



# Exploring the Implications of Handheld Augmented Reality in Collaborative Practices

**Thomas James Wells, BSc (Hons), MSc**

School of Computing and Communication

Lancaster University

A thesis submitted for the degree of

*Doctor of Philosophy*

August, 2025

# Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of 80,000 words including appendices and footnotes, but excluding the bibliography. A rough estimate of the word count is: 59426

Thomas James Wells



# Exploring the Implications of Handheld Augmented Reality in Collaborative Practices

Thomas James Wells, BSc (Hons), MSc.

School of Computing and Communication, Lancaster University

A thesis submitted for the degree of *Doctor of Philosophy*. August, 2025

## Abstract

It is well known that Augmented Reality (AR), particularly handheld mobile AR, holds great potential for enhancing co-located collaborative experiences. However, despite the significant advances in AR, there is still a fundamental gap in the absence of a unifying framework that addresses the challenges in understanding group dynamics, device configurations, and interaction techniques. This thesis addresses the overarching problem of a lack of a framework for designing and evaluating collaborative AR experiences by systematically investigating the cognitive and physical demands of handheld AR on users. This research explores how these factors influence group behaviours. Through empirical studies, this work aims to contribute a deeper understanding of the interplay between AR technology and collaborative work, paving the way for more effective and intuitive collaborative AR applications.

Building on the technological advancements of handheld mobile AR, this research aims to shift the focus from individual user-device interactions to a broader interaction space of co-located collaborative applications. We investigate how different device configurations affect user engagement and collaboration, revealing that larger devices like tablets enhance interaction through effective screen sharing and shoulder surfing. In comparison, smaller devices like smartphones impose higher cognitive and physical loads. Additionally, the research highlights the need for AR systems to support seamless transitions between AR and face-to-face interactions to maintain collaboration efficacy and reduce these negative impositions.

By integrating key Human-Computer Interaction (HCI) concepts such as proxemics, territoriality, and shared interaction spaces, this thesis presents a comprehensive framework for designing collaborative AR experiences. The findings underscore the importance of ergonomic and social considerations in developing AR interfaces that are not only functional but also intuitive and accessible for users.

This thesis contributes to the field of HCI and CSCW by providing empirical evidence, design principles, and a conceptual framework that guides the future development of AR applications. It addresses core challenges in collaborative AR, such as managing cognitive and physical demands, optimising device configurations, and developing effective interaction techniques that stretch beyond the physical reality. The ultimate aim is to enhance handheld mobile AR's effectiveness and user experience, making it a more intuitive tool for collaborative interactions.

# Acknowledgements

I would first like to express my deepest gratitude to my supervisor, Dr. Steven Houben, for his exceptional patience, guidance, and engaging discussions throughout my research journey. Steven, your mentorship and general advice over the past 4-5 years have been pivotal in my development, not just as an academic, but also as a person. Your support has been second to none, and I am incredibly grateful for your unwavering belief in me and my work.

I would also like to thank my examiners, Dr. Nicolai Marquardt and Dr. Abe Karnik, for taking the time to read and evaluate my thesis, and for their insightful feedback and engaging discussion during the viva. I truly appreciate your thoughtful comments, which have helped me to view my work from new perspectives.

My heartfelt thanks also go to the (honorary) members of the wider research group – Kim, Ludwig, Dom, Chris, Boyeun, Argenis, Claudia, Hayat, Franzi, Josh, James, Hans, Olaf, Dimitra, and Bianca – for fostering such a supportive and intellectually stimulating environment. Thank you for the memories, the baked goods, and the shared moments of both frustration and laughter that made this journey all the more enriching.

I'm also grateful to Martin, whose support and encouragement beyond my academic life gave me the confidence to grow and take on new challenges. Your trust and willingness to take a chance on me gave me opportunities that profoundly shaped how I approach challenges, people, and responsibility — and left a lasting mark on how I think, work, and lead.

A special thank you needs to go to Katie and Danielle, who have been there for me since the very first week of my university career. From fresher's week to countless nights in the library, through to final submission, your friendship has been one of the few constants in this long and winding journey. Thank you for the laughter, the venting, the encouragement, and the unwavering support – I honestly wouldn't have done this without you.

This journey has been long, challenging, and at times unpredictable — but never solitary. I'm deeply grateful to everyone who has supported, encouraged, and believed in me over the past five years. Thank you for being part of this chapter of my life.

# Publications

The following publications have been generated while developing this thesis, and to an extent have guided the thesis into what it has become:

Thomas Wells and Steven Houben. “CollabAR - Investigating the Mediating Role of Mobile AR Interfaces on Co-Located Group Collaboration”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–13. DOI: [10 . 1145 / 3313831 . 3376541](https://doi.org/10.1145/3313831.3376541)

Thomas Wells, Dominic Potts, and Steven Houben. “A Study into the Effect of Mobile Device Configurations on Co-Located Collaboration Using AR”. in: *Proceedings of the ACM on Human-Computer Interaction* 6 (MHCI Sept. 20, 2022), 200:1–200:23. DOI: [10 . 1145 / 3546735](https://doi.org/10.1145/3546735)

Mille Skovhus Lunding, Jens Emil Sloth Grønbæk, Nicolai Grymer, Thomas Wells, Steven Houben, and Marianne Graves Petersen. “Reality and Beyond: Proxemics as a Lens for Designing Handheld Collaborative Augmented Reality”. In: *Proc. ACM Hum.-Comput. Interact.* 7 (ISS Nov. 1, 2023), 427:21–427:40. DOI: [10.1145/3626463](https://doi.org/10.1145/3626463)

# Contribution Statements

## Chapter 3

Parts of this chapter are published in the proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20.. As the first and sole author of this publication, I was entirely responsible for the conception, design, and technical implementation of the CollabAR system. I also led the empirical study to investigate its impact on co-located group collaboration, including the collection and analysis of data related to cognitive load, group dynamics, and device handling. This work provided the foundational insights for RQ1.

## Chapter 4

Parts of this chapter are published in the Proceedings of the ACM on Human-Computer Interaction 6 (MHCI Sept. 20, 2022). As the first author of this publication, I was primarily responsible for the design and execution of the experimental study investigating how various mobile device configurations influence collaborative strategies and user behaviours in Augmented Reality environments. My contributions encompassed the entire study methodology, the comprehensive data collection, and the in-depth analysis of task efficacy, communication patterns, and interaction behaviours, directly addressing RQ2. My co-authors provided valuable guidance and feedback throughout the research process.

## Chapter 5

Parts of this chapter are published in Proc. ACM Hum.-Comput. Interact. 7 (ISS Nov. 1, 2023). This chapter establishes a foundational lens for understanding socio-spatial interactions within handheld co-located AR environments, using proxemics as a guiding theory. My core contributions to this publication focused on the design and implementation of the interaction techniques discussed, such as Freeze Frame, Remote Peeking, Remote Sketch, Proximity Threshold, and Privacy Fence. I also developed the application scenarios that illustrate their practical utility in domains

like architectural design and classroom group work. Furthermore, I contributed to the construction of the underlying AR system that allowed for the exploration of these techniques. My co-authors were instrumental in introducing and refining the conceptual lens of proxemics, which provided the theoretical framework for articulating and exploring these techniques in relation to aligning with and going beyond physical reality. This work directly informs RQ3 by expanding the design space for collaborative AR beyond traditional physical limitations.

# Contents

|                                                                                              |              |
|----------------------------------------------------------------------------------------------|--------------|
| <b>List of Figures</b>                                                                       | <b>xiii</b>  |
| <b>List of Tables</b>                                                                        | <b>xviii</b> |
| <b>List of Abbreviations</b>                                                                 | <b>xix</b>   |
| <b>1 Introduction</b>                                                                        | <b>1</b>     |
| 1.1 Research Challenges . . . . .                                                            | 4            |
| 1.1.1 Co-Located Collaboration in AR . . . . .                                               | 5            |
| 1.1.2 User Interactions . . . . .                                                            | 6            |
| 1.1.3 Group Dynamics in Mobile AR Collaboration . . . . .                                    | 8            |
| 1.1.4 Describing Multi-Device Mobile AR Interactions . . . . .                               | 9            |
| 1.1.5 Central Challenge: Designing Effective Co-Located Mobile AR<br>Collaboration . . . . . | 10           |
| 1.2 Research Questions . . . . .                                                             | 10           |
| 1.3 Research Context . . . . .                                                               | 13           |
| 1.4 Methodology . . . . .                                                                    | 14           |
| 1.5 Contributions . . . . .                                                                  | 17           |
| 1.6 Thesis Structure . . . . .                                                               | 19           |
| <b>2 Background</b>                                                                          | <b>22</b>    |
| 2.1 Augmented Reality Displays . . . . .                                                     | 23           |
| 2.1.1 Classifications of Augmented Reality . . . . .                                         | 24           |
| 2.1.2 Handheld Augmented Reality . . . . .                                                   | 26           |
| 2.1.3 Collaborative Handheld AR . . . . .                                                    | 27           |
| 2.1.4 Application Areas of Co-Located Collaborative Augmented<br>Reality . . . . .           | 27           |
| 2.1.5 Challenges in Handheld AR . . . . .                                                    | 30           |
| 2.2 Collaboration . . . . .                                                                  | 33           |
| 2.2.1 Computer-Supported Cooperative Work . . . . .                                          | 35           |
| 2.2.2 Awareness . . . . .                                                                    | 37           |

|          |                                                                                                          |           |
|----------|----------------------------------------------------------------------------------------------------------|-----------|
| 2.2.3    | Proxemic Interaction . . . . .                                                                           | 38        |
| 2.2.4    | Affordances for Collaboration . . . . .                                                                  | 39        |
| 2.2.5    | Group Dynamics . . . . .                                                                                 | 42        |
| 2.3      | Cross-Device Computing in Collaborative Augmented Reality (AR) .                                         | 45        |
| 2.4      | Conclusion . . . . .                                                                                     | 48        |
| <b>3</b> | <b>Exploring the Dynamics of Group Interaction in AR</b>                                                 | <b>50</b> |
| 3.1      | Introduction . . . . .                                                                                   | 51        |
| 3.2      | Exploring the CollabAR System . . . . .                                                                  | 52        |
| 3.2.1    | User Interface . . . . .                                                                                 | 52        |
| 3.2.2    | Gesture-Based Interaction . . . . .                                                                      | 53        |
| 3.2.3    | Technical Implementation . . . . .                                                                       | 55        |
| 3.2.4    | Virtual Models . . . . .                                                                                 | 56        |
| 3.3      | Study Methodology . . . . .                                                                              | 56        |
| 3.3.1    | Participants . . . . .                                                                                   | 57        |
| 3.3.2    | Apparatus . . . . .                                                                                      | 58        |
| 3.3.3    | Tasks . . . . .                                                                                          | 59        |
| 3.3.4    | Data Collection and Analysis . . . . .                                                                   | 60        |
| 3.4      | Findings . . . . .                                                                                       | 61        |
| 3.4.1    | Utilisation Patterns . . . . .                                                                           | 61        |
| 3.4.2    | Task Completion Patterns . . . . .                                                                       | 62        |
| 3.4.3    | Usability Analysis . . . . .                                                                             | 64        |
| 3.4.4    | User Experience and Observations . . . . .                                                               | 66        |
| 3.4.5    | Summary of Findings . . . . .                                                                            | 69        |
| 3.5      | Discussion . . . . .                                                                                     | 69        |
| 3.5.1    | Context Switching and Device Handling . . . . .                                                          | 69        |
| 3.5.2    | Impact of Model Complexity on Collaboration . . . . .                                                    | 70        |
| 3.5.3    | Mobility in Tabletop AR . . . . .                                                                        | 71        |
| 3.5.4    | Design Recommendations . . . . .                                                                         | 71        |
| 3.6      | Limitations and Future Work . . . . .                                                                    | 73        |
| 3.7      | Conclusion . . . . .                                                                                     | 74        |
| <b>4</b> | <b>A Study into the Effect of Mobile Device Configurations on Co-<br/>Located Collaboration using AR</b> | <b>75</b> |
| 4.1      | Introduction . . . . .                                                                                   | 76        |
| 4.2      | Motivation . . . . .                                                                                     | 77        |
| 4.2.1    | Device Symmetry (H1) . . . . .                                                                           | 78        |
| 4.2.2    | Device Size (H2) . . . . .                                                                               | 78        |

