ1 by showing that OUT-OF-BODY VIEW can increase perceived workload and fatigue for guests, whereas hosts experience more fatigue under EMBEDDED ANCHORED VIEW, particularly in navigation-intensive tasks. This highlights a role-dependent divergence in how perspective affects burden.

Fatigue (VAS-F). For *guests*, fatigue was higher in OUT-OF-BODY VIEW than in EMBEDDED ANCHORED VIEW in the *Factory* task ($p < .05$; $\Delta = 3.25$, 95% CI [0.05, 6.45]), and showed

the same direction without reaching significance in *Transportation* ($p > 0.10$; $\Delta = 2.67$, 95% CI [0.25, 5.09]), consistent with Fig. 5.7(b)(c). For *hosts*, fatigue was higher in EMBEDDED ANCHORED VIEW than in OUT-OF-BODY VIEW during *Transportation* ($p < .05$; $\Delta = -16.46$, 95% CI [-19.84, -13.08]), and showed no reliable difference in *Factory* ($p > 0.10$; $\Delta = -12.96$, 95% CI [-15.29, -10.63]), matching Fig. 5.7(a)(d).

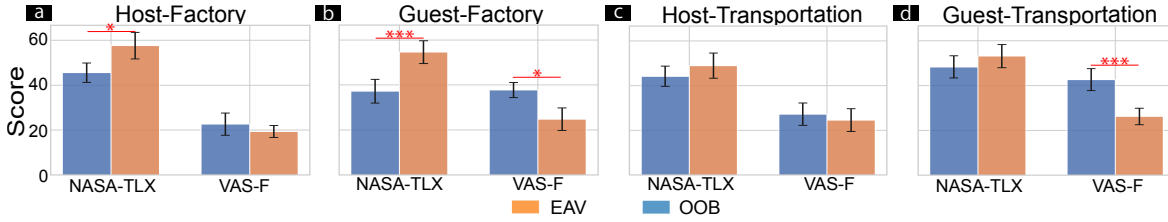


FIGURE 5.7. Subjective measures (VAS-F and NASA-TLX) for *hosts* and *guests* across tasks: (a) *Host* scores for Factory Task, (b) *Guest* scores for Factory Task, (c) *Host* scores for Transportation Task, and (d) *Guest* scores for Transportation Task. Scores are compared between EMBEDDED ANCHORED VIEW and OUT-OF-BODY VIEW. * indicates $p < .05$, ** indicates $p < .01$ and *** indicates $p < .001$.

5.7.3 Embodiment Questionnaire Analysis (RQ2)

RQ2 examines how perspective-switching strategies shape subjective embodiment. AEQ ratings were collected immediately before and after each perspective-switch condition, following standard practice in VR embodiment research [23]. Factor scores were computed as

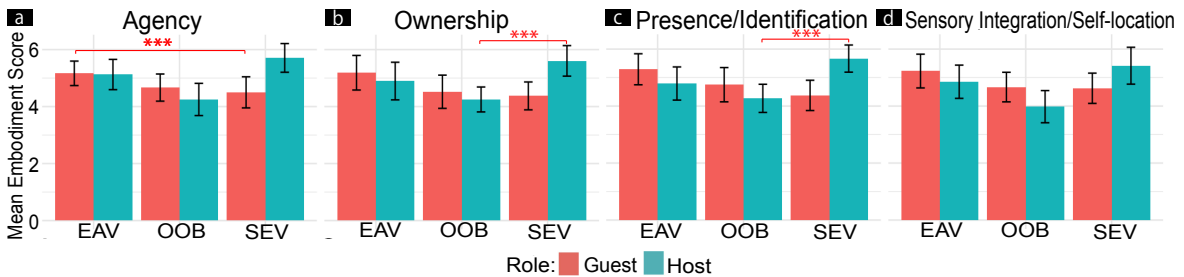


FIGURE 5.8. Mean embodiment scores ($\pm$ SE) across three perspective conditions (EMBEDDED ANCHORED VIEW, OUT-OF-BODY VIEW, and SHARED EMBODIED VIEW), separated by user role (Host and Guest). Subplots represent AEQ factors: (a) Agency, (b) Body Ownership, (c) Presence/Identification, and (d) Sensory Integration/Self-location. * indicates $p < .05$, ** indicates $p < .01$ and *** indicates $p < .001$.

the mean of their constituent items:

$$\text{Factor Score} = \frac{\sum \text{Item Scores}}{\text{Number of Items per Factor}}.$$

We analysed factor scores with linear mixed-effects models (lme4 [387]; lmerTest [388]), including fixed effects of Perspective, User Role (host vs. guest), and Embodiment Factor, and by-participant random intercepts with random slopes for within-subject factors. Degrees of freedom for fixed effects were estimated via the Kenward–Roger method [389]; omnibus tests are reported as Type-III F statistics, and post-hoc pairwise contrasts were computed with *emmeans* [390] using Tukey adjustment [391].

We observed a main effect of Perspective ($p < .001$) and Embodiment Factor ($p < .01$), and a Perspective $\times$ User Role interaction ($p < .001$). Post-hoc comparisons aligned with Fig. 5.8: across factors, SHARED EMBODIED VIEW yielded the highest embodiment, EMBEDDED ANCHORED VIEW was intermediate, and OUT-OF-BODY VIEW was lowest. For *guests*, body ownership and presence were higher in SHARED EMBODIED VIEW than in OUT-OF-BODY VIEW ($M_{\text{SEV}} = 4.47$, $SD = 0.97$ vs. $M_{\text{OOB}} = 3.52$, $SD = 1.04$; both Tukey $p < .001$). corresponding to a large standardised mean difference of approximately $d = 0.94$, 95% CI [0.34, 1.53]. For *hosts*, agency was higher in SHARED EMBODIED VIEW than in EMBEDDED ANCHORED VIEW ($M_{\text{SEV}} = 4.26$, $SD = 0.94$ vs. $M_{\text{EAV}} = 3.72$, $SD = 1.01$; Tukey $p < .001$); corresponding to a moderate standardised mean difference of approximately $d = 0.60$, 95% CI [0.37, 0.82].

These findings answer RQ2 by showing that self-embodiment is perspective- and role-dependent: OUT-OF-BODY VIEW reduces both presence and ownership, particularly for guests, while EMBEDDED ANCHORED VIEW improves embodiment relative to OUT-OF-BODY VIEW, but remains lower than the SHARED EMBODIED VIEW.

5.7.4 Task Performance Analysis (RQ3)

Here, we answer RQ3 by examining task performance metrics across different perspective modes. Task completion times and error rates were analyzed using GLMMs with gamma

and Poisson distributions respectively (random intercept for participant ID). Results are summarized in Figure 5.9 (a,b). Participants completed the Transportation task significantly faster in OUT-OF-BODY VIEW ($M = 15.5$, s , $SE = 1.2$) than in EMBEDDED ANCHORED VIEW ($M = 22.6$, s , $SE = 1.7$; $z = 6.64$, $p < .001$). with the GLMM indicating an estimated time ratio of ≈ 0.69 , 95% CI $[0.55, 0.84]$ (i.e., about 31% faster in OUT-OF-BODY VIEW). In contrast, completion times for the Factory task did not differ reliably between perspectives ($p = .07$). In the Factory task, significantly fewer errors were observed in OUT-OF-BODY VIEW ($M = 3.4$, $SE = 0.5$) than in EMBEDDED ANCHORED VIEW ($M = 4.4$, $SE = 0.6$; $z = -2.49$, $p < .05$). The Transportation task yielded no statistically reliable difference in error rate.

In answering RQ3, our results suggest that the observer perspective (OUT-OF-BODY VIEW) offers performance advantages under different demands: faster navigation in dynamic tasks (Transportation) and fewer errors in precision tasks (Factory).

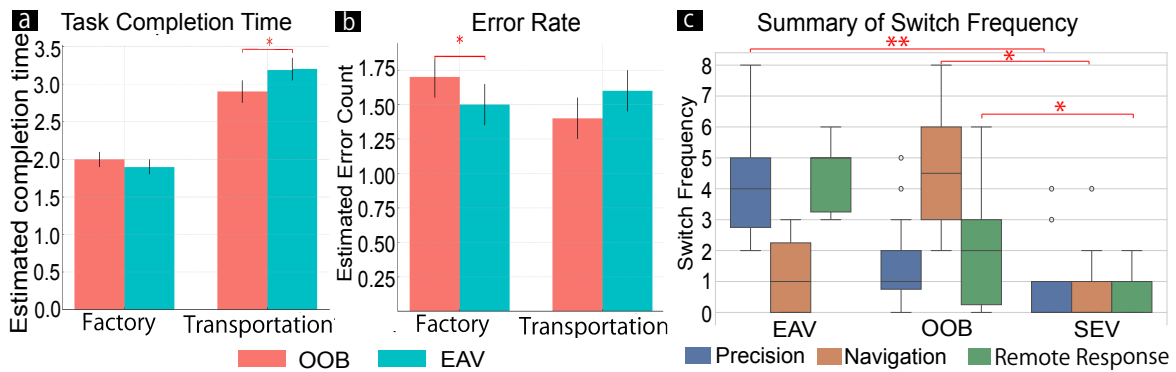


FIGURE 5.9. Effects of perspective mode and task phase. (a) Completion time: OUT-OF-BODY VIEW led to significantly faster task completion in the Transportation task. (b) Error count: OUT-OF-BODY VIEW resulted in significantly fewer errors in the Factory task. (c) Switch frequency by task phase. Switching was infrequent in the baseline SHARED EMBODIED VIEW condition (median ≤ 1 across phases). In comparison, participants switched more frequently to EMBEDDED ANCHORED VIEW during precision subtasks, and to OUT-OF-BODY VIEW during navigation and remote-response phases. * indicates $p < .05$, ** $p < .01$, and *** $p < .001$.

5.7.5 Perspective Switching Behaviors (RQ4)

To address RQ4, we examined how participants strategically switched perspectives during tasks. As shown in Figure 5.9(c), only *guests* could initiate perspective switches. Following prior work in collaborative HRI/VR [392, 393, 394], each trial was segmented (from logs) into three phases: *precision* (near-body manipulation), *navigation* (locomotion/wayfinding), and *remote response* (reacting to host movement or environment events).

- **Navigation phase:** *host* locomotion (>1.5 s continuous motion), no active VSL control from *guest*. This phase aligns with “approach” in handover models [392].
- **Precision-focused subtasks:** stationary *host*, active VSL grasp/place attempts by *guest*. These represent the “manipulation” phase typical in object transfer tasks [393].
- **Remote-response events:** short *guest*-initiated viewpoint switches in response to unexpected *host* movement (e.g., rotation, misalignment). These reflect reactive perspective adjustments during miscoordination or disorientation [395].

We modeled switch frequency with a GLMM (Poisson; random intercepts for participant), revealing strong modulation by task phase. Guests switched predominantly to EMBEDDED ANCHORED VIEW during *precision* subtasks ($M = 4.20$, $SE = 0.30$, 95% CI [3.61, 4.79], $p < .01$), and to OUT-OF-BODY VIEW during *navigation* ($M = 4.89$, $SE = 0.31$, 95% CI [4.28, 5.50], $p < .05$). For *remote-response* moments, guests again preferred OUT-OF-BODY VIEW ($M = 3.17$, $SE = 0.32$, 95% CI [2.54, 3.80]; $p < .05$ versus other views). Switching while in SHARED EMBODIED VIEW was rare (median ≈ 1 across phases), indicating little perceived need to change away from the baseline shared first-person view except when task demands shifted.

These findings answer RQ4 by showing that guests adaptively switch perspectives based on phase-specific demands, preferring EMBEDDED ANCHORED VIEW for close manipulation and OUT-OF-BODY VIEW for navigation and rapid corrections. This indicates that switching behaviors are tightly coupled to functional needs and task context.

5.7.6 Qualitative Findings

We conducted an inductive thematic analysis [352] to explore participants' experiences with VSLs operation and perspective-switching during collaborative tasks. This analysis aimed to identify interaction strategies, user preferences, and challenges in managing viewpoint shifts and VSLs control. Two independent coders analyzed interview transcripts, achieving good inter-rater reliability (Cohen's Kappa, $\kappa = 0.78$).

Following Braun and Clarke's framework [353], we initially familiarized ourselves with the data by reviewing interview transcripts and noting key insights. During initial coding, we labeled relevant segments related to strategies, cognitive load, viewpoint preferences, spatial awareness, and collaborative challenges. Independent coding was followed by joint discussions to consolidate these codes into initial themes.

Iterative refinement grouped codes into four broader themes: "Visual Decoupling and Embodiment", "Prior Experience and Perspective Adaptation", "Strategies for Managing Perspective Switching", and "Emergent Communication through Gestures". These themes structured our understanding of participants' adaptive behaviors, cognitive demands, and collaboration strategies during dynamic perspective switching in VSLs supported VR teleoperation.

The following sections detail these themes, illustrating participants' approaches to viewpoint control, collaboration strategies, and adaptations to dynamic teleoperation demands.

5.7.6.1 Role-asymmetric Embodiment Trade-offs

As previously introduced, *Visual Decoupling* refers to the *guest*'s adoption of an independent external perspective (OUT-OF-BODY VIEW), decoupled from the *host*'s visual alignment. Initially, based on prior research [18], we hypothesized that visual decoupling would negatively impact users' sense of embodiment due to the visual and proprioceptive mismatch. However, *hosts* overwhelmingly reported an enhanced sense of embodiment when their remote partners adopted the OUT-OF-BODY VIEW, emphasizing a feeling of freedom and control over their movements.

Several *hosts* explicitly mentioned that they felt less inhibited when their partners were not visually coupled to their head movements. For example, one participant stated: “I felt comfortable moving naturally without constantly worrying” (*host*, P17). Specifically, *hosts* were concerned that their own head movements or sudden turns could cause visual instability or discomfort for their remote partners, which often made them overly cautious and hesitant during collaboration. When freed from these concerns, they reported increased confidence and spontaneity in their movements, contributing to a stronger sense of ownership over the shared body.

In contrast, *guests* generally reported reduced self-embodiment in the OUT-OF-BODY VIEW, often describing a sense of detachment: “felt like watching myself from outside” (*guest*, P19). This was attributed to the lack of visual alignment with their avatar’s first-person perspective, which made it harder to associate with the embodied body. One *guest* described: “When I was not seeing through the avatar’s eyes, it didn’t feel like it was me moving.” We interpret this divergence as a role-dependent trade-off: for *hosts*, the removal of visual coupling alleviated their self-imposed constraints and enabled greater natural movement and self-ownership; for *guests*, the visual decoupling created psychological distance and disrupted the integration of the avatar into their body schema. These asymmetrical embodiment effects reveal that OUT-OF-BODY VIEW can simultaneously enhance host autonomy while disrupting guest self-identification with the avatar. Consistent with this interpretation, RMSSD-based HRV did not show large role-specific divergences across perspectives, suggesting that these embodiment trade-offs are primarily driven by visual coupling and perceived responsibility rather than gross differences in physiological arousal.

5.7.6.2 Importance of Gaming Experience over VR Experience

Participants with extensive gaming experience adapted more rapidly to OUT-OF-BODY VIEW than those with mainly non-gaming VR backgrounds. Several participants linked this to prior fluency in third-person or drone-style games. One noted, “This view feels just like controlling a game character in third-person games” (*guest*, P8), while another commented, “I like playing games. The drone one felt like a game” (*guest*, P10). These users consistently

described OUT-OF-BODY VIEW as intuitive and familiar, citing spatial awareness and mental models shaped by gaming.

In contrast, participants with strong VR experience but limited gaming backgrounds struggled more with the externally detached spatial perspective. One such participant shared, “I kept getting disoriented” (*guest*, P15), attributing their discomfort to the lack of visual alignment and direct embodiment. These findings suggest that gaming, rather than general VR use, better equips users to operate within non-egocentric, third-person viewpoints, as common in the OUT-OF-BODY VIEW condition.

These contrasting patterns highlight that prior gaming experience, particularly with third-person perspectives, plays a more critical role than general VR familiarity in shaping users’ ability to navigate detached viewpoints like OUT-OF-BODY VIEW.

5.7.6.3 Strategic Minimization of Perspective Switching

Though we initially hypothesized that frequent perspective switching would improve task performance, several participants ($N=6$) reported deliberately minimizing their switches to maintain focus and reduce mental fatigue. As one *guest* explained, “I found fewer switches kept me focused” (P3). These participants described a strategy of limiting perspective changes to key task transitions, citing benefits such as reduced cognitive load and sustained concentration.

In contrast, a smaller group adopted a high-frequency switching strategy, believing it improved their situational awareness. One such participant remarked, “I needed different perspectives to manage different tasks better” (*guest*, P11). Although these users acknowledged the added cognitive effort and occasional disorientation, they perceived the benefits of enhanced task visibility as outweighing the costs.

Together, these divergent strategies highlight individual differences in managing cognitive resources and suggest that perspective-switching preferences may depend on thresholds for mental load and spatial tracking. Consistent with this, our HRV analysis (Sec. 5.7.1) shows

perspective-dependent differences between trials with more versus fewer switches, suggesting that these strategies are associated with differences in physiological effort.

5.7.6.4 Emergence of Natural Gesture-Based Communication Strategies

A spontaneous, non-verbal communication strategy emerged in some participant pairs (N=18), particularly those who relied heavily on OUT-OF-BODY VIEW. “I use robotic arm gestures to explain” (*guest*, P30). These pairs developed gesture-based conventions to complement or replace verbal instructions, often describing this as more efficient and natural than speech alone.

However, not all dyads adopted this strategy. “We kept misunderstanding each other, so we stuck with verbal instructions” (*guest*, P14). Some participants found gestures ambiguous or lacked confidence in their clarity. Thus, while gesture-based interaction emerged organically under visual decoupling, it was not universal.

Taken together, these findings suggest that gesture-based communication strategies may emerge organically, shaped by spatial alignment, mutual visibility, and the degree of interpersonal trust.

5.8 DISCUSSION

Our study examined how guest-driven perspective switching (SHARED EMBODIED VIEW ↔ EMBEDDED ANCHORED VIEW, SHARED EMBODIED VIEW ↔ OUT-OF-BODY VIEW) strategies influence performance, embodiment, workload, and adaptive user behaviors in collaborative VR teleoperation. Results revealed clear task-dependent performance advantages (RQ1), distinct role-specific embodiment preferences (RQ2), and notable effects on workload and physiological responses (RQ3). Additionally, we identified strategic viewpoint-switching patterns aligned with evolving task demands (RQ4).

5.8.0.1 Reducing Workload and Physiological Strain through Perspective Design (RQ1)

Our findings clearly demonstrate that perspective choice significantly influences subjective workload and physiological strain. The OUT-OF-BODY VIEW effectively reduces users' perceived fatigue (VAS-F) and workload (NASA-TLX) during dynamic tasks, confirmed by elevated RMSSD and reduced heart rate. Conversely, despite enhancing embodiment, the EMBEDDED ANCHORED VIEW unexpectedly increased cognitive and physiological burden, likely due to continuous visual stabilization demands. These results extend previous findings [396] by providing integrated subjective and physiological evidence. Our study highlights a key design trade-off between spatial embodiment and physiological comfort, suggesting designers dynamically balance perspectives, prioritizing the OUT-OF-BODY VIEW during intensive tasks and carefully managing usage duration of the EMBEDDED ANCHORED VIEW. Additionally, as several guests reported avoiding perspective switches to stay focused (Section 5.7.6.3), we propose intelligent guidance mechanisms to help minimize unnecessary switching. Future work should evaluate such mechanisms in realistic settings.

5.8.0.2 Balancing Embodiment and Spatial Awareness: Role-specific Considerations (RQ2)

Our findings reveal significant role-specific differences regarding embodiment and spatial awareness across perspectives. *Hosts* experienced heightened agency and ownership primarily within the SHARED EMBODIED VIEW, benefiting from a coherent first-person grounding, whereas EMBEDDED ANCHORED VIEW aided fine control near the body. *Guests* reported stronger embodiment in the SHARED EMBODIED VIEW, but preferred the OUT-OF-BODY VIEW for better spatial awareness and stable control during navigation. This extends prior work [397], explicitly identifying distinct user needs based on collaborative roles. Critically, we found a clear design tension: the OUT-OF-BODY VIEW enhanced spatial cognition but diminished *guest's* sense of embodiment and increased perceived effort in some tasks, consistent with prior studies highlighting the impact of visual-body mismatches [398]. To mitigate this trade-off, we recommend adaptive, role-specific viewpoint strategies to dynamically balance

embodiment and spatial awareness. Future work should validate role-dependent effects in more complex, real-world scenarios.

TABLE 5.2. Design Suggestions for Perspective Switching and Embodiment Cues. Rows marked [†] are directly grounded in our study findings, whereas rows marked [‡] are forward-looking design opportunities extrapolated from these findings.

Adaptive Viewpoint Recommendations		
Design Focus	Main Finding	Design Suggestions
Task-based Adaptive Switching [†] (Section 5.8.0.3)	OUT-OF-BODY VIEW aided navigation; in precision phases guests preferentially chose EMBEDDED ANCHORED VIEW; frequent switching increased cognitive load for some.	Automatically recommend optimal viewpoints based on real-time task context. Provide minimal yet clear alerts for viewpoint transitions.
Intelligent Switching Cost Alerts [‡] (Section 5.8.0.1)	Users were unaware of cognitive costs incurred by frequent viewpoint switching, leading to unnecessary cognitive strain.	Introduce subtle, predictive visual or auditory indicators informing users about switching costs; ensure alerts are concise and unobtrusive.
Role-Specific Embodiment Optimization		
Embodiment Interfaces [†] (Section 5.8.0.2)	Hosts benefited when guests used anchored views; guests needed flexible views (SHARED EMBODIED VIEW or OUT-OF-BODY VIEW)	Provide differentiated interfaces and distinct embodiment cues tailored to <i>host</i> and <i>guest</i> . Clearly differentiate avatar alignment visuals.
Embodiment Reinforcement Cues [†] (Section 5.8.0.2)	Users experienced reduced embodiment due to inconsistent avatar-body alignment cues, especially during transitions.	Implement dynamic real-time IK-based visualizations of limb alignment. Use selective limb highlighting to reinforce embodiment and user agency.
Visual Stability and Emergency Response		
Visual-Stability Enhancement [†] (Section 5.8.0.1)	Rapid or frequent viewpoint transitions caused visual instability and discomfort, negatively impacting user comfort and performance.	Utilize adaptive smoothing techniques during perspective transitions; dynamically adjust smoothing intensity according to user physiology (e.g., HRV).
Rapid Emergency Switch [‡] (Section 5.8.1)	Situational hazards required rapid spatial reassessment; delayed perspective switches negatively impacted safety and performance.	Design intuitive rapid perspective-switching gestures or single-click mechanisms for emergency use; ensure clear visual affordances minimize false activations.

5.8.0.3 Optimizing Collaborative Performance through Adaptive Perspective Switching (RQ3, RQ4)

Our results demonstrated clear task-dependent advantages for each perspective mode. Specifically, the OUT-OF-BODY VIEW significantly improved performance in navigation tasks compared to the EMBEDDED ANCHORED VIEW (RQ3), confirming prior insights regarding third-person benefits for spatial awareness [399, 400]. However, unlike previous studies, we explicitly quantified these perspective effects across diverse collaborative scenarios. Role-specific embodiment differences emerged, with *host* users reporting greater agency and ownership in the SHARED EMBODIED VIEW, and *guest* experiencing stronger embodiment in the SHARED EMBODIED VIEW but better spatial orientation in the OUT-OF-BODY VIEW (RQ2), extending previous research [118]. Moreover, our subjective and physiological data indicate a trade-off: OUT-OF-BODY VIEW was associated with *lower physiological stress for guests* (higher RMSSD) but *higher subjective fatigue/workload in certain contexts* (e.g., higher NASA-TLX in Factory, and higher host workload during Transportation), while EMBEDDED ANCHORED VIEW supported near-body stability without the same spatial overview. Qualitative analysis also identified strategic user-driven viewpoint switching aligned with dynamic task demands (RQ4).

5.8.1 Implications for the Design of Collaborative Teleoperation Systems

Our findings provide clear practical implications for designing collaborative VR teleoperation systems, emphasizing adaptive perspective selection to balance performance, user comfort, and cognitive load. Specifically, the OUT-OF-BODY VIEW improves spatial navigation and reduces cognitive strain, while the EMBEDDED ANCHORED VIEW enhances precision and embodiment. We specifically highlight previously overlooked cognitive and physiological costs associated with frequent viewpoint switching, recommending adaptive guidance to address these trade-offs. Our practical guidelines are outlined in the Table 5.2.

Consider future high-stakes settings where space is constrained but expert intervention is critical, such as tele-supported surgery [401, 402] or in-situ aircraft maintenance [403, 404].

A local generalist could maintain the mobility and support functions while a remote specialist “inhabits” the same body for specialized tasks manipulation, relying on an EMBEDDED ANCHORED VIEW-like stabilized view that supports tasks that needs precision. In such cases, periodically switching to the OUT-OF-BODY VIEW may help maintain spatial awareness and reduce strain, making interfaces proactively recommend viewpoint transitions based on task demands. In sports coaching, such as dynamic rugby drills, the OUT-OF-BODY VIEW can effectively reduce discomfort during rapid movements, while detailed skill training benefits from stabilized visual feedback provided by the EMBEDDED ANCHORED VIEW. Designing seamless, context-aware switching interfaces can minimize cognitive load, improving coaching effectiveness. Similarly, in industrial safety contexts, adaptive perspective selection can significantly enhance hazard detection and precise manipulation tasks. Rapid, intuitive perspective switching combined with context-sensitive hazard alerts can optimize both safety and efficiency.

Finally, although our study operates in a purely virtual environment, the perspective trade-offs we identify are directly relevant to emerging teleoperation systems built on AI-based 3D world models and digital twins [296, 297]. As AI-based 3D world models and digital twins mature, remote operators will work inside persistent virtual replicas of physical workspaces, moving through the scene independently of the robot’s physical camera pose [295, 296, 297, 298]. In such systems, OUT-OF-BODY VIEW-like views can be implemented as world-locked virtual cameras, while SHARED EMBODIED VIEW and EMBEDDED ANCHORED VIEW correspond to avatar-locked and portal-like views. Our results suggest that these interfaces should expose multiple coupled and decoupled viewpoints aligned with task phases, and manage switching costs through predictive cues or semi-automatic camera positioning, even when specific robot kinematics are abstracted away.

5.8.2 Limitations and Future Work

Our study was conducted within a controlled laboratory environment, which may limit the generalizability of findings to real-world scenarios characterized by more unpredictable factors and complex interactions. Additionally, our participant pool predominantly consisted

of university students, who might possess relatively uniform demographic characteristics, familiarity with VR systems, and cognitive strategies. Future research should include a broader and more diverse participant population, along with testing under more ecologically valid conditions, to better assess the robustness and applicability of our perspective-switching approach.

We also did not administer the Simulator Sickness Questionnaire (SSQ), relying instead on workload (NASA-TLX), fatigue (VAS-F), and HRV as proxies for discomfort in shared viewpoints. We interpret these discomfort-related findings as subjective and physiological strain rather than a standardized measure of simulator sickness, because SSQ was not collected. Future work should incorporate standardized cybersickness measures to facilitate finer-grained comparison across VR studies. We did not run an a priori power analysis for our mixed-effects design; a conservative dyad-level sensitivity check (paired comparison, two-sided $\alpha=.05$, $N=24$) suggests the study is powered ($\sim 80\%$) to detect effects around $d \approx 0.60$, so smaller effects may be underpowered. A further limitation concerns the formative study: its short interaction duration (two five-minute trials) primarily surfaces early coordination breakdowns. Longer usage periods may reduce role confusion and improve synchronisation as partners adapt to one another. In extended pilot testing, we observed similar adaptation effects, with guest users becoming more anticipatory of the host's movement patterns over time.

While VR provided an accessible [196] and controlled environment for investigating perspective switching interactions in shared-body teleoperation scenarios involving VSLs, we acknowledge that our findings remain speculative regarding applicability to physical robotic implementations. Although previous research suggests transferability from virtual to real-world systems [23, 405], the specific dynamics of perspective-switching behaviors and collaborative interactions identified here may differ significantly when physical robotic limbs and real-world constraints are introduced. Therefore, our results should be validated through future studies using physical prototypes before generalizing these insights to practical teleoperation applications.

5.9 CONCLUSION

We explored dynamic, guest-driven perspective switching in shared-body VR teleoperation with virtual supernumerary limbs (VSLs). Building on a formative study that identified locomotion- and coordination-related challenges under fixed first-person viewing, we implemented two switching conditions: SHARED EMBODIED VIEW $\leftrightarrow$ EMBEDDED ANCHORED VIEW and SHARED EMBODIED VIEW $\leftrightarrow$ OUT-OF-BODY VIEW. A within-subjects study (N=48) showed that OUT-OF-BODY VIEW supports navigation and reduces errors, while EMBEDDED ANCHORED VIEW is preferred for near-body manipulation phases, with *hosts* reporting stronger agency. Completion times showed no reliable difference across views, but OUT-OF-BODY VIEW yielded fewer errors in the Factory task. Together, these findings yield practical, role- and phase-aware design guidelines for optimizing collaborative performance, comfort, and embodiment in mobile shared-avatar VSL systems.

5.9.1 Contribution to the Thesis

This chapter makes three thesis-level contributions.

First, it establishes **perspective configuration** as a central design problem in collaborative embodiment. By showing how viewpoint shapes coordination, comfort, workload, and embodied experience, it demonstrates that collaborative VR cannot be understood only in terms of control allocation. It must also account for how perception is distributed across collaborators.

Second, it shows that perspective in collaborative VR is best understood as a problem of **distributed perception**. Rather than asking which view is universally optimal, this chapter shows that different perspectives support different roles, task phases, and coordination demands.

Third, it identifies a key design principle for later chapters and for the thesis's framework synthesis. **Perspective flexibility is necessary, but continuous switching is not ideal**. Users benefit from the ability to reconfigure perspective, but tend to prefer stable, phase-sensitive

viewpoint arrangements. In this way, the chapter contributes the empirical basis for the perspective dimension of the Interaction Elements Matrix.

CHAPTER 6

Juggling Extra Limbs

Juggling Extra Limbs: User Control Strategies for Semi-Autonomous Supernumerary Arms in VR

This chapter addresses the third core dimension of collaborative embodiment in this thesis, namely **autonomy and delegation**. Whereas prior work on supernumerary limbs has largely emphasized additional effectors and control techniques, the present chapter focuses on how users engage with partially delegated control. In collaborative embodied systems, action may no longer be exhausted by direct manipulation. It may also involve assigning subtasks to semi-autonomous limbs, monitoring their behaviour, and intervening when adjustment is required. Accordingly, this chapter approaches supernumerary limb interaction not only as a problem of augmentation, but also as a problem of how embodied action is reorganized under conditions of partial autonomy.

Empirically, the chapter examines how users coordinate with virtual supernumerary limbs across different levels of autonomy. The central concern is therefore not only whether additional limbs are usable, but how control shifts between direct action, delegated action, and selective intervention. The findings indicate that autonomy is not valuable in itself. Rather, its effects depend on whether users can interpret system behaviour, recover from errors, and retain a meaningful sense of agency while relinquishing parts of the task.

Within the broader argument of this thesis, the chapter extends collaborative embodiment beyond shared control among human users alone. It shows that embodied coordination may also involve semi-autonomous system partners and thereby provides a basis for the later

discussion of controllable autonomy, delegation, and forms of human–AI and human–robot co-embodiment.