|          |                                                                                           |            |
|----------|-------------------------------------------------------------------------------------------|------------|
| 4.2.3    | Device Quantity (H3) . . . . .                                                            | 78         |
| 4.2.4    | Model Occlusion (H4) . . . . .                                                            | 79         |
| 4.2.5    | AR Activity (H5) . . . . .                                                                | 79         |
| 4.3      | Study Detail . . . . .                                                                    | 79         |
| 4.3.1    | Device Ecology &Interaction Techniques . . . . .                                          | 80         |
| 4.3.2    | AR Tasks &Virtual Models . . . . .                                                        | 80         |
| 4.3.3    | Apparatus &Participants . . . . .                                                         | 82         |
| 4.3.4    | Data Collection &Analysis . . . . .                                                       | 83         |
| 4.4      | Results . . . . .                                                                         | 85         |
| 4.4.1    | Efficacy of Collaboration . . . . .                                                       | 85         |
| 4.4.2    | Task Errors . . . . .                                                                     | 88         |
| 4.4.3    | Nature of Collaboration . . . . .                                                         | 89         |
| 4.4.4    | Focus . . . . .                                                                           | 90         |
| 4.4.5    | Communication . . . . .                                                                   | 92         |
| 4.4.6    | Interaction . . . . .                                                                     | 95         |
| 4.4.7    | Strategies of Collaboration . . . . .                                                     | 96         |
| 4.4.8    | Frequency of strategies across device configurations . . . . .                            | 97         |
| 4.5      | Discussion . . . . .                                                                      | 98         |
| 4.5.1    | Task Performance . . . . .                                                                | 99         |
| 4.5.2    | Impact of Device on Collaboration . . . . .                                               | 100        |
| 4.5.3    | Collaborative Strategies . . . . .                                                        | 100        |
| 4.6      | Limitations and Future Work . . . . .                                                     | 101        |
| 4.7      | Conclusion . . . . .                                                                      | 102        |
| <b>5</b> | <b>A Conceptual Lens on Handheld AR Collaboration</b>                                     | <b>103</b> |
| 5.1      | Introduction . . . . .                                                                    | 105        |
| 5.2      | A Conceptual Lens for Co-Located Collaborative Handheld Augmented Reality . . . . .       | 107        |
| 5.2.1    | The Interplay Between F-Formations, Collaboration Styles, and Physical Features . . . . . | 107        |
| 5.2.2    | Cross-Device AR for Aligning With and Going Beyond Reality                                | 109        |
| 5.3      | Research Through Design: Prototyping AR Interaction Techniques . .                        | 111        |
| 5.4      | Reality &Beyond: Designing Collaborative AR Interaction Techniques                        | 111        |
| 5.4.1    | Configuring Perception Proxemics . . . . .                                                | 114        |
| 5.4.2    | Configuring Deixis Proxemics . . . . .                                                    | 116        |
| 5.4.3    | Configuring Control Proxemics . . . . .                                                   | 117        |
| 5.5      | Example Applications Scenarios . . . . .                                                  | 119        |
| 5.5.1    | Architectural Design Meetings . . . . .                                                   | 119        |



|          |                                                                                                                              |            |
|----------|------------------------------------------------------------------------------------------------------------------------------|------------|
| 5.5.2    | Classroom Group Work . . . . .                                                                                               | 121        |
| 5.6      | Discussion . . . . .                                                                                                         | 123        |
| 5.6.1    | AR for Aligning With and Going Beyond Reality . . . . .                                                                      | 123        |
| 5.6.2    | A Many-To-Many Relationship Between Collaboration Styles<br>and F-Formations . . . . .                                       | 124        |
| 5.6.3    | Designing for Awareness and Control in Flexible Proxemic<br>Interactions . . . . .                                           | 125        |
| 5.6.4    | Conceptual Framework for Real-World Application Scenarios .                                                                  | 125        |
| 5.6.5    | Insights on Trade-Offs in Flexibility and Intuitiveness . . . . .                                                            | 126        |
| 5.6.6    | Potential for Extending Proxemics Theory to AR . . . . .                                                                     | 126        |
| 5.6.7    | Limitations and Future Work . . . . .                                                                                        | 126        |
| 5.7      | Conclusion . . . . .                                                                                                         | 127        |
| <b>6</b> | <b>Designing Collaborative Spaces in Mobile Augmented Reality: In-<br/>tegrating Proxemics and Cross-Device Interactions</b> | <b>129</b> |
| 6.1      | Introduction . . . . .                                                                                                       | 131        |
| 6.2      | CARIS Framework: Operationalising Proxemics in AR . . . . .                                                                  | 132        |
| 6.2.1    | Foundational Elements of CARIS . . . . .                                                                                     | 134        |
| 6.2.2    | Interplay of CARIS Components . . . . .                                                                                      | 136        |
| 6.2.3    | Inspiring Design Guidelines . . . . .                                                                                        | 137        |
| 6.3      | System Implementation . . . . .                                                                                              | 139        |
| 6.3.1    | System Architecture and Component Structure . . . . .                                                                        | 139        |
| 6.3.2    | Flow of Information in CARIS . . . . .                                                                                       | 141        |
| 6.3.3    | Interaction Space . . . . .                                                                                                  | 143        |
| 6.4      | Study Design . . . . .                                                                                                       | 146        |
| 6.4.1    | Participants . . . . .                                                                                                       | 146        |
| 6.4.2    | Apparatus . . . . .                                                                                                          | 147        |
| 6.4.3    | Tasks . . . . .                                                                                                              | 147        |
| 6.4.4    | Data Collection and Analysis . . . . .                                                                                       | 147        |
| 6.5      | Results . . . . .                                                                                                            | 148        |
| 6.5.1    | Interactions . . . . .                                                                                                       | 148        |
| 6.5.2    | Context Switching . . . . .                                                                                                  | 149        |
| 6.5.3    | System Usability . . . . .                                                                                                   | 153        |
| 6.5.4    | NASA-TLX Analysis . . . . .                                                                                                  | 155        |
| 6.6      | Discussion . . . . .                                                                                                         | 156        |
| 6.6.1    | Perceived Workload of CARIS . . . . .                                                                                        | 157        |
| 6.6.2    | System Usability . . . . .                                                                                                   | 158        |
| 6.6.3    | Context Switches & Immersion . . . . .                                                                                       | 159        |
| 6.7      | Conclusion . . . . .                                                                                                         | 160        |

|          |                                                                    |            |
|----------|--------------------------------------------------------------------|------------|
| <b>7</b> | <b>Discussion</b>                                                  | <b>161</b> |
| 7.1      | Revisiting Research Questions . . . . .                            | 163        |
| 7.2      | Broader Implications and Relation to Prior Work . . . . .          | 166        |
| 7.2.1    | Proxemics in Collaborative AR . . . . .                            | 166        |
| 7.2.2    | Implications for Mobile Handheld AR . . . . .                      | 168        |
| 7.2.3    | Implications for Cross-Device Collaboration . . . . .              | 170        |
| 7.2.4    | Implications for Group Collaboration Strategies . . . . .          | 172        |
| 7.2.5    | Implications for Methodologies . . . . .                           | 173        |
| 7.3      | Implications for Future Design in Collaborative Settings . . . . . | 175        |
| 7.3.1    | Technical Implications . . . . .                                   | 176        |
| 7.3.2    | Empirical Implications . . . . .                                   | 177        |
| 7.3.3    | Design Ideas . . . . .                                             | 178        |
| 7.3.4    | Design Guidelines . . . . .                                        | 180        |
| 7.4      | Summary and Key Takeaways . . . . .                                | 184        |
| 7.5      | Future Work . . . . .                                              | 186        |
| 7.5.1    | Limitations . . . . .                                              | 186        |
| 7.5.2    | Open Questions and Concerns . . . . .                              | 186        |
| 7.5.3    | Emerging Ideas and New Directions . . . . .                        | 187        |
| <b>8</b> | <b>Conclusion</b>                                                  | <b>189</b> |
|          | <b>References</b>                                                  | <b>191</b> |

# List of Figures

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                     |    |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.1  | Milgram and Kishino’s Mixed Reality on the Reality-Virtuality Continuum [133] . . . . .                                                                                                                                                                                                                                                                                                                                                             | 24 |
| 3.1  | The three views of the CollabAR Interface: nothing in view (left), model in view (middle) and the locking border (right). . . . .                                                                                                                                                                                                                                                                                                                   | 53 |
| 3.2  | Basic interactions with CollabAR . . . . .                                                                                                                                                                                                                                                                                                                                                                                                          | 54 |
| 3.3  | Technical architecture of CollabAR . The system is implemented as a web application accessed through a WebGL- and WebRTC-enabled browser on a user’s device. 3D objects are rendered locally upon detecting a marker. Real-time collaboration is enabled via Firebase, which synchronises object data across clients. The three user interaction phases – Configuration, Interaction, and Disengagement – are illustrated along the bottom. . . . . | 55 |
| 3.4  | The six models of varying complexity used for the tasks. . . . .                                                                                                                                                                                                                                                                                                                                                                                    | 57 |
| 3.5  | The study set up in a meeting room . . . . .                                                                                                                                                                                                                                                                                                                                                                                                        | 58 |
| 3.6  | The timeline visualises the coding of the video recording per group according to observed categories. Tasks are labelled using a session-based scheme (e.g., 1.1, 1.2, 1.3) where the first digit indicates the model level and the second digit the task number. Tasks x.1 and x.2 correspond to counting tasks. Tasks x.3 represent comparison tasks, which introduced a second model. . . . .                                                    | 59 |
| 3.7  | Task completion time for each task. . . . .                                                                                                                                                                                                                                                                                                                                                                                                         | 62 |
| 3.8  | Overall occurrences for each group. . . . .                                                                                                                                                                                                                                                                                                                                                                                                         | 63 |
| 3.9  | Statistics of the inspection tasks shown as figures. The top figure is a graph showing the pairwise comparisons and significance between each level. The bottom figure is the table of the mean values and the F-statistic for each metric. . . . .                                                                                                                                                                                                 | 64 |
| 3.10 | Statistics of the comparison tasks shown as figures. The top figure is a graph showing the pairwise comparisons and significance between each level. The bottom figure is the table of the mean values and the F-statistic for each metric. . . . .                                                                                                                                                                                                 | 65 |

|      |                                                                                                                                                                                                                            |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.11 | Disjoint and Distributed View was the second most common collaboration style. . . . .                                                                                                                                      | 67  |
| 4.1  | The five research categories looked into throughout this chapter are demonstrated in the study set-up. . . . .                                                                                                             | 77  |
| 4.2  | This figure illustrates the different combinations of device configurations between participants, including pairings of phones, tablets, and no device. . . . .                                                            | 80  |
| 4.3  | Basic interactions with the system. Rotation is controlled by swiping the object and scaling via pinch gestures, as is common in many commercial AR apps. . . . .                                                          | 81  |
| 4.4  | The two different models used for each task. One with more occlusion (Cube, Left) and one with less occlusion (Sphere, Right). . . . .                                                                                     | 82  |
| 4.5  | The study was set up in a circular meeting room portraying a pair of participants with the Phone + Phone configuration. Virtual models were projected using standard marker tracking. . . . .                              | 84  |
| 4.6  | Median responses for the NASA-TLX questionnaire metrics and overall perceived workload, along with K-Related samples chi-square results, where the p-values report the significance between device configurations. . . . . | 86  |
| 4.7  | The mean task time for each device configuration to complete the tasks. . . . .                                                                                                                                            | 88  |
| 4.8  | A total of all errors across all device configurations and models occlusions on the ‘Identify’ task. . . . .                                                                                                               | 89  |
| 4.9  | A timeline detailing how focus, communication and task completion time varied across groups. Different shades of the same colour represent the two different participants. . . . .                                         | 90  |
| 4.10 | The average percentage of participant focus during all tasks. ‘Focus on other’ in this case refers to participants being idle, such as looking around the room. . . . .                                                    | 91  |
| 4.11 | The average percentage of time participants spent talking in each task. . . . .                                                                                                                                            | 92  |
| 4.12 | Median scores on a 5-point Likert scale of participants’ individual self-reflection on their communication. A dashed border is present to show significant differences in ‘Consider’ and ‘Understood’ questions. . . . .   | 94  |
| 4.13 | The average amount of rotations per task on occluded and non-occluded models. . . . .                                                                                                                                      | 95  |
| 4.14 | The average amount of collaboration strategies adopted over each device configuration. . . . .                                                                                                                             | 97  |
| 5.1  | The design space of collaborative AR interaction techniques. The 10 interaction techniques explored range from aligning with reality to beyond reality. . . . .                                                            | 112 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |     |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.2 | In Freeze Frame Alice freezes a perspective on the AR content (A1). Alice can now use micro-mobility to share the perception of the AR object with Beth with flexibility in interpersonal distance and orientation (A2). . . . .                                                                                                                                                                                                                                                     | 114 |
| 5.3 | In Remote Peeking, Beth can peek at Alice's perspective (A1) by selecting the position of Alice's device (B1). Beth can now swap between seeing Alice's perspective (B2) and her own perspective (B3) in focus. . . . .                                                                                                                                                                                                                                                              | 115 |
| 5.4 | Alice can create a Spatial Reference at a desired perspective (A1). An AR object appears pointing from the previously selected perspective (A2). Beth also sees the spatial reference (B2), and when Alice points towards the AR object, Beth knows what she is referencing with help from the spatial reference (B2). . . . .                                                                                                                                                       | 117 |
| 5.5 | In Remote Sketch, Alice can freeze a frame and share it with everybody (A1). She sketches the frame in order to explain a detail (A2). Beth sees the frame and follows Alice's sketch in real-time. . . . .                                                                                                                                                                                                                                                                          | 117 |
| 5.6 | In Proximity Threshold, users are able to control AR objects when being close by. Beth is manipulating an object in front of her (B1). Alice only sees that some object is there from a distance, but she does not see details and has no access to manipulate the object (A1). When she moves closer, the object appears, and she can now manipulate the object (A2). . . . .                                                                                                       | 118 |
| 5.7 | In Privacy Fence, users can create their own private territory anywhere. Alice creates a Privacy Fence next to a shared object (A1-A2). Alice works on her private object (A3), while Beth still only sees the shared object (B1). . . . .                                                                                                                                                                                                                                           | 119 |
| 5.8 | A closely coupled collaboration between architects in an internal design meeting. A) The architects view and discuss the overall design of the building. B) They shift into discussing a detail, one peeking over the shoulder (Basic AR Perception), while the others use Remote Peeking to see the detail. C) One creates a Spatial Reference highlighting the detail. D) Later, one of the architects points towards the Spatial Reference to reference the detail again. . . . . | 120 |
| 5.9 | A meeting between architects (green) and clients (yellow). A) The architects and clients discuss the overall design of a building. B) One of the architects shows details to everyone using Perspective Broadcasting. One of the architects breaks out of the closely coupled collaboration and creates a Privacy Fence to work on a new design without interrupting the others' conversation. . . . .                                                                               | 121 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |     |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.10 | Classroom scenario where a teacher (yellow) walks around to groups of students to help with their AR exercises. A-B) The teacher walks within the Proximity Threshold to a group to see their AR content. C) The teacher uses Freeze Frame to show a problem to a student. D) The teacher shows the problem to the entire class by broadcasting a live sketch (Remote Sketch) of the problem to the class while explaining.                                                                                                                                                                                                                                                                                      | 122 |
| 6.1  | A sketch of the conceptual Collaborative Augmented Reality Interaction Space (CARIS) with each of the components highlighted and labelled. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 133 |
| 6.2  | The three foundational elements of CARIS: Physical Features (left), Interaction Portals (middle), and People (right). . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 134 |
| 6.3  | Two of the core features of the CARIS system implementation. The left figure shows a full and empty plot, and the right figure shows the Geofence placement. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 140 |
| 6.4  | Example flow of information during block placement in CARIS . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 142 |
| 6.5  | Illustration of input actions in the CARIS system. (a) <b>Point and Click</b> : Users interact by aligning their device's crosshair with an object or another portal and tapping, allowing for selection without direct touch input. (b) <b>Bring Portals Together</b> : Users initiate interaction by moving two AR portals into close proximity, triggering a spatial event. (c) <b>Bring Together and Shake</b> : In addition to proximity-based interaction, shaking the device while portals are near each other activates a distinct input event. These techniques leverage spatial positioning, device movement, and proxemic interactions to enable more natural and intuitive user experiences. . . . . | 144 |
| 6.6  | Demonstration of Perspective Peeking. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 145 |
| 6.7  | Distribution of "Perspective Peek" Interactions Over Time. Each dot represents a unique instance of the interaction, with its vertical position indicating the duration. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | 148 |
| 6.8  | Boxplot visualising the distribution of immersion percentages across all participants. The central line within the box represents the median immersion level. The interquartile range (IQR), indicated by the width of the box, demonstrates the range within which the middle 50% of participants' immersion levels fall. The whiskers of the boxplot extend to the minimum and maximum values, providing a comprehensive overview of immersion variations. . . . .                                                                                                                                                                                                                                             | 149 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 6.9  | Timeline illustrating the context switches of four participants (P1 to P4). Each horizontal bar indicates a period during which a participant disengaged from the immersive experience, while the crosses represent momentary switches. The highlighted regions, shown in a light yellow overlay, indicate periods with a significantly high number of context switches based on a 2-minute rolling average. . . . .                                     | 150 |
| 6.10 | Rolling average of context switches over a 2-minute window for each group (G1 to G5) during the AR task. The graph illustrates the frequency of context switches, with each group represented by a distinct line. The rolling average offers insights into the periodicity and consistency of participants' engagement and transitions during the task. . . . .                                                                                          | 151 |
| 6.11 | System Usability Scale (SUS) Analysis. (a) Average adjusted scores for each of the ten SUS statements, highlighting the variability in participant responses to different usability aspects of the system. S5, related to the integration of functions, stands out with a notably high score. (b) Distribution of overall SUS scores across participants, showcasing a varied user experience with notable peaks around scores of 32.5 and 87.5. . . . . | 154 |
| 6.12 | Distribution of NASA-TLX scores across six dimensions of perceived workload during an AR-based co-creation task. Each bar represents the mean score, with error bars indicating the standard deviation. This visualisation highlights variations in participants' experiences across different workload dimensions. . . . .                                                                                                                              | 155 |

# List of Tables

- 2.1 CSCW Time-Space Matrix based on Johansen’s framework [98]. . . . 34
- 4.1 Logistic Regression Predicting Likelihood of Task Error based on Device Configuration and Model Occlusion. . . . . 89
- 6.1 Non-Immersion Times and Immersion Percentages for Each Participant 152
- 6.2 Summary of System Usability Scale (SUS) Scores, showing the average score and standard deviation for each statement. . . . . 153



# List of Abbreviations

|             |           |                                     |
|-------------|-----------|-------------------------------------|
| <b>MR</b>   | . . . . . | Mixed Reality                       |
| <b>AR</b>   | . . . . . | Augmented Reality                   |
| <b>VR</b>   | . . . . . | Virtual Reality                     |
| <b>CSCW</b> | . . . .   | Computer-Supported Cooperative Work |
| <b>HCI</b>  | . . . . . | Human-Computer Interaction          |
| <b>HMD</b>  | . . . . . | Head-Mounted Display                |

# 1

## Introduction

### Contents

---

|     |                     |    |
|-----|---------------------|----|
| 1.1 | Research Challenges | 4  |
| 1.2 | Research Questions  | 10 |
| 1.3 | Research Context    | 13 |
| 1.4 | Methodology         | 14 |
| 1.5 | Contributions       | 17 |
| 1.6 | Thesis Structure    | 19 |

---

Augmented Reality (AR) displays are devices that allow users to superimpose computer-generated images and information onto the real world. Given the widely accepted definition by Azuma [5], AR displays are defined by three main characteristics: (i) they combine the real and virtual worlds, (ii) they are interactive in real-time, and (iii) they are registered in a 3D environment [5]. Traditionally, HCI has focused on the ways humans interact with computers and digital interfaces, focusing on usability, user experience, and cognitive aspects of interaction. AR, however, extends these paradigms by introducing a dynamic interplay between the physical and virtual worlds, thereby reshaping our understanding of interaction spaces and modalities. As Tönnis et al. [202] articulates, AR involves a continuous dynamic interplay between two distinct but intertwined worlds: the physical and the virtual. Each of these worlds has its own geometric and photometric properties and evolves over time. Their relationship is intricate and reciprocal, and one world often adapts or influences the other [202]. In AR, intricate interactions and information exchanges between

individuals not only give meaning to the virtual components of the environment but also play a role in shaping the physical aspects, highlighting the symbiotic relationship between the virtual and the real [175, 197]. Contrary to traditional user interfaces, this interaction is not a mere juxtaposition of virtual elements onto the real world but weaves interconnections that create an enhanced, more informative reality. This allows AR to provide more information about the user’s physical environment than human senses would otherwise detect and offers additional support through spatial, visual, and auditory cues [18, 28, 166]. Ultimately, AR acts as an “intelligence amplifying system” by Azuma’s standards, highlighting the symbiotic relationship between the virtual and the real in shaping our environment and interactions [5].

The integration of virtual information within the physical world allows for a seamless blend, where interactions between the two realms flow naturally, enhancing user experience. Augmented Reality (AR) is most effective when users perceive the physical and virtual worlds as a cohesive whole, as if they were one [109]. Recent advancements in digital technologies have facilitated new methods for blending these worlds, improving the realism and integration of virtual content with real environments [100]. This highlights the bidirectional relationship between the virtual and physical realms, where actions and behaviours in one domain influence experiences in the other [53, 168]. Furthermore, the concept of the metaverse exemplifies this seamless integration, merging the physical world and virtual spaces through sophisticated computing and communication infrastructures [86, 215]. This convergence underscores the transformative potential of AR in creating interactive, informative, and immersive environments.

Interactivity in immersive experiences holds the potential to enhance the user experience by reducing the cognitive load associated with transitioning between realities. Immersive experiences involve deep engagement with virtual environments, which promotes a sense of presence and increased interaction [148, 194]. This high level of immersion can improve the user experience by enveloping them within a digital interface, as observed in virtual reality applications [47]. However, one of the distinct advantages of AR over Virtual Reality (VR) is the reduced need for users to learn to navigate a new environment [82]. AR allows users to interact with content added in real or virtual environments, leading to a more positive perception of AR content [16, 90]. In AR, the environment is familiar and tangible, filled with objects that users can physically see and interact with. This familiarity lowers the barrier to entry, making AR more accessible to a wider range of users [37, 141]. In recent years,

marked by rapid technological advancements, AR has emerged as a potent tool in the digital landscape. Technology corporations such as Microsoft, Meta, and Apple are investing heavily in this domain, launching increasingly sophisticated hardware to support the ever-expanding AR market. Their efforts reflect the growing recognition of the potential of AR to revolutionise various aspects of our lives, from work [2, 21] and education [203, 225] to entertainment [110, 112] and social interaction [174].

Ad hoc situations that use AR can be developed and deployed on almost any smart handheld device, broadening the potential user base [57]. This thesis focuses on handheld mobile AR displays due to their wide accessibility and versatility in various applications. Handheld mobile AR refers to AR systems that operate on widely accessible mobile devices, particularly smartphones and tablets. While AR research dates back to the 1960s [5], advancements in mobile technology have been instrumental in the rise of handheld AR. The proliferation of touchscreen smartphones has made this the most prevalent form of AR display today [106], driving its widespread adoption among the general public [57]. This is supported by a growing number of users of mobile Augmented Reality AR worldwide, which is estimated to have been 1.03 billion in 2024.<sup>1</sup>

Given that handheld mobile AR is already supported by widely available devices, it is important to explore how this technology extends beyond individual use, enabling new forms of interaction and collaboration. One of AR’s most promising applications is its ability to facilitate shared experiences and collective work. As Ens et al. states, “Among the possible applications in AR, it is widely viewed that collaborative systems are to be among its killer applications” [46]. This highlights AR’s potential in transforming collaboration by integrating virtual elements into the physical world, creating shared spaces where information is not just presented but actively interacted with, enhancing collective understanding and engagement.

In this context, some foundational concepts within HCI such as proxemics [75], territoriality [184], and shared interaction spaces [125] gain new significance. Proxemics have become crucial in designing AR experiences that respect personal space while enhancing interaction. Territoriality, or how people use physical space to establish personal boundaries, could inform the design of AR interfaces that seamlessly integrate into the user’s environment without infringing on their sense of personal space. Cross-device computing and shared interaction spaces are also essential for enabling collaborative experiences in AR, using technology to break down barriers between individual and collective digital interactions.

---

<sup>1</sup><https://www.statista.com/statistics/1098630/global-mobile-augmented-reality-ar-users/>

In this thesis, I aim to gain a deeper understanding of how mobile handheld AR is experienced and interacted with in a co-located group scenario. By exploring the dynamic interplay between the physical and virtual worlds within group settings, we can identify key factors that enhance collaborative efforts and user engagement. This approach will provide invaluable insights for developers and designers, enabling them to create more effective and engaging AR applications. Through this research, we aspire to push the boundaries of AR technology, fostering innovations that transform how we perceive and interact with both our immediate individual environment and each other in co-located group scenarios.

## 1.1 Research Challenges

In contrast to other established areas of co-located collaboration, such as tabletop computing, there are few established guidelines or best practices when it comes to designing and developing co-located collaborative Augmented Reality (AR) scenarios [117, 120]. Although previous research suggests that AR interfaces can support collaborative work [71, 188], most handheld mobile AR interfaces are built primarily with a focus on individuals rather than collaboration. Some work explores co-located collaboration for mobile AR, but there is generally a lack of understanding of how mobile AR affects collaboration and group interactions. This lack of established guidelines highlights the need for further research and development in this area to enhance the effectiveness and user experience of co-located collaborative AR applications.

This thesis addresses this overarching challenge by building on a set of four core research challenges initially motivated by prior research on general co-located collaboration in HCI as well as observations of the usage of typical co-located AR systems in group settings. These research challenges are:

1. **Co-Located Collaboration in AR** – Understanding how AR shapes collaboration and group interactions in co-located settings.
2. **User Interactions** – Investigating how interaction design in mobile AR influences collaborative experiences and cognitive load.
3. **Group Dynamics in Mobile AR Collaboration** – Exploring the social and spatial factors that impact collaborative engagement in AR environments.

4. **Describing Multi-Device Mobile AR Interactions** – Developing conceptual frameworks to describe and optimise interactions in multi-device AR settings.

These four challenges are interrelated, as they collectively address different aspects of collaboration in mobile AR. Co-located collaboration in AR (Challenge 1) establishes the foundation by exploring AR’s role in enhancing collaborative interactions. User interactions (Challenge 2) build upon this by examining the specific mechanisms that shape these interactions. Group dynamics (Challenge 3) further contextualises collaboration by focusing on how spatial and social factors influence user engagement. Finally, multi-device mobile AR interactions (Challenge 4) extend these ideas by considering how multiple AR-enabled devices can be used to enhance or complicate collaboration.

### 1.1.1 Co-Located Collaboration in AR

While AR holds immense potential to enhance face-to-face collaborative experiences [15, 120], the dimension of collaboration and, more specifically, the effects that AR has on group collaboration remains under explored. Although AR has been shown to support both co-located and distributed collaboration, there is a need to delve deeper into its role in co-located collaborative scenarios. Collaborative AR has been found to reduce cognitive load, task durations, and errors while enhancing learning and facilitating training, maintenance, and education [80, 97, 116, 142]. Studies on synchronous collaboration in mixed reality settings suggest that users prefer synchronous, co-located collaboration [71].

Early systems like TransVision [165] and Studierstube [199] demonstrated AR’s capacity to facilitate collaborative design and visualisation using handheld displays and AR glasses. Butz et al. [27] further explored co-located interaction with shared virtual information using head-worn displays. More recently, Benko et al. [11] developed a projection-based AR system that overlays 3D objects into the real world, highlighting the evolving nature of AR in collaborative contexts. Handheld mobile AR systems such as AR Pad [137] and game scenarios explored by Henrysson et al. [81] underscore AR’s potential in face-to-face collaboration. However, some commercial mobile AR applications, such as Ingress [170] and IKEA Place [115], primarily involve implicit and secondary collaborative aspects, often leading to isolated user experiences. This underlines the need to better understand AR’s mediating role in co-located collaborative scenarios.

Touch-based interfaces on phones, tablets, and tabletops offer a more advanced understanding of co-located interactions and collaboration practices. Early work in this area examined closely coupled collaborations using shared stationary displays [42, 77]. Interactive tabletops, which naturally afford group work, have led to guidelines for constructing collaborative work environments [184], and studies have explored concepts like territoriality [186], group territory [185], and collaborative coupling strategies [200]. Mobile devices bring significant advantages for users requiring mobility, as demonstrated by Lucero et al. [118, 119], who showed how these devices enable co-located collaboration for tasks like photo sharing and brainstorming. Research indicates that collaboration and communication positively impact team performance [95], and mobile devices can enhance group collaboration [187]. However, in small group settings, tablets may negatively affect communication [74]. Technologies like HuddleLamp [161] and Connichiwa [182] reflect efforts to better understand co-located interaction [49, 121], emphasising the need to examine how device size and configuration influence collaboration in handheld mobile AR.