The rest of the text of this chapter is a published paper: Zhou, H., Kip, T., Bianchi, A., Sarsenbayeva, Z., Withana, A. Juggling Extra Limbs: Identifying Control Strategies for Supernumerary Multi-Arms in Virtual Reality. In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI 2025), and is kept unchanged to reflect the university thesis requirements.

6.1 Abstract

Using supernumerary multi-arms to carry out complex tasks is becoming a popular research domain in application areas such as virtual reality and robotics. Understanding how users can collectively utilise extra arms together with their own to achieve shared goals is critically important to develop efficient supernumeraries. This paper presents results from an exploratory user study (N=14) to understand user strategies for simultaneous control and coordination of virtual supernumerary limbs with varying autonomy levels in VR object manipulation tasks. We used a Wizard-of-Oz approach to simulate semi-autonomous extra arms and collected qualitative data through think-aloud and semi-structured interviews. We also gathered quantitative data on task performance and embodiment. Results reveal that participants adapted their control strategies according to the task complexity and system autonomy, affecting task delegation, coordination, and their sense of body ownership. From these findings, we propose guidelines to improve multi-arm interaction design in future systems.

6.2 INTRODUCTION

Leveraging multiple supernumerary limbs to perform complex tasks is becoming a popular research domain in application areas such as Virtual Reality (VR) and robotics [406, 335, 336, 337]. To harness the potential of supernumerary limbs in addressing complex tasks, prior research has predominantly employed fixed control mappings for operating the supernumerary

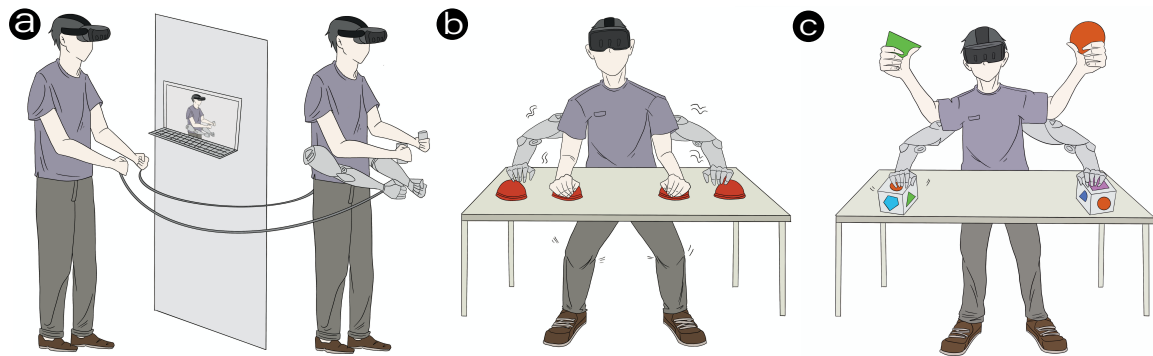


FIGURE 6.1. Wizard-of-Oz approach and experimental tasks. (a) The participants' VSLs were controlled by a human operator behind a partition to simulate semi-autonomous supernumerary limbs. (b) Basic Control Task: In this task, participants used both their own arms and the VSLs to press buttons. (c) Factory Task: Participants used their own arms and the VSLs to collaboratively insert found objects into the appropriate holes.

limbs, where specific commands are assigned to the movement of particular body parts [331, 343]. However, while fixed control mappings can be effective for certain tasks, they lack the flexibility required for dynamic interactions in complex environments, especially when supernumerary limbs are given *a level of autonomy*.

To facilitate the exploration of dynamic control mappings, it is crucial to understand how users would effectively coordinate these semi-autonomous supernumerary limbs with their own limbs to expand human capabilities and address the unique challenges of interaction design and user experience [407] in a multiple supernumerary limbs setting. In this study, we examine the use of supernumerary limbs within a VR environment to simulate interactions that could inform future applications. VR provides a valuable platform for studying and evaluating the integration of supernumerary limbs, offering a controlled, immersive space to simulate complex scenarios that are difficult to replicate in real-world robotic applications [46, 23, 405]. This environment allows users to engage in a “suspension of disbelief [408],” which supports the study of embodied interaction in novel body configurations. In this paper, we deal with “virtual supernumerary limbs (VSLs)” to explore VR-based interactions rather than real-world robotic applications. Although there are clear differences between VR and real-world environments, insights gained from VR-based studies can potentially inform future

developments not only of VR applications, but also physical mechanical limbs, as shown in previous work [409, 410].

Particularly, this paper focuses on how VSLs, when equipped with different and increasing degrees of autonomy, can assist users in managing non-repetitive, multi-layered tasks in VR environments. Instead of simply mapping another limb (e.g. foot movement [8]) or executing predefined commands [289], autonomous VSLs in our study learn, adapt, and act semi-independently to respond to changing task requirements and conditions. This adaptability is especially crucial for scenarios involving real-time decision-making and task variations, where the user needs more responsive control than what fixed mappings can provide [8]. This shift from fixed mappings to autonomous, context-sensitive control can elicit unique user interaction strategies, highlighting the influence of different levels of autonomy in VSLs on user behavior.

To this end, we implemented a Wizard-of-Oz setup to simulate autonomous VSLs that respond to user input and task complexity. Integrating these autonomous capabilities allowed us to freely explore the design space in breadth and depth without restrictions imposed by limitations of today's technologies. Our study focused on key research questions surrounding interaction strategies and system autonomy of multiple VSLs control: **RQ1:** *What control strategies do users employ when interacting with semi-autonomous VSLs in different scenarios?* **RQ2:** *How does the level of autonomy in VSLs affect users' ability to manage tasks?* and **RQ3:** *How do users' perceptions of embodiment and control change with level of autonomy?*

To capture a comprehensive understanding of user interaction, our study blended qualitative and quantitative research methods. Qualitative data were gathered through think-aloud and semi-structured interviews, providing insights into the participants' thought processes. Quantitative data on task performance and the awareness of embodiment were collected to quantify the effectiveness of different control strategies and their impact on user performance and perception. Furthermore, we analyzed operator (or wizard) reaction patterns to validate the operator's behavior, ensuring their actions were consistent and adhered to predefined patterns. This analysis supports the conclusion that our task performance results are based on

participant behavior, unaffected by operator errors. This mixed approach revealed several key findings: users favored task delegation to the VSLs for repetitive or lower-priority actions, adjusted their strategies to switch between manual control and automation depending on task complexity, and developed more efficient coordination as they grew familiar with the system. Additionally, users found object labeling and categorization to be crucial in managing VSLs effectively, and they experienced shifts in embodiment based on the level of system autonomy. These findings underscore the need for adaptive control mechanisms that help users balance attention, streamline task-switching, and enhance the sense of embodiment, informing design recommendations for future systems employing multiple VSLs.

6.3 RELATED WORK

6.3.1 Virtual Reality and Robotic Body

Recent advancements in robotics and VR technologies have enabled novel interactions where users can experience the detachment, attachment, and exchange of body parts, even while retaining their natural limbs. For instance, Iwasaki et al. [411] introduced the concept of a “detachable body”, where a third and fourth robotic limb can be seamlessly attached and detached, allowing users to operate these limbs remotely across different locations. Additionally, research on supernumerary robotic limbs has explored task-specific customization of limb end-effectors, enhancing adaptability and functionality in varied contexts. VR and wearable Shape-Memory Alloys (SMA) have further expanded these possibilities by enabling shared body experiences, where two individuals can collaboratively control a single body (e.g., [289, 292, 412]). For example, Saraiji et al. created a system where wearable robotic arms are remotely operated by another person, facilitating a shared embodiment [289]. Takizawa et al. extended this exploration into VR, investigating how different perspectives impact the shared body experience [292].

Building on these developments, VR provides unprecedented opportunities to explore more abstract states of being a digital cyborg, such as inhabiting an invisible body [413], existing as

separated body parts [414], or even duplicating oneself [406]. These immersive experiences push the boundaries of body ownership and self-perception, offering new avenues to understand how users adapt to technologically augmented bodies. Recent studies providing the experience of acquiring virtual robotic limbs [415, 315]. Arai et al. [23] has demonstrated that augmenting a participant's avatar with an extra body part in a virtual environment can evoke the sense of embodiment, leading participants to perceive the virtual limb as an extension of their own body.

Inspired by these findings, our research leverages VR technology to investigate how the addition of multiple semi-autonomous VSLs influences user interaction. To further quantify the effects, we employed an additional questionnaire to measure the sense of agency and ownership [23].

6.3.2 Supernumerary Robotic Limbs

Research on supernumerary robotic limbs (SRL) has explored various forms of augmenting the human body [330, 331, 332, 45], such as robotic legs [334], fingers [332], and arms [45]. Our research focuses on virtual supernumerary limbs, which add extra limbs to enhance functionality in a VR environment.

Initially, SRL were developed for industrial applications, such as assisting workers in construction and assembly tasks [335, 336, 337]. The field has since expanded to include designs like robotic tails for balance [338, 339] and additional limbs for complex tasks [340, 330, 341, 342]. For example, Maekawa, Kawamura and Inami [338] introduced a wearable robotic tail to aid in balance, showing its potential in physically demanding activities.

Control mechanisms for SRL have been a focal point of investigation, with various methods proposed to enable intuitive and effective operation. Techniques such as remapping body motions (e.g., foot or shoulder movements) [331, 8, 343] and utilizing brain-computer interfaces for direct limb control [43] have been explored. Despite the promise of these approaches, the majority of studies rely on fixed mapping schemes that limit adaptability in dynamic environments. Moreover, there is a noted lack of comprehensive user studies, with

many systems validated through proof-of-concept demonstrations involving single users [331, 44, 45]

To address these limitations, our research investigates the interaction dynamics involved in controlling multiple supernumerary limbs within a VR environment s) using a non-fixed mapping scheme for supernumerary arms capable of different levels of autonomy. In our study, participants will utilize voice and gestural commands to dynamically switch control between two VSLs while performing semi-realistic tasks. This approach aims to provide new insights into flexible and responsive multiple VSLs control for practical usages, contributing to a deeper understanding of human-machine interaction in increasingly complex contexts.

6.3.3 Mapping Control Strategies in SRL

Existing research has primarily focused on fixed-mapping control schemes, where SRL are governed by pre-determined mappings, such as proprioceptive interfaces like foot pedals [8]. While these methods have demonstrated effectiveness in controlled environments, they struggle to accommodate the dynamic nature of real-world scenarios, where user intentions and tasks may shift rapidly. Similarly, another study on artificial limb substitution added two robotic arms controlled by leg movements, expanding the user's capacity to four arms [289]. However, such systems still lack the flexibility required for real-time adaptation, particularly when users must manage multiple limbs simultaneously. Similarly, efforts to enhance physical capabilities through SRL, such as the development of robotic tails for balance augmentation, underscore their potential but fall short in addressing the complexities of dynamic, multi-limb control [338]. Additionally, studies on wearable robotic arms examined user-preferred interaction styles but focused on controlled environments, leaving the need for flexible, real-time control largely unaddressed [9].

Our study design is inspired by existing research on user-preferred interaction methods and the control of multiple VSLs. Building on these foundations, our research shifts the focus to non-fixed mapping, enabling users to dynamically manage and seamlessly switch control among multiple VSLs in real-time.

6.3.4 Wizard-of-Oz Methodology in Supernumerary Robotic Limbs Research

The Wizard-of-Oz (WoZ) methodology is widely used in human-computer interaction research to simulate autonomous systems, enabling the study of user interactions before full prototypes are available [416, 417]. For example, Diederichs et al. used a WoZ vehicle to explore human interaction with AI-driven cars, providing insights into user trust and adaptability [418]. Similarly, Hu et al. developed “Wizundry,” a platform for simulating speech-based interfaces with multiple wizards, facilitating the study of complex human-AI collaboration [419]. In SRL research, Muehlhaus et al. used the WoZ approach to simulate semi-autonomous functionality, analyzing user strategies for controlling additional limbs [9]. This methodology has also been applied to such contexts as automated vehicles [420], offering insights into user acceptance and trust. This approach is particularly valuable for SRL research, allowing researchers to investigate user behavior and interaction patterns in controlled environments. By leveraging WoZ setups, researchers can collect rich data that informs the design of SRLs and other advanced robotic systems, bridging the gap between early prototypes and user-centered design.

Our research builds on these insights by leveraging the WoZ methodology to simulate varying levels of autonomy in VSLs in a VR environment. This approach enables us to systematically study user behavior, control strategies, and embodiment under controlled conditions, mitigating the technical constraints of fully autonomous systems. By focusing on VR, we extend the applicability of WoZ in SRL research, providing foundational insights into user interaction patterns while acknowledging the need for future validation in physical settings.

6.4 DESIGN AND IMPLEMENTATION

This section outlines the design considerations and implementation details of our Wizard-of-Oz setup for VSLs in a VR.

6.4.1 Design Considerations

In HCI, shared manipulation in virtual environments has been explored through frameworks like the Level of Manipulation (LoM)[253]. In our VR study, we implemented “distributed control across different LoM to form a cohesive entity” [225] using a four-arm setup, where two VSLs were controlled by the participant and two by the human operator. This setup allowed both to independently control separate LoM, collectively contributing to the operation of a single virtual entity for coordinated interaction.

The control mechanism in our study was implemented by mapping the user-controlled virtual avatar’s limbs and the VSLs to separate controllers, with both the participant and the operator wearing VR headsets. The participant directly controlled the avatar’s limbs, while the operator managed the avatar’s VSLs. Non-arm-related actions, such as walking, were excluded to maintain focus on arm manipulation.

To address the concerns of over-reliance on human operators in Wizard-of-Oz setups [421], we introduced *random error injections* and *delayed responses* in the operator-controlled limbs to simulate the unpredictability of imperfect autonomous systems [422]. These errors were designed to mimic real-world limitations in autonomous control, requiring the participant to adapt to system imperfections. Additionally, during critical tasks, the operator introduced *temporary loss of control* in certain limbs to simulate failures in the system, forcing the participant to find alternative strategies, thus enhancing the realism and complexity of the simulated autonomy.

Since both the participant and operator were responsible for controlling a total of four virtual limbs for the same avatar, the question of how virtual vision will be shared in a collaborative setting is an important consideration in the system design. To facilitate seamless collaboration, we synchronized the visual experience within the virtual environment. Both the participant and operator views were anchored to the same avatar, ensuring a unified first-person perspective. The participant’s head movements controlled the viewpoint, minimizing motion sickness and maintaining alignment between the two users’ actions.

The participant's Unity interface was shared with the operator via, allowing the operator to monitor actions in real-time. The operator could also view the participant's perspective directly through the VR headset, ensuring effective coordinated interaction within the virtual environment.

6.4.2 Implementation

We used the Unity platform to implement our study. This setup enabled both the participant and a human operator to share a single virtual environment, with their real-time interactions synchronously reflected on the same avatar. Our setup was based on two Oculus Quest 3 head-mounted displays and controllers. The application was developed using Unity version 2022.3.7f1 and maintained a stable frame rate of 90 Hz throughout the experiment. This allowed us to reduce the delay between the participants' actions to 5ms. In the virtual environment, the participant and the human operator were simultaneously tethered to a single anthropomorphic avatar.

To avoid biases, the participant and the human operator were positioned two meters apart in the same room but separated by a physical partition, ensuring the participant was unaware of the operator's presence, as shown in Figure 6.1. This setup allowed the operator to clearly hear the participant's instructions while remaining undetected, thereby maintaining the integrity of the Wizard-of-Oz methodology. To enhance the realism of autonomous interaction, we implemented *blinded control conditions*, where the operator received incomplete task information, simulating limited system awareness. This forced the participant to adjust for the system's imperfect decision-making. Combined with *error injection and control delays*, this approach captured user behavior in scenarios that mimicked real-world autonomous system challenges. We selected the "Photon Unity Networking" (Photon PUN) framework to implement the task environment due to its robust capabilities in supporting multiplayer interactions. Photon PUN's integration with Photon Cloud, a software-as-a-service (SaaS) solution, facilitated the development of our multiplayer environment. Additionally, we utilized an on-premises Photon Server to host the task locally, which further minimized

network latency and ensured real-time responsiveness between the participant's actions and the system's feedback.

6.5 EXPERIMENTAL METHOD

The aim of this study is to explore interaction strategies for controlling multiple semi-autonomous VSLs in non-fixed mapping VR scenarios. To ensure the tasks reflect real-world use cases for supernumerary limbs, we designed two VR tasks representing distinct categories of interaction: targeted reaching named *Basic Control Task* and object manipulation named *Factory Task*.

6.5.1 Experimental Conditions and Procedure

6.5.1.1 Participants

We invited 14 participants with prior VR experiences (6 female, 8 male; $M = 25.6$ yrs; $SD = 7.9$ yrs; no color vision deficiencies) to participate in the study. This sample size is consistent with the standards used in HCI research [142, 9]. Participants with VR experience were selected to avoid distractions caused by learning controls [423]. 7 out of 14 participants used VR regularly in professional settings, and the other 7 used VR for entertainment or gaming.

6.5.1.2 Experimental Conditions

We employed a within-subjects experimental design. All participants completed the experiment in both autonomy conditions — *low-autonomy level* and *high-autonomy level*. These two autonomy levels, as theoretically outlined by Kim et al. [424], provide the understanding of how different degrees of robot independence influence human-robot interaction. Detailed descriptions of autonomy levels, including participant instructions, system capabilities, and operational guidelines, are provided in Appendix 1. The autonomy settings, which encompass *description*, *capabilities*, and *role*, were designed following the practical framework proposed

by Lima et al. [425]. To minimize any potential fatigue or learning effects, we counterbalanced the order of conditions presented to the participants with a break between each task and condition.

To formulate the differentiation of the two autonomy levels, we classified plausible instructions for VSL control into several categories informed by principles from robot learning and multi-agent coordination frameworks [426, 49, 427, 428]. They start with two high-level categories: *action instructions* and *meta-action instructions*. *Action instructions* are instructions issued by the participants that specify how the VSLs should operate or proceed. These include **Command Instructions**, which involve direct physical actions, such as pressing or rotating objects [426]; **Demonstration Instructions**, involving participant-led imitation, derive from the robot learning from demonstration (LfD) paradigm, where robots replicate observed behaviors [49]; and **Delegation Instructions**, which issue abstract commands for autonomous planning, using conditional generation and reinforcement learning to enable systems to infer strategies for multi-step tasks [427]. *Meta-instructions*, on the other hand, do not specify a course of action but are used to simplify or speed up communication between the operator and the VSLs. For example, these include **Labeling Instructions**, which allow participants to use properties like color or size to specify or indicate objects, hence simplifying naming conventions for future action commands, which is important for contextual recognition in robotic task sequencing [428]. We used this classification to define the two autonomy levels:

Low-autonomy level: At this level, the VSLs require explicit, step-by-step guidance from participants, responding to instructions categorized as **Command**, **Labeling**, or **Demonstration instructions**. Each command from the participant triggers a single action from the operator (controlling the VSLs) without independent planning or task abstraction. For example, a **Command instruction** such as “Press here” results in simple, discrete movements that reflect immediate task execution [426], i.e., pressing a button. A **Labeling instruction**, e.g., “This is a red ball,” involves identifying and categorizing objects based on specified properties, consistent with robotic object classification frameworks [428]. A **Demonstration instruction**, such as rotating an object by 90 degrees, requires the operator to replicate the participant’s

real-time actions, exemplifying robot learning from demonstration [49]. This level limits the VSLs to reactive behaviors, entirely dependent on participant guidance.

High-autonomy level: At this level, the VSLs demonstrate advanced capabilities, including object recognition, multi-step task execution, and autonomous planning. The key difference in this level is the inclusion of ***Delegation instructions***, such as “Sort all blocks by colour” prompts the operator to autonomously identify object properties, develop sorting plans, and execute multi-step sequences, reflecting decentralized task planning strategies [427]. By incorporating planning and abstraction, this level provides enhanced efficiency and flexibility for managing complex tasks.

6.5.1.3 Procedure

The study procedure began by collecting demographic information, including participants’ age, gender, VR proficiency, prior VR gaming experience, and any color vision deficiencies. Participants were then briefed on the experiment’s purpose, and written informed consent was obtained. Participants were explicitly informed that the VSLs were designed to respond to specific verbal commands and demonstrations using hand gestures. Next, we provided participants with detailed descriptions of the two autonomy levels (see Appendix 1) and verbally explained the components of each, covering the *description*, *capabilities*, and *role*. To ensure consistency, the operator’s actions were guided by predefined rules (see Appendix 2), developed during pilot studies to standardize responses across participants. In the low-autonomy condition, the operator performed tasks step-by-step based on explicit participant instructions. In the high-autonomy condition, the operator executed tasks as single actions based on participant commands.

Participants were then assisted with putting on the VR headsets and provided with controllers. Before beginning the main tasks, they engaged in a 5-minute warm-up exercise based on the relaxed think-aloud protocol (RTA) [344, 345].

This exercise encouraged them to verbalize their thought processes, enabling us to gather richer qualitative data on their decision-making and interaction strategies. Throughout the

experiment, participants were encouraged to continue thinking aloud, describing their actions, reasoning, and adjustments during the tasks.

The experimental tasks followed a sequence of *Basic Control Task* → *Factory Task*, conducted once at the *low-autonomy level* and again at the *high-autonomy level*. Each task required the use of both avatar's limbs (controlled by the participant and VSLs, and participants described their strategies and reasoning throughout the tasks. To prevent potential biases, the order of autonomy levels was counterbalanced.

After each autonomy level, participants removed their VR headsets and completed an embodiment questionnaire. Following the *Factory Task*, we conducted a semi-structured interview focusing on their experience of managing the VSLs, their perceptions of task performance under different levels of system autonomy, and their awareness of embodiment. The experiment, including the briefing, training, data collection, and interviews, took approximately 50 minutes. No compensation was offered to participants for their involvement.

6.5.2 Experimental Tasks

The tasks were designed to explore participants' control strategies to employ the VSLs in VR, requiring coordination, precision, and timing. They required participants to simultaneously and collaboratively use the avatar's limbs and VSLs.

6.5.2.1 Basic Control Task

The first task in our study (Figure 6.2(a)) is a typical "reaching" task, commonly used in VR object manipulation [429]. Participants completed basic coordination through the action of three-dimensional reaching out to a button, a fundamental skill in remote robotics manipulation [42]. The reaching task, chosen as the first task for its straightforward objective, singular focus, and absence of distracting elements. It provides a controlled environment to examine basic multiple-limb coordination. As demonstrated by Guterstam et al. [430], reaching tasks are crucial for exploring how users integrate the VSLs into their body schema.

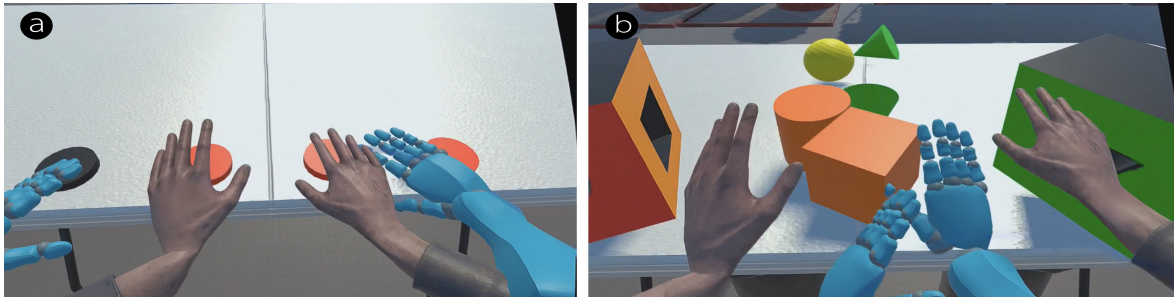


FIGURE 6.2. Two tasks used in the user study: the two avatar's limbs were controlled by the participants, while the VSLs were controlled by a human operator. (a) Basic Control Task: Participants used the avatar's limbs and VSLs to simultaneously actuate illuminated buttons in a sequence. (b) Factory Task: One hand grasps a falling target, while the other rotates a Shape Sorter to match and insert the target into the correct hole based on shape and color.

The task included four stages, with an increasing number of lit buttons at each stage. In stage 1, only one button was lit; Vstage 2 had two; and progressively stage 4 had four buttons. The participants were asked to use both their avatar's limbs and SLs to press lit buttons to turn them off. The task was considered complete once all lit buttons were pressed in each stage, with no repeated attempts or iterations within the stages. Performance metrics such as task completion time and error rate were recorded to assess how effectively participants managed the tasks across different stages. Participants had the freedom to choose which limb to use for each button, but in the later stages, with three and four buttons lit, participants needed to engage both their avatar's limbs and VSLs simultaneously to complete the task.

6.5.2.2 Factory Task

The second task is a fast-paced robotic-control task (Figure 6.2(b)) involving grasping, rotating, and reaching [431]. It builds on earlier reaching task by adding precision, timing, and coordination requirements. To ensure the task reflected natural behavior, it needs bilateral coordination, which involve using both limbs for simultaneous or alternating movements. Specifically, we focused on asymmetrical bilateral coordination, where each limb performs a distinct action, such as one hand grasping while the other rotates or positions an object [432]. Our task design aligns with key insights from the study by Zhang, Long and Luo [433], who emphasize the role of SRLs in high-pressure, multitasking environments such as industrial

settings. This task was chosen as the second task for incorporating these elements, reflecting real-world applications of SRLs in enhancing human performance in repetitive, demanding tasks.

The task mechanism involved a *Shape Sorter*, a cubic object with four sides containing differently shaped holes each. Two Shape Sorters were placed on the table, with the same set of shapes on each, but arranged with different color associations. While the shapes on both sorters were identical, the colors corresponding to each shape varied between the two sorters. For example, a triangular hole might be on the green side of one and the yellow side of the other.

The task structure featured ten objects, with matching shapes and colours of the *Shape Sorter* faces. The objects were dropped one by one on the table every three seconds (Fig. 6.2). The participants were asked to insert the objects to the shape sorter matching the shape (e.g., such as triangles, cylinders, and cones) and color. Participants needed to grab the falling target, rotate the Shape Sorter to align both the shape and the color, and insert the target into the corresponding hole. This design tested participants' ability to coordinate the avatar's limbs and VSLs simultaneously, managing the tasks of grasping, rotating, and inserting the target under time pressure. Participant's earned points for each correct insertion, and task completion time and accuracy measured performance.

6.5.3 Experimental Setup

We conducted the experiment in a 2.5 x 3.5 meter laboratory room, which was divided into two sections by a partition to ensure that the participant was unaware of the human operator's presence, maintaining the integrity of the Wizard-of-Oz setup. A GoPro camera recorded the participant's activities for later analysis. The participant wore a VR headset and provided voice commands, which the operator could hear, but the operator remained silent and did not communicate directly with the participant during the experiment.

6.5.4 Experimental Data

6.5.4.1 Qualitative data

We audio and video recorded the think-aloud session and the answers to the interviews as qualitative data.

6.5.4.2 Performance

We measured the participants' performance using two variables: task completion time and error rate. Specifically, for the *Basic Control Task*, the task completion time was calculated from the moment the button illuminated until the participant successfully actuated the button. For the *Factory Task*, performance was assessed by measuring the time taken for each action, from grasping the target to placing it into the correct hole, along with the number of accurate target insertions. To avoid biases, participants were not shown their scores during the experiment.

6.5.4.3 Embodiment Experience

As a measure of the degree of shared virtual body experience, we utilized the Avatar Embodiment Questionnaire (AEQ) to assess evaluation of embodiment [18]. The participants completed the questionnaire at the end of tasks associated with each level of system autonomy (i.e., after completing the low-autonomy level, and again after completing the high-autonomy level). This approach mirrors the two critical stages identified in [23], which is pre- and post-learning stages. Participants removed their VR headsets to complete the questionnaire.

6.6 RESULTS

In this section, we present both quantitative and qualitative findings, focusing on embodiment, control and task management strategies against different levels of autonomy of multiple VSLs.

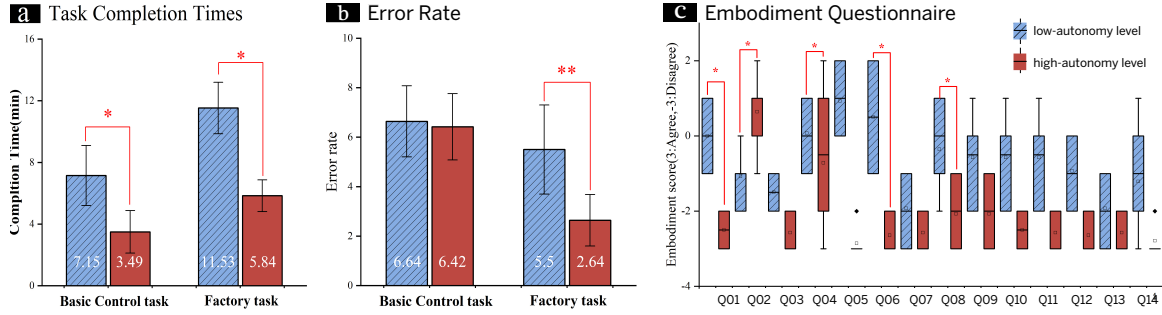


FIGURE 6.3. The figure shows the averages for *low-autonomy* level and *high-autonomy* levels in (a) the task completion time in the basic control task and factory task; (b) the error rates in the basic control task and factory task; (c) the embodiment questionnaire. * indicates $p < .05$ and ** indicates $p < .01$.

6.6.1 Quantitative Findings

Figure 6.3 outlines our quantitative findings in terms of task completion time, error rate and participants response to the Avatar Embodiment Questionnaire (AEQ). Since the data did not follow a normal distribution, we used the Aligned Rank Transform (ART) test to perform non-parametric comparisons.

As shown in Figure 6.3(a), participants in average took significantly longer time to complete the tasks in *low-autonomy level* compared to *high-autonomy level* in both *Basic Control task* ($z = -1.05, p < 0.05$) and *Factory task* ($z = -2.09, p < 0.05$). For the error rate, we could not observe a statistically significant difference between the autonomy levels in the *Basic Control task* ($z = -0.75, p = 0.44$). However, as shown in Figure 6.3(b), in the *Factory task*, error rate in the *low-autonomy level* was significantly higher compared to that of *high-autonomy level* ($z = -2.93, p < 0.01$)

As shown in the Figure 6.3(c), results from the embodiment questionnaire administered after each autonomy level revealed statistically significant differences (using Wilcoxon–Mann–Whitney test) to several questions. Statistically significant differences were observed in several key questions: Specifically, Q1: “I felt as if the virtual extra limbs/arms were my limbs/arms” ($p < 0.05$); Q2: “It felt as if the virtual extra arms/limbs I saw were someone else’s” ($p < 0.05$); Q3, “It seemed as if I might have more than two limbs/arms” ($p < 0.05$); Q4, “It felt like I could control the virtual extra arms as if they were my own arms” ($p < 0.05$);

Q6, “I felt as if the movements of the virtual extra arm were influencing my own movements” ($p < 0.05$); Q8, “I felt as if my arms were located where I saw the virtual extra arms” ($p < 0.05$). As a trend, the participants tend to feel a higher sense of body ownership (Q1, 2), a sense of agency (Q4, 6), and a sense of self-location (Q8) after *low-autonomy level* compare to that of *high-autonomy level*.

To validate the reliability of our quantitative measures, we analyzed completion times for different tasks and how they relate to instruction types to ensure their actions were consistent and adhered to guidelines (Figure 6.4). Since the data conforms to a normal distribution, we employed a One-Way ANOVA to compare the operator’s average reaction time for instructions.