While prior research has identified key aspects of collaboration, there is still a *lack of integration* between findings in HCI, AR, and multi-device interaction research. A practical framework is needed to *translate theoretical insights into actionable design principles* for real-world AR applications.

### 1.1.2 User Interactions

User Interactions play a critical role when it comes to shaping the effectiveness and fluidity of collaborative efforts in Human-Computer Interaction (HCI) and are particularly important in the context of mobile Augmented Reality (AR). Traditionally, user interactions have been explored through a variety of different lenses, such as collaborative behaviours and the transition between individual and group workspaces. For example, Scott et al. emphasised the importance of supporting collaborative behaviours through simultaneous interactions and smooth transitions between personal and group activities in tabletop workspaces [186]. In the realm of larger interactive displays, Rogers and Lindley highlighted how the orientation of a display (horizontal versus vertical) impacts idea exploration, role switching, and collaborative behaviours in group settings [169]. Similarly, Tang et al. focused on how users maintain independent actions while working on shared surfaces, a dynamic known as collaborative coupling, which highlights the fluidity of interactions in co-located settings [200].

These studies collectively contribute to the understanding of user interactions within HCI, particularly in co-located collaborative settings. By exploring various dimensions of collaborative behaviours, display orientations, cross-device interactions, and other nuances, researchers can gain insight into optimising user experiences and interface designs in these collaborative scenarios. However, mobile AR presents distinct challenges in terms of user interactions, particularly because of the nature of handheld devices and the attention splitting required between physical and virtual elements. In mobile AR, attentional tunnelling, where users become overly focused on the virtual content to the detriment of real-world engagement, is of concern [198]. The small screen sizes and the requirement to navigate both real-world and virtual objects can lead to disruptions in the collaborative flow. Research shows that when users frequently switch between real-world interactions and virtual tasks, cognitive load increases, affecting performance [54]. In the realm of Cross-Device and Ubiquitous computing, Marquardt et al. investigated cross-device interaction through micro-mobility and F-formations, introducing the concept of using sociological constructs to facilitate interactions across multiple devices [126]. These concepts are especially relevant to mobile AR, where collaboration often depends on how individuals coordinate between their devices and physical movements in the space.

While there is a wealth of research on various aspects of HCI, including collaborative work and AR interfaces [41, 46, 120], there remains a gap in understanding how mobile AR specifically impacts collaboration and group interactions. Whilst this existing research provides a strong foundation for the understanding of user interactions in collaborative HCI, there remains a gap in how mobile AR specifically mediates these interactions. As AR technology continues to evolve, understanding the intricacies of user interactions within this context will be crucial in designing systems that optimise collaboration and enhance both individual and group performance.

Unlike tabletop or large-screen collaborative systems, mobile AR *introduces a tension between individual and group work*. This tension refers not necessarily to a flaw, but to a design challenge arising from contrast between the personal, device-bound nature of handheld AR and the demands of group collaboration. While mobility offers flexibility, it also *reduces shared awareness and synchronised engagement*, necessitating new interaction paradigms. Additionally, little work explores *how design interventions (for example, interface simplifications and multimodal feedback) might reduce these burdens in collaborative settings*. This stands in contrast to tabletop systems, where shared displays promote collaborative coupling and natural territorial behaviours that support fluid transitions between individual and joint tasks [200].



### 1.1.3 Group Dynamics in Mobile AR Collaboration

Group dynamics in handheld co-located mobile AR collaborations pose intricate challenges and opportunities. Context switching, spatial movements, and the balance between focusing on one's device and engaging with others are examples of key elements that shape group dynamics [54, 198, 214]. For example, context switching in such environments can affect user attention and interactions, and research has shown that context switching and focal distance switching in AR displays can affect human performance and user engagement with collaborative context [54, 198]. Users may face difficulties in managing tasks when encountering virtual content at different distances, leading them to concentrate mainly on their handheld AR devices, potentially decreasing interaction with collaborative content and the other group members. Research on student flow experiences during a mobile AR science game [23] and on human-robot collaboration through an AR approach [61] offer valuable insights into how spatial interactions and collaborative experiences are shaped within group settings. Other research has also emphasised the importance of investigating collaborative work in AR settings to understand the current research status in AR-supported collaborative work [188]. Adding to this challenge, a study on the affordances of AR in science learning provides insights into how technology can influence user interactions and collaborative work in groups, underscoring the intricate web of factors that impact co-located collaboration in AR environments [34].

Investigations of group dynamics in AR environments underscore the importance of considering spatial interactions to enhance collaborative experiences. These studies emphasise the need to explore how group dynamics influence collaborative dynamics within mobile AR contexts. Although these studies contribute to a nuanced understanding of the interplay between group dynamics, such as context switches, spatial movements, and device usage in mobile AR collaborations, there is a notable absence of specific recommendations for improving collaboration within these environments. Existing research has yet to comprehensively explore the *role of device proximity, alignment, and affordances* in shaping collaborative interactions. This research will *propose new models for orchestrating device interactions* based on spatial configurations and contextual awareness.

### 1.1.4 Describing Multi-Device Mobile AR Interactions

The rapid evolution of handheld mobile devices has fundamentally transformed the AR landscape, enabling diverse new interaction possibilities [19, 48, 57, 79]. This evolution has become increasingly pivotal in collaborative work and social interactions, highlighting the need to redefine our understanding of user interactions within multi-device contexts. Studies have shown the benefits of using multiple devices in collaborative decision-making and sense-making activities, leading to shared understanding, mutual awareness, and more closely coupled collaboration [25]. This change shows a need for a revised perspective that considers the interplay between multiple users and devices and the seamless integration of virtual content.

Previously, concepts and terminologies focused on isolated user-device interactions, such as those primarily concerned with single-user usability or device-centric on-screen interaction metaphors, were considered sufficient for describing user engagement with AR. However, with the emergence of collaborative and cross-device engagements, these terms are now inadequate, failing to capture the full complexity of group dynamics and interactions. This inadequacy is evident in issues such as reduced shared awareness, increased cognitive load from continuous device handling, and communication breakdowns in co-located AR settings. To address this gap, it is crucial to develop a vocabulary that can accurately represent the intricacies of these interactions. Additionally, exploring the design space for cross-device mobile handheld AR interactions is essential, and researchers have examined various 3D manipulation techniques in AR interfaces, highlighting the importance of understanding and improving interaction techniques to optimise user experience [57]. The design space encompasses various dimensions, such as the spatial arrangement of devices, the synchronisation of content, and the real-time coordination of user interactions. Understanding these dimensions is vital to creating AR experiences that are not only technologically feasible but also engaging and meaningful for users.

Navigating the complexities of multi-device, handheld mobile AR interactions shows the need for innovative concepts, vocabulary, and frameworks to enhance user experiences and address challenges associated with this technology. For example, Blattgerste et al. introduces the TrainAR system that presents a didactic framework that is tailored for procedural training tasks in handheld AR environments, providing a way to guide interactions in such settings [20].

Although prior research highlights the potential of multi-device environments, *there remains a lack of structured approaches to designing and assessing multi-device*

*interactions in AR.* Current methodologies often overlook *how spatial configuration, real-time synchronisation, and cross-device coordination impact collaboration.* Furthermore, existing evaluation frameworks rely on usability-focused metrics, which fail to capture the unique affordances of collaborative AR experiences. Addressing these gaps is essential to developing a comprehensive understanding of multi-device AR interactions and their role in co-located collaboration.

### 1.1.5 Central Challenge: Designing Effective Co-Located Mobile AR Collaboration

From these four challenges, it is clear that there is an *absence of a structured framework for designing and understanding collaborative AR experiences.* More specifically, a central overarching challenge emerges: *how can we design and implement effective co-located mobile AR collaboration that balances interaction design, group dynamics, and multi-device integration while mitigating cognitive and spatial constraints?*

This central challenge encapsulates the need for a holistic approach to AR design that considers technological, cognitive, and social dimensions simultaneously. By addressing this challenge, we aim to bridge the gap between individual-focused AR applications and fully collaborative AR environments that enhance group interactions and productivity. Furthermore, developing a framework for designing collaborative AR experiences will provide researchers and practitioners with guidelines to create more effective and immersive co-located AR interactions.

## 1.2 Research Questions

The research challenges outlined in this thesis highlight several gaps in our understanding of co-located mobile AR collaboration. They also indicate a need to reconceptualise our understanding of mobile AR interactions within collaborative settings. Insights from seminal studies (e.g. [169, 186, 200]) have shed light on the nuances of supporting collaborative behaviours and the effects of display orientations on collaborative processes. Moreover, research by Marquardt et al. has deepened our understanding of cross-device interactions, emphasising the role of sociological constructs in enhancing device interplay [126]. Despite these advances, the specific impact of mobile AR on collaboration introduces new complexities that challenge our existing research and terminologies.

The integration of mobile AR in co-located collaborations presents unique challenges, particularly in navigating the intricacies of group dynamics. Studies have

illustrated how context switches and spatial movements influence user engagement and collaborative efficacy [54, 198], highlighting the need for a refined understanding of how AR can facilitate and hinder co-located group interactions. Furthermore, exploration of the benefits of AR in learning and collaborative work settings [34, 188] underlines the role of physicality and contextual factors in shaping these interactions.

While these challenges span multiple dimensions, ranging from user interactions to multi-device integration, not all aspects can be addressed within the scope of this thesis. To ensure a focused investigation that meaningfully advances the field, three areas where empirical insights and theoretical contributions are most needed have been prioritised:

- **Understanding Group Dynamics in Mobile AR Collaboration** – This is directly motivated by **Challenge 3: Group Dynamics in Mobile AR Collaboration**, which emphasises the need to explore how spatial and social factors influence user engagement. Existing research highlights context switching and attentional shifts as key challenges in co-located mobile AR interactions, yet little is known about how these manifest in real-world collaborative settings. This understanding also fundamentally informs **Challenge 1: Co-Located Collaboration in AR**, by detailing how AR shapes group interactions.
- **Exploring Device Size and Configuration in AR Collaboration** – This stems from **Challenge 4: Describing Multi-Device Mobile AR Interactions**, which highlights the need to understand the spatial and technological configuration of AR systems in collaboration. Prior work in cross-device computing has demonstrated that device characteristics (e.g., size, orientation, input methods) significantly impact collaborative behaviours, but these effects remain largely unexplored in AR. This investigation also contributes to **Challenge 2: User Interactions**, by examining how device characteristics influence collaborative experiences and cognitive load.
- **Developing a Design Framework for Co-Located Mobile AR** – This area directly addresses the **Central Challenge: Designing Effective Co-Located Mobile AR Collaboration**, which highlights the absence of a structured framework for designing collaborative AR experiences. It aims to synthesise these empirical insights and theoretical contributions into a practical, structured framework, ensuring that the findings can be translated into actionable design principles to guide future AR system development. By

focusing on this framework, the thesis provides a concrete solution to a core aspect of the Central Challenge, making its broader implications actionable.

Building on these focus areas, in this thesis, I aim to address key gaps in our ability to describe, evaluate, and enhance co-located mobile AR collaboration. Specifically, this thesis contributes to fulfilling these needs by providing:

1. Empirical research demonstrating how handheld mobile AR can effectively mediate collaboration in co-located settings.
2. A vocabulary that captures the complexities of AR-enhanced collaboration across multiple devices.
3. A framework that translates theoretical insights into actionable design principles to guide future AR system development.

By addressing these aspects, this thesis moves beyond device- or domain-specific investigations towards a more holistic understanding of the collaborative ecosystem shaped by AR technologies. As such, the superordinate research question driving this work is:

*How can insights from co-located mobile Augmented Reality collaboration inform the design of future Augmented Reality systems?*

The question seeks to bridge the identified gaps by focusing on the conceptual and empirical dimensions of AR technology’s influence on collaboration, aiming to enrich our understanding of AR’s role in collaborative environments and the current landscape of AR research and application in co-located collaborative settings. To achieve this, I define three sub-research questions, each building on insights from prior research and advancing our understanding of AR’s impact on co-located collaboration. The research questions are as follows:

**RQ1 How do group dynamics and practices affect co-located collaborations enhanced by handheld mobile AR applications?**

The first research question seeks to assess the impact of handheld AR on group dynamics and user practices in co-located collaborative settings. Handheld AR applications provide affordances that support shared views and the blending of physical and virtual content; however, a detailed understanding of AR’s effects on natural co-located group interactions remains underexplored. RQ1

investigates how group behaviours unfold in a basic handheld mobile AR environment, identifying potential obstacles to collaboration and offering initial design recommendations to enhance user experiences.

**RQ2 How does device size and configuration affect the facilitation of co-located mobile AR collaboration?**

As the use of handheld mobile devices is the current primary way that users can access AR experiences, device sizes play a crucial role in facilitating co-located mobile AR collaborations. No studies currently look at the role that devices play in guiding group collaborations in this type of AR, and we do not understand how various handheld devices affect group work. Inspired by work in cross-device computing, RQ2 aims to examine the group coordination dynamics and user interactions that are influenced by device characteristics such as device size, quantity and symmetry, and how these could affect the efficiency of collaborative work and collaborative behaviours.

**RQ3 How can we formulate design principles into a design space that enhances co-located mobile AR collaborations?**

The third research question seeks to bridge the gap between theoretical insights and practical applications in co-located mobile AR by exploring how design principles might be operationalised into a structured design space. This approach considers how proxemic concepts and other spatial dynamics might shape collaborative interactions, as well as how flexible, user-responsive interfaces could better support varying group sizes, roles, and evolving collaboration needs. By examining these aspects, RQ3 aims to lay a foundation for a design framework that could guide the development of AR systems optimised for diverse collaborative tasks, ultimately contributing to a more nuanced understanding of the interactions enabled by mobile AR.