Across the study, participants issued a total of 672 commands. Specifically, **Command** - 192 times, **Labeling** - 160, **Demonstration** - 176, and **Delegation** - 144 times.

For **Command Instructions**, shorter commands (1–3 seconds) resulted in an average operator reaction time of $M = 2.50$ ($SD = 0.31$) seconds, while longer commands (4–6 seconds) required significantly more time at $M = 4.53$ ($SD = 0.40$) seconds ($z = -3.21, p < 0.01$). Similarly, for **Labeling Instructions**, shorter commands (1–2 seconds) corresponded to a reaction time of $M = 2.02$ ($SD = 0.23$) seconds, whereas longer commands (3–5 seconds) took $M = 3.84$ ($SD = 0.31$) seconds. For more complex **Demonstration Instructions**, the operator’s reaction time also scaled proportionally with command length. Commands lasting 3–4 seconds elicited an average reaction time of $M = 4.52$ ($SD = 0.39$), while those lasting 6–8 seconds increased significantly to $M = 8.03$ ($SD = 0.48$) seconds ($z = -3.52, p < 0.001$). Lastly, **Delegation Instructions** exhibited the longest reaction times, with shorter commands (2–3 seconds) averaging $M = 3.46$ ($SD = 0.30$) seconds and longer commands (6–8 seconds) reaching $M = 9.04$ ($SD = 0.61$) seconds.

Notably, as shown in Figure 6.4, **Command**, **Demonstration**, and **Delegation Instructions** were consistently followed by an operator reaction, with reaction times roughly aligning with the length of the commands. In contrast, **Labeling Instructions** were followed by shorter reaction times since no physical movement was required. Additionally, under Low Autonomy,

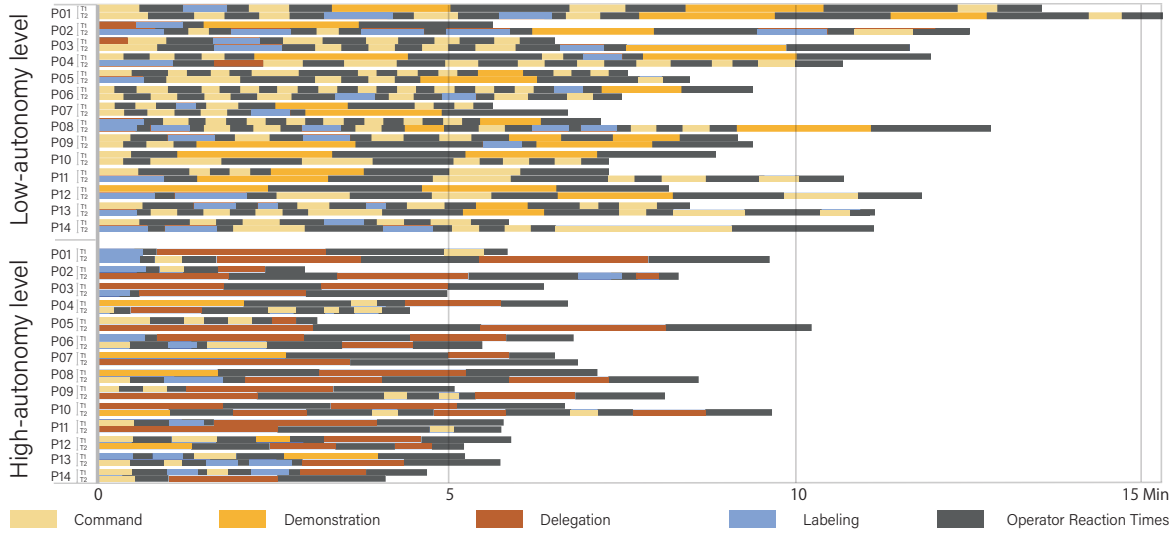


FIGURE 6.4. The figure illustrates the participant-issued instructions, including *Command Instructions*, *Demonstration Instructions*, *Delegation Instructions* and *Labeling Instructions*, along with the WoZ operator’s reaction times to these four instruction types. A single colored segment may represent multiple instructions of the same type. It also compares the reaction times across two distinct autonomy levels for both the Basic Control Task and the Factory Task.

Delegation Instructions did not trigger any reaction as this feature was unavailable, while in high Autonomy, *Delegation Instructions* consistently resulted in reactions of equivalent durations.

6.6.2 Qualitative Findings

We conducted an inductive thematic analysis [352] to explore how participants controlled the VSLs and how their experiences varied with different autonomy levels. This analysis aimed to understand participants’ interaction strategies, focusing on how they balanced manual and autonomous control, as well as the challenges in managing multiple VSLs. The analysis was conducted by two independent coders, ensuring diverse perspectives and reducing individual bias. To enhance consistency, inter-rater reliability was measured using Cohen’s Kappa ($\kappa = 0.74$).

Following Braun and Clarke's framework [353], we began by familiarizing ourselves with the data through a review of all interview transcripts to identify key elements of participants' feedback. In the initial coding phase, we assigned descriptive labels to relevant passages, focusing on body ownership, agency, task delegation, labeling strategies, and managing multiple tasks. Independent coding was followed by collaborative discussions to refine and consolidate codes into initial categories reflecting participants' strategies for handling task complexity and adapting control methods.

We refined and grouped the codes into higher-order themes as patterns emerged, such as "Body Ownership, Agency, and Control Embodiment" and "Control Strategies and Their Perceived Performance". These themes helped us build a framework for understanding participants' experiences in adapting to multitasking demands with VSLs. Our iterative coding approach consistently identified key themes such as "Balancing Cognitive Load and Task Complexity" and "Trust, Ownership, and Control Embodiment" across participant data. We followed reflexive thematic analysis principles [353], allowing ongoing reflection and refinement to shape the analysis.

The following sections outlines the key themes that emerged from participants' experiences, providing insights into how they adapted their control strategies and perceptions across different tasks and levels of autonomy.

6.6.2.1 "It felt weird, like they were someone else's arms" — Body Ownership, Agency, and Control Embodiment

Having two additional limbs was an unusual experience for participants, with many (N = 8) initially describing a sense of disconnection. As P11 remarked, "*It felt weird, like they were someone else's arms*", especially when the VSLs didn't fully align with their intended movements. However, participants were surprisingly able to develop a sense of autonomy over time: "*At first, they felt awkward, but the more I used them, the more natural*", noted P08.

Sense of Body Ownership: 8 out of 14 participants felt a stronger connection to the VSLs when they had direct control, particularly during low-autonomy level. “*When I was guiding the arm, it made me feel more real*”, P02 said. This manual control reinforced their sense of ownership, particularly in VR settings, where repeated interactions allowed participants to adapt to the virtual embodiment of the arms, rather than perceiving them as physical robotic systems. “*The more I used it, the more natural it felt*”, explained P05.

In contrast, 9 out of 14 participants found that as the system’s autonomy increased, their sense of ownership diminished. “*When the arms acted on their own, it was more like watching but not controlling*”, P05 said. Although the system efficiently performed tasks, the lack of direct involvement left participants feeling disconnected. “*It didn’t feel like my arm anymore*”, noted P07. This is also mirrored in our quantitative findings in the Section 6.6.1.

Sense of Agency: Participants (10 out of 14) reported that direct control enhanced their sense of agency, providing a feeling of mastery, especially during precision tasks. “*I felt more control when I moved the arms*”, said P03. Conversely, increased autonomy reduced participants’ sense of being in control, with 8 out of 14 participants reporting a shift to an observer role. “*When the arms worked, I felt like more of an observe*”, shared P02. The quantitative findings (Section 6.6.1) also mirrored these. While autonomy allowed efficiency, it diminished participants’ perception of control.

Sense of Self-Location: 9 out of 14 participants experienced a sense of self-location when the VSLs closely synchronized with their real body. “*I didn’t really feel like they were extra or separate*”, said P07. However, when the system’s movements failed to meet expectations, it disrupted this connection. “*If the arms didn’t follow my action in time, I felt disconnected*”, noted P02. 7 out of 14 participants, managing multiple VSLs simultaneously challenged their spatial awareness, particularly during complex tasks.

6.6.2.2 “I felt like I was juggling!” — Control Strategies and Their Perceived Performance

Controlling a single limb in virtual reality allows participants to intuitively manage actions. However, controlling multiple VSLs highlights how often coordination is required and how difficult it is to maintain. 11 out of 14 participants in our study highlighted how managing several limbs at once required them to constantly shift focus and attention, making it difficult to allow smooth coordination between VSLs. “*It felt like juggling tasks*”, P05 remarked, emphasizing the complexity of maintaining precise, synchronized movements when operating multiple VSLs simultaneously.

Direct Control and Autonomous Repetition: 8 out of 14 participants used various strategies to delegate tasks to the VSLs. Some (N=13) preferred for direct control, where the VSLs mirrored their movements, especially at *low autonomy* level. “*Left arm, mirror my movements*”, P11 said, who employed a specific strategy to teach the VSLs to follow their avatar’s limbs’ actions. P03 described their approach, saying, “*I choose to do a simple action, and ask it to learn and follow*”. “*It is easier*”, they added, “*It [is] like another arm that did learn what I was doing, but I need to look at two arms*”. 6 out of 14 participants named the action before demonstrating it. For example, P08 said “*This is rotate, rotate this way*” to teach the VSLs. This allowed for real-time adjustments, but it became mentally exhausting when managing multiple VSLs. “*It required so much more effort and guidance, which was tiring*”, noted P02.

6 out of 14 participants opted to teaching the VSLs a set of actions to repeat autonomously. “*Once I showed the arm how to do something, I didn’t have to guide it anymore—it just kept going*”, P05 said. While this method was efficient, it sometimes required several tries to ensure correct execution. P06 clearly described this idea: “*I wanted to demonstrate a movement and let left arm learn it. I taught it a simple action, then keep it do it again and again. Just kept going until I say stop*”. In contrast, at higher levels of autonomy, all participants opted for more direct command strategies rather than demonstrations to control the VSLs. As P04 issued the command, “*Robotic [virtual supernumerary] Arm, press the red*

button”, the system executed the task autonomously. Instead of showing the VSLs how to perform an action, 8 out of 14 participants simply gave verbal commands and let the system execute the task autonomously.

Organizing Strategies for Efficiency: Pre-planning and labelling tasks by object properties played a crucial role in improving task management for 13 out of 14 participants. 10 out of 14 participants in the factory task tried labeling objects by color or shape to improve task management and streamline interactions with the VSLs. P06 said, “*Robotic [virtual super-numerary] arms, this is light [Illuminated button], always press the light one [Illuminated button]*”, this participant explained it by adding, “*I [labelled] the target how I felt like, and yeah, it’s red, but I’m ... used to calling stuff that lights up a ‘light button,’ so that’s what I said*”. However, 4 out of 14 participants felt that too much pre-planning disrupted their natural workflow. “*I spent more time trying to label thing[s]*”, explained P09, indicating that excessive organization could slow down task execution.

In addition to labeling, 5 out of 14 participants categorized objects by size or function, which enhanced coordination. “*Once everything was sorted, I didn’t have to keep thinking about it*” (P08). This reduced decision-making during the task, but 4 out of 14 participants found it challenging to balance planning with keeping their momentum. One participant remarked, “*[Planning] helped, but I sometimes lost track of the main task when I was too focused on organizing*” said P11.

At the *high-autonomy* level, 6 out of 14 participants chose not to label targets themselves. Instead, they followed conventional rules when issuing commands, rather than assigning custom names to the targets. “*I could say something like ‘rotate the box 180 degrees’ or ‘press the red button’, and they’d just do it*”, noted P04.

Managing Task Control Approaches: Sequential handling allowed for better control but took more time, especially at *low-autonomy* levels. “*Doing one thing at a time... , it gave me more control, but it just slowed me down*”, noted P08. In contrast, multitasking offered speed but led to more errors. “*When I tried doing both at once, it felt faster, but..., I made way more mistakes*” (P01). At *high-autonomy* level, 7 out of 14 participants demonstrated different

strategies for managing control between themselves and the VSLs. For high-priority, delicate tasks, 4 out of 14 participants preferred keeping full control to ensure accuracy. *“I just liked handling those tasks myself, the ones that needed constant adjusting”* said P03. Others found it more efficient to delegate even complex tasks to the autonomous VSLs, especially when juggling multiple tasks. *“I’d let the arms handle the boring stuff, and even some of the tricky bits, just so I could focus on hard thing, like find the shape”*, noted P05.

Switching between sequential and simultaneous actions was tough for 5 out of 14 participants, while 6 out of 14 participants adapted by letting the VSLs handle repetitive tasks. 6 out of 14 participants often switched control modes based on task complexity. *“For the simpler stuff, I’d just ... teach it to take over... but when it was more detailed, I’d switch back to doing it manually”*, explained P09.

6.6.2.3 Summary

Our study examined participant interactions with semi-autonomous VSLs across two autonomy levels. Qualitative results showed that participants preferred direct control in *low-autonomy* level, which strengthened their sense of body ownership, agency, and self-location. At this level, participants used strategies like having the VSLs follow their actions, repeat simple tasks, or even labeling objects themselves to improve task management. In *high-autonomy* level, participants found it more efficient to delegate tasks to the system, relying on verbal commands and pre-defined rules rather than manual control or custom labeling. However, this shift led to a diminished sense of control and connection to the limbs. This study answered RQ1 by identifying the primary strategies participants used to balance cognitive load (Section 6.6.2.2) and task complexity (Section 6.6.2.1 during multiple VSLs control). RQ2 was addressed by detailing how participants adapted their control strategies based on perceived task demands (Section 6.6.2.2) and autonomy levels (Section 6.6.2.1). Finally, RQ3 was explored through insights on how control embodiment influenced participant interactions with the VSLs (Section 6.6.2.1).

Control strategies and VSLs Control (Section 6.7.1)		
Strategists	Main Finding	Design Suggestions
Verbal Control	Frequent switching between VSLs interrupted workflow and added cognitive load.	Enable smoother switching between control modes. Provide flexible task-role reassignment mechanisms.
Task Demonstration	Direct mirroring required high user attention and frequent recalibration.	Incorporate hybrid control (manual and automated) systems. Offer real-time feedback on positioning.
Task Management	Labeling improved coordination but pre-defined categories reduced flexibility.	Support dynamic object recognition. Allow on-the-fly adjustments to categories and labels.
Control and Performance in Multiple VSLs Systems (Section 6.7.2)		
Body Ownership	Participants felt disconnected from VSLs during high autonomy and poor alignment with their body perception.	Provide real-time visual feedback. Gradually adjust control autonomy.
Agency	Users preferred manual control in critical tasks, and labeling improved trust in autonomous actions.	Balance autonomy and manual control with user-defined labels. Offer options for customizable task assignment.
Task Efficiency	Synchronizing multiple VSLs was challenging, and interruptions impacted performance.	Introduce visual and auditory cues for better synchronization. Enable dynamic task distribution based on VSLs usage.
Challenges and Opportunities of Non-Fixed Mapping (Section 6.7.3)		
Workload	Manual control in <i>low-autonomy</i> settings increased workload; <i>high-autonomy</i> shifted users to a supervisory role.	Provide proactive feedback mechanisms. Let users toggle between manual and autonomous modes.
Autonomy	High autonomy caused uncertainty and reduced user confidence, while manual control was seen as more predictable.	Ensure autonomy is predictable and transparent. Implement clear trust-building mechanisms.

TABLE 6.2. Summary of control strategies, performance findings and design suggestions for VSL systems.

6.7 DISCUSSION

Building on our findings, we have identified key implications for designing and improving autonomous VSLs in VR. This section explores how users adapt control strategies, manage task complexity, and experience embodiment when coordinating VSLs. Additionally, we discuss challenges and opportunities in integrating system autonomy to enhance user experience and system responsiveness in future VR applications.

6.7.1 Implications for Control strategies and VSLs Control

Our findings reveal that selecting effective control strategies is critical for optimizing the use of VSLs. Participants adapted their strategies based on task complexity and autonomy levels, shifting between verbal commands, teaching by demonstration, and manual control to best suit task demands. While low autonomy settings benefited from precise, step-by-step instructions, higher autonomy allowed for efficient task delegation. However, balancing manual input and autonomy presented challenges, particularly in multitasking scenarios, where task-switching disrupted workflow.

We selected VR as a research platform for simulating and evaluating VSLs interactions due to its capacity to safely and flexibly replicate scenarios that are challenging to achieve with physical robotic systems. Prior research suggests that VR can serve as an effective prototyping environment for understanding body schema extension and control strategies, with potential applications in physical settings [23, 405]. Although VR-based embodiment involves a level of user “suspension of disbelief” [408], it still enables valuable insights into user adaptation to novel body configurations. While our findings focus on VR, they provide foundational insights that may guide future exploration with physical robotic limbs, assuming that VR-to-reality transferability is feasible.

While previous research has explored multimodal interaction and task management [434, 435], further development is needed in the following directions to enable smoother transitions of control.

Verbal Control: Our results demonstrate that continuous verbal control was most effective in *low-autonomy* settings, where participants needed precise, step-by-step instructions to guide the VSLs, aligned with prior research [436]. However, as task complexity or autonomy increased, continuous verbal control became less efficient, reflecting the challenges of maintaining precision when managing VSLs in real-time.

With higher autonomy, participants adopted more efficient strategies such as task delegation or pre-labeled ensemble actions (Figure 6.6), which reduced reliance on continuous verbal

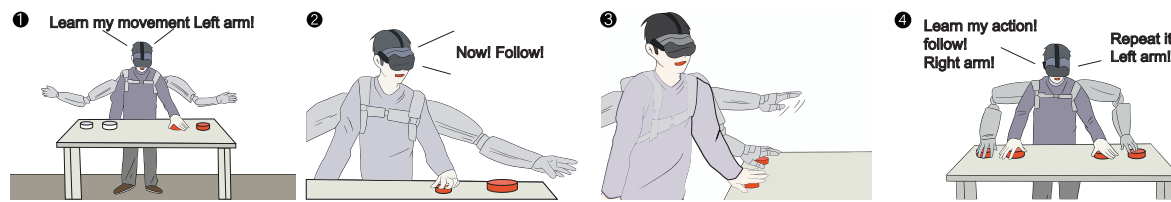


FIGURE 6.5. This figure shows the flow of the follow and repeat strategies performed by participants.

input by allowing the system to execute tasks autonomously after initial instructions. Previous studies on multimodal interaction similarly suggest that as autonomy increases, users prefer intuitive, low-effort methods like task abstraction or labeling [434]. Controlling VSLs under non-fixed mapping further exacerbated challenges, as simultaneous tracking and the lack of consistent mapping made precise verbal input difficult.

Participants often combined labeling with control by example to simplify task management, streamlining control and enabling autonomous handling of multi-step tasks. However, unlike studies employing pre-defined control schemes [437, 438], the reliance on customized labels introduced challenges, particularly the repeated guidance needed when the system failed to learn correctly. To address these issues, future systems should reduce user intervention required for correcting learning failures.

Control Choice as a Function of Task or Autonomy: Participants adapted their control strategies based on task complexity, with simpler, repetitive tasks delegated to the autonomous VSLs and more complex, high-stakes tasks managed manually, as shown in figure 6.5. Research in task management, such as Rasmussen’s Skill-Rule-Knowledge Framework, supports this approach, where simple tasks are categorized as skill-based and easily automated, while complex tasks, which are knowledge-based, require manual control for greater precision [439]. The consistent preference for manual control in critical tasks, even when autonomy was available, reflects the importance of precision and control for participants.

Task-switching, intended to facilitate multitasking, often ended up disrupting the flow of tasks instead. Switching frequently between manual and autonomous control, especially in complex tasks, interrupted their workflow, making it harder to maintain smooth task execution. While autonomy improved efficiency for simpler tasks, more complex tasks required careful

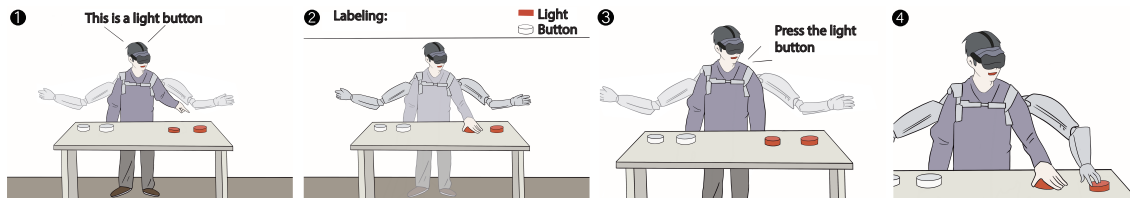


FIGURE 6.6. This figure shows the organizing strategies performed by participants.

oversight. Rather than treating autonomy and manual control as separate modes, future systems should integrate these approaches more seamlessly, allowing for dynamic shifts based on task complexity without disrupting the user's workflow.

6.7.2 Optimizing Control and Performance in Multiple VSLs Systems

Managing multiple VSLs revealed insights into the intersection of body ownership, agency, and task efficiency. While participants faced challenges in coordination across autonomy levels, the study highlighted opportunities for optimization. Participants noted the potential of multiple VSLs when designed with a balance of manual and autonomous control. However, achieving optimal efficiency and engagement requires careful attention to synchronization, attention allocation, and user feedback mechanisms. This section explores these factors and identifies strategies for improving multiple VSLs systems.

Sense of Body Ownership in Autonomous Multiple VSLs Systems: In *low-autonomy* conditions, participants felt a stronger sense of body ownership, perceiving the VSLs as extensions of their body. Coordinated control reinforced this connection, but higher autonomy levels introduced a disconnect, especially when autonomous actions disrupted limb coordination, as similarly observed in previous research [440, 16]. Participants favored partial autonomy, which allowed manual control of key movements while delegating routine tasks, balancing control and assistance. Visual alignment between the avatar's limbs and VSLs remained crucial for maintaining ownership across all autonomy levels. Future systems should leverage real-time visual feedback to reinforce this connection.

Sense of Agency in Autonomous Multiple VSLs Systems: Higher autonomy shifted participants' roles from direct control to supervision, transforming agency into a strategic interaction—a contrast to the findings of Firlej and Taeihagh [441]. While partial control over one VSL, combined with monitoring the other, helped maintain participants' sense of influence even in *high-autonomy* conditions, agency was not perceived as a simple binary state [442]. Instead, it emerged as a dynamic experience, with participants continually balancing delegation and responsibility. However, unpredictability in the system's responses introduced cautious delegation, which, in turn, affected participants' confidence in the autonomous features.

Attention Management and Task Efficiency in Multiple VSLs Control: Task efficiency depended on how participants allocated attention [443, 444, 445, 445]. Attention theories, such as Kahneman's model of attention [446], suggest that humans allocate limited mental resources to tasks based on demand and importance. Dividing attention between manual and autonomous tasks optimized performance, allowing participants to focus on complex actions while VSLs handled routine tasks [447]. However, frequent task-switching disrupted flow and reduced efficiency [448]. Sustained attention became challenging in high-demand tasks, with monitoring autonomous VSLs leading to fatigue. Future designs should prioritize attention management by assigning tasks based on graded attention levels.

Opportunities of Using Multiple VSLs: The potential of multiple VSLs systems spans various fields as supernumerary limbs approach real-world applications [449, 450, 451]. Prior studies have explored augmenting human capabilities through robotic systems in controlled settings like manufacturing and healthcare [452, 453, 454]. Our research builds on this by showing how VR-developed control strategies can guide VSLs implementation in dynamic scenarios. In industries like manufacturing and logistics, delegating repetitive tasks to VSLs allows workers to focus on precision-driven actions [455, 456, 457], reflecting our findings where participants assigned simpler tasks to autonomous VSLs while managing complex ones manually. This approach could enhance productivity by aligning with selective attention needs in high-demand environments. In healthcare, multiple supernumerary limbs systems could support injury recovery, physical therapy, and assistive technologies for individuals with

disabilities [458]. Studies highlight adaptability as crucial in rehabilitation technologies [459], and our findings suggest VSLs could further personalize therapy.

6.7.3 Challenges and Opportunities of Non-Fixed Mapping

The non-fixed mapping of control in multi-arm systems posed both challenges and opportunities. *High-autonomy* conditions reduced mental load by allowing VSLs to handle simpler tasks, but the lack of fixed associations introduced uncertainty. In *low-autonomy* settings, frequent switching and manual coordination increased cognitive strain. Even in *high-autonomy* scenarios, participants hesitated to fully trust the system's autonomous decisions, especially when precision was needed. This section explores the impact of non-fixed mapping on workload and autonomy.

Workload Distribution and Cognitive Strain in Autonomous Non-Fixed Mapping Systems:

In *low-autonomy* conditions, participants experienced higher cognitive load due to frequent task switching and manual coordination between the VSLs. The absence of fixed task-to-limb associations required continuous monitoring and role adjustments, increasing mental strain and reducing efficiency. Conversely, *high-autonomy* conditions alleviated much of this burden by delegating simpler tasks to the VSLs, allowing participants to focus on more complex actions, consistent with prior findings [460]. However, participants still had to monitor the VSLs, and the non-fixed mapping introduced uncertainty in balancing manual control and trust in the system.

Trust and Control in Autonomous Multiple VSLs Systems:

Managing multiple VSLs in *low-autonomy* settings was challenging, particularly when precision was required, as participants had to manually coordinate each limb. In *high-autonomy* settings, trust became a key issue as participants were hesitant to rely fully on the system's autonomous decision-making when control switched between VSLs without direct input. While autonomy reduced the manual burden, unpredictable decisions led to reduced confidence. As emphasized by Hancock et al. [460], mismatches between user trust and system reliability can result in either overtrust or

distrust. Future systems should offer more transparent and predictable autonomous decisions to address these challenges in non-fixed mapping scenarios.

6.8 Limitations and Future Work

Our study has several limitations that should be noted to inform future research. In this study, the autonomy of the VSLs was simulated using a Wizard-of-Oz approach, where the system's autonomous behavior was covertly controlled. Although participants believed they were interacting with fully autonomous VSLs, the system's responses were manually guided. This allowed us to explore a range of interaction scenarios, but the simulation may not fully capture how users would behave with truly autonomous systems. Specifically, participants were limited to interacting with the VSLs through mediated commands without the ability for direct physical interaction. The absence of real-time system errors or unexpected actions typically encountered with fully autonomous VSLs could influence the user's experience. As a result, these findings may not entirely generalize to real-world autonomous systems. Future studies with actual autonomous systems are needed to further validate how users would engage with and trust such technologies in uncontrolled environments.

While VR served as an accessible and controlled platform for studying supernumerary limb interactions, we recognize that our findings are speculative regarding their application to physical robotic systems. Although past studies indicate potential transferability from VR to real-world settings [23, 405], our findings require further validation with physical prototypes before they can be confidently applied to actual supernumerary robotic limbs. Consequently, while this study provides initial insights, future work is necessary to assess the relevance of these insights in real-world contexts.

We focused on two distinct levels of autonomy to examine how users interact with the VSLs, offering valuable insights into basic control strategies. However, this limited scope may not capture the full range of system autonomy or task complexity that could arise in more dynamic, real-world scenarios. Future work should explore a broader spectrum of autonomy levels and task complexities to better understand how users adapt to varying system autonomy. Similarly,

while our tasks were designed to isolate specific coordination and control challenges, they may not represent the full variety of real-world applications for supernumerary limbs. Expanding the scope of tasks in future studies will help validate findings across a wider range of use cases.

While the sample size meets the typical standards for exploratory studies in Human-Computer Interaction (HCI) to gather qualitative insights and identify key interaction patterns [142], it may still limit the generalizability of the findings. Similar sample sizes have been used in comparable exploratory studies, such as those by Muehlhaus et al. [9], to effectively investigate interaction strategies. The study's small sample size, while suitable for qualitative exploration, limits the generalizability of the quantitative findings. The quantitative metrics, such as task completion time and error rate, should be interpreted as exploratory trends rather than definitive results. Furthermore, certain user diversity factors, such as color blindness or other visual impairments, were not considered. Future studies should incorporate a larger, more diverse participant pool to address potential accessibility challenges and ensure the system's inclusivity.

6.9 CONCLUSION

In this work, we explored how varying levels of autonomy in VSLs affected user interaction, control strategies, and embodiment in a VR environment. Through two task-based VR activities—the *Basic Control Task* and the *Factory Task*—we examined how users adapted to low and high autonomy using the VSLs across two tasks. Our findings demonstrate that participants preferred assigning repetitive or lower-priority tasks to the VSLs while maintaining direct control over more critical actions. Additionally, participants reported a stronger embodiment and control with a *low-autonomy* level, while a *high-autonomy* level allowed for greater task efficiency but a weakened embodiment of the VSLs. We have discussed the key implications of our themes on future research and design. In addition, we have provided a discussion of use cases for control strategies and a table of concrete suggestions. Our findings are based on VR-based simulations of VSLs and provide insights

into virtual embodiments. While these insights offer potential directions for real-world robotic systems, further research is needed to validate these results in physical environments. This understanding is crucial for designing autonomous systems that enhance user experience, improve task performance, and maintain a strong embodiment, with potential applications in gaming, rehabilitation, and assistive technologies.

PART III

Findings, Reflections, and Future Directions

CHAPTER 7

Discussion

This thesis contributes to the fields of Human–Computer Interaction (HCI) and computer-supported cooperative Work by advancing a novel lens of **Synergistic Embodiment** for the design and analysis of complex multi-user interaction in Virtual Reality (VR). Rather than treating collaborative VR as a straightforward extension of single-user embodiment, this thesis has examined how shared avatars, supernumerary limbs, perspective switching, and varying levels of autonomy reconfigure coordination, agency, workload, and the lived experience of collaboration. Across the studies presented in this thesis, we identified recurring design tensions, user strategies, and methodological considerations that together form a foundational account of synergistic embodiment. While this account is necessarily shaped by the systems and contexts investigated here, it provides a principled basis for future research and design in collaborative embodiment.

7.1 Limits of Existing Approaches

The literatures reviewed in Chapters 2 and 3 provide essential foundations for this thesis, but they do not in themselves yield an adequate account of collaborative embodiment. Existing work has made substantial progress on embodiment, shared control, perspective design, and supernumerary limb interaction. However, these areas have often developed in parallel, each emphasizing a different part of the problem while leaving their interdependence undertheorized. As a result, the field still lacks a framework for understanding how collaborative embodiment is configured across multiple users, viewpoints, control channels, and system processes.

Research on embodiment has been especially important for establishing the roles of ownership, agency, and self-location in VR [106]. However, much of this literature begins from the assumption of a single user engaging with a single body through a relatively unified perception-action loop [22]. In collaborative settings, this assumption becomes unstable. Agency may be shared, ownership may be partial or distributed, and perceptual access may be unevenly allocated across collaborators. What is therefore required is not simply an extension of single-user embodiment theory, but a framework capable of accounting for how embodied experience changes when control and perception are distributed across multiple actors.

Research on shared control has likewise provided important insights, particularly into input allocation, coordination burden, and conflict management. However, this literature is often primarily control-centric [232, 233, 234, 235, 236]. It tends to focus on who controls what, how conflicts are resolved, and how shared action can be made operationally effective. Less attention is typically given to how these mappings interact with perception, embodiment, or autonomy. Consequently, shared control is often treated as a problem of coordination architecture rather than as part of a broader problem of collaborative embodiment.

Research on perspective design has shown that viewpoint affects comfort, awareness, and task performance. Yet this literature is frequently camera-centric. It often compares first-person, third-person, and hybrid views as isolated display conditions, without fully situating perspective within a broader account of role asymmetry, shared agency, or distributed perception [302]. This makes it difficult to explain why a viewpoint that appears suboptimal from the perspective of individual embodiment may nonetheless be effective in a collaborative context.

Research on supernumerary limbs and augmentation has expanded the field's understanding of how users adapt to additional effectors and varying control schemes [330, 331, 332, 45]. However, this work is often augmentation-centric. It tends to emphasize usability, mapping strategies, or the effectiveness of additional limbs for a single user, rather than examining how autonomy, delegation, and interpretability reshape embodied collaboration. In particular, it rarely addresses how autonomous or semi-autonomous limbs alter not only performance, but also the user's relation to action, responsibility, and intervention.

Taken together, these literatures remain insufficient because they do not adequately explain how control allocation, perceptual access, feedback legibility, and autonomy jointly structure collaborative embodiment. They are valuable as partial views, but not as a unified account. What remains under-specified is how these dimensions interact in practice, especially under conditions of role asymmetry, uneven expertise, task transitions, and delegated action. It is precisely this gap that the empirical studies in this thesis begin to address.

7.2 Design Tensions Across the Empirical Studies

The three empirical studies in this thesis do not merely report separate findings about shared control, perspective switching, and supernumerary limbs. Taken together, they reveal a set of recurring design tensions that cut across systems and tasks. These tensions are important because they indicate that collaborative embodiment is not best understood as a collection of isolated technical features. Rather, it is organized around a small number of interdependent design questions that repeatedly shaped performance, workload, agency, comfort, and coordination.

Chapter 4 (*CoplayingVR*) makes clear that collaborative embodiment cannot be understood only in terms of whether multiple users can act through the same avatar. The more consequential issue is how control is partitioned once a single virtual body is no longer tied to a single source of action. In this setting, interaction ceases to be a matter of individual input alone and becomes a matter of negotiated timing, conflict management, and role clarity. The findings from *CoplayingVR* show that shared control can indeed scaffold less experienced users and improve engagement under appropriate conditions. At the same time, these benefits are inseparable from the additional coordination burden introduced by distributed action. What emerges here is not simply a result about shared gameplay, but a broader tension between redistributed support and the effort required to sustain intelligible shared agency.

A different but related issue emerged in Chapter 5 (*One Body, Two Minds*), where the central question was no longer who controls the body, but how access to the task world is perceptually organized across collaborators. The study showed that viewpoint in collaborative VR is not

well described as a matter of display preference alone. Different perspectives make different forms of information available: local bodily alignment, partner-relevant cues, and wider spatial structure are not equally visible from any single view. The findings further suggest that users do not necessarily benefit from maximal flexibility in viewpoint selection. What appeared to matter more was the availability of stable but reconfigurable perspectives that matched the phase of the task. This shifts the interpretation of perspective from camera choice to a broader problem of distributed perceptual access, and with it reveals a recurring tension between local embodiment and the demands of global coordination.

Chapter 6 (*Juggling Extra Limbs*) extends this line of argument by showing that autonomy changes not only how much users can do, but also how they relate to their own action. Once semi-autonomous limbs are introduced, the interaction is no longer exhausted by direct manipulation. Users begin to move between direct control, delegated action, and selective intervention, and the quality of the experience depends increasingly on whether system behaviour remains intelligible and recoverable. Higher autonomy can reduce manual burden and support parallel action, but it may also weaken agency if users can no longer understand what is happening or when intervention is possible. The central tension here is therefore not simply between manual and automatic control, but between the efficiency afforded by delegation and the risk that control becomes experientially opaque.

Taken together, the three studies suggest that collaborative embodiment is shaped less by any single design choice than by the relation among several recurring concerns: how control is distributed, how perceptual access is organized, and how much action is delegated under autonomy. Across the studies, these concerns did not operate independently. Rather, they became consequential insofar as they either supported or destabilized coordinated action. What therefore recurs across the empirical chapters is not a set of isolated findings, but a common structural problem: collaborative embodiment depends on how these dimensions are configured together in practice.

A further point follows from this observation. For control distribution, perceptual access, and autonomy to remain workable in practice, users must be able to interpret the state of the system and of one another's actions. This makes **feedback legibility** a cross-cutting

requirement rather than an additional feature alongside the other dimensions. In different ways, each study showed that users relied on stable cues to determine who controlled what, which perspective was currently operative, whether an action was user-driven or system-driven, and when adjustment or intervention was needed. Although feedback was not always the manipulated variable, it repeatedly emerged as a condition under which the other design dimensions became intelligible and therefore usable. In this sense, legibility is not auxiliary to collaborative embodiment, but one of the conditions that makes coordinated embodied action possible.

These recurring tensions suggest that the problem space of collaborative embodiment can be organized around four interdependent design concerns: how control is allocated, how perception is distributed, how system state is made legible, and how autonomy is introduced without undermining agency. The Interaction Elements Matrix proposed in the next section is derived from this observation. It does not replace the empirical findings of Chapters 4–6. Rather, it provides a way of relating them to one another by identifying the principal dimensions through which these recurring tensions become designable.

7.3 Interaction Elements Matrix

The recurring tensions identified above suggest that collaborative embodiment cannot be adequately described through any single variable taken in isolation. Across the empirical studies, difficulties and successes did not arise from control, perspective, feedback, or autonomy alone, but from the ways in which these factors were configured together. This makes it necessary to move from a chapter-by-chapter summary of findings toward a more explicit design-space representation.

The *Interaction Elements Matrix* (IEM) is introduced for this purpose. It is intended neither as an exhaustive taxonomy nor as a predictive model of all collaborative VR behaviour. Rather, it serves as a structured basis for design reasoning. Its role is to identify the principal dimensions that repeatedly shaped coordination, comfort, agency, and task performance across

the systems studied in this thesis, and to show how these dimensions become consequential only in relation to one another.

More specifically, the IEM is built around a distinction between **primary design dimensions** and **cross-cutting moderators**. The primary dimensions capture aspects of system design that can be directly configured by a designer and that were repeatedly manipulated, foregrounded, or rendered consequential across Chapters 4–6. The moderators do not describe design dimensions of the same kind. Instead, they refer to contextual factors that influence how particular configurations are likely to function in practice.

7.3.1 Primary Design Dimensions

The first level of the IEM consists of four primary and analytically separable design dimensions.

- **Control Mapping:** how control is allocated, shared, or reassigned across users and limbs, ranging from fixed assignment to shared control and dynamic redistribution.
- **Perspective Strategy:** how visual access to the body, partner, and task space is configured, including egocentric, exocentric, anchored, hybrid, or selectively switchable views.
- **Feedback Design:** how the system communicates control state, stability, and autonomy through cues that make control locus, viewpoint status, and delegated action legible. These cues may be visual, auditory, or haptic.
- **Autonomy Level:** the extent to which actions are manually controlled, delegated, or semi-autonomous.

These four dimensions recur across the empirical chapters for different reasons. **Control Mapping** becomes central when embodied action is distributed across multiple users, as seen in *CoplayingVR*, where the main issue was not only whether users could share a body, but how shared action remained coordinated and intelligible. **Perspective Strategy** becomes central when collaborators require different forms of perceptual access to support local manipulation and global coordination, as shown in *One Body, Two Minds*. **Feedback Design** emerges as

necessary because none of the other dimensions remains workable unless users can interpret who controls what, what the current viewpoint is, or when delegated action is taking place. **Autonomy Level** becomes central once the system takes over part of the action, as in Juggling Extra Limbs, where the key issue was not only efficiency, but whether delegation remained understandable and agency-preserving.

For this reason, these dimensions are best understood as *analytically separable but inter-dependent in practice*. A given perspective strategy may alter how a control mapping is experienced. An autonomy setting may only remain acceptable when paired with sufficiently legible feedback. A mapping that works well in one task may become unstable in another. The value of the IEM lies precisely in making these interactions available for analysis.

7.3.2 Cross-Cutting Moderators

The second level of the IEM consists of cross-cutting moderators that shape how the primary dimensions operate in practice.

- **Task Context:** whether the task emphasizes local manipulation, global coordination, repetition, time pressure, or multi-step procedure.
- **Role Asymmetry:** whether collaborators occupy similar or differentiated functional roles.
- **Expertise Distribution:** whether collaborators bring comparable or uneven levels of skill and prior knowledge.
- **Transition Policy:** how and when the system supports changes in control, viewpoint, or autonomy across phases of activity.

These moderators are not additional dimensions of the same kind as mapping, perspective, feedback, and autonomy. Rather, they help explain why the same configuration may function differently across settings. A shared control arrangement that scaffolds a novice user may be effective when expertise is unevenly distributed, but unnecessary or disruptive when collaborators are equally skilled. A perspective strategy that supports stable coordination in a multi-step task may become too restrictive under rapid transitions. Likewise, a degree

of autonomy that reduces burden in repetitive work may undermine agency in tasks that require frequent correction or improvisation. The moderators therefore indicate that the design problem is not only which configuration to choose, but under what conditions a given configuration remains intelligible and effective.

7.3.3 Interaction Elements as Configurations

Within this design space, each configuration can be understood as an *interaction element*: a modular combination of control mapping, perspective strategy, feedback design, and autonomy level, shaped by contextual moderators. In this sense, the IEM does not prescribe a single optimal arrangement. Instead, it provides a way of reasoning about why different constellations of design choices give rise to different forms of collaborative action and different qualities of embodied experience.

The three prototypes developed in this thesis occupy different but connected regions of this space. **CoplayingVR** primarily examined shared but relatively fixed control mappings under predominantly egocentric interaction, thereby foregrounding the relation between redistributed control and shared agency. **One Body, Two Minds** extended the space by introducing alternative perspective strategies and role-asymmetric perceptual access, thereby foregrounding the relation between viewpoint and distributed awareness. **Juggling Extra Limbs** explored increasing autonomy together with the need for feedback that made delegated action legible, thereby foregrounding the relation between delegation, interpretability, and agency.

Taken together, these systems do not exhaust the IEM. Rather, they provide empirically grounded instances through which its structure became visible. Figure 7.1 presents the IEM as a conceptual map of collaborative embodiment design possibilities. The purpose of the figure is not to close the design space, but to render its principal dimensions and their relations available for future comparison, refinement, and extension.

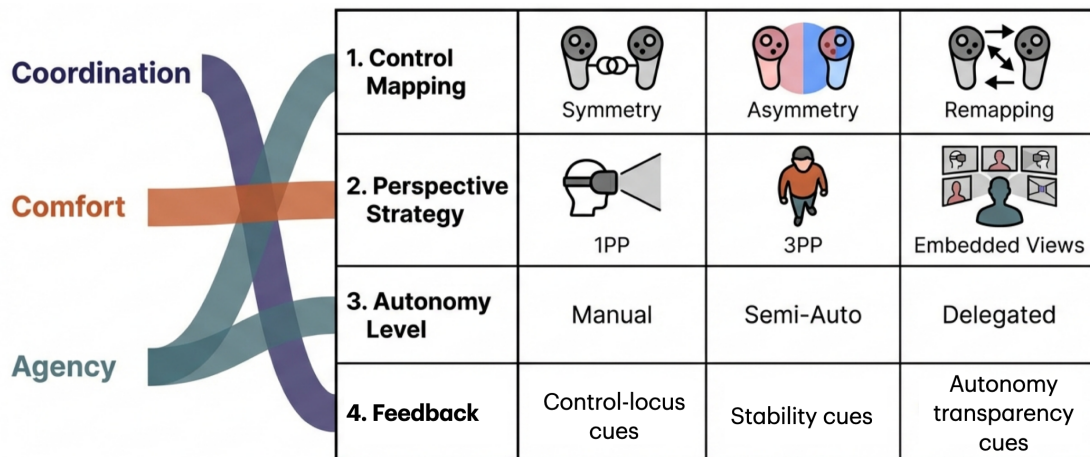


FIGURE 7.1. Interaction Elements Matrix (IEM): a first-order design space for collaborative embodiment in VR, organized around control mapping, perspective strategy, feedback design, and autonomy level. The matrix summarizes the principal dimensions surfaced across Chapters 4–6 and relates the studied prototypes to adjacent design possibilities.

7.4 The Explanatory Value of the IEM

The value of the IEM lies less in adding another classificatory scheme to the literature than in making visible a set of relations that existing approaches tend to treat separately. The framework does not replace prior work on embodiment, shared control, perspective, or supernumerary limbs. Rather, it provides a way of understanding how these concerns become interdependent once embodied action is distributed across multiple users, viewpoints, and system processes.

A first implication concerns the status of *collaborative embodiment* itself. Much existing research on embodiment begins from a relatively unified relation between one user and one virtual body. Within that analytic frame, ownership, agency, and perception are primarily treated as properties of a single user-body loop. The findings of this thesis suggest that such a frame becomes insufficient once action is shared, delegated, or perceptually distributed. In collaborative settings, agency may be negotiated, ownership may be partial or uneven, and perceptual access may differ systematically across participants. What the IEM makes visible, then, is that collaborative embodiment is not simply individual embodiment extended

to more than one person. It is a differently organized problem in which control, perception, and responsibility must be configured across a distributed system of action.

A second implication concerns the role of asymmetry. Prior work often treats symmetry as the implicit baseline for good collaboration, whether in relation to viewpoint, access, or control. The studies in this thesis do not support symmetry as a universal design ideal. In several cases, collaboration was improved not when users occupied identical positions within the system, but when they occupied differentiated and complementary ones. This was especially apparent in the perspective study, where the requirements of local manipulation and global coordination did not always align within a single view. The significance of the IEM here is that it provides a way of treating role asymmetry not as a deviation from an ideal arrangement, but as a potentially productive feature of collaborative embodiment, provided that such asymmetries remain intelligible and appropriate to the task.

A third implication follows from the findings on viewpoint switching. Existing discussions of flexibility in immersive systems often imply that more reconfigurability is inherently better. The empirical results of this thesis suggest a more qualified picture. Users did value the availability of alternative perspectives, but they did not generally benefit from constant or maximal switching. Instead, they tended to settle into perspectives that were sufficiently stable for the current phase of activity and only reconfigured when task demands changed. The IEM helps explain this by showing that flexibility is not valuable in abstraction. Its value depends on how it interacts with task phase, role structure, and the cognitive cost of reorientation. In this sense, the framework shifts attention from flexibility as abundance of options to flexibility as appropriately timed reconfiguration.

A fourth implication concerns autonomy. Research on augmentation and autonomous assistance often emphasizes what automation can do in performance terms. The findings from Chapter 6 suggest that this is only part of the issue. Increased autonomy may indeed reduce burden, but it may also weaken agency when users can no longer determine what the system is doing, when intervention is possible, or where control currently resides. The IEM brings this into focus by treating autonomy and feedback design as inseparable. What matters is not autonomy alone, but whether autonomous behaviour remains legible, recoverable, and

compatible with a meaningful sense of involvement in action. This helps explain why the design problem is not simply how to automate more, but how to make delegated action experientially intelligible.

Taken together, these points indicate that the IEM contributes not merely by organizing dimensions of system variation, but by clarifying the form of the design problem itself. It shows that collaborative embodiment cannot be adequately analysed through isolated questions of mapping, viewpoint, or automation alone. What must instead be understood is how these dimensions are configured together, and how their configuration shapes the possibility of coordinated, intelligible, and sustainable embodied action.

7.5 Research Propositions Emerging from the IEM

If the IEM is to function as more than a retrospective synthesis, it should also support the formulation of future research questions. Its value, in other words, lies not only in organizing the findings of this thesis, but in indicating what kinds of relations among control, perception, feedback, and autonomy ought to be examined more systematically. The propositions below are offered in this spirit. They are not intended as fixed laws, but as empirically tractable claims that follow from the recurring tensions identified across the thesis.

A first set of propositions concerns the relation between **task structure** and **perspective**. The studies in this dissertation suggest that viewpoint preferences are unlikely to remain stable across all forms of activity. Where task demands shift from fine-grained local manipulation to broader coordination across space, users may increasingly prefer exocentric or hybrid perspective strategies over tightly coupled first-person views. Relatedly, the findings suggest that the value of perspective flexibility lies less in constant switching than in the availability of stable, phase-sensitive reconfiguration.

A second set of propositions concerns the relation between **control distribution** and **perceptual access**. The empirical studies suggest that collaboration quality depends not only on whether control is clearly mapped, but also on whether collaborators are given access to the

forms of information most relevant to their role. This implies that, in role-asymmetric tasks, complementary rather than symmetric configurations of control and perspective may often produce better coordination and lower workload.

A third set of propositions concerns **autonomy**, **feedback**, and **agency**. The findings of Chapter 6 suggest that increasing autonomy may reduce interaction burden, but that such reductions are unlikely to remain beneficial unless system behaviour is transparent, recoverable, and legible in terms of control locus. Future work should therefore examine not autonomy in isolation, but the extent to which different forms of feedback allow delegated action to remain intelligible and agency-preserving.

A final proposition concerns **expertise distribution**. The findings from CoplayingVR suggest that shared control may be particularly effective as a scaffolding mechanism when expertise is unevenly distributed and when task pacing allows time for coordination and adaptation. This implies that the value of shared control may depend not simply on its mechanics, but on the social and temporal conditions under which it is introduced.

7.6 Addressing the Research Questions

To guide this investigation, this thesis addressed three overarching research questions introduced in Chapter 1:

- **RQ1:** How do different visual perspectives and control mechanisms influence user performance, experience (including comfort and workload), and the sense of embodiment in collaborative VR tasks involving shared avatars and supernumerary limbs?
- **RQ2:** What interaction strategies do users develop and prefer when co-controlling a single avatar, especially with varying levels of limb autonomy or when distinct user roles are implied by the control scheme?

- **RQ3:** How can insights from these user experiences be synthesized into a design framework to guide the creation of more effective and intuitive collaborative VR systems?

Each question offers a different analytic entry point into the thesis' contributions. **RQ1** concerns the experiential and performance consequences of different configurations of perspective, mapping, and autonomy. **RQ2** concerns the adaptive strategies users develop to manage collaborative complexity. **RQ3** concerns how these findings can be abstracted into a framework capable of informing future design.

7.6.1 RQ1: How do different visual perspectives and control mechanisms influence user performance, experience (including comfort and workload), and the sense of embodiment in collaborative VR tasks involving shared avatars and supernumerary limbs?

The findings across Part II show that there is no universally optimal configuration of perspective and control. Instead, the most appropriate configuration depends on the structure of the task, the distribution of roles, and the degree of coordination required between collaborators.

With respect to **visual perspective**, the thesis identified a persistent trade-off among immersion, situational awareness, and comfort. Tightly coupled first-person viewpoints support immediacy and sensorimotor alignment, but in mobile or role-asymmetric scenarios they can also produce disorientation, instability, and motion discomfort. By contrast, anchored and third-person perspectives improve awareness and reduce discomfort, but often at the cost of reduced direct embodiment. These observations suggest that perspective should be treated not merely as a viewing condition, but as a mechanism for distributing perceptual access according to task demands and user roles.

With respect to **control mechanisms**, the thesis similarly found that redistributing control has both benefits and costs. In CoplayingVR, shared control improved novice performance and supported engagement, suggesting that control sharing can act not only as a collaborative

mechanism but also as a pedagogical and assistive one [16]. At the same time, poorly articulated mappings increased ambiguity, frustration, and coordination overhead. In this sense, mapping and perspective are best understood as interacting rather than isolated design variables.

Taken together, these findings motivate the treatment of control mapping and perspective strategy as co-dependent dimensions within the IEM. They also suggest that collaborative embodiment should not be optimized for immersion, performance, or embodiment in isolation, since these qualities often come into tension under distributed action.

7.6.2 RQ2: What interaction strategies do users develop and prefer when co-controlling a single avatar, especially with varying levels of limb autonomy or when distinct user roles are implied by the control scheme?

Across the three empirical studies, users did not behave as passive recipients of system logic. They actively developed strategies to manage complexity, reduce uncertainty, and stabilize collaboration.

In relation to **semi-autonomous supernumerary limbs**, participants moved away from direct control alone and adopted strategies of task delegation. Repetitive or peripheral subtasks were increasingly handed to the additional limbs in order to preserve cognitive resources for more demanding actions. Participants also developed organizational strategies such as mentally labeling objects, spatial regions, or limb roles. These patterns suggest that increasing autonomy does not simply reduce manual effort. It reorganizes action by moving users toward more strategic forms of coordination.

In relation to **perspective switching**, participants were flexible but selective. They rarely switched viewpoints continuously. Instead, they tended to select a viewpoint that was “good enough” for the current phase of the task and remain there until conditions changed. This indicates that users value stability as much as flexibility. It also supports the view that

reconfiguration is most useful when it is aligned with task phases and role requirements rather than treated as a constant resource to be exploited.

More broadly, users developed both explicit and implicit coordination strategies. Explicit strategies included verbal negotiation, turn-taking, and instruction-giving. Implicit strategies emerged through rhythm, anticipation, and complementary division of labour. These observations suggest that collaborative embodiment must be understood procedurally, as users continually negotiate roles, delegation, and perceptual access over time.

7.6.3 RQ3: How can insights from these user experiences be synthesized into a design framework to guide the creation of more effective and intuitive collaborative VR systems?

The answer developed in this thesis is that collaborative embodiment can be productively organized as a design problem structured by control mapping, perspective strategy, feedback design, and autonomy level, with task context, role asymmetry, expertise distribution, and transition policy acting as cross-cutting moderators. The IEM captures this structure as a design space, while the SEF derives design principles from it.

The resulting framework contributes more than a post hoc summary of individual studies. It offers a basis for explanation, comparison, and hypothesis generation. For example, the principle of *Perspective Flexibility* follows from the demonstrated trade-off between immersion and situational awareness. The principle of *Controllable Autonomy* follows from the observation that delegation is productive only when users can still interpret and recover from system action [17]. The principle of *Asymmetric Support* follows from the finding that shared control can scaffold novice users when expertise is unevenly distributed [16]. Likewise, the importance of feedback legibility and role clarity follows from repeated evidence that users rely on stable cues to interpret who controls what, what the system is doing, and when intervention is possible.

The value of the framework therefore lies not only in consolidating prior results, but in showing that collaborative embodiment can be treated as a coherent HCI design problem rather than as a set of isolated technical techniques. By identifying its recurring dimensions and their interactions, the framework provides a basis for more generalizable design reasoning in collaborative VR.

7.7 Methodological Discussion

Building on the systematic survey in Chapter 3 and the three empirical studies in Chapters 4–6, this section reflects on the methodological lessons that emerged across the thesis. These reflections are not offered as generic recommendations for VR experimentation. Rather, they arise from a more specific observation: collaborative embodiment is difficult to study through any single metric, task format, or prototype in isolation. Because the phenomenon spans coordination, agency, workload, perceptual access, adaptation, and delegation, the methodological problem is itself one of integration.

7.7.1 Mixed Methods as a Requirement Rather Than an Option

One of the clearest lessons from this thesis is that collaborative embodiment in VR cannot be adequately captured through any single class of measure. Objective performance, subjective embodiment, verbalized strategy, and behavioural adaptation do not move in lockstep. A configuration that improves speed may simultaneously reduce ownership. A condition that increases perceived agency may also increase workload. A task that appears smooth in quantitative terms may nevertheless involve substantial hidden coordination effort. For this reason, mixed methods were not simply useful in this dissertation. They were methodologically necessary.

Across the studies, different forms of evidence revealed different parts of the phenomenon:

- *Quantitative performance measures* such as completion time, error rate, and task success were essential for identifying how mappings, perspectives, and autonomy conditions affected observable effectiveness.
- *Workload and embodiment scales* made it possible to distinguish between systems that were merely operable and systems that remained experientially sustainable, comfortable, or agency-preserving.
- *Qualitative protocols*, including think-aloud and post-task interviews, exposed strategy formation, role negotiation, and delegation practices that were not inferable from performance data alone.
- *Behavioural and physiological proxies*, where used, helped reveal adaptation, arousal, and interaction fluency at levels that sat between explicit report and task outcome.

The methodological implication is that future research on collaborative embodiment should not treat mixed methods as an optional supplement to “core” performance measures. In this domain, triangulation is often the only way to understand whether a configuration is effective because it supports collaboration well, or merely because users are compensating for system deficiencies through effortful strategy.

7.7.2 Progressive Complexity as an Analytic Device

A second lesson concerns the role of progressive complexity in experimental design. In this thesis, task structure was not only a means of participant onboarding. It also functioned as an analytic device for revealing how collaborative strategies changed as demands increased.

Across the empirical studies, tasks were typically arranged so that participants first encountered relatively simple or bounded forms of action before moving into more dynamic, multi-step, or time-sensitive scenarios. This progression served at least three purposes.

First, it reduced early confusion and made novel mappings or perspective strategies learnable. Participants needed an initial period in which to establish a baseline relation to the system before higher-order strategy could emerge.

Second, progressive complexity made it possible to identify thresholds at which collaboration changed in kind. For example, some mappings remained workable under slow and structured conditions but broke down once timing pressure increased. Similarly, some perspective arrangements supported initial orientation but became insufficient once partners had to coordinate across larger spatial fields.

Third, this structure made visible the temporal character of adaptation. Users rarely entered a condition with a stable strategy already in place. Instead, they developed ways of working with the system over time, often moving from explicit verbal coordination to more implicit, rhythm-based, or delegated interaction. Without progressively staged tasks, such transitions would have been much harder to detect.

The broader methodological point is that collaborative embodiment should not always be studied as a single steady-state condition. It often unfolds through phases of orientation, negotiation, stabilization, and reconfiguration. Progressive complexity offers a way of making those phases empirically visible.

7.7.3 Prototypes as Instruments of Conceptual Discovery

A third lesson concerns the role of prototyping in the development of the thesis. The prototypes in this dissertation did not function only as vehicles for implementing pre-specified hypotheses. They also served as instruments of conceptual discovery. Each system surfaced interaction dynamics that were not fully foreseeable at the level of abstract design reasoning alone.

In *CoplayingVR*, the practical difficulties of assigning different hands to different users made clear that control allocation was not merely a technical mapping problem. It was also a problem of role clarity, conflict management, and shared agency. In *One Body, Two Minds*, the comparison among perspective conditions revealed that the central issue was not camera preference in itself, but how perceptual access should be distributed across collaborators. In *Juggling Extra Limbs*, the staged variation of autonomy exposed that the critical issue was not simply whether extra limbs could help, but how delegation altered users' relation to action, monitoring, and intervention.

These observations suggest that prototyping in collaborative embodiment research can serve a dual function. It tests ideas, but it also reorganizes the conceptual vocabulary through which the design space is understood. In this sense, the prototypes in this thesis did not merely instantiate the later framework. They helped bring its dimensions into view. Prototyping therefore functioned here not only as a method of evaluation, but also as a method of conceptual articulation.

7.7.4 Methodological Implications for Future Research

Taken together, these reflections suggest three methodological implications for future work on collaborative embodiment.

First, experimental designs should assume from the outset that collaboration will involve multiple simultaneous layers of experience, including performance, agency, interpretability, adaptation, and role negotiation. No single metric should be expected to stand in for these collectively.

Second, studies should be designed to capture change over time rather than only final outcomes. Collaborative embodiment is often unstable at first and only becomes intelligible through repeated interaction. Methods that capture transition, adaptation, and strategy shift are therefore particularly valuable.

Third, system-building should not be treated only as a downstream implementation activity. In this domain, prototypes are often among the most powerful means of discovering what the actual design problem is. Future research would benefit from embracing this iterative relation among prototype, empirical observation, and conceptual synthesis more explicitly.

7.8 Boundary Conditions of the Framework

The Interaction Elements Matrix and the broader synergistic embodiment perspective are intended as first-order tools for design reasoning rather than as exhaustive or universal models. Their value lies in organizing recurring tensions that emerged across the survey and empirical

studies in this thesis. At the same time, it is important to specify the conditions under which the framework was developed and the limits of what it presently claims.

First, the framework is grounded primarily in dyadic collaboration. Although some of its dimensions may generalize to larger groups, the studies in this dissertation focused mainly on two-user configurations. Questions of turn-taking, coalition formation, and multi-party control may therefore require further dimensions or different emphases.

Second, the framework is derived largely from VR-based embodied interaction. While its implications likely extend to teleoperation, robotics, and other human–AI or human–robot systems, it was not built from those domains directly. Its application outside VR should therefore be understood as a theoretically motivated extension rather than as an already validated generalization.

Third, the framework does not claim that the same dimension will carry equal weight in all domains. In some systems, control allocation may dominate design reasoning. In others, autonomy or perceptual access may be the more consequential issue. The framework is intended to support structured comparison and reflection, not to prescribe a single sequence of design priorities.

Finally, the IEM is not an optimization tool and does not provide a complete predictive account of collaborative embodiment. It identifies recurring dimensions and moderators, but does not by itself determine which combination will be best in a given setting. That question remains empirical and context-dependent.

7.9 Limitations and Future Directions

While this thesis advances understanding of collaborative embodiment in VR, several limitations remain and important avenues for future research are still open.

7.9.1 Limitations

Methodological reporting and published-chapter constraints. A further cross-cutting limitation concerns methodological reporting across the empirical chapters, particularly because Chapters 5 and 6 reproduce published or accepted conference papers. To preserve the published record, the main article text in these chapters has been retained, and methodological concerns that cut across the studies are therefore addressed here at the thesis level. Although the studies used counterbalancing and familiarization procedures where possible, they were short, repeated-measures laboratory studies; residual learning, fatigue, and order effects may therefore have influenced participants' strategies, especially in complex tasks involving perspective switching or semi-autonomous limbs. Similarly, while Chapter 5 strengthens statistical reporting through estimated marginal means, confidence intervals, standardized effect sizes, and sensitivity analysis, the level of effect-size and sample-size reporting is not fully uniform across all published chapters. These limitations do not alter the central empirical claims of the thesis, but they constrain the precision with which smaller effects and between-condition differences can be interpreted. Future studies of collaborative embodiment should report counterbalancing procedures, recruitment channels, dyad familiarity, training and readiness criteria, effect sizes, confidence intervals, and a priori power or sensitivity analyses more systematically, particularly when empirical findings are intended to support reusable design frameworks.

Scope of the Survey. The PRISMA-based survey presented in Chapter 3 catalogued 150 studies of collaborative VR, but it was necessarily limited by the databases searched and by its focus on English-language publications. Relevant work in non-English venues, technical reports, and industry-facing publications may therefore have been excluded [461]. This matters not only for completeness, but also because alternative application domains or design traditions may have foregrounded dimensions of collaborative embodiment less visible in the corpus reviewed here.

Experimental Contexts. All three prototypes—CoplayingVR (Chapter 4), One Body, Two Minds (Chapter 5), and Juggling Extra Limbs (Chapter 6)—were evaluated in controlled laboratory settings and with relatively small samples. These designs were appropriate for isolating key interaction effects, but they constrain the ecological validity of the findings. In particular, the temporal development of trust, strategy, and co-adaptation remains underexplored. The framework proposed in this chapter should therefore be read as grounded in controlled empirical evidence, but not yet as validated through long-term real-world deployment [26, 462, 17].

Participant Diversity. Participants were primarily university students and local volunteers. As a result, the studies captured only limited demographic, cultural, and experiential diversity. Other user populations may bring different expectations, comfort thresholds, and strategies to collaborative embodiment systems, especially in relation to role asymmetry, authority, autonomy, and control sharing. The framework would therefore benefit from validation across more diverse participant groups and contexts.

Wizard-of-Oz Autonomy. In Juggling Extra Limbs, semi-autonomy was simulated through a Wizard-of-Oz procedure rather than implemented algorithmically. This allowed careful exploration of the autonomy design space before committing to a fixed system implementation, but it also means that participant responses may differ from those elicited by a fully realized autonomous system. In particular, the interpretability and trust dynamics of algorithmic autonomy remain to be studied more directly.