## 1.3 Research Context

The research context of this thesis is situated within the field of Human-Computer Interaction (HCI), “a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them” [192]. The work within this thesis falls within the area of CSCW, a research area and discipline that falls within Computer Science and HCI that focuses on the design and implementation of systems that facilitate collaboration

and coordination among individuals or groups [181]. The field of CSCW considers the technical, human, and social factors that impact the development and use of collaborative systems [69]. Research within the CSCW area seeks to understand how computer systems can simplify the complexity of coordinating cooperative activities. Over the years, it has dedicated itself to exploring and creating technologies that unite people in partnerships, teams, crowds, communities, and other group formations [78].

This thesis aims to identify the gaps in *co-located handheld mobile Augmented Reality environments*, expanding our knowledge of how people use these technologies and the frictions that present themselves and proposes recommendations on how to improve them in future designs and developments.

## 1.4 Methodology

The goal of this thesis is to advance the understanding of how users collaborate within co-located AR scenarios and to introduce, design, and develop new interaction techniques and design spaces that would enhance these collaborations. The introduced research questions are primarily exploratory in nature and aim to uncover the socio-spatial dynamics of user interaction and collaboration within AR environments. To address these research questions, I adopt a mixed-methods approach [36], which combines technical, qualitative, and quantitative methodologies to provide a comprehensive analysis of collaboration in co-located AR. This approach is informed by the foundational work of Creswell and Clark [36] on mixed-methods research.

The mixed-methods research approach combines qualitative insights, such as user observations, interviews, and thematic analysis, with quantitative evaluations, including immersion scores, context-switch analysis, and usability metrics. Initially, qualitative data are gathered through user observations, interviews, and interaction logs to identify themes and patterns related to collaboration dynamics. These data are analysed through thematic coding to inform subsequent iterations of system design. In parallel, the thesis employs quantitative techniques to measure task performance, immersion, and user workload using tools like NASA-TLX and SUS. The research found within this thesis also includes the design, prototyping, and empirical evaluation of the Collaborative Augmented Reality Interaction Space (CARIS) framework. This framework serves as a testbed for the iterative development of interaction techniques, informed by both qualitative findings and quantitative data. For instance, participant feedback on interaction techniques is thematically coded to refine design principles,

while immersion and context-switch metrics provide measurable indicators of system performance and user engagement.

This mixed-methods research approach offers a holistic understanding of collaboration in co-located AR by systematically combining qualitative and quantitative methods. It ensures that both users' subjective experiences and objective performance metrics are considered, resulting in a robust foundation for developing interaction techniques and a design space that aligns closely with user needs and behaviours. This approach supports a comprehensive exploration of co-located AR collaboration, bridging theoretical insights and practical applications.

**A. CollabAR** – A preliminary study to investigate the impact of handheld mobile AR interfaces on ad-hoc co-located group practises. By conducting an observational study using a simple exemplar AR system based on current AR applications, we examined the interactions of co-located groups engaging with AR technology, focusing on how different complexities of virtual models influence collaborative practices, cognitive load, and user experience. The participants, grouped into teams of four, were observed performing a series of tasks using the AR system. This allowed the capture of rich, qualitative data on their interaction patterns, collaboration styles, and the challenges encountered. The methodological approach facilitated a grounded understanding of the actual use and implications of AR in these collaborative settings, ensuring that the findings were deeply rooted in observed behaviours and interactions. Through detailed analysis, including the examination of video and speech data, the study offers a nuanced exploration of the affordances and limitations of current AR technologies in supporting effective co-located collaboration.

**Methods:** *Observational study, system design and implementation, video recording and analysis, task analysis, surveys, semi-structured interviews, video coding, statistical analysis.*

**B. DeviceAR** – In exploring the influence of device configurations and sizes on co-located collaboration within mobile AR settings, the methodology incorporated a multifaceted experimental design. This design was structured to assess collaborative strategies, behaviours, and task efficacy across various device setups. The participants, organised into dyads, engaged in specifically designed AR tasks that required both cooperative and individual efforts. These tasks were designed to simulate common collaborative scenarios and were executed across a



range of device configurations and sizes to systematically evaluate their impact on collaboration. Data collection was comprehensive, including pre-study and post-task questionnaires to gauge participants' background and perceived workload. Semi-structured interviews were conducted to capture qualitative insights into the participants' experiences and strategies. The study sessions were recorded on video, providing a rich dataset for subsequent behaviour analysis. This dataset was coded and analysed, focusing on collaboration dynamics, including focus, communication patterns, and interaction with AR models. Statistical analysis, including Friedman tests and ANOVA, was used to interpret the data and identify trends and differences between conditions.

**Methods:** *Experimental study design, task analysis, surveys, semi-structured interviews, video recording and analysis, statistical analysis, coding scheme application.*

**C. A Proxemics-Informed Lens on Handheld Collaborative AR** – This chapter examines the socio-spatial dynamics of collaborative AR through a proxemics-informed conceptual framework, establishing the foundation for understanding co-located interaction with handheld devices in shared spaces. By synthesising insights from prior chapters on system design and device configuration in collaborative AR, we present proxemics as a lens to address socio-spatial relations, enabling a more nuanced understanding of collaborative engagement. The exploration involved mapping interaction techniques within a continuum from “aligning with reality” to “going beyond reality.” We created a series of interaction prototypes based on this framework, tested them through scenarios in architectural design and educational group work, and illustrated the many-to-many relationship between collaboration styles and spatial arrangements. This was followed by data collection, involving video analysis and user feedback, capturing how proxemic-aware techniques impact user perception, deixis, and control in AR environments. Our findings support the potential of proxemics to enrich handheld AR design by guiding configurations that foster flexible, meaningful interactions among co-located users.

**Methods:** *Proxemics-based framework development, interaction technique exploration, prototype implementation, design scenarios, user feedback collection, and video analysis.*

**D. Collaborative Augmented Reality Interaction Space** – This chapter introduces the Collaborative Augmented Reality Interaction Space framework, drawing from existing literature on co-located collaborative augmented reality (AR) and findings from studies of group interactions through handheld mobile interfaces and from a research-through-design methodology using proxemics as a lens. Using a robust methodology that integrates theoretical frameworks with empirical evaluations, the research commences with a design space exploration. These guide the creation of CARIS – a conceptual framework addressing the main challenges in cross-device, multi-user AR systems. A prototype system based on CARIS was developed and tested through application scenarios, followed by participant engagement in one such scenario: a collaborative architectural design task. Data collection included video recordings of interactions, pre- and post-task questionnaires on user experience, cognitive load assessments, and group interviews to capture qualitative insights. The data analysis, using both statistical and thematic methods, revealed collaborative dynamics within AR, assessing patterns of engagement and the effectiveness of proxemic design principles in enhancing collaboration across handheld mobile AR scenarios.

**Methods:** *Literature review, framework development, design space exploration, prototype system implementation, user study, group interviews, quantitative and qualitative data analysis.*

## 1.5 Contributions

The work presented in this thesis makes the following conceptual, methodological, and empirical contributions to HCI research:

**CollabAR: an investigation into the dynamics of collaborative practices facilitated by mobile AR technology.** By examining the impact of mobile AR interfaces on co-located group practices, this research offers insights into the challenges and limitations of current AR interfaces in supporting effective collaboration. The study identifies a nuanced relationship between AR model complexity and collaborative engagement, suggesting a possible decrease in group interaction as complexity increases. Through this analysis, the chapter contributes a set of design principles aimed at enhancing collaborative experiences in AR, addressing the identified challenges without prescribing specific technological solutions. These principles offer strategic recommendations for future developments in AR technologies, ensuring a more collaborative and user-centric approach. This contribution is pivotal

in advancing our understanding of AR’s role in collaborative settings and provides a foundational basis for improving AR interface design for co-located group work.

**DeviceAR: Examining the Impact of Device Configurations on Co-located AR Collaboration.** Following the preliminary CollabAR study, this research is grounded in a comprehensive study that explores how the use of mobile AR interfaces across various device sizes and configurations influences collaboration strategies, participant behaviour, and overall task efficacy. Through an experimental setup involving dyads of participants and employing tasks of varying complexities, the study systematically examined the subtleties of collaborative dynamics in AR-enhanced environments. Key findings of this research illuminate the nuanced differences in how participants approach and interact with virtual models, depending on the size and configuration of the devices. This contribution is notable for its detailed analysis and insightful observations on the role that device symmetry, size, and quantity play in guiding group collaboration within mobile AR contexts.

**Proxemics in AR: Investigating Socio-Spatial Dynamics in Co-located AR Collaboration.** Building on the insights from CollabAR and DeviceAR, this thesis introduces an exploration of socio-spatial dynamics in co-located, collaborative handheld AR environments through the conceptual lens of proxemics. By examining how spatial configurations – particularly perception, deixis, and control – influence co-located collaborative practices, this thesis identifies the crucial role of AR system design in shaping both individual experiences and group collaborative dynamics. The findings highlight how proxemic dimensions could impact user engagement, spatial arrangements, and interaction styles in AR, offering insights for designing systems that foster more natural, intuitive collaboration. This contribution establishes a theoretical and practical foundation for understanding how spatial relationships in AR can enhance or hinder effective collaboration, paving the way for the development of more responsive, socio-spatially aware AR systems.

**Framework and Application of Collaborative Augmented Reality Interaction Space (CARIS) in Co-located Collaborative Settings.** This chapter presents a framework – CARIS – that redefines the interaction paradigms for AR in co-located collaborative environments. By focusing on the integration of physical and digital interaction spaces, this research explores how AR can enhance collaborative dynamics beyond traditional interaction methods. A key contribution is the development of a detailed conceptual framework that includes design guidelines and interaction techniques tailored for enhancing collaborative AR experiences. These

contributions address existing gaps in AR research by providing a structured approach to designing AR systems that support complex, intuitive, and effective collaborative activities. Furthermore, the application of the CARIS framework through practical scenarios in architecture and education demonstrates its utility and adaptability in real-world settings. This research advances our understanding of AR’s capabilities in collaborative contexts and offers actionable insights for future AR system designs, aiming to improve user engagement and collaborative efficiency.

## 1.6 Thesis Structure

This thesis serves as a monograph in which I present all of my work on cross-device handheld mobile AR co-located collaboration.

**Chapter 2** provides an overview of the background of research within CSCW and AR. It provides an overview of different AR displays, as well as existing empirical and conceptual work within the field of AR research. It also discusses insights from related fields, identifying four research challenges for co-located handheld mobile AR collaboration.

**Chapter 3** addresses RQ1 by examining the role of handheld mobile AR in facilitating co-located group collaboration. This chapter introduces and explores the “CollabAR” system, a basic AR interface for mobile devices, which is designed with the basic interactions of AR systems in mind. It is used to study the dynamics of group interaction when engaging with virtual models of varying complexity. The insights gained from this empirical investigation are used to propose design recommendations aimed at enhancing collaborative experiences in mobile AR environments. The work presented in this chapter was originally published in the Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems [217].

**Chapter 4** delves into RQ2 by investigating the impact of mobile device configurations on co-located collaboration in AR environments. This chapter explores how differences in device size and configuration influence collaborative strategies, behaviours, and effectiveness through a controlled experimental study. The insights gathered contribute to a deeper understanding of the role that device characteristics play in enhancing or hindering effective collaboration in AR settings. The work presented in this chapter was originally published in [218].

**Chapter 5** partly addresses RQ3 by establishing a foundational lens for understanding socio-spatial interactions within handheld co-located AR environments. By using proxemics as a guiding theory, this chapter introduces an analytical approach

to examine the spatial, perceptual, and social dynamics essential to collaborative AR settings. The chapter systematically explores how dimensions such as perception, deixis, and control manifest in co-located interactions, particularly focusing on how these elements shape user engagement and interaction flow. This conceptual lens is then applied to inform design principles that balance individual and group engagement, providing a structured approach to facilitating effective multi-user AR experiences. Through this exploration, the chapter builds a theoretical framework that lays the groundwork for designing collaborative AR systems, offering insights that bridge theoretical foundations with practical design considerations, thus directly addressing the critical elements of RQ3.

**Chapter 6** addresses RQ3 by building upon the socio-spatial dynamics outlined in the conceptual lens to introduce a dedicated framework for enhancing collaborative interactions with AR technologies. Drawing on proxemic principles, this chapter presents a framework that redefines traditional user interactions by embedding collaborative, contextual, and spatial elements into AR design. Named the Collaborative Augmented Reality Interaction Space (CARIS), the framework transforms individual AR experiences into dynamic, multi-user engagements within shared physical and virtual environments. This approach aims to yield improvements in user engagement, communicative clarity, and collaborative effectiveness. The chapter thoroughly examines the core elements of the CARIS framework – People, Interaction Portals, and Physical Features – and illustrates its practical application through its system implementations and study. The investigation includes a discussion of deployment strategies and the potential future impact of this framework on AR systems, offering critical insights that advance the comprehension and practical development of AR for co-located collaborative use cases. This work directly addresses the essential aspects of RQ3 by building a robust structure for AR system design that fosters genuine collaborative engagement.

**Chapter 7** revisits the research questions and discusses the overall contributions of this thesis, reflecting on both the theoretical and empirical findings. This chapter describes how handheld mobile AR impacts co-located collaboration by examining the intersection of user interactions, device configurations, and spatial dynamics. It highlights how mobile AR introduces both opportunities and challenges in collaborative settings, particularly in balancing individual and group engagement within shared augmented environments. A key takeaway is the role of proxemics in structuring collaborative AR interactions, offering a socio-spatial lens that informs

new design considerations. Additionally, this discussion explores how the findings extend beyond immediate research questions, addressing broader implications for cross-device collaboration, interface adaptation, and cognitive load management in AR systems. Finally, the chapter identifies gaps in current knowledge. It outlines future research directions, particularly in refining proxemic-aware AR frameworks, enhancing multi-user interaction techniques, and scaling AR collaboration beyond handheld devices.

**Chapter 8** concludes the work of this thesis.

# 2

## Background

### Contents

---

|            |                                                                                 |           |
|------------|---------------------------------------------------------------------------------|-----------|
| <b>2.1</b> | <b>Augmented Reality Displays . . . . .</b>                                     | <b>23</b> |
| <b>2.2</b> | <b>Collaboration . . . . .</b>                                                  | <b>33</b> |
| <b>2.3</b> | <b>Cross-Device Computing in Collaborative Augmented Reality (AR) . . . . .</b> | <b>45</b> |
| <b>2.4</b> | <b>Conclusion . . . . .</b>                                                     | <b>48</b> |

---

Augmented Reality (AR) has evolved over the past few decades, moving from early conceptualisations of blending the real and virtual worlds to its current widespread use across various industries. The integration of AR into fields such as education, healthcare, and entertainment has highlighted its potential to enhance user interaction, collaboration, and engagement. This chapter delves into the foundational aspects of AR technology, examining its classification, the development of different display types, and its application in co-located collaborative environments. By exploring the intersection of AR and Computer-Supported Cooperative Work (CSCW), this chapter aims to provide a comprehensive understanding of how AR can support collaboration while also identifying the challenges and limitations that continue to shape its practical use in real-world scenarios. This context sets the stage for the research questions and investigations undertaken in this thesis, which aim to push the boundaries of AR's role in facilitating effective collaborative experiences.

## 2.1 Augmented Reality Displays

Augmented Reality (AR) displays enable users to overlay digital content onto their physical environment, enhancing perception and interaction with the real world. Azuma’s widely accepted definition of AR highlights three fundamental characteristics: AR integrates virtual and real-world elements, operates interactively in real-time, and aligns virtual objects within a three-dimensional space [5]. These displays come in various forms, including handheld devices, head-mounted displays (HMDs), and projection-based systems, each offering different affordances for user interaction and collaboration. In their paper, Azuma mentions that “it would appear to the user that the virtual and real objects coexist in the same space” [5]. These displays can be used to create immersive and interactive experiences that can enhance collaboration and communication among groups of people.

The idea of a superimposition of objects onto a real-world plane is not a new one. It dates back to the 16th century, when Neapolitan scientist Giambattista della Porta had a similar vision and created the optical illusion technique “How we may see in a Chamber things that are not” [157] in which della Porta wrote “The variety of the images that appear, proceed either from the matter or form of the glass. Crystal must be clear, transparent, and exactly made plain on both sides. And if one or both of these be wanting, they will represent diverse and deformed apparitions to our sight”. Moving forward a few centuries, Sutherland presented a Head-Mounted Display (HMD) that was able to create virtual objects using computing power. This implementation, the Sword of Damocles [196], was one of the main cornerstones of AR history. This is because objects were no longer dependent on tangibility, meaning the physical laws no longer apply to these objects. The systems required to create this illusion – even from the time of Sutherland’s prototype – have also become smaller, adaptable, mobile, and affordable. So much so that in modern times, devices and systems that allow access to AR are in the pockets of the average user at all times.

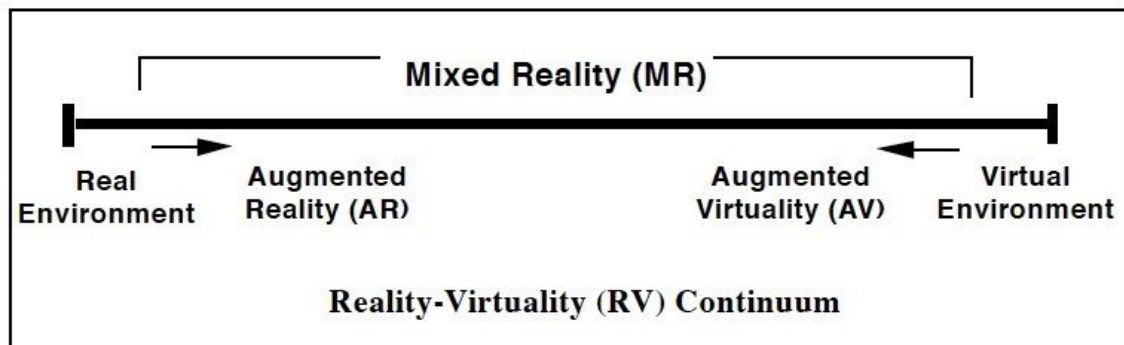
Within the last two decades, the AR industry has experienced huge growth, with large companies heavily investing in AR. In 2014, we saw the introduction of *Google Glass* [193], which, despite its quite limited success, demonstrated the potential of AR in the hands of users. Just a few years later, Microsoft announced its release of the first iteration of the *HoloLens* [132]. Whilst these are two of the most popular early iterations of AR displays, many other companies such as Magic Leap [113], Epson [139], and Vuzix [210] released AR displays that were applicable for both business and consumer applications. In addition to the advancement of



hardware is the rapid development of software and toolkits that were created, such as Vuforia [209], which allows developers to create AR applications easily. Even more recently, the introduction of software frameworks for iOS in the form of ARKit [1] and Android in the form of ARCore [59] introduce even more complex algorithms that allow for person tracking and occlusion handling based on machine learning techniques. These have led to a new wave of mobile AR applications and have made AR ubiquitous and a hot topic for research and businesses.

Today, there are different types of AR displays, including head-mounted displays (HMDs) and handheld displays. HMDs, such as the Microsoft HoloLens, Apple Vision Pro, and the Magic Leap, are worn on the head and allow users to see AR images overlaid on the real world. Handheld displays, such as smartphones and tablets, are held in the user's hand, providing a smaller and more portable AR experience. AR displays can be used in various applications, including games, entertainment, education, healthcare, and many others [44, 85]. For example, by leveraging AR technology, people who work in education can create interactive lessons that engage students in ways traditional methods can not [102]. In another example, within the healthcare sector, AR displays aid surgeons with real-time data and visual overlays during operations [131], providing touchless interaction technologies.

### 2.1.1 Classifications of Augmented Reality



**Figure 2.1:** Milgram and Kishino's Mixed Reality on the Reality-Virtuality Continuum [133]

There are numerous different ways in which AR can be classified. One of the most popular and widely used models to classify AR displays is the reality-virtuality continuum provided by Milgram and Kishino [133]. “Mixed Reality” (MR) enables dynamic blending of real and virtual elements, where users can interact with both

physical and digital objects in a seamless, integrated manner, allowing for more immersive and flexible experiences than traditional AR or VR. They suggest that AR display technologies can be classified in relation to the proportions of real and virtual world components that they present to the user (Figure 2.1). This spectrum definition places Virtual Reality (VR) in the far rightmost position because it represents an exclusively virtual environment to the user with no visual connections to the real world. Whereas, more broadly, displays that contain both real-world and virtual components are referred to as the aforementioned MR displays. One of the subtypes that arise from Mixed Reality displays is AR displays, which refer to “all cases in which the display of an otherwise real environment is augmented by means of virtual (computer graphic) objects” [133]. There has also been some research that has proposed the idea of *transitional interfaces*. These are interfaces that enable the users to easily switch between Augmented and Virtual reality displays [16, 108, 183]. These transitional interfaces allow users to move along this continuum freely during collaboration, providing different perspectives on the shared view that is provided [16].

Returning to Azuma’s view of AR, they say it should abide by three key technical characteristics. Firstly they combine the real and virtual worlds, secondly they are interactive in real time, and thirdly they are registered in a 3D environment [5]. Like Milgram and Kishino, he considers that AR is a variation of VR and lives within a reality continuum, noting that VR completely replaces the real world, whereas AR augments and enhances it. In this situation, AR gives the display more information about the user’s physical environment that human senses would not otherwise detect. By Azuma’s standards, AR systems can then be used as “intelligence amplifying systems” [24].

Other classifications [18, 28, 166] pay particular attention to how AR can alter as the user’s perspective of their own physical environment changes. According to this research, AR technologies actually *mediate* reality, allowing the virtual environment to modify the existing physical reality environment. AR has been classified in its ability to help mediate collaborative activities. Billinghurst and Kato in their research has emphasised two central advantages of using AR over other interface types within the reality-virtuality continuum. Firstly, the technology that AR provides enables seamless Computer-Supported Cooperative Work (CSCW), where users of the technology can shift their focus between their shared workspace and also their communication space in the physical environment. The assumed strength of AR builds on the same notion as Rekimoto and Nagao’s view of *augmented interaction*,

wherein a user can see the world through a portable device to “enhance the ability of the user to interact with the real world environment” [166], including the interaction with other users. Secondly, this augmented reality allows for the enhancement of the physical shared workspace to support the collaborators’ conversations. It is thus assumed that co-located collaborations can be enhanced with this augmentation, enhancing face-to-face collaborations by providing additional support through spatial, visual, and auditory cues [12, 13].

### 2.1.2 Handheld Augmented Reality

Handheld augmented reality (AR) is an evolving field that capitalises on mobile devices such as smartphones and tablets to provide accessible AR experiences, blending digital and real-world content. Unlike head-mounted displays (HMDs), handheld AR platforms offer portability and user familiarity, making AR technologies available to a broad audience. These advantages have accelerated the adoption of AR in fields such as education, entertainment, and industrial collaboration [176]. Despite its advantages, handheld AR also presents challenges in user experience and interaction design, particularly compared to the immersive potential of more specialised AR platforms.

The development of handheld AR systems has been driven by advances in mobile device capabilities, such as increased processing power, improved camera resolution, and more efficient graphics rendering. One of the foundational steps in this field was the creation of the ARToolKit, which enabled AR experiences on mobile phones by providing the essential tools for marker-based tracking and integration of 3D models into the real world [81]. This milestone allowed developers to create applications where users could manipulate and interact with virtual objects overlaid onto their physical environment.

Over time, handheld AR systems have integrated more sophisticated tracking methods, such as simultaneous localisation and mapping (SLAM), which allows for more accurate and reliable tracking of the device’s position in the environment without requiring external markers [176]. This improvement has enhanced the usability of AR applications, as users no longer need to rely on specific physical markers to trigger digital content. Mobile AR frameworks like Apple’s ARKit and Google’s ARCore have further streamlined the development of AR applications by providing standardised libraries for creating spatially aware mobile applications that integrate seamlessly with existing hardware [128].

### 2.1.3 Collaborative Handheld AR

Collaboration is a key aspect of handheld AR, which offers unique advantages. In contrast to HMDs, which may isolate users from one another, handheld devices allow for easier face-to-face interaction while simultaneously engaging with digital content. The AR Pad, one of the early handheld AR systems, exemplified this approach by using a camera and a handheld display to enable collaborative interactions around shared AR models [137]. In this system, users could interact with virtual objects within a shared physical space, making collaborative tasks more intuitive and engaging.

Recent studies on collaborative AR have highlighted the effectiveness of handheld devices in promoting group interactions. In museum settings, handheld AR has been used to foster collaboration among visitors by allowing them to explore augmented exhibits together. For example, the system designed for the Acropolis Museum adapts its content dynamically based on user interactions and behaviours, enabling groups to co-experience and discuss the AR-enhanced exhibits [103]. Similarly, educational systems have leveraged handheld AR to support collaborative learning, with students working together to solve problems or complete tasks in AR-enhanced environments [110].

### 2.1.4 Application Areas of Co-Located Collaborative Augmented Reality

Co-located collaborative augmented reality (AR) has opened up new opportunities for shared interactive experiences where multiple users interact with digital content while maintaining direct interaction with their physical environment and one another. This form of AR leverages handheld devices or other AR systems to enable users to collaborate in the same physical space, making it particularly effective in settings that benefit from face-to-face communication and shared task execution. Several key application areas have emerged as prominent users of co-located collaborative AR.

#### Education and Learning Environments

One of the most prominent areas for co-located collaborative AR is education. Handheld AR systems, such as mobile phones and tablets, allow students to interact with digital content in a collaborative setting, enhancing learning outcomes through shared exploration and problem-solving. For instance, AR-based learning tools like ScholAR allow students to visualise complex topics, such as three-dimensional

shapes and mathematical concepts, in a collaborative environment, where multiple students can view and interact with the same AR models [171]. This collaborative interaction encourages students to work together to solve problems, promoting a deeper understanding of the subject matter and improving retention of information.

In formal and informal education environments, such as museums and science centres, co-located collaborative AR enables groups of students or visitors to engage with educational content in a more interactive and immersive way. At the Acropolis Museum, for example, handheld AR is used to allow visitors to collectively explore historical reconstructions, where they can view and discuss virtual models overlaid on the physical exhibits [103]. This co-located collaboration enhances the learning experience by encouraging peer-to-peer discussion and reflection on the presented content.

### **Professional and Industrial Collaboration**

In professional settings, co-located collaborative AR has significant potential for improving teamwork and efficiency. In industrial environments, for example, AR can be used to assist with maintenance, repair, and training tasks by providing real-time information overlaid on machinery or workspaces. Technicians can collaborate using handheld AR devices to view the same digital overlays on equipment, guiding them through complex procedures and reducing the risk of errors [116]. This type of collaboration is particularly useful in industries where precision and coordination are critical, such as aerospace or automotive manufacturing.

In addition to maintenance and repair tasks, co-located collaborative AR is useful for collaborative design and prototyping. Designers can use AR to project virtual models of products or structures in a shared physical space, allowing team members to view and manipulate the same 3D models from different angles. This facilitates more effective brainstorming sessions and design iterations, as team members can physically interact with the virtual objects while discussing potential improvements in real-time [81].