7.9.2 Future Directions

Real-World and Longitudinal Studies. A key next step is to deploy collaborative embodiment systems in operational contexts such as industrial training, remote maintenance, healthcare, education, or rehabilitation, and to study them longitudinally. Such work would make it possible to examine how collaborative skill, trust, adaptation, and role stabilization evolve over time, and whether the tensions identified in this thesis remain stable under repeated use.

Adaptive and Intelligent Systems. Future systems are likely to incorporate more explicit forms of adaptation. By sensing workload, coordination breakdown, uncertainty, or task phase, systems could dynamically adjust mappings, perspectives, or autonomy levels. The challenge for such systems will not be adaptation alone, but *intelligible* adaptation: users must remain able to understand why a reconfiguration occurs, what it changes, and how intervention remains possible.

Multi-Modal Embodiment. This thesis has focused primarily on visual, control, and autonomy-related factors. Future work should integrate haptic, auditory, and physiological feedback more deeply, including cues such as force feedback, heartbeat, or breathing synchrony. Such modalities may not merely enrich immersion. They may also help stabilize role understanding, shared agency, and perceptual access across collaborators.

Scalability to Larger Groups. The studies in this thesis focus mainly on dyadic collaboration. Extending collaborative embodiment to three or more users raises questions not only of scale, but of structure: how control, viewpoint, and autonomy should be coordinated when collaboration is no longer pairwise. Multi-party limb sharing, role hierarchies, and distributed perceptual pooling remain especially promising areas for investigation.

Cultural and Demographic Variation. Future studies should examine how cultural norms, age, expertise, disability, and individual differences shape preferences for ownership, agency, perspective switching, and collaboration strategies. This is important not only for generalizability, but for determining whether the framework's moderators, especially role asymmetry and expertise distribution, take different forms across populations.

Physical Supernumerary Limbs and Hybrid Systems. An important next step is to connect the virtual systems studied here with robotic or wearable supernumerary limbs in the physical world. Such work would enable comparative analysis of digital and physical augmentation, and would test whether the tensions identified in VR—especially those involving interpretability, delegation, and control-locus awareness—remain similar when additional limbs produce real bodily and environmental consequences.

Framework Generalization and Refinement. Finally, the IEM and the broader synergistic embodiment perspective should be tested and refined in domains beyond those studied here, including telepresence, entertainment, education, assistive technologies, and human–robot collaboration. Such work is necessary if the framework is to function not only as a synthesis of this thesis, but as a genuinely portable guide for future collaborative embodiment research.

7.10 Conclusion

This chapter has argued that the principal contribution of the thesis lies not only in the individual findings of Chapters 4–6, but in the way those findings can be reorganized into a more coherent account of collaborative embodiment. The Interaction Elements Matrix was introduced for this purpose, not as an exhaustive taxonomy, but as a structured design space through which recurring tensions among control allocation, perceptual access, feedback legibility, and autonomy can be understood and compared.

The broader claim that emerges from this synthesis is that collaborative embodiment should not be treated as a special case of embodiment, nor merely as coordination layered onto immersive interaction. Once control, perception, and agency are distributed across multiple users, limbs, and system processes, the unit of design changes. The question is no longer simply how a single user relates to a virtual body, but how embodied action is jointly organized, interpreted, and stabilized across a distributed system.

Seen in this way, collaborative embodiment becomes a distinct HCI design problem. It requires attention not only to what users can do, but to how control is allocated, how perception is distributed, how system state is made legible, and how autonomy is introduced without eroding intelligibility or agency. By identifying these dimensions and their interactions, this thesis offers a foundation for thinking about future immersive systems not simply as interfaces for individual users, but as environments for distributed action, shared agency, and negotiated embodied cooperation.

Future work will determine how far this account generalizes beyond the systems and contexts studied here. Nevertheless, the argument developed in this dissertation is that collaborative embodiment is already substantial enough to warrant its own design language, its own methodological commitments, and its own conceptual frameworks. The aim of this thesis has been to provide an initial basis for that work.

CHAPTER 8

Conclusion

Drawing together the threads of conceptual grounding (Chapter 2), systematic survey (Chapter 3) and exploratory prototypes (Chapters 4–6), this thesis has forged a cohesive framework for co-embodying multiple users and supernumerary limbs in immersive environments.

We began this journey by articulating the latent potential of *collaborative embodiment* and *shared control* in VR, and by refining the constructs of ownership and agency beyond the solitary body to encompass dyadic and multi-limb configurations. Our domain-wide review distilled the prevailing task archetypes, mapping conventions and evaluation measures, revealing both promise and lacunae in current practice. Through three bespoke systems: CoplayingVR’s four-hand sharing, One Body Two Minds’ perspective modulation, and a semi-autonomous factory scenario, we brought theory to life, showing how control mappings, visual vantage and autonomy calibration each sculpt performance, immersion and felt togetherness. These empirical touchstones underpinned the synthesis of nine strategic design principles into a unified Interaction Elements Matrix (Chapter 7), offering a practical guide for orchestrating limbs, viewpoints, feedback and role transitions.

Classical embodiment theories have long privileged the unitary body; our contribution lies in extending these insights to the polyphony of multi-user bodies and the novelty of supernumerary limbs. By embedding phenomenological tenets into concrete VR prototypes, we have demonstrated that the same sensorimotor alignments that underpin ownership and agency can be harnessed, recast and choreographed to support collaborative action and emergent intercorporeality.

Our investigations revealed three complementary patterns. First, sharing a virtual body can heighten engagement, participants report heightened presence when they co-own limbs, but only when arbitration mechanisms prevent input conflicts. Second, shifting perspectives between embedded insets and external viewpoints allows users to balance immersion with spatial awareness, enabling both fluid coordination and a stable sense of “we.” Third, tuning the autonomy of extra arms exposes a delicate curve: too little autonomy burdens users with micromanagement, while too much leaves them feeling disempowered. Together, these findings sketch a dynamic continuum of collaborative embodiment, where control, view and autonomy must be carefully balanced.

From these insights emerges a three-phase model of boundary dissolution. We begin with *self-embodiment*, where each participant establishes a reliable sense of ownership and agency. We then progress to *shared attention*, orienting not only to one’s own limbs but to those of the collaborator. Finally, we arrive at *intercorporeal emergence*, in which the interplay between bodies, limbs and perspectives fuses participants into a coherent joint agent, dissolving the lines between self and other.

This model invites us to move beyond the linear sender–receiver metaphor toward an embodied, somatic paradigm in which interaction is a co-experienced action perception loop. Here, technology and bodies are not separate endpoints but coalesce into a single intercorporeal system, an ongoing dialogue of movement and sensation.

In practice, we confront three creative tensions: the pull between individual agency and collective fluency; the choice of fixed versus dynamic control mappings; and the balance of autonomy support against user control. Our framework offers concrete tactics, such as in-situ cues for role transitions, adjustable autonomy sliders, and synchronized multi-view displays to navigate these tensions and invite harmonious collaboration.

Looking ahead, we envision adaptive systems that respond in real time to evolving collaboration needs, longitudinal studies that trace the cultivation of shared-embodiment skills, and integrations with physical supernumerary limbs that bridge the virtual and tangible. By designing for the *in-between* the liminal space where bodies, limbs and perspectives converge,

we can cultivate VR experiences that transcend mere functionality to become transformative encounters of shared presence and collective resonance.

Proofs in Chapter 5

In this section, we show all the proofs in Chapter 5.

1 Physiological Data Processing Details

Signal source and export. Heart activity was recorded continuously with a Polar H10 chest strap. We logged the device-provided inter-beat interval (IBI/RR) stream via Bluetooth Low Energy throughout each trial and exported the timestamped RR series for offline processing. Analyses were performed on normal-to-normal (NN) intervals derived from the RR stream.

Artefact handling and NN reconstruction. RR series were preprocessed following recommended HRV workflows [383]. We removed (i) biologically implausible intervals and dropouts (e.g., extremely short/long RR values indicating missed or extra detections) and (ii) transient artefacts identified as outliers relative to local beat-to-beat dynamics. Short gaps created by artefact removal were linearly interpolated to preserve continuity for window-based HRV estimation, while longer discontinuities were excluded from HRV computation. All HRV metrics reported in the paper were computed only from contiguous NN segments that satisfied the minimum window-length requirement described below.

Windowing for HRV. HRV was computed over active-task periods using contiguous analysis windows of at least 60 s (minimum duration per window), consistent with common practice for time-domain HRV when task segments are short [383]. When a trial segment exceeded 60 s, we computed HRV on consecutive windows and then aggregated to one value per participant $\times$ condition by averaging across valid windows within that segment.

RMSSD computation. We used RMSSD as the primary time-domain HRV index:

$$\text{RMSSD} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (NN_{i+1} - NN_i)^2}.$$

RMSSD was chosen because it is widely used in psychophysiology, is comparatively robust to slow trends in heart rate, and reflects short-term vagal modulation [383]. Higher RMSSD values indicate lower physiological stress/arousal.

Peak heart rate and event-triggered analysis. Peak heart rate was extracted within the same active-task segments as the maximum observed heart rate during the segment. For event-triggered analyses, we time-aligned the heart-rate stream to perspective-switch timestamps from system logs and summarized heart rate in a short post-switch window; these event-level summaries were then modeled as a function of switch direction (e.g., EMBEDDED ANCHORED VIEW → OUT-OF-BODY VIEW vs. OUT-OF-BODY VIEW → EMBEDDED ANCHORED VIEW) and role.

Proofs in Chapter 6

In this section, we show all the proofs in Chapter 6.

1 Autonomy level details

Before starting the study, we introduced the participants to an action guide detailing the operational behaviors of the virtual supernumerary limbs. The operator adhered to this predefined guide, which specified responses to participant commands across different autonomy levels.

1.1 Low-autonomy level

Description: At this level, the virtual supernumerary limbs have basic capabilities, including the ability to perceive what is happening in their environment. They can understand and respond to simple instructions and learn from user interactions. However, while they can recognize general elements in the environment, they cannot automatically determine where or how to place specific objects without explicit user guidance. The arms rely on you to guide them step by step on how to complete tasks, categorizing items based on your instructions.

Capabilities: While they understand their surroundings, they depend on you to guide their actions. They respond well to step-by-step instructions where you guide each action.

Role: You will need to provide detailed, step-by-step guidance through verbal or gestural commands on how the virtual supernumerary limbs should sort and handle objects. These arms can only recognize and act upon the tasks you have specifically taught them. They

cannot automatically identify or categorize items they haven't learned before, so they rely entirely on your instructions to understand and execute each action.

1.2 High-autonomy level

Description: This version of the virtual supernumerary limb system is significantly more advanced, with improved language comprehension and the ability to understand complex tasks. It also automatically recognizes things in the environment, reducing the need for teach them, and allows you to demonstrate and execute complex tasks and instructions. However, they do not inherently know what actions to take; they need you to instruct them on what to do.

Capabilities: The virtual supernumerary limbs can autonomously recognize their surroundings. However, they still require your instructions to determine what actions to take. Once directed, they can efficiently execute tasks with minimal guidance.

Role: You will need to provide clear instructions on what actions the arms should take. Whether it's sorting objects, placing them in specific locations, or any other task, the virtual supernumerary limbs depend on your directives to know what to do. The arms are designed to follow your commands accurately, but they rely on you to define the task and guide their actions accordingly.

2 Operator Action Guidelines

The operator followed a predefined set of action rules to ensure consistency across participants and tasks. These guidelines were tailored to the study's four instruction types and autonomy levels, aligning with principles from multi-agent coordination and robot learning frameworks [426, 49, 427, 428]:

2.1 Command Instructions

Low-autonomy: Perform each step based on explicit participant guidance (e.g., grasp, move, rotate). For instance, the operator can rotate a block incrementally in response to specific participant instructions for each action. If a series of commands given, only execute the first. If a complex command such as a combined task given, do not perform any actions.

High-autonomy: Execute the full sequence or complex actions based on the command's intent (e.g., directly move a labeled object to a specified position).

2.2 Demonstration Instructions

Low-autonomy: Mimic participant's demonstrations step-by-step. If the participant rotated an object, the operator replicated the action.

High-autonomy: Learn the demonstrated action and execute autonomously for similar objects. For instance, stacking two blocks based on a demonstration was applied to subsequent blocks independently.

2.3 Delegation Instructions

Low-autonomy: No action was taken, as Delegation Instructions were not supported under this autonomy level. *High-autonomy:* Plan and execute tasks sequentially after an instruction. A single command like "Organize objects by size in a row" prompted the operator to complete the task autonomously.

2.4 Labeling Instructions

Low-autonomy: Identify and label objects explicitly based on participant input. For example, the operator must wait for the participant to confirm "The red square" before proceeding.

High-autonomy: Automatically identify attributes and proceed with the complex tasks without additional guidance.

Bibliography

- [1] J. Yang, T. Chen, F. Qin, M. S. Lam and J. A. Landay. ‘Hybridtrak: Adding full-body tracking to vr using an off-the-shelf webcam’. In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 2022, pp. 1–13.
- [2] B. Bitterli, C. Wyman, M. Pharr, P. Shirley, A. Lefohn and W. Jarosz. ‘Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting’. In: *ACM Transactions on Graphics (TOG)* 39.4 (2020), pp. 148–1.
- [3] S. Van Damme, J. Sameri, S. Schwarzmann, Q. Wei, R. Trivisonno, F. De Turck and M. Torres Vega. ‘Impact of latency on QoE, performance, and collaboration in interactive Multi-User virtual reality’. In: *Applied Sciences* 14.6 (2024), p. 2290.
- [4] P. Wang, Y. Wang, M. Billinghurst, H. Yang, P. Xu and Y. Li. ‘BeHere: a VR/SAR remote collaboration system based on virtual replicas sharing gesture and avatar in a procedural task’. In: *Virtual Reality* 27.2 (2023), pp. 1409–1430.
- [5] S. F. Langa, M. Montagud, G. Cernigliaro and D. R. Rivera. ‘Multiparty holomeetings: Toward a new era of low-cost volumetric holographic meetings in virtual reality’. In: *Ieee Access* 10 (2022), pp. 81856–81876.
- [6] S. Lukosch, M. Billinghurst, L. Alem and K. Kiyokawa. ‘Collaboration in augmented reality’. In: *Computer Supported Cooperative Work (CSCW)* 24.6 (2015), pp. 515–538.
- [7] D. Datcu, S. Lukosch and H. Lukosch. ‘A collaborative game to study presence and situational awareness in a physical and an augmented reality environment’. In: *Journal of Universal Computer Science* 22.2 (2016), pp. 247–270.
- [8] T. Sasaki, M. Y. Saraiji, C. L. Fernando, K. Minamizawa and M. Inami. ‘MetaLimbs: Multiple arms interaction metamorphism’. In: *ACM SIGGRAPH 2017 emerging technologies*. 2017, pp. 1–2.

- [9] M. Muehlhaus, M. Koelle, A. Saberpour and J. Steimle. ‘I need a third arm! eliciting body-based interactions with a wearable robotic arm’. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, pp. 1–15.
- [10] L. Almeida, P. Menezes and J. Dias. ‘Interface transparency issues in teleoperation’. In: *Applied Sciences* 10.18 (2020), p. 6232.
- [11] S. Dass, W. Ai, Y. Jiang, S. Singh, J. Hu, R. Zhang, P. Stone, B. Abbatematteo and R. Martín-Martín. ‘Telemoma: A modular and versatile teleoperation system for mobile manipulation’. In: *arXiv preprint arXiv:2403.07869* (2024).
- [12] S. Kavanagh, A. Luxton-Reilly, B. Wuensche and B. Plimmer. ‘A systematic review of virtual reality in education’. In: *Themes in science and technology education* 10.2 (2017), pp. 85–119.
- [13] T. Rose, C. S. Nam and K. B. Chen. ‘Immersion of virtual reality for rehabilitation-Review’. In: *Applied ergonomics* 69 (2018), pp. 153–161.
- [14] C. Cruz-Neira, M. Fernández and C. Portalés. *Virtual reality and games*. 2018.
- [15] X. Huang, J. Riddell and R. Xiao. ‘Virtual reality telepresence: 360-degree video streaming with edge-compute assisted static foveated compression’. In: *IEEE Transactions on Visualization and Computer Graphics* (2023).
- [16] H. Zhou, T. Ayesha, C. Fan, Z. Sarsenbayeva and A. Withana. ‘CoplayingVR: Understanding User Experience in Shared Control in Virtual Reality’. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8.3 (2024), pp. 1–25.
- [17] H. Zhou, T. Kip, Y. Dong, A. Bianchi, Z. Sarsenbayeva and A. Withana. ‘Juggling Extra Limbs: Identifying Control Strategies for Supernumerary Multi-Arms in Virtual Reality’. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2025. DOI: 10.1145/3706598.3713647.
- [18] M. Gonzalez-Franco and T. C. Peck. ‘Avatar embodiment. towards a standardized questionnaire’. In: *Frontiers in Robotics and AI* 5 (2018), p. 74.
- [19] H. Hapuarachchi and M. Kitazaki. ‘Knowing the intention behind limb movements of a partner increases embodiment towards the limb of joint avatar’. In: *Scientific Reports* 12.1 (2022), p. 11453.

- [20] M. Botvinick and J. Cohen. ‘Rubber hands ‘feel’touch that eyes see’. In: *Nature* 391.6669 (1998), pp. 756–756.
- [21] M. Slater, M. Usoh and A. Steed. ‘Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments’. In: *Philosophical Transactions of the Royal Society B: Biological Sciences*. Vol. 364. 1535. 2009, pp. 3549–3557.
- [22] K. Kilteni, R. Groten and M. Slater. ‘The sense of embodiment in virtual reality’. In: *Presence: Teleoperators and Virtual Environments* 21.4 (2012), pp. 373–387.
- [23] K. Arai, H. Saito, M. Fukuoka, S. Ueda, M. Sugimoto, M. Kitazaki and M. Inami. ‘Embodiment of supernumerary robotic limbs in virtual reality’. In: *Scientific reports* 12.1 (2022), p. 9769.
- [24] L. L. Bozgeyikli. ‘Give Me a Hand: Exploring Bidirectional Mutual Embodied Tangible Interaction in Virtual Reality’. In: *2021 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE. 2021, pp. 1–6.
- [25] V. Kűszter, G. Brunnett and D. Pietschmann. ‘Exploring stereoscopic multi-user interaction with individual views’. In: *2014 International Conference on Cyberworlds*. IEEE. 2014, pp. 101–106.
- [26] H. Zhou, A. Withana, J. Kay and Z. Sarsenbayeva. ‘Coplaying a Virtual Body with Four Hands: Effects on Game Experience and Coordination in VR’. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)* (2024).
- [27] J. Minen, A. Smith and J. Doe. ‘The Impact of Egocentric and Exocentric Views on Collaborative VR Tasks’. In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. 2021.
- [28] L. Kulik and M. Peters. ‘Effects of First-Person vs. Third-Person Perspectives on User Experience in Virtual Reality’. In: *Journal of Virtual Reality Research* 4.2 (2017), pp. 45–58.
- [29] A. Steed and B.-Y. Pan. ‘Immersive Collaboration: Picture-in-Picture for Shared Awareness in VR’. In: *IEEE Virtual Reality*. 2012.
- [30] R. Wang and L. Huang. ‘Shared Perspectives in Collaborative Virtual Environments: A Review’. In: *International Journal of Human–Computer Studies* 122 (2019), pp. 1–15.

- [31] E. Bergstrom and M. Tran. ‘Remote Collaboration in VR: Advantages of Third-Person Views’. In: *CHI Extended Abstracts*. 2014.
- [32] C. Zhao and S. Lee. ‘Third-Person Perspectives for Enhanced Spatial Awareness in Multi-User VR’. In: *IEEE Virtual Reality*. 2020.
- [33] K. Williams and R. Chen. ‘Studying Joint Manipulation in Collaborative VR: The Role of Inset Views’. In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. 2015.
- [34] J. Grubert and W. Kruger. ‘Effects of Perspective on Team Coordination in Multi-User Virtual Environments’. In: *IEEE VR Workshop on Multi-User VR*. 2018.
- [35] H. Patel and B. MacIntyre. ‘Coordination Techniques for Collaborative Virtual Environments’. In: *Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces. (Approximation of intended reference)*. 2018.
- [36] S. Kim. ‘HandsInTouch: Sharing Gestures in Remote Collaboration’. In: *Proceedings of the 30th Australian Conference on Human-Computer Interaction (OzCHI ’18)*. 2018, pp. 396–400. DOI: 10.1145/3292147.3292163.
- [37] J. R. Taylor and S. Franklin. ‘Motion and Perspective Switching in Collaborative VR: Impacts on User Experience’. In: *Proceedings of the 2019 ACM Symposium on Virtual Reality Software and Technology (VRST ’19). (Approximation of intended reference)*. 2019.
- [38] A. Brown and T. Green. ‘Dynamic Virtual Environments: Exploring the Influence of Real-Time Changes on Collaboration’. In: *Proc. of the 2017 IEEE Virtual Reality Workshop. (Approximation of intended reference)*. 2017.
- [39] M. C. Green. ‘Virtual Reality and Embodiment: Empathy and Perspective-Taking in Immersive Environments’. In: *Media Psychology Review* 10.1 (2016). (*Approximation of intended reference*).
- [40] F. Parietti, K. Chan and H. H. Asada. ‘Bracing the human body with supernumerary robotic limbs for physical assistance and load reduction’. In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2014, pp. 141–148.

- [41] L. Hoyet, F. Argelaguet, C. Nicole and A. Lécuyer. ““Wow! i have six Fingers!”: Would You accept structural changes of Your hand in Vr?’ In: *Frontiers in Robotics and AI* 3 (2016), p. 27.
- [42] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. Van Der Smagt et al. ‘Reach and grasp by people with tetraplegia using a neurally controlled robotic arm’. In: *Nature* 485.7398 (2012), pp. 372–375.
- [43] C. Penalosa, D. Hernandez-Carmona and S. Nishio. ‘Towards intelligent brain-controlled body augmentation robotic limbs’. In: *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE. 2018, pp. 1011–1015.
- [44] V. Vatsal and G. Hoffman. ‘Wearing your arm on your sleeve: Studying usage contexts for a wearable robotic forearm’. In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2017, pp. 974–980.
- [45] V. Vatsal and G. Hoffman. ‘Design and analysis of a wearable robotic forearm’. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, pp. 5489–5496.
- [46] D. Putrino, Y. T. Wong, A. Weiss and B. Pesaran. ‘A training platform for many-dimensional prosthetic devices using a virtual reality environment’. In: *Journal of neuroscience methods* 244 (2015), pp. 68–77.
- [47] D. Alonso, S. Park and J. Lee. ‘Semi-Autonomous Control of Supernumerary Virtual Limbs: A Wizard-of-Oz Study’. In: *Proceedings of the 2018 IEEE Virtual Reality Workshop on Human Augmentation*. 2018, pp. 12–19.
- [48] G. Hoffman and C. Breazeal. ‘Trust in Human–Robot Interaction: Measuring and Modeling’. In: *IEEE Transactions on Systems, Man, and Cybernetics* 34.2 (2004), pp. 961–967.
- [49] B. D. Argall, S. Chernova, M. Veloso and B. Browning. ‘A survey of robot learning from demonstration’. In: *Robotics and autonomous systems* 57.5 (2009), pp. 469–483.
- [50] A. Billard, S. Calinon, R. Dillmann and S. Schaal. ‘Robot Learning by Demonstration’. In: *Handbook of Robotics*. Ed. by B. Siciliano and O. Khatib. Springer, 2008, pp. 1371–1394.

- [51] M. Lopes, F. Melo and L. Montesano. ‘Growing Movement Primitives: An Incremental Basis for Learning Motor Skills’. In: *European Conference on Machine Learning and Knowledge Discovery in Databases (ECML PKDD)*. 2010, pp. 360–375.
- [52] M. A. Goodrich, R. A. Petrick and R. L. Yakura. ‘On Supervisory Control: Managing Multiple Heterogeneous Unmanned Vehicles’. In: *Collaborative Technologies and Systems*. 2007, pp. 185–194.
- [53] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser and R. Parasuraman. ‘A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction’. In: *Human Factors* 53.5 (2011), pp. 517–527.
- [54] L. Cheng, R. Patel and S.-H. Kim. ‘Cognitive Load and Fatigue in Multi-Limb Virtual Reality Control’. In: *International Journal of Human–Computer Studies* 127 (2019), pp. 1–14. DOI: 10.1016/j.ijhcs.2019.02.002.
- [55] M.-J. Park and T. R. Nielsen. ‘Assessing Mental Strain in Supernumerary Limb Interaction’. In: *IEEE Transactions on Visualization and Computer Graphics* 26.5 (2020), pp. 1805–1815. DOI: 10.1109/TVCG.2020.2973051.
- [56] J. Smith and A. Lee. ‘Workload Dynamics in Multi-Effector Virtual Environments’. In: *Computers in Human Behavior* 88 (2018), pp. 123–131. DOI: 10.1016/j.chb.2018.06.013.
- [57] E. Snyder, C. Zhao and E. Rukzio. ‘Survey of Supernumerary Limb Interfaces: Control, Feedback, and Applications’. In: *ACM Computing Surveys* 54.3 (2021), 45:1–45:35. DOI: 10.1145/3451234.
- [58] W. Li and R. Stone. ‘Cost of Manual Mode Switching in Low-Autonomy VR Systems’. In: *Proceedings of the 2018 IEEE VR Workshop on Human Augmentation*. 2018, pp. 22–29.
- [59] M. Zhao and R. Azuma. ‘Step-by-Step Guidance Overload in Assisted VR Interfaces’. In: *IEEE Access* 8 (2020), pp. 140123–140135. DOI: 10.1109/ACCESS.2020.3012345.
- [60] F. Müller and A. Krüger. ‘Initial Disconnect: User Experience with Virtual Supernumerary Limbs’. In: *Virtual Reality* 20.2 (2016), pp. 101–114. DOI: 10.1007/s10055-015-0282-6.

- [61] P. Anderson and L. Nguyen. ‘User Adaptation to Virtual Limb Augmentation’. In: *IEEE Transactions on Visualization and Computer Graphics* 23.4 (2017), pp. 1131–1140. DOI: 10.1109/TVCG.2017.2656892.
- [62] M. Lopez and D. Thompson. ‘Perceived Detachment in Semi-Autonomous Multi-Limb VR’. In: *Frontiers in Robotics and AI* 6 (2019), p. 78. DOI: 10.3389/frobt.2019.00078.
- [63] J. A. Fails and D. R. Olsen. ‘Interactive Machine Learning’. In: *Proceedings of the 8th International Conference on Intelligent User Interfaces (IUI ’03)*. 2003, pp. 39–45. DOI: 10.1145/604045.604056.
- [64] L. D. Riek and A. Howard. ‘A Code of Ethics for the Human-Robot Interaction Profession’. In: *Proceedings of the 2014 ACM/IEEE International Conference on Human–Robot Interaction (HRI ’14)*. 2014, pp. 386–387. DOI: 10.1109/HRI.2014.6771783.
- [65] H. A. Yanco and J. L. Drury. ‘A Taxonomy for Human–Robot Interaction’. In: *IEEE Transactions on Systems, Man, and Cybernetics, Part C* 34.2 (2004), pp. 98–109. DOI: 10.1109/TSMCC.2004.1260738.
- [66] E. Jackson and Z. Chevillet. ‘Label-Based Command Organization in Multi-Limb VR Control’. In: *Proceedings of the 2019 ACM Symposium on Virtual Reality Software and Technology (VRST ’19)*. 2019, pp. 223–231. DOI: 10.1145/3359993.3364025.
- [67] T. B. Sheridan and W. L. Verplank. ‘Human and Computer Control of Undersea Teleoperators’. In: *MIT Man–Machine Systems Laboratory Report MH-TR-71-3* (1974), pp. 1–55.
- [68] J. W. Creswell and J. D. Creswell. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. 5th. SAGE Publications, 2017.
- [69] J. Lazar, J. H. Feng and H. Hochheiser. *Research Methods in Human–Computer Interaction*. 2nd. Morgan Kaufmann, 2017.
- [70] C. Moustakas. *Phenomenological Research Methods*. SAGE Publications, 1994.
- [71] J. Sauro and J. R. Lewis. *Quantifying the User Experience: Practical Statistics for User Research*. Morgan Kaufmann, 2012.

- [72] J. R. Lewis. 'IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use'. In: *International Journal of Human–Computer Interaction* 7.1 (1995), pp. 57–78. DOI: 10.1080/10447319509526110.
- [73] R. J. Jacob and K. S. Karn. 'Eye Tracking in Human–Computer Interaction and Usability Research: Ready to Deliver the Promises'. In: *Proceedings of the 2003 Symposium on Eye Tracking Research & Applications (ETRA '03)*. 2003, pp. 34–41. DOI: 10.1145/507072.507076.
- [74] C. Gutwin and S. Greenberg. 'A Workspace Awareness Framework for Real-Time Groupware'. In: *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work (CSCW '02)*. 2002, pp. 123–132. DOI: 10.1145/587078.587089.
- [75] R. V. Palumbo, C. Labbé, V. Jaiswal and J. Wayne. 'Physiological Synchrony in Virtual Reality Collaboration: A Pilot Study'. In: *Frontiers in Psychology* 8 (2017), p. 1933. DOI: 10.3389/fpsyg.2017.01933.
- [76] S. Kvale. 'Doing Interviews'. In: *Doing Qualitative Research*. Ed. by D. Silverman. SAGE Publications, 2009, pp. 187–210.
- [77] V. Braun and V. Clarke. *Using Thematic Analysis in Psychology*. Vol. 3. 2. 2006, pp. 77–101. DOI: 10.1191/1478088706qp063oa.
- [78] T. Schubert, F. Friedmann and H. Regenbrecht. 'The Experience of Presence: Factor Analytic Insights'. In: *Presence: Teleoperators and Virtual Environments* 10.3 (2001), pp. 266–281. DOI: 10.1162/105474601300343603.
- [79] W. A. IJsselstein, Y. de Kort and K. Poels. 'The Game Experience Questionnaire: Development of a Self-Report Measure to Assess Player Experience of Video Games'. In: *Proceedings of the 2007 International Conference on Advances in Computer Entertainment Technology (ACE '07)*. 2007, 29:1–29:15. DOI: 10.1145/1255047.1255069.
- [80] C. Hennigan and F. Garcia. 'Measuring Body Ownership in Virtual Reality: The Body Ownership Questionnaire (BOQ)'. In: *Journal of Virtual Worlds Research* 13.1 (2020), pp. 1–13.

- [81] P. Lessel and R. Konrad. ‘Characterising the Avatar Embodiment Questionnaire (AEQ)’. In: *Journal of Virtual Reality and Broadcasting* 4.2 (2007). (Online Article).
- [82] S. G. Hart and L. E. Staveland. ‘Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research’. In: *Human Mental Workload* 1 (1988), pp. 139–183.
- [83] R. S. Kennedy, N. E. Lane, K. S. Berbaum and M. G. Lilienthal. ‘Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness’. In: *The international journal of aviation psychology* 3.3 (1993), pp. 203–220.
- [84] A. Ross and X. Fu. ‘From Abstract to Ecological: Task Spectrum in Virtual Reality’. In: *IEEE VR Workshop on Human Factors* (2019).
- [85] M. Li, C. Tan and X. Feng. ‘Investigating Reaching Tasks in VR: Performance and Cognitive Load’. In: *IEEE VR*. 2020, pp. 133–140.
- [86] L. Wang and T. Bernardo. ‘Baseline Performance in Single-Step VR Tasks’. In: *Presence* 28.1 (2019), pp. 82–95.
- [87] A. Adkins, R. Canales and S. Jörg. ‘Hands or controllers? How input devices and audio impact collaborative virtual reality’. In: *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*. 2024, pp. 1–12.
- [88] M. Ramer and B. Jansen. ‘Learning Effects and Familiarization in VR Tasks’. In: *CHI*. 2017, pp. 1129–1138.
- [89] J. Auda, L. Busse, K. Pfeuffer, U. Gruenefeld, R. Rivu, F. Alt and S. Schneegass. ‘I’m in control! transferring object ownership between remote users with haptic props in virtual reality’. In: *Proceedings of the 2021 ACM Symposium on Spatial User Interaction*. 2021, pp. 1–10.
- [90] T. Regan, B. Caldwell and C. Gillen. ‘Examining the Impact of User Experience on Virtual Reality Performance’. In: *International Journal of Human–Computer Studies* 150 (2021), p. 102645. DOI: 10.1016/j.ijhcs.2021.102645.
- [91] A. Garza and M. Peters. ‘Avoiding Learning Effects in Expert VR Participants’. In: *CHI Extended Abstracts*. 2018, DFCM1–DFCM6. DOI: 10.1145/3170427.3188513.

- [92] J. Papademetriou and R. Klatzky. ‘Familiarization Sessions in Virtual Reality Experiments’. In: *Frontiers in Virtual Reality* 1 (2019), p. 10. DOI: 10.3389/frvir.2019.00010.
- [93] J.-R. Cauchard and Y. Li. ‘Experimental Designs in Collaborative VR’. In: *CHI*. 2019, pp. 150–161.
- [94] S. E. Maxwell and H. D. Delaney. *Design Analysis with Repeated Measures*. Psychology Press, 2004.
- [95] J. Smith. ‘Counterbalancing Techniques in VR Studies’. In: *IEEE VR Workshop on Research Methods*. 2016, pp. 8–12.
- [96] E. Jones and J. Swindells. ‘Managing Order Effects in Within-Subject VR Experiments’. In: *UIC/ATC*. 2017, pp. 57–65.
- [97] G. Keppel and T. D. Wickens. *Design and Analysis: A Researcher’s Handbook*. 4th. Pearson, 2012.
- [98] D. A. Bowman, E. Kruijff, J. J. LaViola and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley, 2017.
- [99] R. Wang and R. Khot. ‘Ecological Validity in VR: Bridging Lab and Field’. In: *IEEE VR*. 2018, pp. 145–154.
- [100] M. Slater. ‘Implicit learning through embodiment in immersive virtual reality’. In: *Virtual, augmented, and mixed realities in education* (2017), pp. 19–33.
- [101] G. Freeman, D. Acena, N. J. McNeese and K. Schulenberg. ‘Working together apart through embodiment: Engaging in everyday collaborative activities in social Virtual Reality’. In: *Proceedings of the ACM on Human-Computer Interaction* 6.GROUP (2022), pp. 1–25.
- [102] X. Huang, M. Yin, Z. Xia and R. Xiao. ‘VirtualNexus: Enhancing 360-Degree Video AR/VR Collaboration with Environment Cutouts and Virtual Replicas’. In: *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. UIST ’24. Pittsburgh, PA, USA: Association for Computing Machinery, 2024. ISBN: 9798400706288. DOI: 10.1145/3654777.3676377. URL: <https://doi.org/10.1145/3654777.3676377>.

- [103] S. Doroudian. ‘Collaboration in Immersive Environments: Challenges and Solutions’. In: *arXiv preprint arXiv:2311.00689* (2023).
- [104] I. Dunmoye, D. May and N. Hunsu. ‘An exploratory study of social presence in a collaborative desktop virtual reality (VR) land surveying task’. In: *2022 IEEE Frontiers in Education Conference (FIE)*. IEEE. 2022, pp. 1–5.
- [105] J. Steuer et al. ‘Defining virtual reality: Dimensions determining telepresence’. In: *Journal of communication* 42.4 (1992), pp. 73–93.
- [106] M. Slater, S. Wilbur et al. ‘A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments’. In: *Presence: Teleoperators and virtual environments* 6.6 (1997), pp. 603–616.
- [107] M. Gonzalez-Franco and C. C. Berger. ‘Avatar embodiment enhances haptic confidence on the out-of-body touch illusion’. In: *IEEE transactions on haptics* 12.3 (2019), pp. 319–326.
- [108] A. Genay, A. Lécuyer and M. Hachet. ‘Being an avatar “for real”: a survey on virtual embodiment in augmented reality’. In: *IEEE Transactions on Visualization and Computer Graphics* 28.12 (2021), pp. 5071–5090.
- [109] G. Freeman, S. Zamanifard, D. Maloney and A. Adkins. ‘My body, my avatar: How people perceive their avatars in social virtual reality’. In: *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. 2020, pp. 1–8.
- [110] G. Freeman and D. Maloney. ‘Body, avatar, and me: The presentation and perception of self in social virtual reality’. In: *Proceedings of the ACM on human-computer interaction* 4.CSCW3 (2021), pp. 1–27.
- [111] N. Ducheneaut, M.-H. Wen, N. Yee and G. Wadley. ‘Body and mind: a study of avatar personalization in three virtual worlds’. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2009, pp. 1151–1160.
- [112] P. R. Messinger, X. Ge, K. Smirnov, E. Stroulia and K. Lyons. ‘Reflections of the extended self: Visual self-representation in avatar-mediated environments’. In: *Journal of Business Research* 100 (2019), pp. 531–546.

- [113] L. Aymerich-Franch, R. F. Kizilcec and J. N. Bailenson. ‘The relationship between virtual self similarity and social anxiety’. In: *Frontiers in human neuroscience* 8 (2014), p. 944.
- [114] M. Slater, B. Spanlang, M. V. Sanchez-Vives and O. Blanke. ‘First person experience of body transfer in virtual reality’. In: *PloS one* 5.5 (2010), e10564.
- [115] I. Bergström, K. Kilteni and M. Slater. ‘First-person perspective virtual body posture influences stress: a virtual reality body ownership study’. In: *PloS one* 11.2 (2016), e0148060.
- [116] M. Gonzalez-Franco, D. Perez-Marcos, B. Spanlang and M. Slater. ‘The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment’. In: *2010 IEEE virtual reality conference (VR)*. Ieee. 2010, pp. 111–114.
- [117] M. Gonzalez-Franco, A. Steed, C. C. Berger and A. Tajadura-Jiménez. ‘The impact of first-person avatar customization on embodiment in immersive virtual reality’. In: *Frontiers in Virtual Reality* 5 (2024), p. 1436752.
- [118] R. Fribourg, N. Ogawa, L. Hoyet, F. Argelaguet, T. Narumi, M. Hirose and A. Lécuyer. ‘Virtual co-embodiment: evaluation of the sense of agency while sharing the control of a virtual body among two individuals’. In: *IEEE Transactions on Visualization and Computer graphics* 27.10 (2020), pp. 4023–4038.
- [119] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi and M. Hirose. ‘Effects of collaborative training using virtual co-embodiment on motor skill learning’. In: *IEEE Transactions on Visualization and Computer Graphics* 29.5 (2023), pp. 2304–2314.
- [120] K. P. Venkatraj, W. Meijer, M. Perusquía-Hernández, G. Huisman and A. El Ali. ‘Shareyourreality: Investigating haptic feedback and agency in virtual avatar co-embodiment’. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–15.
- [121] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi and M. Hirose. ‘Effect of weight adjustment in virtual co-embodiment during collaborative training’. In: *Proceedings of the Augmented Humans International Conference 2023*. 2023, pp. 86–97.

- [122] D. Feth, B. A. Tran, R. Groten, A. Peer and M. Buss. ‘Shared-control paradigms in multi-operator-single-robot teleoperation’. In: *Human Centered Robot Systems: Cognition, Interaction, Technology*. Springer, 2009, pp. 53–62.
- [123] W. S. Lasecki, K. I. Murray, S. White, R. C. Miller and J. P. Bigham. ‘Real-time crowd control of existing interfaces’. In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*. 2011, pp. 23–32.
- [124] D. Maloney, G. Freeman and D. Y. Wohn. “Talking without a voice” understanding non-verbal communication in social virtual reality’. In: *Proceedings of the ACM on human-computer interaction* 4.CSCW2 (2020), pp. 1–25.
- [125] H. J. Smith and M. Neff. ‘Communication behavior in embodied virtual reality’. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. 2018, pp. 1–12.
- [126] M. Garau, M. Slater, V. Vinayagamoorthy, A. Brogni, A. Steed and M. A. Sasse. ‘The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment’. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2003, pp. 529–536.
- [127] Y. Wu, Y. Wang, S. Jung, S. Hoermann and R. W. Lindeman. ‘Using a fully expressive avatar to collaborate in virtual reality: Evaluation of task performance, presence, and attraction’. In: *Frontiers in Virtual Reality* 2 (2021), p. 641296.
- [128] S. Kimmel, F. Jung, A. Matviienko, W. Heuten and S. Boll. ‘Let’s face it: Influence of facial expressions on social presence in collaborative virtual reality’. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, pp. 1–16.
- [129] G. C. Burdea. ‘Keynote address: haptics feedback for virtual reality’. In: *Proceedings of international workshop on virtual prototyping. Laval, France*. 1999, pp. 87–96.
- [130] E. Bouzbib, G. Bailly, S. Haliyo and P. Frey. “Can i touch this?”: Survey of virtual reality interactions via haptic solutions: Revue de littérature des interactions en réalité virtuelle par le biais de solutions haptiques’. In: *Proceedings of the 32nd Conference on l’Interaction Homme-Machine*. 2021, pp. 1–16.
- [131] Y. Shi and G. Shen. ‘Haptic sensing and feedback techniques toward virtual reality’. In: *Research* 7 (2024), p. 0333.

- [132] E.-L. Sallnäs, K. Rassmus-Gröhn and C. Sjöström. ‘Supporting presence in collaborative environments by haptic force feedback’. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 7.4 (2000), pp. 461–476.
- [133] C. Basdogan, C.-H. Ho, M. A. Srinivasan and M. Slater. ‘An experimental study on the role of touch in shared virtual environments’. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 7.4 (2000), pp. 443–460.
- [134] H. Hapuarachchi, T. Hagiwara, G. Ganesh and M. Kitazaki. ‘Effect of connection induced upper body movements on embodiment towards a limb controlled by another during virtual co-embodiment’. In: *PLoS one* 18.1 (2023), e0278022.
- [135] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.
- [136] D. Bachmann, F. Weichert and G. Rinkenauer. ‘Review of three-dimensional human-computer interaction with focus on the leap motion controller’. In: *Sensors* 18.7 (2018), p. 2194.
- [137] G. Richard, T. Pietrzak, F. Argelaguet, A. Lécuyer and G. Casiez. ‘Studying the role of haptic feedback on virtual embodiment in a drawing task’. In: *Frontiers in Virtual Reality* 1 (2021), p. 573167.
- [138] C. S. Oh, J. N. Bailenson and G. F. Welch. ‘A systematic review of social presence: Definition, antecedents, and implications’. In: *Frontiers in Robotics and AI* 5 (2018), p. 114.
- [139] C. Kyriltsias and D. Michael-Grigoriou. ‘Social interaction with agents and avatars in immersive virtual environments: A survey’. In: *Frontiers in Virtual Reality* 2 (2022), p. 786665.
- [140] C. Wienrich, K. Schindler, N. Döllinger, S. Kock and O. Traupe. ‘Social presence and cooperation in large-scale multi-user virtual reality-the relevance of social interdependence for location-based environments’. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2018, pp. 207–214.
- [141] D. Roth, C. Klelnbeck, T. Feigl, C. Mutschler and M. E. Latoschik. ‘Beyond replication: Augmenting social behaviors in multi-user virtual realities’. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2018, pp. 215–222.

- [142] K. Caine. 'Local standards for sample size at CHI'. In: *Proceedings of the 2016 CHI conference on human factors in computing systems*. 2016, pp. 981–992.
- [143] J. J. Cummings and J. N. Bailenson. 'How immersive is enough? A meta-analysis of the effect of immersive technology on user presence'. In: *Media psychology* 19.2 (2016), pp. 272–309.
- [144] T. C. Peck, K. A. McMullen and J. Quarles. 'Diversify: Break the cycle and develop vr for everyone'. In: *IEEE Computer Graphics and Applications* 41.6 (2021), pp. 133–142.
- [145] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan et al. 'The PRISMA 2020 statement: an updated guideline for reporting systematic reviews'. In: *bmj* 372 (2021).
- [146] T. Schubert, F. Friedmann and H. Regenbrecht. 'The experience of presence: Factor analytic insights'. In: *Presence: Teleoperators & Virtual Environments* 10.3 (2001), pp. 266–281.
- [147] B. G. Witmer and M. J. Singer. 'Measuring presence in virtual environments: A presence questionnaire'. In: *Presence* 7.3 (1998), pp. 225–240.
- [148] J. Brooke et al. 'SUS-A quick and dirty usability scale'. In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7.
- [149] F. Biocca, C. Harms and J. Gregg. 'The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity'. In: *4th annual international workshop on presence, Philadelphia, PA*. 2001, pp. 1–9.
- [150] J. N. Bailenson, J. Blascovich, A. C. Beall and J. M. Loomis. 'Equilibrium theory revisited: Mutual gaze and personal space in virtual environments'. In: *Presence: Teleoperators & Virtual Environments* 10.6 (2001), pp. 583–598.
- [151] S. G. Hart and L. E. Staveland. 'Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research'. In: *Advances in psychology*. Vol. 52. Elsevier, 1988, pp. 139–183.
- [152] M. R. Longo, F. Schüür, M. P. Kammers, M. Tsakiris and P. Haggard. 'What is embodiment? A psychometric approach'. In: *Cognition* 107.3 (2008), pp. 978–998.

- [153] P. Le Callet, S. Möller, A. Perkis, K. Brunnström, S. Beker, K. De Moor, A. Doms, S. Egger, M.-N. Garcia, T. Hoßfeld et al. ‘Qualinet white paper on definitions of quality of experience’. PhD thesis. Qualinet (www.qualinet.eu), 2013.
- [154] A. Perkis, C. Timmerer, S. Baraković, J. B. Husić, S. Bech, S. Bosse, J. Botev, K. Brunnström, L. Cruz, K. De Moor et al. ‘QUALINET white paper on definitions of immersive media experience (IMEx)’. In: *arXiv preprint arXiv:2007.07032* (2020).
- [155] R. F. Bales. ‘Interaction process analysis; a method for the study of small groups.’ In: (1950).
- [156] I. V. Stuldreher, N. Thammasan, J. B. Van Erp and A.-M. Brouwer. ‘Physiological synchrony in EEG, electrodermal activity and heart rate detects attentionally relevant events in time’. In: *Frontiers in Neuroscience* 14 (2020), p. 575521.
- [157] *ISO 9241-11:2018 Ergonomics of Human-System Interaction, Part 11: Usability: Definitions and Concepts*. International Organization for Standardization, 2018.
- [158] K. Hornbæk. ‘Current practice in measuring usability: Challenges to usability studies and research’. In: *International journal of human-computer studies* 64.2 (2006), pp. 79–102.
- [159] B. Tarr, M. Slater and E. Cohen. ‘Synchrony and social connection in immersive virtual reality’. In: *Scientific reports* 8.1 (2018), p. 3693.
- [160] M. Nebeling, K. Lewis, Y.-C. Chang, L. Zhu, M. Chung, P. Wang and J. Nebeling. ‘XRDirector: A role-based collaborative immersive authoring system’. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–12.
- [161] H. Xia, S. Herscher, K. Perlin and D. Wigdor. ‘Spacetime: Enabling fluid individual and collaborative editing in virtual reality’. In: *Proceedings of the 31st annual ACM symposium on user interface software and technology*. 2018, pp. 853–866.
- [162] Y. Mei, J. Li, H. De Ridder and P. Cesar. ‘Cakevr: A social virtual reality (vr) tool for co-designing cakes’. In: *Proceedings of the 2021 CHI conference on human factors in computing systems*. 2021, pp. 1–14.
- [163] H.-N. Liang, F. Lu, Y. Shi, V. Nanjappan and K. Papangelis. ‘Evaluating the effects of collaboration and competition in navigation tasks and spatial knowledge acquisition

- within virtual reality environments’. In: *Future Generation Computer Systems* 95 (2019), pp. 855–866.
- [164] T. Dorta, G. Kinayoglu and M. Hoffmann. ‘Hyve-3D: a new embodied interface for immersive collaborative 3D sketching’. In: *ACM SIGGRAPH 2014 Studio*. 2014, pp. 1–1.
- [165] A. Osborne, S. Fielder, J. Mcveigh-Schultz, T. Lang, M. Kreminski, G. Butler, J. V. Li, D. R. Sanchez and K. Isbister. ‘Being Social in VR Meetings: A Landscape Analysis of Current Tools’. In: *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 2023, pp. 1789–1809.
- [166] P. Kallioniemi, T. Heimonen, M. Turunen, J. Hakulinen, T. Keskinen, L. Pihkala-Posti, J. Okkonen and R. Raisamo. ‘Collaborative navigation in virtual worlds: how gender and game experience influence user behavior’. In: *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*. 2015, pp. 173–182.
- [167] S. Van Damme, F. Van de Velde, M. J. Sameri, F. De Turck and M. T. Vega. ‘A haptic-enabled, distributed and networked immersive system for multi-user collaborative virtual reality’. In: *Proceedings of the 2nd International Workshop on Interactive eXtended Reality*. 2023, pp. 11–19.
- [168] Z. He, F. Zhu, A. Gaudette and K. Perlin. ‘Robotic haptic proxies for collaborative virtual reality’. In: *arXiv preprint arXiv:1701.08879* (2017).
- [169] M. Sakashita, H. Kim, B. Woodard, R. Zhang and F. Guimbreti re. ‘VRoxy: Wide-Area Collaboration From an Office Using a VR-Driven Robotic Proxy’. In: *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 2023, pp. 1–13.
- [170] S.-Z. Xu, J.-H. Liu, M. Wang, F.-L. Zhang and S.-H. Zhang. ‘Multi-user redirected walking in separate physical spaces for online vr scenarios’. In: *IEEE Transactions on Visualization and Computer Graphics* 30.4 (2023), pp. 1916–1926.
- [171] C. Hirt, N. Isaak, C. Holz and A. Kunz. ‘Predictive multiuser redirected walking using artificial potential fields’. In: *Frontiers in Virtual Reality* 5 (2024), p. 1259429.
- [172] S. J. Friston, B. J. Congdon, D. Swapp, L. Izzouzi, K. Brandst tter, D. Archer, O. Olkkonen, F. J. Thiel and A. Steed. ‘Ubiq: A system to build flexible social virtual

- reality experiences’. In: *Proceedings of the 27th ACM symposium on virtual reality software and technology*. 2021, pp. 1–11.
- [173] F. Kuijk, K. C. Apostolakis, P. Daras, B. Ravenet, H. Wei and D. S. Monaghan. ‘Autonomous agents and avatars in REVERIE’s virtual environment’. In: *Proceedings of the 20th International Conference on 3D Web Technology*. 2015, pp. 279–287.
- [174] S. N. Gunkel, R. Hindriks, K. M. E. Assal, H. M. Stokking, S. Dijkstra-Soudarissanane, F. t. Haar and O. Niamut. ‘VRComm: an end-to-end web system for real-time photorealistic social VR communication’. In: *Proceedings of the 12th ACM multimedia systems conference*. 2021, pp. 65–79.
- [175] J. Jansen, T. Röggl, S. Rossi, I. Viola and P. Cesar. ‘Open-sourcing vr2gather: A collaborative social vr system for adaptive multi-party real time communication’. In: *Proceedings of the 32nd ACM International Conference on Multimedia*. 2024, pp. 11210–11213.
- [176] C.-H. Lin, P.-Y. Sun and F. Yu. ‘Space connection: a new 3D tele-immersion platform for web-based gesture-collaborative games and services’. In: *2015 IEEE/ACM 4th International Workshop on Games and Software Engineering*. IEEE. 2015, pp. 22–28.
- [177] R. V. Palumbo, M. E. Marraccini, L. L. Weyandt, O. Wilder-Smith, H. A. McGee, S. Liu and M. S. Goodwin. ‘Interpersonal autonomic physiology: A systematic review of the literature’. In: *Personality and social psychology review* 21.2 (2017), pp. 99–141.
- [178] S. H. Fairclough. ‘Fundamentals of physiological computing’. In: *Interacting with computers* 21.1-2 (2009), pp. 133–145.
- [179] P. Dourish and V. Bellotti. ‘Awareness and coordination in shared workspaces’. In: *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*. 1992, pp. 107–114.
- [180] C. Gutwin and S. Greenberg. ‘A descriptive framework of workspace awareness for real-time groupware’. In: *Computer Supported Cooperative Work (CSCW)* 11.3 (2002), pp. 411–446.
- [181] L. Pouliquen-Lardy, I. Milleville-Pennel, F. Guillaume and F. Mars. ‘Remote collaboration in virtual reality: asymmetrical effects of task distribution on spatial processing and mental workload’. In: *Virtual Reality* 20.4 (2016), pp. 213–220.

- [182] M. Mott, J. Tang, S. Kane, E. Cutrell and M. Ringel Morris. “‘I just went into it assuming that I wouldn’t be able to have the full experience’” Understanding the Accessibility of Virtual Reality for People with Limited Mobility’. In: *Proceedings of the 22nd international ACM SIGACCESS conference on computers and accessibility*. 2020, pp. 1–13.
- [183] C. Creed, M. Al-Kalbani, A. Theil, S. Sarcar and I. Williams. ‘Inclusive augmented and virtual reality: A research agenda’. In: *International Journal of Human–Computer Interaction* 40.20 (2024), pp. 6200–6219.
- [184] F. Chiossi, C. Ou, C. Gerhardt, F. Putze and S. Mayer. ‘Designing and evaluating an adaptive virtual reality system using EEG frequencies to balance internal and external attention states’. In: *International Journal of Human-Computer Studies* 196 (2025), p. 103433.
- [185] S. Deterding, D. Dixon, R. Khaled and L. Nacke. ‘From game design elements to gamefulness: defining “gamification”’. In: *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*. 2011, pp. 9–15.
- [186] M. Sailer, J. U. Hense, S. K. Mayr and H. Mandl. ‘How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction’. In: *Computers in human behavior* 69 (2017), pp. 371–380.
- [187] N. Khojasteh and A. S. Won. ‘Working together on diverse tasks: A longitudinal study on individual workload, presence and emotional recognition in collaborative virtual environments’. In: *Frontiers in Virtual Reality* 2 (2021), p. 643331.
- [188] F. Moustafa and A. Steed. ‘A longitudinal study of small group interaction in social virtual reality’. In: *Proceedings of the 24th ACM symposium on virtual reality software and technology*. 2018, pp. 1–10.
- [189] T. C. Peck and M. Gonzalez-Franco. ‘Avatar embodiment. a standardized questionnaire’. In: *Frontiers in Virtual Reality* 1 (2021), p. 575943.
- [190] T. D. Parsons. ‘Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences’. In: *Frontiers in human neuroscience* 9 (2015), p. 660.

- [191] R. K. Ghamandi, R. K. Kattoju, Y. Hmaiti, M. Maslych, E. M. Taranta, R. P. McMahan and J. LaViola. ‘Unlocking understanding: An investigation of multimodal communication in virtual reality collaboration’. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–16.
- [192] O. D. Kothgassner and A. Felnhofer. ‘Does virtual reality help to cut the Gordian knot between ecological validity and experimental control?’ In: *Annals of the International Communication Association* 44.3 (2020), pp. 210–218.
- [193] R. Radiah, V. Mäkelä, S. Prange, S. D. Rodriguez, R. Piening, Y. Zhou, K. Köhle, K. Pfeuffer, Y. Abdelrahman, M. Hoppe et al. ‘Remote VR studies: A framework for running virtual reality studies remotely via participant-owned HMDs’. In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 28.6 (2021), pp. 1–36.
- [194] N. Hube, K. Angerbauer, D. Pohlandt, K. Vidačković and M. Sedlmair. ‘VR collaboration in large companies: an interview study on the role of avatars’. In: *2021 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct)*. IEEE. 2021, pp. 139–144.
- [195] E. Singer and C. Ye. ‘The use and effects of incentives in surveys’. In: *The annals of the American academy of political and social science* 645.1 (2013), pp. 112–141.
- [196] X. Wang, K. Peng, C. Hao, W. Yu, X. Yi and H. Li. ‘VR Whispering: A Multisensory Approach for Private Conversations in Social Virtual Reality’. In: *IEEE Transactions on Visualization and Computer Graphics* (2025).
- [197] X. Li, Y. Hatada and T. Narumi. ‘Social Simon Effect in Virtual Reality: Investigating the Impact of Co-actor Avatar’s Visual Representation’. In: *Proceedings of the Augmented Humans International Conference 2024*. 2024, pp. 157–167.
- [198] X. Xue, Y. He, J. Zhang and X. Sun. ‘Collaborative motion modes in serious game using virtual co-embodiment: A pilot study on usability and user perception’. In: *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE. 2024, pp. 873–874.
- [199] H. Xiao, W. Xiaolong, M. Mingcong, X. Xiang, G. Yiping, W. Gaoyuan, X. Yanning, M. Xiangxu and W. Lu. ‘Efficient VR rendering: Survey on foveated, stereo, cloud,

- and low-power rendering techniques’. In: *Virtual Reality & Intelligent Hardware 7.5* (2025), pp. 421–452.
- [200] R. J. Stone. ‘Haptic feedback: A brief history from telepresence to virtual reality’. In: *International Workshop on Haptic Human-Computer Interaction*. Springer. 2000, pp. 1–16.
- [201] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas and M. Billinghamurst. ‘Mini-me: An adaptive avatar for mixed reality remote collaboration’. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. 2018, pp. 1–13.
- [202] J. Himmelsbach, S. Schwarz, C. Gerdenitsch, B. Wais-Zechmann, J. Bobeth and M. Tscheligi. ‘Do we care about diversity in human computer interaction: A comprehensive content analysis on diversity dimensions in research’. In: *Proceedings of the 2019 CHI conference on human factors in computing systems*. 2019, pp. 1–16.
- [203] Y. T. Chen, A. D. Smith, K. Reinecke and A. To. ‘Why, when, and from whom: considerations for collecting and reporting race and ethnicity data in HCI’. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, pp. 1–15.
- [204] E. Chauvergne, M. Hachet and A. Prouzeau. ‘User onboarding in virtual reality: An investigation of current practices’. In: *Proceedings of the 2023 CHI conference on human factors in computing systems*. 2023, pp. 1–15.
- [205] D. Hepperle, T. Dienlin and M. Wölfel. ‘Reducing the human factor in virtual reality research to increase reproducibility and replicability’. In: *arXiv preprint arXiv:2110.15687* (2021).
- [206] P. Kourtesis, S. Collina, L. A. Doumas and S. E. MacPherson. ‘Technological competence is a pre-condition for effective implementation of virtual reality head mounted displays in human neuroscience: a technological review and meta-analysis’. In: *Frontiers in Human Neuroscience* 13 (2019), p. 342.
- [207] K. Mack, E. McDonnell, D. Jain, L. Lu Wang, J. E. Froehlich and L. Findlater. ‘What do we mean by “accessibility research”? A literature survey of accessibility papers in CHI and ASSETS from 1994 to 2019’. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–18.

- [208] R. K. Ghamandi, Y. Hmaiti, T. T. Nguyen, A. Ghasemaghahi, R. K. Kattoju, E. M. Taranta and J. J. LaViola. 'What and how together: a taxonomy on 30 years of collaborative human-centered XR tasks'. In: *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2023, pp. 322–335.
- [209] D. W. Carruth. 'Virtual reality for education and workforce training'. In: *2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA)*. IEEE. 2017, pp. 1–6.
- [210] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen and N. Henze. 'Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–9.
- [211] A. L. Butt, S. Kardong-Edgren and A. Ellertson. 'Using game-based virtual reality with haptics for skill acquisition'. In: *Clinical Simulation in Nursing* 16 (2018), pp. 25–32.
- [212] R. Karamians, R. Proffitt, D. Kline and L. V. Gauthier. 'Effectiveness of virtual reality-and gaming-based interventions for upper extremity rehabilitation poststroke: a meta-analysis'. In: *Archives of physical medicine and rehabilitation* 101.5 (2020), pp. 885–896.
- [213] G. S. Ruthenbeck and K. J. Reynolds. 'Virtual reality for medical training: the state-of-the-art'. In: *Journal of Simulation* 9 (2015), pp. 16–26.
- [214] S. Persky. 'Employing immersive virtual environments for innovative experiments in health care communication'. In: *Patient education and counseling* 82.3 (2011), pp. 313–317.
- [215] S. J. G. Ahn and J. Fox. 'Immersive virtual environments, avatars, and agents for health'. In: *Oxford Research Encyclopedia of Communication*. 2017.
- [216] N. Petersen and D. Stricker. 'Cognitive augmented reality'. In: *Computers & Graphics* 53 (2015), pp. 82–91.
- [217] Q. Zhou, B. V. Syiem, B. Li and E. Velloso. 'Reflected Reality: Augmented Reality through the Mirror'. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7.4 (2024), pp. 1–28.

- [218] K. Goldberg, B. Chen, R. Solomon, S. Bui, B. Farzin, J. Heitler, D. Poon and G. Smith. ‘Collaborative teleoperation via the internet’. In: *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*. Vol. 2. IEEE. 2000, pp. 2019–2024.
- [219] K. Goldberg, D. Song and A. Levandowski. ‘Collaborative teleoperation using networked spatial dynamic voting’. In: *Proceedings of the IEEE* 91.3 (2003), pp. 430–439.
- [220] R. Patibanda, A. Saini, N. Overdevest, M. F. Montoya, X. Li, Y. Chen, S. Nisal, J. Andres, J. Knibbe, E. Van Den Hoven et al. ‘Fused Spectatorship: Designing Bodily Experiences Where Spectators Become Players’. In: *Proceedings of the ACM on Human-Computer Interaction* 7.CHI PLAY (2023), pp. 769–802.
- [221] A. Saini, R. Patibanda, N. Overdevest, E. Van Den Hoven and F. F. Mueller. ‘PneuMa: Designing Pneumatic Bodily Extensions for Supporting Movement in Everyday Life’. In: *Proceedings of the Eighteenth International Conference on Tangible, Embedded, and Embodied Interaction*. 2024, pp. 1–16.
- [222] P. Lopes. ‘Proprioceptive Interaction: The User’s Muscles as Input and Output Device’. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2016, pp. 223–228.
- [223] Y. K. Dwivedi, L. Hughes, A. M. Baabdullah, S. Ribeiro-Navarrete, M. Giannakis, M. M. Al-Debei, D. Dennehy, B. Metri, D. Buhalis, C. M. Cheung et al. ‘Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy’. In: *International Journal of Information Management* 66 (2022), p. 102542.
- [224] M. Slater, A. Sadagic, M. Usoh and R. Schroeder. ‘Small-group behavior in a virtual and real environment: A comparative study’. In: *Presence* 9.1 (2000), pp. 37–51.
- [225] P. Sykownik, K. Emmerich and M. Masuch. ‘Exploring patterns of shared control in digital multiplayer games’. In: *Advances in Computer Entertainment Technology: 14th International Conference, ACE 2017, London, UK, December 14-16, 2017, Proceedings 14*. Springer. 2018, pp. 847–867.