### **Healthcare and Medical Training**

In healthcare, co-located collaborative AR has emerged as a valuable tool for training and surgical planning. Medical professionals can use AR to collaborate in the same physical space while viewing virtual models of anatomy or surgical procedures. For example, surgeons can use handheld AR devices to project 3D models of patient

anatomy during pre-operative planning, allowing them to collaboratively assess the best surgical approach [117]. This enhances communication and decision-making among medical teams, leading to more precise and effective treatments.

Medical training is another area where co-located AR shines, allowing students and professionals to engage with virtual simulations of medical procedures collaboratively. Interacting with AR models in a shared space allows trainees to practice procedures and discuss techniques with their peers, leading to more hands-on learning experiences than traditional methods.

### **Entertainment and Gaming**

Entertainment and gaming represent one of the most popular application areas of co-located collaborative AR, particularly due to the rise of mobile AR games like Pokémon Go. These games encourage players to collaborate or compete with others in the same physical environment while interacting with digital elements overlaid onto the real world. Co-located collaborative AR adds a social dimension to gaming, as players can engage in real-time interactions and experiences that blend the physical and virtual worlds.

Beyond gaming, AR experiences in museums and theme parks also leverage co-located collaboration. These experiences allow visitors to collectively explore AR-enhanced narratives or challenges, turning passive observation into active participation. For example, visitors can engage with virtual artefacts or characters that are superimposed onto physical environments, providing an entertaining and educational experience that can be shared with friends and family [110].

### **Cultural Heritage and Museums**

Museums and cultural heritage sites are increasingly adopting co-located collaborative AR to enhance visitor engagement. By allowing visitors to view historical reconstructions, annotations, and other digital content overlaid on real-world artefacts, museums can provide a richer narrative experience. Visitors can interact with the same virtual objects, discuss their significance, and gain a deeper understanding of the exhibits [103].

Overall, collaborative co-located AR has many potential applications in a wide range of fields, such as education, healthcare, entertainment, manufacturing, and many others. By providing a shared, immersive, and interactive environment for collaboration, co-located collaborative AR has the potential to support and enhance collaboration among people who are physically co-located in the same space.

### 2.1.5 Challenges in Handheld AR

Despite its many advantages, handheld AR faces challenges, particularly in user interaction and immersion. Unlike HMDs, which provide a hands-free, immersive experience, handheld AR requires users to hold and manoeuvre their devices, which can detract from the sense of presence and engagement. Studies have shown that handheld AR experiences may suffer from lower levels of immersion compared to head-mounted systems, as users are constantly reminded of the device as a mediator between the digital and physical worlds [117].

The limited field of view on mobile devices also poses a challenge, as users must continuously adjust the device’s position to view the augmented content. This can lead to user fatigue during extended AR sessions and may reduce the overall effectiveness of the AR experience. Moreover, the need to focus on both the screen and the physical environment simultaneously can create a cognitive burden, especially in tasks that require precise interaction with digital objects overlaid on real-world elements [171].

In addition, designing intuitive and efficient interaction methods for handheld AR remains an ongoing challenge. While touch-based interaction is the most common method, it is often less suitable for complex AR tasks that involve manipulating 3D objects or navigating large augmented environments. Researchers continue to explore alternative interaction techniques, such as gesture-based controls or hybrid methods that combine touch and spatial input, to improve the usability of handheld AR systems [175].

Handheld mobile AR represents AR systems that are based on handheld mobile devices, especially smartphones or tablets, which are popular and daily communication devices that the majority of the population has immediate access to. Early research into AR can be traced back to the 1960s [5], but it is the rise of mobile technology that has significantly propelled the field of handheld AR. Handheld AR leverages the rapid development and ubiquity of smartphones, making it the most popular type of AR display today [106]. Recent handheld mobile AR applications, such as *Pokemon Go* and *Ikea Place*, promote the use of AR, and these show the potential of handheld mobile AR in the consumer markets. Many different research papers refer to handheld mobile AR in different ways. Some use the word “mobile” AR to represent mobile/smartphone or tablet-based AR in their papers, while some other studies use the word “handheld” AR to represent mobile/smartphone or tablet-based AR systems. To avoid confusion, we will use the combination of the two, “handheld” and “mobile” as “handheld mobile” AR to represent and differentiate from “mobile”



AR that not only represents handheld mobile AR displays but also Head-Mounted Display (HMD) based AR systems [17, 87] which is not the main scope of this thesis.

Handheld AR displays utilise the device's back-facing camera to capture the physical environment, which is then augmented with virtual objects on the screen. This method, known as video see-through, provides a digitised representation of the real environment, allowing users to interact with both the physical and virtual worlds seamlessly [165, 199]. Not only does this allow for virtual presence through ubiquitous devices, but it also provides scalability, giving many users access to the virtual space synchronously. Handheld mobile devices provide different advantages that support AR applications as these devices are easy to use, cost-effective, and portable [57]. They also provide a high level of social interactivity and independent operability [57]. This makes them more useful for a variety of different applications, and the flexibility and portability of mobile devices make them ideal for AR applications, allowing users to engage with AR content on the go [211], contributing to the user's collaboration skills. Unlike traditional desktop/tabletop AR interfaces, handheld mobile AR lacks traditional input devices such as keyboards and mice [17]. The limited screen size of smartphones and tablets constrains the field of view and, consequently, the level of immersion users can experience. Additionally, the battery life of these devices poses a restriction on usage time, especially during power-intensive AR operations [11, 27]. Unlike HMDs, where the display is worn on the head, handheld AR systems require users to hold the device, integrating both display and input components into a single unit. This configuration necessitates new interaction techniques that are intuitive and effective within the constraints of handheld devices [17].

Handheld mobile AR has found applications in various industries, including tourism, maintenance, education, and medical practices. These applications showcase the versatility and practical benefits of handheld AR in real-world scenarios. For instance, handheld AR is used in educational settings to provide interactive learning experiences and in medical practices to assist in visualising complex anatomical structures [32, 105, 138, 155, 167, 190, 207].

One of the significant challenges in handheld AR is the accurate and efficient manipulation of 3D objects. This includes tasks such as selection, translation, and rotation of virtual objects within the physical environment. Accurate 3D object manipulation is crucial for applications like assembly tasks and spatially consistent alignment of virtual and real objects [22, 43, 51, 160, 216].

The focus on handheld mobile devices within this thesis is driven by their robustness and greater potential for mass-market adoption than HMDs. Handheld AR



displays are socially accepted and ergonomically manageable for casual use compared to bulkier HMDs, which are not yet fully accepted in most social environments [17]. The integrated nature of handheld devices makes them suitable for a wide range of applications, from gaming to industrial uses.

From 2005 to 2014, there was a notable shift in AR research, with a decline in studies focusing on HMDs and desktop/tabletop AR interfaces and an increase in those employing handheld mobile displays. This shift can be attributed to the growing popularity and ubiquity of smartphones [46]. Handheld AR offers unique challenges and opportunities, such as the need for innovative interaction techniques that leverage the integrated display and input components of mobile devices.

Given the widespread availability and acceptance of smartphones and tablets, handheld AR represents a promising platform for large-scale deployment and user adoption. Future research should continue to address the limitations of handheld AR, such as improving battery life, enhancing environmental tracking, and developing more sophisticated interaction techniques. This will help realise the full potential of handheld AR, making it more intuitive and accessible for a broader range of applications.

In summary, handheld AR has emerged as a dominant form of augmented reality due to the ubiquity and versatility of mobile devices. This section has highlighted the technical characteristics, challenges, and potential of handheld AR, setting the stage for further research and development in this dynamic field.

Generally, whilst there have been many advancements in the display quality and the capabilities of AR displays, more obstacles to widespread adoption have been encountered than in VR. Further, whilst both AR and VR have been ever-present in the public consciousness for a long time, the integration of AR specifically into everyday work practices has not been as successful as the adoption of personal computing devices, such as smartphones and the Internet of Things. There are many factors that have impacted the adoption of AR, but one of the main factors found, and the focus of this thesis, is the challenges and frustrations that users experience when they are interacting with AR. For all the strides that have been made through moving to markerless augmentation, the enhanced display capabilities, and the tracking available, interactions and collaborative opportunities for AR remain in their infancy.

## 2.2 Collaboration

Collaboration is everywhere and is an inherent aspect of human society. Whether it is two people trying to move a table from A to B or a group of people working on a single research project, it is the art of working together to achieve a common goal rooted in communication, trust, and shared objectives. Collaboration refers to synchronised activity that takes place in a shared context, typically toward the pursuit of shared objectives. Looking back in history, collaboration has always been a cornerstone of human civilisation. The survival and growth of ancient societies were based on their ability to collaborate effectively, whether in hunting, farming, or defence. They understood that a group of individuals working together could achieve feats that would be impossible for a single individual. This form of collaboration was primarily direct and involved face-to-face interactions and physical tasks.

One of the most significant advancements in collaboration came with the development of computer technologies. While using computers can help mediate collaboration, we still need to understand the underlying principles. A case described in Chapter 1 of Schmidt’s 2011 book *Cooperative Work and Coordinative Practices* involves moving a table from one side of the room to the other [180]. Some parts can be done individually, such as moving chairs out of the way, but moving a large table requires at least two people. He describes the coordinated actions of the two people and their need to synchronise their respective actions, for example, when to pick up the table and when to place it down. In a more defined context, collaboration involves multiple participants working together, following the same rules, norms, and structures to achieve a common goal.

With the emergence of 3D input and output devices, the perspectives of HCI design have shifted. Collaborative Virtual Environments, as well as Virtual Reality and Augmented Reality technologies, allow their users to interact with digital content in three dimensions. Because not everyone has the same skills, they must collaborate to share their skills. Collaboration is a key aspect in many fields, such as architectural design, construction, engineering, and health care. Information technology has been increasingly used to support collaboration in these domains. Within this context, the CSCW community has studied a variety of groupware. Johnson-Lenz and Johnson-Lenz defined this as “intentional group processes plus software to support them” [99]. At the turn of 1989, the term *Computer-Supported Cooperative Work* (CSCW) was defined as a “scientific discipline guiding the design and development of groupware in a meticulous and appropriate way” by Greenberg and Chang [62]. CSCW is a

multidisciplinary research focus that now encompasses the technical, human, and social factors that can vary within its implementations by the context, culture, organisation, and country that it applies to [69].

**Table 2.1:** CSCW Time-Space Matrix based on Johansen’s framework [98].

|                | Same Place                            | Different Place                   |
|----------------|---------------------------------------|-----------------------------------|
| Same Time      | Co-located synchronous collaboration  | Remote synchronous collaboration  |
| Different Time | Co-located asynchronous collaboration | Remote asynchronous collaboration |

The CSCW community has explored ways to support collaboration across different scenarios, considering both temporal and spatial dimensions. Johansen et al. [98] introduced the time-space matrix, classifying collaborative activities based on whether participants share the same time and/or the same location. This model defines four quadrants: same time, same place (co-located, synchronous collaboration), same time, different place (remote, synchronous collaboration), different time, same place (asynchronous collaboration in shared physical environments), and different time, different place (fully distributed, asynchronous collaboration). These classifications are summarised in Table 2.1.

In the context of Mixed Reality (MR) environments, where multiple users interact with both physical and virtual content, these classifications become particularly relevant. Synchronous collaboration in co-located MR settings allows users to share the same augmented space in real-time, integrating their physical presence with virtual interactions. Meanwhile, asynchronous MR collaboration leverages memory storage and decentralised networks to enable users to contribute to shared virtual spaces at different times. This thesis focuses on awareness-supporting techniques in co-located, synchronous MR collaboration, aligning with Johansen’s framework’s “same time, same place” quadrant.

Several dimensions of a collaboration environment, such as the mobility of its users, its virtual content, the roles of its users, and the visualisation hardware, are critical to consider. These dimensions are discussed by Guerrero et al. [70], who emphasise the importance of contextual attributes in mobile collaborative scenarios, including environmental factors and user goals. Understanding these aspects helps design effective collaborative systems that can adapt to various user needs and contexts, whether in co-located or remote scenarios. Additionally, work by Isenberg et al. [94] provides a comprehensive overview of collaborative visualisation, highlighting the unique challenges and requirements that arise from the intersection of computer-supported cooperative work (CSCW) and visualisation research. Combining user mobility, roles, and visualisation technology is essential for designing effective collaborative systems adapting to various user needs and contexts.

### 2.2.1 Computer-Supported Cooperative Work

Computer-Supported Cooperative Work (CSCW) is a multidisciplinary field that focuses on designing, implementing, and studying tools and technologies that facilitate collaborative work. As defined by Carstensen and Schmidt, CSCW studies “how collaborative activities and their coordination can be supported by means of computer systems” [29]. In this sense, CSCW bridges the gap between technical development and the social dynamics of collaboration, ensuring that tools designed to support teamwork are aligned with how people naturally interact and communicate.

In addition to distinguishing the cooperative arrangement between two or more people, we also distinguish the work itself. In Schmidt’s example, it is the work of moving the table set [180]. As defined by Strauss in Work and the Division of Labour, articulation work refers to the distribution and integration of tasks between groups [195]. A major reason for the wide adoption of the concept of collective intelligence is the encapsulation of superior outcomes that result from collaboration. Sociological approaches to collaboration translate into the field of computing, eventually ushering in Computer-Supported Cooperative Work (CSCW) aimed at understanding the nature of collaborative work and how computer systems could facilitate it. The development of this field was driven by the recognition that many work environments involve collaborative efforts and that computer systems could potentially support such work in powerful ways.