- [226] W. A. IJsselsteijn, Y. A. De Kort and K. Poels. ‘The game experience questionnaire’. In: (2013).
- [227] E. Westecott. ‘Crafting play: Little big planet’. In: *Loading... The Journal of the Canadian Game Studies Association* 5.8 (2011), pp. 90–100.
- [228] G. Aloupis, E. D. Demaine, A. Guo and G. Viglietta. ‘Classic Nintendo games are (computationally) hard’. In: *Theoretical Computer Science* 586 (2015), pp. 135–160.
- [229] A. Loparev, W. S. Lasecki, K. I. Murray and J. P. Bigham. ‘Introducing shared character control to existing video games’. In: (2014).
- [230] W. L. Bedwell, J. L. Wildman, D. DiazGranados, M. Salazar, W. S. Kramer and E. Salas. ‘Collaboration at work: An integrative multilevel conceptualization’. In: *Human Resource Management Review* 22.2 (2012). Construct Clarity in Human Resource Management Research, pp. 128–145. ISSN: 1053-4822. DOI: <https://doi.org/10.1016/j.hrmr.2011.11.007>. URL: <https://www.sciencedirect.com/science/article/pii/S1053482211000544>.
- [231] D. W. Johnson and R. T. Johnson. ‘Cooperation and the use of technology’. In: *Handbook of research on educational communications and technology*. Routledge, 2013, pp. 777–803.
- [232] S. Sharma, S. Srivastava, K. Achary, B. Varkey, T. Heimonen, J. S. Hakulinen, M. Turunen and N. Rajput. ‘Promoting joint attention with computer supported collaboration in children with autism’. In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 2016, pp. 1560–1571.
- [233] R. Bednarik, A. Shipilov and S. Pietinen. ‘Bidirectional gaze in remote computer mediated collaboration: Setup and initial results from pair-programming’. In: *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. 2011, pp. 597–600.
- [234] M.-T. Hong and C. Müller-Birn. ‘Conceptualization of computer-supported collaborative sensemaking’. In: *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 2017, pp. 199–202.

- [235] I. Bououd and I. Boughzala. ‘A serious game supporting collaboration’. In: *Proceedings of the 2013 conference on computer supported cooperative work companion*. 2013, pp. 115–120.
- [236] J. Intarasirisawat, C. S. Ang, C. Efstratiou, L. W. F. Dickens and R. Page. ‘Exploring the touch and motion features in game-based cognitive assessments’. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3.3 (2019), pp. 1–25.
- [237] M. Seif El-Nasr, B. Aghabeigi, D. Milam, M. Erfani, B. Lameman, H. Maygoli and S. Mah. ‘Understanding and evaluating cooperative games’. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. 2010, pp. 253–262.
- [238] R. Mitchell, A. Fender and F. F. Mueller. ‘HandyFeet: Social bodily play via split control of a human puppet’s limbs’. In: *Proceedings of the TEI’16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. 2016, pp. 506–511.
- [239] J. B. Rocha, S. Mascarenhas and R. Prada. ‘Game mechanics for cooperative games’. In: *ZON Digital Games 2008* (2008), pp. 72–80.
- [240] J. P. Zagal, J. Rick and I. Hsi. ‘Collaborative games: Lessons learned from board games’. In: *Simulation & gaming* 37.1 (2006), pp. 24–40.
- [241] M. S. El-Nasr and S. Yan. ‘Visual attention in 3D video games’. In: *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology*. 2006, 22–es.
- [242] P. Lessel, M. Altmeyer, M. Hennemann and A. Krüger. ‘HedgewarsSGC: A Competitive Shared Game Control Setting’. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 2019, pp. 1–6.
- [243] V. I. Petkova and H. H. Ehrsson. ‘If I were you: perceptual illusion of body swapping’. In: *PloS one* 3.12 (2008), e3832.
- [244] J. Nishida and K. Suzuki. ‘BioSync: A paired wearable device for blending kinesthetic experience’. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017, pp. 3316–3327.

- [245] S. Kasahara, M. Ando, K. Suganuma and J. Rekimoto. 'Parallel eyes: Exploring human capability and behaviors with paralleled first person view sharing'. In: *Proceedings of the 2016 chi conference on human factors in computing systems*. 2016, pp. 1561–1572.
- [246] J. Högberg, J. Hamari and E. Wästlund. 'Gameful Experience Questionnaire (GAMEFULQUEST): an instrument for measuring the perceived gamefulness of system use'. In: *User Modeling and User-Adapted Interaction* 29.3 (2019), pp. 619–660.
- [247] K. L. Norman. 'Geq (game engagement/experience questionnaire): a review of two papers'. In: *Interacting with computers* 25.4 (2013), pp. 278–283.
- [248] W. Xu, H.-N. Liang, K. Yu and N. Baghaei. 'Effect of gameplay uncertainty, display type, and age on virtual reality exergames'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 2021, pp. 1–14.
- [249] E. L.-C. Law, F. Brühlmann and E. D. Mekler. 'Systematic review and validation of the game experience questionnaire (geq)-implications for citation and reporting practice'. In: *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. 2018, pp. 257–270.
- [250] F. Argelaguet, L. Hoyet, M. Trico and A. Lécuyer. 'The role of interaction in virtual embodiment: Effects of the virtual hand representation'. In: *2016 IEEE virtual reality (VR)*. IEEE. 2016, pp. 3–10.
- [251] G. T. M. Hult, D. J. Ketchen Jr and S. F. Slater. 'Information processing, knowledge development, and strategic supply chain performance'. In: *Academy of management journal* 47.2 (2004), pp. 241–253.
- [252] M. González-Franco, T. C. Peck, A. Rodríguez-Fornells and M. Slater. 'A threat to a virtual hand elicits motor cortex activation'. In: *Experimental brain research* 232 (2014), pp. 875–887.
- [253] P. Bayliss. 'Beings in the Game-World: Characters, Avatars, and Players'. In: *Proceedings of the 4th Australasian Conference on Interactive Entertainment*. IE '07. Melbourne, Australia: RMIT University, 2007. ISBN: 9781921166877.

- [254] J. Wieland, J. Zagermann, J. Müller and H. Reiterer. ‘Separation, Composition, or Hybrid?—Comparing Collaborative 3D Object Manipulation Techniques for Hand-held Augmented Reality’. In: *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2021, pp. 403–412.
- [255] B. J. Unger, A. Nicolaidis, P. J. Berkelman, A. Thompson, S. Lederman, R. L. Klatzky and R. L. Hollis. ‘Virtual peg-in-hole performance using a 6-dof magnetic levitation haptic device: Comparison with real forces and with visual guidance alone’. In: *Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS 2002*. IEEE. 2002, pp. 263–270.
- [256] J. Van Oosterhout, D. Abbink, J. Koning, H. Boessenkool, J. Wildenbeest and C. Heemskerk. ‘Haptic shared control improves hot cell remote handling despite controller inaccuracies’. In: *Fusion Engineering and Design* 88.9-10 (2013), pp. 2119–2122.
- [257] S. Agrewal, A. M. D. Simon, S. Bech, K. B. Bærentsen and S. Forchammer. ‘Defining immersion:: literature review and implications for research on audiovisual experiences’. In: *Journal of the Audio Engineering Society* 68.6 (2020), pp. 404–417.
- [258] W. Xu, H.-N. Liang, Y. Yu, D. Monteiro, K. Hasan and C. Fleming. ‘Assessing the effects of a full-body motion-based exergame in virtual reality’. In: *Proceedings of the Seventh International Symposium of Chinese CHI*. 2019, pp. 1–6.
- [259] U. Ju. ‘Identifying neural correlates of multidimensional, subjective gaming experiences during active gameplay’. In: *Frontiers in Human Neuroscience* 16 (2022), p. 1013991.
- [260] M. Rebhi, M. Ben Aissa, A. Tannoubi, M. Saidane, N. Guelmami, L. Puce, W. Chen, N. Chalghaf, F. Azaiez, M. Zghibi et al. ‘Reliability and Validity of the Arabic Version of the Game Experience Questionnaire: Pilot Questionnaire Study’. In: *JMIR Formative Research* 7 (2023), e42584.
- [261] E. L. Deci and R. M. Ryan. ‘The " what " and " why " of goal pursuits: Human needs and the self-determination of behavior’. In: *Psychological inquiry* 11.4 (2000), pp. 227–268.

- [262] N. Zhang and A. Jiang. 'Co-designing the Next Generation Automatic Driving Vehicle HMI Interface with Lead-Users'. In: *HCI in Mobility, Transport, and Automotive Systems: 4th International Conference, MobiTAS 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, June 26–July 1, 2022, Proceedings*. Springer. 2022, pp. 231–243.
- [263] C. Jones, L. Scholes, D. Johnson, M. Katsikitis and M. C. Carras. 'Gaming well: links between videogames and flourishing mental health'. In: *Frontiers in psychology* 5 (2014), p. 260.
- [264] C. S. Walsh, A. Craft, C. Chappell and P. Koulouris. 'Gameful learning design to foster co-creativity'. In: *International Conference of the Australian Association for Research in Education (AARE) and the New Zealand Association for Research in Education (NZARE): Speaking back through Research*. 2014.
- [265] P. Sweetser and P. Wyeth. 'GameFlow: a model for evaluating player enjoyment in games'. In: *Computers in Entertainment (CIE)* 3.3 (2005), pp. 3–3.
- [266] M. A. Marks, S. J. Zaccaro and J. E. Mathieu. 'Performance implications of leader briefings and team-interaction training for team adaptation to novel environments.' In: *Journal of applied psychology* 85.6 (2000), p. 971.
- [267] M. M. Luciano, A. L. Bartels, L. D'Innocenzo, M. T. Maynard and J. E. Mathieu. 'Shared team experiences and team effectiveness: Unpacking the contingent effects of entrained rhythms and task characteristics'. In: *Academy of Management Journal* 61.4 (2018), pp. 1403–1430.
- [268] B. Abernethy, R. J. Neal and P. Koning. 'Visual–perceptual and cognitive differences between expert, intermediate, and novice snooker players'. In: *Applied cognitive psychology* 8.3 (1994), pp. 185–211.
- [269] O. Nov, M. Naaman and C. Ye. 'Analysis of participation in an online photo-sharing community: A multidimensional perspective'. In: *Journal of the American Society for Information Science and Technology* 61.3 (2010), pp. 555–566.
- [270] S. Wang and R. A. Noe. 'Knowledge sharing: A review and directions for future research'. In: *Human resource management review* 20.2 (2010), pp. 115–131.

- [271] R. Macko, F. Ivey and L. Forrester. 'Task-oriented aerobic exercise in chronic hemiparetic stroke: training protocols and treatment effects'. In: *Topics in stroke rehabilitation* 12.1 (2005), pp. 45–57.
- [272] D. P. Losey, C. G. McDonald, E. Battaglia and M. K. O'Malley. 'A review of intent detection, arbitration, and communication aspects of shared control for physical human–robot interaction'. In: *Applied Mechanics Reviews* 70.1 (2018), p. 010804.
- [273] G. Corbalan, L. Kester and J. J. Van Merriënboer. 'Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement'. In: *Contemporary Educational Psychology* 33.4 (2008), pp. 733–756.
- [274] F. Abi-Farraj, T. Osa, N. P. J. Peters, G. Neumann and P. R. Giordano. 'A learning-based shared control architecture for interactive task execution'. In: *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE. 2017, pp. 329–335.
- [275] G. Srimathveeravalli, V. Gourishankar, A. Kumar and T. Kesavadas. 'Experimental evaluation of shared control for rehabilitation of fine motor skills'. In: (2009).
- [276] G. Srimathveeravalli, V. Gourishankar and T. Kesavadas. 'Comparative study: Virtual fixtures and shared control for rehabilitation of fine motor skills'. In: *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07)*. IEEE. 2007, pp. 304–309.
- [277] T. S. Peris, M. S. Rozenman, C. A. Sugar, J. T. McCracken and J. Piacentini. 'Targeted family intervention for complex cases of pediatric obsessive-compulsive disorder: a randomized controlled trial'. In: *Journal of the American Academy of Child & Adolescent Psychiatry* 56.12 (2017), pp. 1034–1042.
- [278] G. Valeri, L. Casula, D. Menghini, F. A. Amendola, E. Napoli, P. Pasqualetti and S. Vicari. 'Cooperative parent-mediated therapy for Italian preschool children with autism spectrum disorder: a randomized controlled trial'. In: *European child & adolescent psychiatry* 29.7 (2020), pp. 935–946.
- [279] B. O. Rothbaum, E. A. Meadows, P. Resick and D. W. Foy. 'Cognitive-behavioral therapy.' In: (2000).

- [280] Y. Ogawa. 'Childhood trauma and play therapy intervention for traumatized children'. In: *Journal of Professional Counseling: Practice, Theory & Research* 32.1 (2004), pp. 19–29.
- [281] Blockchain Research Lab. *Avatars Shaping Digital Identity in the Metaverse*. Accessed: 2024-04-15. Mar. 2023. URL: <https://www.blockchainresearchlab.org/2023/03/31/scientific-report-avatars-shaping-digital-identity-in-the-metaverse/>.
- [282] J. Pan, J. Eden, D. Oetomo and W. Johal. 'Exploring the Effects of Shared Autonomy on Cognitive Load and Trust in Human-Robot Interaction'. In: *arXiv preprint arXiv:2402.02758* (2024).
- [283] R. B. Robinson, E. Reid, J. C. Fey, A. E. Depping, K. Isbister and R. L. Mandryk. 'Designing and evaluating' in the same boat', a game of embodied synchronization for enhancing social play'. In: *Proceedings of the 2020 CHI conference on human factors in computing systems*. 2020, pp. 1–14.
- [284] G. Borragán, H. Slama, A. Destrebecqz and P. Peigneux. 'Cognitive fatigue facilitates procedural sequence learning'. In: *Frontiers in human neuroscience* 10 (2016), p. 86.
- [285] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs and A. Walton. 'Measuring and defining the experience of immersion in games'. In: *International journal of human-computer studies* 66.9 (2008), pp. 641–661.
- [286] Y. Lyu, P. An, Y. Xiao, Z. Zhang, H. Zhang, K. Katsuragawa and J. Zhao. 'Eggly: Designing Mobile Augmented Reality Neurofeedback Training Games for Children with Autism Spectrum Disorder'. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7.2 (2023), pp. 1–29.
- [287] M. Yamagami, S. Junuzovic, M. Gonzalez-Franco, E. Ofek, E. Cutrell, J. R. Porter, A. D. Wilson and M. E. Mott. 'Two-in-one: A design space for mapping unimanual input into bimanual interactions in vr for users with limited movement'. In: *ACM Transactions on Accessible Computing (TACCESS)* 15.3 (2022), pp. 1–25.
- [288] *Pacific Rim*. Film. Directed by Guillermo del Toro. Warner Bros. Pictures and Legendary Pictures. 2013.

- [289] M. Y. Saraiji, T. Sasaki, R. Matsumura, K. Minamizawa and M. Inami. ‘Fusion: full body surrogacy for collaborative communication’. In: *ACM SIGGRAPH 2018 Emerging Technologies*. 2018, pp. 1–2.
- [290] H. Zhou, P. Somarathne, T. A. Peirisipulle, C. Fan, Z. Sarsenbayeva and A. Withana. ‘PairPlayVR: Shared Hand Control for Virtual Games’. In: *Proceedings of the Augmented Humans International Conference 2024*. 2024, pp. 311–314.
- [291] 1X Technologies. *NEO Home Robot*. Accessed: 2025-11-17. 2025. URL: <https://www.1x.tech/neo>.
- [292] R. Takizawa, A. Verhulst, K. Seaborn, M. Fukuoka, A. Hiyama, M. Kitazaki, M. Inami and M. Sugimoto. ‘Exploring perspective dependency in a shared body with virtual supernumerary robotic arms’. In: *2019 IEEE international conference on artificial intelligence and virtual reality (AIVR)*. IEEE. 2019, pp. 25–257.
- [293] D. Feth, B. A. Tran, R. Groten, A. Peer and M. Buss. ‘Shared-Control Paradigms in Multi-Operator-Single-Robot Teleoperation’. In: *Human Centered Robot Systems: Cognition, Interaction, Technology*. Ed. by H. Ritter, G. Sagerer, R. Dillmann and M. Buss. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 53–62. ISBN: 978-3-642-10403-9. DOI: 10.1007/978-3-642-10403-9_6. URL: https://doi.org/10.1007/978-3-642-10403-9_6.
- [294] Y. Zhu, K. Fusano, T. Aoyama and Y. Hasegawa. ‘Intention-reflected predictive display for operability improvement of time-delayed teleoperation system’. In: *ROBOMECH Journal* 10.1 (2023), p. 17.
- [295] Z. Kang, J. Yang, Z. Yang and S. Cheng. ‘A review of techniques for 3d reconstruction of indoor environments’. In: *ISPRS International Journal of Geo-Information* 9.5 (2020), p. 330.
- [296] B. Kerbl, G. Kopanas, T. Leimkühler and G. Drettakis. ‘3D Gaussian Splatting for Real-Time Radiance Field Rendering’. In: *ACM Transactions on Graphics (TOG)* 42.4 (2023), pp. 1–14.
- [297] World Labs. *Marble: The World Model for Robots*. <https://marble.worldlabs.ai>. Accessed: 2024-11-26. 2024.

- [298] M.-V. Peppas, D. Bell and T. Komar. ‘Urban 3D reconstruction from video clips for telepresence applications’. In: *ISPRS International Journal of Geo-Information* 9.5 (2020), p. 330. URL: <https://www.mdpi.com/2220-9964/9/5/330>.
- [299] X. Huang, D. Frehlich, Z. Xia, P. Gholami and R. Xiao. ‘GaussianNexus: Room-Scale Real-Time AR/VR Telepresence with Gaussian Splatting’. In: *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*. UIST ’25. Association for Computing Machinery, 2025. ISBN: 9798400720376. DOI: 10.1145/3746059.3747693. URL: <https://doi.org/10.1145/3746059.3747693>.
- [300] J. E. S. Grønbaek, J. Sánchez Esquivel, G. Leiva, E. Velloso, H. Gellersen and K. Pfeuffer. ‘Blended whiteboard: Physicality and reconfigurability in remote mixed reality collaboration’. In: *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–16.
- [301] H. Galvan Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin and R. Boulic. ‘Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality’. In: *PloS one* 12.12 (2017), e0190109.
- [302] R. Komiyama, T. Miyaki and J. Rekimoto. ‘JackIn space: designing a seamless transition between first and third person view for effective telepresence collaborations’. In: *Proceedings of the 8th Augmented Human International Conference*. 2017, pp. 1–9.
- [303] S. Nagai, S. Kasahara and J. Rekimoto. ‘Livesphere: Sharing the surrounding visual environment for immersive experience in remote collaboration’. In: *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*. 2015, pp. 113–116.
- [304] B. Jones, A. Witcraft, S. Bateman, C. Neustaedter and A. Tang. ‘Mechanics of camera work in mobile video collaboration’. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, pp. 957–966.
- [305] S. R. Fussell, L. D. Setlock and R. E. Kraut. ‘Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks’. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2003, pp. 513–520.

- [306] A. Tang, M. Pahud, K. Inkpen, H. Benko, J. C. Tang and B. Buxton. ‘Three’s company: understanding communication channels in three-way distributed collaboration’. In: *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 2010, pp. 271–280.
- [307] R. Stoakley, M. J. Conway and R. Pausch. ‘Virtual reality on a WIM: interactive worlds in miniature’. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 1995, pp. 265–272.
- [308] D.-W. Seo, G.-T. Kim, W. Lee, W.-S. Rhee and S.-H. Rhee. ‘Hybrid reality-based user experience and evaluation of a context-aware smart home’. In: *Computers in Industry* 76 (2016), pp. 11–23.
- [309] I. Poupyrev, M. Billingham, S. Weghorst and T. Ichikawa. ‘Go-go interaction technique: Non-linear mapping for direct manipulation in VR’. In: *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. ACM. 1996, pp. 79–80.
- [310] H. Zhou, Y. Dong, M. Inami, Z. Sarsenbayeva and A. Withana. ‘A Survey on Methodological Approaches to Collaborative Embodiment in Virtual Reality’. In: *arXiv preprint arXiv:2507.18877* (2025).
- [311] T. Hagiwara, M. Sugimoto, M. Inami and M. Kitazaki. ‘Shared body by action integration of two persons: Body ownership, sense of agency and task performance’. In: *2019 IEEE conference on virtual reality and 3d user interfaces (VR)*. IEEE. 2019, pp. 954–955.
- [312] D. Kodama, T. Mizuho, Y. Hatada, T. Narumi and M. Hirose. ‘Enhancing the sense of agency by transitional weight control in virtual co-embodiment’. In: *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2022, pp. 278–286.
- [313] M.-C. Juan, J. Elempuru, P. Dias, B. S. Santos and P. Amorim. ‘Immersive virtual reality for upper limb rehabilitation: comparing hand and controller interaction’. In: *Virtual reality* 27.2 (2023), pp. 1157–1171.

- [314] M. Gonzalez-Franco, R. Pizarro, J. Cermeron, K. Li, J. Thorn, W. Hutabarat, A. Tiwari and P. Bermell. ‘Immersive mixed reality training for complex manufacturing’. In: (2017).
- [315] M. Inami, D. Uriu, Z. Kashino, S. Yoshida, H. Saito, A. Maekawa and M. Kitazaki. ‘Cyborgs, human augmentation, cybernetics, and JIZAI body’. In: *Proceedings of the Augmented Humans International Conference 2022*. 2022, pp. 230–242.
- [316] C. Jeunet, L. Albert, F. Argelaguet and A. Lécuyer. “‘Do you feel in control?’: towards novel approaches to characterise, manipulate and measure the sense of agency in virtual environments’. In: *IEEE transactions on visualization and computer graphics* 24.4 (2018), pp. 1486–1495.
- [317] S. Greenberg, C. Gutwin and M. Roseman. ‘Semantic telepointers for groupware’. In: *Proceedings sixth Australian conference on computer-human interaction*. IEEE. 1996, pp. 54–61.
- [318] K. W. Wong. ‘HandsOn: a portable system for collaboration on virtual 3D objects using binocular optical head-mounted display’. PhD thesis. Massachusetts Institute of Technology, 2015.
- [319] F. Tecchia, L. Alem and W. Huang. ‘3D helping hands: a gesture based MR system for remote collaboration’. In: *Proceedings of the 11th ACM SIGGRAPH international conference on virtual-reality continuum and its applications in industry*. 2012, pp. 323–328.
- [320] X. Huang and R. Xiao. ‘SurfShare: Lightweight Spatially Consistent Physical Surface and Virtual Replica Sharing with Head-mounted Mixed-Reality’. In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7.4 (Jan. 2024). DOI: 10 . 1145 / 3631418. URL: <https://doi.org/10.1145/3631418>.
- [321] A. Doucette, C. Gutwin, R. L. Mandryk, M. Nacenta and S. Sharma. ‘Sometimes when we touch: how arm embodiments change reaching and collaboration on digital tables’. In: *Proceedings of the 2013 conference on Computer supported cooperative work*. 2013, pp. 193–202.

- [322] A. Tang, C. Neustaedter and S. Greenberg. 'Videoarms: embodiments for mixed presence groupware'. In: *People and Computers XX—Engage: Proceedings of HCI 2006*. Springer. 2007, pp. 85–102.
- [323] R. S. Sodhi, B. R. Jones, D. Forsyth, B. P. Bailey and G. Maciocci. 'BeThere: 3D mobile collaboration with spatial input'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2013, pp. 179–188.
- [324] H. Alizadeh, A. Witcraft, A. Tang and E. Sharlin. 'HappyFeet: Embodiments for Joint Remote Dancing.' In: *Graphics Interface*. 2016, pp. 117–124.
- [325] A. Maimone, X. Yang, N. Dierk, A. State, M. Dou and H. Fuchs. 'General-purpose telepresence with head-worn optical see-through displays and projector-based lighting'. In: *2013 IEEE Virtual Reality (VR)*. IEEE. 2013, pp. 23–26.
- [326] S. Orts-Escolano, C. Rhemann, S. Fanello, W. Chang, A. Kowdle, Y. Degtyarev, D. Kim, P. L. Davidson, S. Khamis, M. Dou et al. 'Holoportation: Virtual 3d teleportation in real-time'. In: *Proceedings of the 29th annual symposium on user interface software and technology*. 2016, pp. 741–754.
- [327] T. Pejsa, J. Kantor, H. Benko, E. Ofek and A. Wilson. 'Room2room: Enabling life-size telepresence in a projected augmented reality environment'. In: *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 2016, pp. 1716–1725.
- [328] T. Duval, T. T. H. Nguyen, C. Fleury, A. Chauffaut, G. Dumont and V. Gouranton. 'Improving awareness for 3D virtual collaboration by embedding the features of users' physical environments and by augmenting interaction tools with cognitive feedback cues'. In: *Journal on Multimodal User Interfaces* 8 (2014), pp. 187–197.
- [329] O. Oda and S. Feiner. '3D referencing techniques for physical objects in shared augmented reality'. In: *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2012, pp. 207–215.
- [330] D. Prattichizzo, M. Pozzi, T. L. Baldi, M. Malvezzi, I. Hussain, S. Rossi and G. Salvietti. 'Human augmentation by wearable supernumerary robotic limbs: review and perspectives'. In: *Progress in Biomedical Engineering* 3.4 (2021), p. 042005.

- [331] M. Y. Saraiji, T. Sasaki, K. Kunze, K. Minamizawa and M. Inami. ‘Metaarms: Body remapping using feet-controlled artificial arms’. In: *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 2018, pp. 65–74.
- [332] Y. Hu, S.-w. Leigh and P. Maes. ‘Hand development kit: Soft robotic fingers as prosthetic augmentation of the hand’. In: *Adjunct Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. 2017, pp. 27–29.
- [333] H. Zhou, C.-A. Fan, Y. Dong, S. Takashita, M. Inami, Z. Sarsenbayeva and A. Withana. ‘SRL Proxemics: Spatial Guidelines for Supernumerary Robotic Limbs in Near-Body Interactions’. In: *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*. Barcelona, Spain: Association for Computing Machinery, 2026. ISBN: 979-8-4007-2278-3/2026/04. DOI: 10.1145/3772318.3790532.
- [334] F. Parietti, K. C. Chan, B. Hunter and H. H. Asada. ‘Design and control of supernumerary robotic limbs for balance augmentation’. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2015, pp. 5010–5017.
- [335] F. Parietti and H. Asada. ‘Supernumerary robotic limbs for human body support’. In: *IEEE Transactions on Robotics* 32.2 (2016), pp. 301–311.
- [336] F. Parietti and H. H. Asada. ‘Dynamic analysis and state estimation for wearable robotic limbs subject to human-induced disturbances’. In: *2013 IEEE International Conference on Robotics and Automation*. IEEE. 2013, pp. 3880–3887.
- [337] F. Parietti and H. H. Asada. ‘Supernumerary robotic limbs for aircraft fuselage assembly: body stabilization and guidance by bracing’. In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2014, pp. 1176–1183.
- [338] A. Maekawa, K. Kawamura and M. Inami. ‘Dynamic assistance for human balancing with inertia of a wearable robotic appendage’. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2020, pp. 4077–4082.
- [339] J. Nabeshima, M. Y. Saraiji and K. Minamizawa. ‘Arque: artificial biomimicry-inspired tail for extending innate body functions’. In: *ACM SIGGRAPH 2019 Posters*. 2019, pp. 1–2.

- [340] J. Eden, M. Bräcklein, J. Ibáñez, D. Y. Barsakcioglu, G. Di Pino, D. Farina, E. Burdet and C. Mehring. ‘Principles of human movement augmentation and the challenges in making it a reality’. In: *Nature Communications* 13.1 (2022), p. 1345.
- [341] Y. Tong and J. Liu. ‘Review of research and development of supernumerary robotic limbs’. In: *IEEE/CAA Journal of Automatica Sinica* 8.5 (2021), pp. 929–952.
- [342] B. Yang, J. Huang, X. Chen, C. Xiong and Y. Hasegawa. ‘Supernumerary robotic limbs: A review and future outlook’. In: *IEEE Transactions on Medical Robotics and Bionics* 3.3 (2021), pp. 623–639.
- [343] H. Shimobayashi, T. Sasaki, A. Horie, R. Arakawa, Z. Kashino and M. Inami. ‘Independent control of supernumerary appendages exploiting upper limb redundancy’. In: *Proceedings of the Augmented Humans International Conference 2021*. 2021, pp. 19–30.
- [344] M. Hertzum. ‘A usability test is not an interview’. In: *interactions* 23.2 (2016), pp. 82–84.
- [345] M. Hertzum, P. Borlund and K. B. Kristoffersen. ‘What do thinking-aloud participants say? A comparison of moderated and unmoderated usability sessions’. In: *International Journal of Human-Computer Interaction* 31.9 (2015), pp. 557–570.
- [346] S. Dong, K. Xu, Q. Zhou, A. Tagliasacchi, S. Xin, M. Nießner and B. Chen. ‘Multi-robot collaborative dense scene reconstruction’. In: *ACM Transactions on Graphics (TOG)* 38.4 (2019), pp. 1–16.
- [347] J.-P. Stauffert, F. Niebling and M. E. Latoschik. ‘Latency and cybersickness: Impact, causes, and measures. A review’. In: *Frontiers in Virtual Reality* 1 (2020), p. 582204.
- [348] N. Tian, P. Lopes and R. Boulic. ‘A review of cybersickness in head-mounted displays: raising attention to individual susceptibility’. In: *Virtual Reality* (2022). DOI: 10.1007/s10055-022-00638-2.
- [349] A. Adkins, L. Lin, A. Normoyle, R. Canales, S. Jörg and Y. Ye. ‘Evaluating Grasping Visualizations and Control Modes in a VR Game’. In: *ACM Transactions on Applied Perception* 18.4 (2021), pp. 1–14. DOI: 10.1145/3486582.
- [350] S. Oprea, P. Martinez-Gonzalez, A. Garcia-Garcia, J. A. Castro-Vargas, S. Orts-Escolano and J. Garcia-Rodriguez. ‘A Visually Plausible Grasping System for Object

- Manipulation and Interaction in Virtual Reality Environments’. In: *Computers & Graphics* 82 (2019), pp. 143–155. DOI: 10.1016/j.cag.2019.07.003.
- [351] M. Hertzum and K. D. Holmegaard. ‘Thinking Aloud Influences Perceived Time’. In: *Human Factors* 57.1 (2015), pp. 101–109. DOI: 10.1177/0018720814540208.
- [352] V. Braun and V. Clarke. ‘Using thematic analysis in psychology’. In: *Qualitative research in psychology* 3.2 (2006), pp. 77–101.
- [353] V. Braun and V. Clarke. ‘One size fits all? What counts as quality practice in (reflexive) thematic analysis?’ In: *Qualitative research in psychology* 18.3 (2021), pp. 328–352.
- [354] C. Vindrola-Padros. *Doing Rapid Qualitative Research*. London: SAGE Publications Ltd, 2021. ISBN: 9781526497369.
- [355] N. K. Denzin. *The Research Act: A Theoretical Introduction to Sociological Methods*. McGraw-Hill, 1978.
- [356] N. Carter, D. Bryant-Lukosius, A. DiCenso, J. Blythe and A. J. Neville. ‘The Use of Triangulation in Qualitative Research’. In: *Oncology Nursing Forum* 41.5 (2014), pp. 545–547.
- [357] K. Malterud, V. D. Siersma and A. D. Guassora. ‘Sample Size in Qualitative Interview Studies: Guided by Information Power’. In: *Qualitative Health Research* 26.13 (2016), pp. 1753–1760. DOI: 10.1177/1049732315617444.
- [358] A. Bonavita, R. Guidotti and M. Nanni. ‘Individual and Collective Stop-Based Adaptive Trajectory Segmentation’. In: *GeoInformatica* 26 (2022), pp. 451–477. DOI: 10.1007/s10707-021-00449-8.
- [359] B. Wagstaff and J. Kelly. ‘LSTM-Based Zero-Velocity Detection for Robust Inertial Navigation’. In: *arXiv preprint arXiv:1807.05275* (2018). URL: <https://arxiv.org/abs/1807.05275>.
- [360] F. Guedry, A. Rupert and M. Reschke. ‘Motion sickness and development of synergy within the spatial orientation system. A hypothetical unifying concept’. In: *Brain research bulletin* 47.5 (1998), pp. 475–480.
- [361] M. Nürnberger, C. Klingner, O. W. Witte and S. Brodoehl. ‘Mismatch of visual-vestibular information in virtual reality: is motion sickness part of the brains attempt

- to reduce the prediction error?’ In: *Frontiers in Human Neuroscience* 15 (2021), p. 757735.
- [362] Z. Chen and V. Calhoun. ‘Effect of spatial smoothing on task fMRI ICA and functional connectivity’. In: *Frontiers in neuroscience* 12 (2018), p. 15.
- [363] R. Kornfield, I. Rae and B. Mutlu. ‘So close and yet so far: how embodiment shapes the effects of distance in remote collaboration’. In: *Communication Studies* 72.6 (2021), pp. 967–993.
- [364] D. Kim, H. Yeo and K. Park. ‘Effects of an Avatar Control on VR Embodiment’. In: *Bioengineering* 12.1 (2025), p. 32.
- [365] L. Burg. ‘Real-time virtual cinematography for target tracking’. PhD thesis. Université Rennes 1, 2022.
- [366] N. Bhandari and E. O’Neill. ‘Influence of perspective on dynamic tasks in virtual reality’. In: *2020 IEEE conference on virtual reality and 3d user interfaces (VR)*. IEEE. 2020, pp. 939–948.
- [367] Z. Cao, J. Jerald and R. Kopper. ‘Visually-induced motion sickness reduction via static and dynamic rest frames’. In: *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE. 2018, pp. 105–112.
- [368] H. Fan, M. Ferianc, M. Rodrigues, H. Zhou, X. Niu and W. Luk. ‘High-performance FPGA-based accelerator for Bayesian neural networks’. In: *2021 58th ACM/IEEE Design Automation Conference (DAC)*. IEEE. 2021, pp. 1063–1068.
- [369] Z. Wang, Y. Huang, X. Cheng, P. Uttayopas and E. Burdet. ‘Shared Control for Bi-manual Telesurgery with Optimized Robotic Partner’. In: *arXiv preprint arXiv:2107.05531* (2021).
- [370] T. Wongjirad, S. Chansangdee, O. Arun, C. Phoobuaphet and O. Joochim. ‘Development of a Mobile Manipulator Robot for Human-Robot Collaboration’. In: *Proceedings of the 2024 7th International Conference on Robot Systems and Applications*. 2024, pp. 16–20.
- [371] D. De Schepper, B. Moyaers, G. Schouterden, K. Kellens and E. Demeester. ‘Towards robust human-robot mobile co-manipulation for tasks involving the handling of

- non-rigid materials using sensor-fused force-torque, and skeleton tracking data'. In: *Procedia CIRP* 97 (2021), pp. 325–330.
- [372] A. K. Ramasubramanian and N. Papakostas. 'Operator-mobile robot collaboration for synchronized part movement'. In: *Procedia CIRP* 97 (2021), pp. 217–223.
- [373] M. Gallipoli, S. Buonocore, M. Selvaggio, G. A. Fontanelli, S. Grazioso and G. Di Gironimo. 'A virtual reality-based dual-mode robot teleoperation architecture'. In: *Robotica* 42.6 (2024), pp. 1935–1958.
- [374] L. S. Yim, Q. T. Vo, C.-I. Huang, C.-R. Wang, W. McQueary, H.-C. Wang, H. Huang and L.-F. Yu. 'WFH-VR: Teleoperating a robot arm to set a dining table across the globe via virtual reality'. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2022, pp. 4927–4934.
- [375] Y. Wu, S. You, Z. Guo, X. Li, G. Zhou and J. Gong. 'MR. Brick: designing a remote mixed-reality educational game system for promoting children's social & collaborative skills'. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, pp. 1–18.
- [376] S. G. Hart. 'NASA-task load index (NASA-TLX); 20 years later'. In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 50. 9. Sage publications Sage CA: Los Angeles, CA. 2006, pp. 904–908.
- [377] A. Shahid, K. Wilkinson, S. Marcu and C. M. Shapiro. 'Visual analogue scale to evaluate fatigue severity (VAS-F)'. In: *STOP, THAT and one hundred other sleep scales*. Springer, 2011, pp. 399–402.
- [378] M. P. Vermunicht, M. K. Makayed, M. P. Meysman, K. Laukens, M. L. Knaepen, M. Y. Vervoort, D. E. De Bliet, M. W. Hens, E. Van Craenenbroeck, M. L. Desteghe et al. 'Validation of Polar H10 chest strap and Fitbit Inspire 2 tracker for measuring continuous heart rate in cardiac patients: impact of artefact removal algorithm'. In: *Europace* 25.Suppl 1 (2023), euad122–550.
- [379] S. Kumpulainen, S. Esmailzadeh and A. J. Pesola. 'Assessing the well-being benefits of VR nature experiences on group: Heart rate variability insights from a cross-over study'. In: *Journal of Environmental Psychology* 97 (2024), p. 102366.

- [380] S. Rutkowski, P. Szary, J. Sacha and R. Casaburi. ‘Immersive virtual reality influences physiologic responses to submaximal exercise: A randomized, crossover trial’. In: *Frontiers in physiology* 12 (2021), p. 702266.
- [381] E. I. Alshanskaia, N. A. Zhozhikashvili, I. S. Polikanova and O. V. Martynova. In: *Frontiers in Psychiatry* 15 (2024), p. 1355846.
- [382] S. Immanuel, M. N. Teferra, M. Baumert and N. Bidargaddi. ‘Heart rate variability for evaluating psychological stress changes in healthy adults: a scoping review’. In: *Neuropsychobiology* 82.4 (2023), pp. 187–202.
- [383] S. Laborde, E. Mosley and J. F. Thayer. ‘Heart rate variability and cardiac vagal tone in psychophysiological research—recommendations for experiment planning, data analysis, and data reporting’. In: *Frontiers in psychology* 8 (2017), p. 213.
- [384] T. Borrowdale. *Zen Mode Soundtrack. Alto’s Adventure*. <https://www.youtube.com/watch?v=R4Cv-yAbTKM>. Accessed: 2025-04-04. 2016.
- [385] T. Luo, G. Chen, Y. Wen, P. Wang, Y. Fan, T. Han and F. Tian. ‘Exploring the Effects of Sensory Conflicts on Cognitive Fatigue in VR Remappings’. In: *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 2024, pp. 1–16.
- [386] F. Faul, E. Erdfelder, A. Buchner and A.-G. Lang. ‘Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses’. In: *Behavior Research Methods* 41.4 (2009), pp. 1149–1160. DOI: 10.3758/BRM.41.4.1149.
- [387] D. Bates, M. Mächler, B. Bolker and S. Walker. ‘Fitting Linear Mixed-Effects Models Using lme4’. In: *Journal of Statistical Software* 67.1 (2015), pp. 1–48. DOI: 10.18637/jss.v067.i01.
- [388] A. Kuznetsova, P. B. Brockhoff and R. H. B. Christensen. ‘lmerTest Package: Tests in Linear Mixed Effects Models’. In: *Journal of Statistical Software* 82.13 (2017), pp. 1–26. DOI: 10.18637/jss.v082.i13.
- [389] M. G. Kenward and J. H. Roger. ‘Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood’. In: *Biometrics* 53.3 (1997), pp. 983–997. DOI: 10.2307/2533558.

- [390] R. V. Lenth. ‘Estimated Marginal Means, aka Least-Squares Means’. In: *Journal of Statistical Software* 69.1 (2018), pp. 1–33. DOI: 10.18637/jss.v069.i01.
- [391] T. Hothorn, F. Bretz and P. Westfall. ‘Simultaneous inference in general parametric models’. In: *Biometrical Journal* 50.3 (2008), pp. 346–363. DOI: 10.1002/bimj.200810425.
- [392] M. Cakmak and S. S. Srinivasa. ‘Using spatial and temporal contrast for fluent robot-human hand-overs’. In: *Proceedings of the 6th international conference on Human-robot interaction*. ACM. 2011, pp. 489–496.
- [393] S. Nikolaidis, S. Nath, A. D. Procaccia and S. S. Srinivasa. ‘Human-robot mutual adaptation in collaborative tasks: Models and experiments’. In: *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction*. IEEE. 2013, pp. 1–8.
- [394] A. Steed, Y. Pan, F. Zisch and W. Steptoe. ‘The impact of a self-avatar on cognitive load in immersive virtual reality’. In: *2016 IEEE Virtual Reality (VR)*. IEEE. 2016, pp. 67–76.
- [395] G.-J. M. Kruijff, M. Janíček, T. López, N. Smets, T. Mioch, J. van Diggelen, P. Groenewegen, A. Visser, R. Looije, T. Bosse et al. ‘Experience in system design for human-robot teaming in urban search and rescue’. In: *Field and Service Robotics*. Springer, 2010, pp. 111–125.
- [396] J. M. Juliano, N. Schweighofer and S.-L. Liew. ‘Increased cognitive load in immersive virtual reality during visuomotor adaptation is associated with decreased long-term retention and context transfer’. In: *Journal of NeuroEngineering and Rehabilitation* 19.1 (2022), p. 106.
- [397] R. Fribourg, F. Argelaguet, L. Hoyet and A. Lécuyer. ‘Studying the sense of embodiment in VR shared experiences’. In: *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE. 2018, pp. 273–280.
- [398] D. Cui and C. Mousas. ‘Evaluating the sense of embodiment through out-of-body experience and tactile feedback’. In: *Proceedings of the 18th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*. 2022, pp. 1–7.

- [399] G. Gorisse, O. Christmann, E. A. Amato and S. Richir. ‘First-and third-person perspectives in immersive virtual environments: presence and performance analysis of embodied users’. In: *Frontiers in Robotics and AI* 4 (2017), p. 33.
- [400] H. Iriye and P. L. St. Jacques. ‘Memories for third-person experiences in immersive virtual reality’. In: *Scientific reports* 11.1 (2021), p. 4667.
- [401] E. Rojas-Muñoz, M. E. Cabrera, C. Lin, D. Andersen, V. Popescu, K. Anderson, B. L. Zarzaur, B. Mullis and J. P. Wachs. ‘The System for Telementoring with Augmented Reality (STAR): A head-mounted display to improve surgical coaching and confidence in remote areas’. In: *Surgery* 167.4 (2020), pp. 724–731.
- [402] F. Nickel, A. Cizmiciu and M. Chand. ‘Telestration and augmented reality in minimally invasive surgery: an invaluable tool in the age of COVID-19 for remote proctoring and telementoring’. In: *JAMA surgery* 157.2 (2022), pp. 169–170.
- [403] S. Hongli, W. Qingmiao, Y. Weixuan, L. Yuan, C. Yihui and W. Hongchao. ‘Application of AR technology in aircraft maintenance manual’. In: *Journal of Physics: Conference Series*. Vol. 1738. 1. IOP Publishing. 2021, p. 012133.
- [404] Y. Jiang, T. H. Tran and L. Williams. ‘Machine learning and mixed reality for smart aviation: Applications and challenges’. In: *Journal of Air Transport Management* 111 (2023), p. 102437.
- [405] Z. Jiang, Y. Huang, J. Eden, E. Ivanova, X. Cheng and E. Burdet. ‘A virtual reality platform to evaluate the effects of supernumerary limbs’ appearance’. In: *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE. 2023, pp. 1–5.
- [406] R. Miura, S. Kasahara, M. Kitazaki, A. Verhulst, M. Inami and M. Sugimoto. ‘Multi-soma: Distributed embodiment with synchronized behavior and perception’. In: *Proceedings of the Augmented Humans International Conference 2021*. 2021, pp. 1–9.
- [407] H. Elsayed, K. Kartono, D. Schön, M. Schmitz, M. Mühlhäuser and M. Weigel. ‘Understanding Perspectives for Single-and Multi-Limb Movement Guidance in Virtual 3D Environments’. In: *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*. 2022, pp. 1–10.

- [408] J. H. Murray. ‘Virtual/reality: how to tell the difference’. In: *Journal of visual culture* 19.1 (2020), pp. 11–27.
- [409] E. Wade and C. J. Winstein. ‘Virtual reality and robotics for stroke rehabilitation: where do we go from here?’ In: *Topics in stroke rehabilitation* 18.6 (2011), pp. 685–700.
- [410] E. Prati, V. Villani, M. Peruzzini and L. Sabattini. ‘An approach based on VR to design industrial human-robot collaborative workstations’. In: *Applied Sciences* 11.24 (2021), p. 11773.
- [411] Y. Iwasaki, K. Ando, S. Iizuka, M. Kitazaki and H. Iwata. ‘Detachable Body: The impact of binocular disparity and vibrotactile feedback in co-presence tasks’. In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 3477–3484.
- [412] H. Zhou, T. Ayes, C. Fan, Z. Sarsenbayeva and A. Withana. ‘CoplayingVR: Understanding User Experience in Shared Control in Virtual Reality’. In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8.3 (Sept. 2024). DOI: 10.1145/3678508. URL: <https://doi.org/10.1145/3678508>.
- [413] R. Kondo, S. Ueda, M. Sugimoto, K. Minamizawa, M. Inami and M. Kitazaki. ‘Invisible Long Arm Illusion: Illusory Body Ownership by Synchronous Movement of Hands and Feet.’ In: *ICAT-EGVE*. Limassol. 2018, pp. 21–28.
- [414] R. Kondo, Y. Tani, M. Sugimoto, M. Inami and M. Kitazaki. ‘Scrambled body differentiates body part ownership from the full body illusion’. In: *Scientific reports* 10.1 (2020), p. 5274.
- [415] M. Fukuoka, A. Verhulst, F. Nakamura, R. Takizawa, K. Masai and M. Sugimoto. ‘FaceDrive: Facial expression driven operation to control virtual supernumerary robotic arms’. In: *SIGGRAPH Asia 2019 XR*. 2019, pp. 9–10.
- [416] A. Weiss, R. Bernhaupt, D. Schwaiger, M. Altmaninger, R. Buchner and M. Tscheligi. ‘User experience evaluation with a wizard of oz approach: Technical and methodological considerations’. In: *2009 9th IEEE-RAS International Conference on Humanoid Robots*. IEEE. 2009, pp. 303–308.

- [417] G. Alce, M. Wallergård and K. Hermodsson. ‘WozARd: a wizard of Oz method for wearable augmented reality interaction—a pilot study’. In: *Advances in Human-Computer Interaction* 2015.1 (2015), p. 271231.
- [418] F. Diederichs, L.-A. Mathis, V. Bopp-Bertenbreiter, B. Bednorz, H. Widloither and F. Flemisch. ‘A Wizard-of-Oz vehicle to investigate human interaction with AI-driven automated cars’. In: *Proceedings of the DSC*. 2021.
- [419] S. Hu, H. C. Yen, Z. Yu, M. Zhao, K. Seaborn and C. Liu. ‘Wizundry: A cooperative Wizard of Oz platform for simulating future speech-based interfaces with multiple Wizards’. In: *Proceedings of the ACM on Human-Computer Interaction* 7.CSCW1 (2023), pp. 1–34.
- [420] A. Habibovic, J. Andersson, M. Nilsson, V. M. Lundgren and J. Nilsson. ‘Evaluating interactions with non-existing automated vehicles: three Wizard of Oz approaches’. In: *2016 IEEE intelligent vehicles symposium (IV)*. IEEE. 2016, pp. 32–37.
- [421] K. Gu, M. Grunde-McLaughlin, A. McNutt, J. Heer and T. Althoff. ‘How do data analysts respond to ai assistance? a wizard-of-oz study’. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 2024, pp. 1–22.
- [422] J. S. Pettersson and M. Wik. ‘Perspectives on Ozlab in the cloud: A literature review of tools supporting Wizard-of-Oz experimentation, including an historical overview of 1971-2013 and notes on methodological issues and supporting generic tools’. In: (2014).
- [423] A. Tong, P. Perera, Z. Sarsenbayeva, A. McEwan, A. C. De Silva and A. Withana. ‘Fully 3D-printed dry EEG electrodes’. In: *Sensors* 23.11 (2023), p. 5175.
- [424] S. Kim, J. R. Anthis and S. Sebo. ‘A taxonomy of robot autonomy for human-robot interaction’. In: *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 2024, pp. 381–393.
- [425] E. S. De Lima and B. Feijó. ‘Artificial intelligence in human-robot interaction’. In: *Emotional Design in Human-Robot Interaction: Theory, Methods and Applications* (2019), pp. 187–199.

- [426] P. N. Hung and T. Yoshimi. ‘Extracting actions from instruction manual and testing their execution in a robotic simulation’. In: *ASEAN Engineering Journal* 6.1 (2016), pp. 47–58.
- [427] M. Zhang, G. Tian, Y. Zhang and P. Duan. ‘Service skill improvement for home robots: Autonomous generation of action sequence based on reinforcement learning’. In: *Knowledge-Based Systems* 212 (2021), p. 106605.
- [428] L. Fernandes and B. Shivakumar. ‘Identification and Sorting of Objects based on Shape and Colour using robotic arm’. In: *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*. IEEE. 2020, pp. 866–871.
- [429] I. Sharp, F. Huang and J. Patton. ‘Visual error augmentation enhances learning in three dimensions’. In: *Journal of neuroengineering and rehabilitation* 8 (2011), pp. 1–6.
- [430] K. Umezawa, Y. Suzuki, G. Ganesh and Y. Miyawaki. ‘Bodily ownership of an independent supernumerary limb: an exploratory study’. In: *Scientific reports* 12.1 (2022), p. 2339.
- [431] Q. M. Marwan, S. C. Chua and L. C. Kwek. ‘Comprehensive review on reaching and grasping of objects in robotics’. In: *Robotica* 39.10 (2021), pp. 1849–1882.
- [432] Y. Guiard. ‘Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model’. In: *Journal of motor behavior* 19.4 (1987), pp. 486–517.
- [433] K. Zhang, Y. Long and X. Luo. ‘Review of Supernumerary Robotic Limbs’. In: *Journal of Physics: Conference Series*. Vol. 2456. 1. IOP Publishing. 2023, p. 012004.
- [434] A. Correa, M. R. Walter, L. Fletcher, J. Glass, S. Teller and R. Davis. ‘Multimodal interaction with an autonomous forklift’. In: *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2010, pp. 243–250.
- [435] Y. Hu, S. Zhang, T. Dang, H. Jia, F. D. Salim, W. Hu and A. J. Quigley. ‘Exploring large-scale language models to evaluate eeg-based multimodal data for mental health’. In: *Companion of the 2024 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 2024, pp. 412–417.
- [436] B. House, J. Malkin and J. Bilmes. ‘The VoiceBot: a voice controlled robot arm’. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2009, pp. 183–192.

- [437] G. Choudhary and C. R. BV. 'Real time robotic arm control using hand gestures'. In: *2014 International Conference on High Performance Computing and Applications (ICHPCA)*. IEEE. 2014, pp. 1–3.
- [438] R. Chotikunnan, K. Roongprasert, P. Chotikunnan, P. Imura, M. Sangworasil and A. Srisiriwat. 'Robotic Arm Design and Control Using MATLAB/Simulink'. In: *International Journal of Membrane Science and Technology* 10.3 (2023), pp. 2448–2459.
- [439] C. J. Lin, W.-J. Shiang, C.-Y. Chuang and J.-L. Liou. 'Applying the skill-rule-knowledge framework to understanding operators' behaviors and workload in advanced main control rooms'. In: *Nuclear Engineering and Design* 270 (2014), pp. 176–184.
- [440] P. Marion, M. Fallon, R. Deits, A. Valenzuela, C. P. D'Arpino, G. Izatt, L. Manuelli, M. Antone, H. Dai, T. Koolen et al. 'Director: A user interface designed for robot operation with shared autonomy'. In: *The DARPA Robotics Challenge Finals: Humanoid Robots To The Rescue* (2018), pp. 237–270.
- [441] M. Firlej and A. Taeihagh. 'Regulating human control over autonomous systems'. In: *Regulation & Governance* 15.4 (2021), pp. 1071–1091.
- [442] M. Noorman and D. G. Johnson. 'Negotiating autonomy and responsibility in military robots'. In: *Ethics and Information Technology* 16.1 (2014), pp. 51–62.
- [443] M. Ferianc, H. Fan, D. Manocha, H. Zhou, S. Liu, X. Niu and W. Luk. 'Improving performance estimation for design space exploration for convolutional neural network accelerators'. In: *Electronics* 10.4 (2021), p. 520.
- [444] B. V. Syiem, R. M. Kelly, J. Goncalves, E. Velloso and T. Dingler. 'Impact of task on attentional tunneling in handheld augmented reality'. In: *Proceedings of the 2021 CHI conference on human factors in computing systems*. 2021, pp. 1–14.
- [445] B. V. Syiem, R. M. Kelly, E. Velloso, J. Goncalves and T. Dingler. 'Enhancing visitor experience or hindering docent roles: attentional issues in augmented reality supported installations'. In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE. 2020, pp. 279–288.

- [446] B. Bruya and Y.-Y. Tang. ‘Is attention really effort? Revisiting Daniel Kahneman’s influential 1973 book attention and effort’. In: *Frontiers in psychology* 9 (2018), p. 1133.
- [447] L. Ahmed and J. W. De Fockert. ‘Focusing on attention: The effects of working memory capacity and load on selective attention’. In: (2012).
- [448] H. M. Chen, L. Castelli, M. Ferianc, H. Zhou, S. Liu, W. Luk and H. Fan. ‘Enhancing Dropout-based Bayesian Neural Networks with Multi-Exit on FPGA’. In: *arXiv preprint arXiv:2406.14593* (2024).
- [449] A. Tung, J. Wong, A. Mandlekar, R. Martín-Martín, Y. Zhu, L. Fei-Fei and S. Savarese. ‘Learning multi-arm manipulation through collaborative teleoperation’. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2021, pp. 9212–9219.
- [450] T. L. Bruns, A. A. Ramirez, M. A. Emerson, R. A. Lathrop, A. W. Mahoney, H. B. Gilbert, C. L. Liu, P. T. Russell, R. F. Labadie, K. D. Weaver et al. ‘A modular, multi-arm concentric tube robot system with application to transnasal surgery for orbital tumors’. In: *The International Journal of Robotics Research* 40.2-3 (2021), pp. 521–533.
- [451] K. A. Szczurek, R. M. Prades, E. Matheson, J. Rodriguez-Nogueira and M. Di Castro. ‘Multimodal multi-user mixed reality human–robot interface for remote operations in hazardous environments’. In: *IEEE Access* 11 (2023), pp. 17305–17333.
- [452] B. Wang, P. Zheng, Y. Yin, A. Shih and L. Wang. ‘Toward human-centric smart manufacturing: A human-cyber-physical systems (HCPS) perspective’. In: *Journal of Manufacturing Systems* 63 (2022), pp. 471–490.
- [453] G. Yang, Z. Pang, M. J. Deen, M. Dong, Y.-T. Zhang, N. Lovell and A. M. Rahmani. ‘Homecare robotic systems for healthcare 4.0: Visions and enabling technologies’. In: *IEEE journal of biomedical and health informatics* 24.9 (2020), pp. 2535–2549.
- [454] M. H. Zafar, E. F. Langås and F. Sanfilippo. ‘Exploring the synergies between collaborative robotics, digital twins, augmentation, and industry 5.0 for smart manufacturing: A state-of-the-art review’. In: *Robotics and Computer-Integrated Manufacturing* 89 (2024), p. 102769.

- [455] C. Schmidbauer. ‘Adaptive task sharing between humans and cobots in assembly processes’. In: *TU Wien* (2022).
- [456] W. Dudek and T. Winiarski. ‘Scheduling of a robot’s tasks with the tasker framework’. In: *IEEE Access* 8 (2020), pp. 161449–161471.
- [457] D. ZHOU, Y. DENG, A. GLOBALA, R. DAVIES, Y. HU, R.-X. WU, L. FU and A. J. QUIGLEY. ‘Social XR: Designing An Extended Reality Application for Interaction in Long-Distance Relationships’. In: (2024).
- [458] S. Leone, L. Giunta, V. Rino, S. Mellace, A. Sozzi, F. Lago, E. M. Curcio, D. Pisle and G. Carbone. ‘Design of a Wheelchair-Mounted Robotic Arm for Feeding Assistance of Upper-Limb Impaired Patients’. In: *Robotics* 13.3 (2024), p. 38.
- [459] S. Palmcrantz, A. Wall, K. S. Vreede, P. Lindberg, A. Danielsson, K. S. Sunnerhagen, C. K. Häger and J. Borg. ‘Impact of intensive gait training with and without electromechanical assistance in the chronic phase after stroke—a multi-arm randomized controlled trial with a 6 and 12 months follow up’. In: *Frontiers in Neuroscience* 15 (2021), p. 660726.
- [460] M. Wongphati, Y. Matsuda, H. Osawa and M. Imai. ‘Where do you want to use a robotic arm? And what do you want from the robot?’ In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2012, pp. 322–327.
- [461] H. Zhou and et al. ‘Collaborative Embodiment in VR: A Methodological Review of Shared Control, Perspective, and Interaction Strategies’. In: *ACM Computing Surveys* (under review) (2024).
- [462] H. Zhou and et al. ‘One Body, Two Minds: Shared Perspectives for Multi-Hand Collaborative Embodiment in VR’. In: *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*. 2025.

3 Appendix A

Proofs in Chapter 5

In this section, we show all the proofs in Chapter 5.

1 Physiological Data Processing Details

Signal source and export. Heart activity was recorded continuously with a Polar H10 chest strap. We logged the device-provided inter-beat interval (IBI/RR) stream via Bluetooth Low Energy throughout each trial and exported the timestamped RR series for offline processing. Analyses were performed on normal-to-normal (NN) intervals derived from the RR stream.

Artefact handling and NN reconstruction. RR series were preprocessed following recommended HRV workflows [383]. We removed (i) biologically implausible intervals and dropouts (e.g., extremely short/long RR values indicating missed or extra detections) and (ii) transient artefacts identified as outliers relative to local beat-to-beat dynamics. Short gaps created by artefact removal were linearly interpolated to preserve continuity for window-based HRV estimation, while longer discontinuities were excluded from HRV computation. All HRV metrics reported in the paper were computed only from contiguous NN segments that satisfied the minimum window-length requirement described below.

Windowing for HRV. HRV was computed over active-task periods using contiguous analysis windows of at least 60 s (minimum duration per window), consistent with common practice for time-domain HRV when task segments are short [383]. When a trial segment exceeded 60 s, we computed HRV on consecutive windows and then aggregated to one value per participant $\times$ condition by averaging across valid windows within that segment.

RMSSD computation. We used RMSSD as the primary time-domain HRV index:

$$\text{RMSSD} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (NN_{i+1} - NN_i)^2}.$$

RMSSD was chosen because it is widely used in psychophysiology, is comparatively robust to slow trends in heart rate, and reflects short-term vagal modulation [383]. Higher RMSSD values indicate lower physiological stress/arousal.

Peak heart rate and event-triggered analysis. Peak heart rate was extracted within the same active-task segments as the maximum observed heart rate during the segment. For event-triggered analyses, we time-aligned the heart-rate stream to perspective-switch timestamps from system logs and summarized heart rate in a short post-switch window; these event-level summaries were then modeled as a function of switch direction (e.g., EMBEDDED ANCHORED VIEW $\rightarrow$ OUT-OF-BODY VIEW vs. OUT-OF-BODY VIEW $\rightarrow$ EMBEDDED ANCHORED VIEW) and role.

Proofs in Chapter 6

In this section, we show all the proofs in Chapter 6.

1 Autonomy level details

Before starting the study, we introduced the participants to an action guide detailing the operational behaviors of the virtual supernumerary limbs. The operator adhered to this predefined guide, which specified responses to participant commands across different autonomy levels.

1.1 Low-autonomy level

Description: At this level, the virtual supernumerary limbs have basic capabilities, including the ability to perceive what is happening in their environment. They can understand and respond to simple instructions and learn from user interactions. However, while they can recognize general elements in the environment, they cannot automatically determine where or how to place specific objects without explicit user guidance. The arms rely on you to guide them step by step on how to complete tasks, categorizing items based on your instructions.

Capabilities: While they understand their surroundings, they depend on you to guide their actions. They respond well to step-by-step instructions where you guide each action.

Role: You will need to provide detailed, step-by-step guidance through verbal or gestural commands on how the virtual supernumerary limbs should sort and handle objects. These arms can only recognize and act upon the tasks you have specifically taught them. They

cannot automatically identify or categorize items they haven't learned before, so they rely entirely on your instructions to understand and execute each action.

1.2 High-autonomy level

Description: This version of the virtual supernumerary limb system is significantly more advanced, with improved language comprehension and the ability to understand complex tasks. It also automatically recognizes things in the environment, reducing the need for teach them, and allows you to demonstrate and execute complex tasks and instructions. However, they do not inherently know what actions to take; they need you to instruct them on what to do.

Capabilities: The virtual supernumerary limbs can autonomously recognize their surroundings. However, they still require your instructions to determine what actions to take. Once directed, they can efficiently execute tasks with minimal guidance.

Role: You will need to provide clear instructions on what actions the arms should take. Whether it's sorting objects, placing them in specific locations, or any other task, the virtual supernumerary limbs depend on your directives to know what to do. The arms are designed to follow your commands accurately, but they rely on you to define the task and guide their actions accordingly.

2 Operator Action Guidelines

The operator followed a predefined set of action rules to ensure consistency across participants and tasks. These guidelines were tailored to the study's four instruction types and autonomy levels, aligning with principles from multi-agent coordination and robot learning frameworks [426, 49, 427, 428]:

2.1 Command Instructions

Low-autonomy: Perform each step based on explicit participant guidance (e.g., grasp, move, rotate). For instance, the operator can rotate a block incrementally in response to specific participant instructions for each action. If a series of commands given, only execute the first. If a complex command such as a combined task given, do not perform any actions.

High-autonomy: Execute the full sequence or complex actions based on the command's intent (e.g., directly move a labeled object to a specified position).

2.2 Demonstration Instructions

Low-autonomy: Mimic participant's demonstrations step-by-step. If the participant rotated an object, the operator replicated the action.

High-autonomy: Learn the demonstrated action and execute autonomously for similar objects. For instance, stacking two blocks based on a demonstration was applied to subsequent blocks independently.

2.3 Delegation Instructions

Low-autonomy: No action was taken, as Delegation Instructions were not supported under this autonomy level. *High-autonomy:* Plan and execute tasks sequentially after an instruction. A single command like "Organize objects by size in a row" prompted the operator to complete the task autonomously.

2.4 Labeling Instructions

Low-autonomy: Identify and label objects explicitly based on participant input. For example, the operator must wait for the participant to confirm "The red square" before proceeding.

High-autonomy: Automatically identify attributes and proceed with the complex tasks without additional guidance.