## Theoretical Foundations of CSCW

The roots of CSCW can be traced back to early research on groupware and the sociotechnical systems needed to facilitate collaboration. Seminal figures like Schmidt have significantly contributed to the field with concepts like “articulation work” and “coordination mechanisms” [180], which emphasise the need for understanding how people organise, share, and coordinate their work in complex environments. In particular, Schmidt’s work on coordination mechanisms highlights how shared artefacts, such as notes or diagrams, can serve as important tools for facilitating collaboration by enabling mutual awareness and reducing misunderstandings during task coordination.

Building on these ideas, Greif first coined the term CSCW at a workshop in 1984, marking the formal recognition of a field dedicated to integrating computers into collaborative work practices. Their pioneering work set the stage for the study of the technologies used to support collaboration and the social and organisational factors that influence the effectiveness of these tools [64].

Another major theoretical development in CSCW is the concept of affordances, which refers to the properties of an object that suggest how it should be used. Affordances are essential in understanding how users interact with collaborative tools, as they guide users toward natural ways of using these tools in shared environments [111]. For example, in collaborative AR systems, affordances like visual cues or shared annotations can help users understand how to manipulate digital content effectively.

## Key Developments in Collaborative Software

CSCW has seen a variety of technological advancements that have shaped how people collaborate. Collaborative software, such as Google Docs, has revolutionised how groups work together by allowing multiple users to edit documents in real-time. This technology embodies key CSCW principles such as awareness and shared workspaces, where users are constantly aware of others’ contributions and actions. Schmidt’s work on articulation finds relevance here, as collaborative software often includes features like version control and annotations that help manage and resolve conflicts in a seamless way [181].

Similarly, video conferencing and email have become essential tools for collaboration in distributed teams. These technologies reduce the barriers of time and space, allowing teams to collaborate regardless of geographical location. This

shift to remote collaboration has brought new challenges, such as maintaining awareness and managing the dynamics of turn-taking in conversations, which CSCW researchers have addressed through studies on communication protocols and coordination mechanisms [178].

Collaborative AR systems, like those seen in educational and industrial contexts, are also rooted in CSCW principles. In augmented reality, collaboration extends to virtual objects shared among users. Henrysson et al. explored how mobile devices could support face-to-face collaboration by allowing users to interact with 3D models in real-time [81]. This kind of interaction highlights the importance of coordination and mutual awareness, as users need to be aware of each other's actions to interact smoothly with shared virtual content.

### 2.2.2 Awareness

In the field of CSCW, awareness refers to “a person's being or becoming aware of something” [179], which provides context for your own activities. To perform effective collaboration, collaborators need a sense of awareness of the other participants, in addition to a shared view of a collaborative workspace [179]. Collaborators also need a sense of who is there, what is there, where, when, and how. To conceptualise these information requirements, Gutwin and Greenberg define workspace awareness as “the up-to-the-moment understanding of another person's interaction with the shared workspace” [72].

Supporting workspace awareness is a requirement for effective collaboration. Workspace awareness can help “recognise opportunities for closer coupling, reducing the effort needed for verbal communication, simplifying coordination, allowing people to act in anticipation of others” [72]. In face-to-face collaborations, workspace awareness is maintained by a *perception-action cycle*. Collaborators perceive information from their environment, add it to previously gathered information, and then use it to obtain additional information about their workspace. For example, when users try to help their collaborators, they might seek visual evidence about their collaborators' understanding, especially when their common vocabulary or past interactions do not provide enough information about how people understand and use this help [72].

The literature on HCI has introduced a range of techniques to support awareness. Nam and Sakong used physical turntables whose rotation was synchronised to help users better perceive the intention of their collaborators [140]. In addition, virtual

shadows of the users' hands were projected on the collaborator's table to give them more detailed feedback about their manipulation actions.

While workspace awareness primarily focuses on how collaborators maintain mutual understanding within a shared task space, spatial awareness also plays a critical role in coordinating collaborative efforts. Beyond visual cues, collaborators rely on their physical positioning, movement, and proximity to interpret intentions and facilitate smooth interactions. This notion of proxemic awareness, introduced by Hall [75], suggests that spatial relationships influence social interactions and engagement. Proxemics has increasingly been explored in HCI research, where it has been shown to shape digital interactions by allowing systems to adapt to users' movements and distances [126]. In the context of AR, proxemic-aware systems can dynamically respond to users' spatial behaviours, enabling more fluid collaboration by adjusting interaction modes based on proximity, orientation, and movement. The next section explores the role of proxemic interaction in AR environments, detailing how spatial factors influence collaboration and system responsiveness.

### 2.2.3 Proxemic Interaction

Proxemic interaction in Augmented Reality (AR) has recently become a focus of researchers, exploring how spatial relationships can affect user engagement within digital environments and when using virtual objects. This section looks at key findings from studies examining the role of proxemic interaction within AR contexts.

One of the foundational elements of proxemic interactions is the understanding of spatial dynamics. These dynamics can enhance the user experience and encourage more intuitive interactions with technology. In their research, Marquardt et al. [126] highlight how proxemic interactions draw upon different spatial relationships between users and devices, facilitating smoother and more natural engagement within systems. Looking further into this idea, Daza et al. [40] outline five essential dimensions within their research: Distance, Identity, Location, Movement, and Orientation. These dimensions can be important for designing AR applications that can respond effectively to user behaviour and positioning within the physical space.

Recently, some research has shown that proxemic interaction can be effectively integrated into AR applications. For example, research by Nigro et al. [144] found that using an AR interface that is designed for personalised proxemics modelling significantly boosted the trust in interactions with an AR interface through providing a pleasurable experience, ensuring they knew what they were doing, and improving

confidence in adopting new technology, especially among older adults. It can be implied that these tailored proxemic experiences can enhance user comfort and engagement when engaging with AR systems, which is particularly important for collaboration. Similarly, Pérez et al. [151] propose a framework for mobile application development based on proxemic interactions in their research, which showcases how adaptable these mobile systems can be when deployed across different contexts. Aslan and André [3] also demonstrate how proxemic touch targets can trigger emotional responses, further enriching the collaborative experience.

The social aspect of proxemics in AR has also gathered some attention. For example, Nijholt [145] further investigate social AR as a concept and highlight the importance of the use of non-verbal cues and spatial awareness in facilitating face-to-face group interactions. This also aligns with the work of Petrak et al. [152], who reveals that proxemic-aware robots can improve first impressions by respecting established social norms around personal space.

This research shows that integrating proxemic interaction principles into AR design can improve user experience, foster collaboration, and enrich social interactions.

#### 2.2.4 Affordances for Collaboration

The concept of affordances, introduced by the psychologist Gibson [56], refers to the actionable properties between an object and an agent. In simple terms, affordances describe the potential actions that a user can perform when interacting with an object or environment. When applied to Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW), affordances help us understand how the design of a system can either facilitate or hinder user interactions, particularly in collaborative settings. Norman later popularised the term within HCI, emphasising the importance of perceived affordances, which represent how users perceive the possible actions with a system based on its design [147].

In the context of CSCW, affordances are crucial for designing systems that support effective collaboration between individuals. Collaborative affordances in technologies, particularly in collaborative AR environments, refer to the properties that enable, constrain, or mediate user interactions in co-located or distributed settings. These affordances shape how users engage with both the system and each other, guiding communication, coordination, and cooperation [73, 153]. For instance, an AR interface that provides virtual pointers or shared markers can enable users to draw attention to specific virtual objects or locations, fostering clearer communication and



better coordination [188]. However, if the system lacks intuitive methods for pointing or indicating shared objects, it can limit the effectiveness of the collaboration.

A critical but often overlooked affordance in collaborative environments is territoriality, which describes how users establish, maintain, and negotiate ownership of space and digital resources. Territorial affordances influence how users claim and control interaction zones in shared AR workspaces. Research by Yim et al. [221] on Wikipedia editors highlights that territoriality is not limited to physical spaces but extends to digital environments, where users assert ownership over contributions. Similarly, Tussyadiah [205] explored how location-based media allow users to establish territorial claims in hybrid physical-digital spaces, a concept that is particularly relevant in AR, where virtual overlays can reinforce or disrupt perceptions of territorial boundaries.

Studies on co-located collaboration demonstrate that territorial dynamics shape interaction zones within shared workspaces. Scott [184] found that in tabletop interfaces, users tend to dominate areas directly in front of them, while shared spaces emerge in the centre for group tasks. This finding aligns with work in AR collaboration, where personal and group territories influence how users interact with virtual content. For instance, proxemic-aware AR systems can dynamically adjust territorial boundaries based on user positioning, movement, and gaze, ensuring that virtual objects remain accessible to the appropriate users. The concept of having a personal territory is also explored in educational settings, as can be seen in the work of Modell and Gray [136]. This study indicates that students display group territoriality by using personal items to mark individual boundaries within collaborative spaces. This insight is particularly relevant in AR applications, where understanding how users establish personal and group territories can inform the design of collaborative environments that utilise AR.

Territoriality also plays a key role in real-time collaboration and task allocation. Research by Mahadevan et al. [124] found that territorial and proxemic factors inform how users negotiate control in human-robot interaction. This directly impacts AR collaboration, where multiple users may need to claim exclusive or shared access to digital objects. Understanding territoriality in AR system design can improve multi-user interactions by enabling explicit ownership markers, turn-taking mechanisms, and adaptive control zones that reflect natural human spatial behaviours.

Mobile AR systems offer several affordances that have the potential to enhance collaboration, such as the ability to view and manipulate 3D objects in a shared

space [15, 199]. These systems afford spatial awareness, which is critical in collaborative tasks that require an understanding of object locations and relationships within a physical or virtual environment. Moreover, mobile AR devices afford mobility, allowing users to move freely in their environment while engaging with digital content, which can enhance the fluidity of collaboration [116]. For example, systems like Studierstube [199] enable multiple users to manipulate 3D models in a shared augmented reality space, leveraging spatial affordances to facilitate collaborative design and decision-making.

However, the affordances of handheld mobile AR systems can also introduce limitations that inhibit collaboration. The small screen size of many mobile devices can constrain the amount of visual information presented, making it difficult for users to share and interact with content collaboratively. This is particularly evident when multiple users gather around a single device, reducing engagement due to limited screen real estate [74, 130]. Additionally, handheld AR devices may afford individual actions, such as private viewing of AR content, at the expense of collaborative awareness. This tension between private and public affordances in mobile AR systems can lead to breakdowns in communication, as users may be focused on their individual screens rather than on shared content or interactions [180].

An illustrative example of how affordances can enhance or inhibit collaboration can be found in studies comparing head-worn AR displays with handheld AR systems. Head-worn displays, such as the Microsoft HoloLens, afford hands-free interaction, allowing users to engage more naturally with both the real and virtual environment without the need to hold or operate a device [33, 188]. This hands-free affordance improves the user's ability to focus on the collaborative task and interact with others without the distraction of manipulating a mobile device. In contrast, handheld devices afford portability but may require users to divert attention away from the shared space, diminishing collaborative engagement [123].

In conclusion, affordances in CSCW systems, particularly in mobile AR environments, play a critical role in shaping the effectiveness of collaboration. Designers must carefully consider how different affordances either enable or constrain user interactions, ensuring that technologies provide adequate support for communication, coordination, and cooperation among users. In particular, territorial affordances must be accounted for in AR system design, as they dictate how users claim, share, and negotiate control over virtual objects and spaces. Balancing individual, shared, and dynamic affordances remains a key challenge for researchers and developers aiming to enhance the user experience in co-located and distributed AR collaborations.

### 2.2.5 Group Dynamics

Group dynamics refer to the behaviours, interactions, and roles that emerge when individuals work together towards a common goal. Understanding group dynamics is essential in collaborative efforts, as the success of a group depends on how well participants communicate, coordinate, and share responsibilities. Several factors, including group formation, spatial arrangements, leadership, decision-making, and conflict management, shape the outcomes of group interactions [6, 150]. Effective group dynamics promote collaboration, streamline decision-making, and enhance problem-solving capabilities, while poor group dynamics can result in inefficiencies, miscommunication, and disengagement.

#### Group Formation and Roles

Groups typically form around a shared objective, with members assuming roles based on their expertise or assigned responsibilities. The effectiveness of a group depends not only on the division of labour but also on how roles are distributed and understood [188]. In Computer-Supported Cooperative Work (CSCW), technology can aid role allocation by structuring workflows and clarifying individual responsibilities. Systems such as MobiSurf [187] and GroupTogether [126] facilitate collaboration by enabling participants to contribute to shared tasks through their personal devices while ensuring an organised workflow.

However, technological interventions can also introduce challenges. When system designs do not account for role clarity, users may experience ambiguity in their contributions, leading to uneven participation and task overlap [212]. Research has shown that role ambiguity can significantly hinder collaboration by fostering inefficiencies and disengagement [130]. Therefore, collaborative systems should incorporate mechanisms that reinforce role differentiation and provide users with clear participation guidelines.

#### Spatial Arrangements and F-Formations

Beyond role distribution, the spatial configuration of group members plays a crucial role in shaping interactions. Facing Formations (F-Formations), introduced by Kendon [104], describe how individuals position themselves in relation to one another during social interactions. These spatial patterns define participation zones, affecting individuals' communication and task coordination.

Research in Human-Computer Interaction (HCI) suggests that recognising and accommodating F-Formations can enhance the design of collaborative environments. Marshall et al. [129] argue that spatial formations provide a practical framework for understanding interaction dynamics in physical and digital settings. In Augmented Reality (AR), where virtual content overlaps physical spaces, maintaining appropriate F-Formations is crucial for ensuring optimal information flow and engagement [219].

In Mixed Reality (MR), F-Formations influence collaborative interactions by structuring how users orient themselves around virtual objects. Tong et al. [201] explored the relationship between F-Formations and device usage, showing that spatial transitions impact collaboration, particularly in AR settings where users frequently shift between physical and digital tasks. Designing AR interfaces that adapt dynamically to spatial formations can enhance usability by ensuring virtual content remains visible, accessible, and aligned with users' natural movements.

### **Leadership in Collaborative Groups**

Leadership plays a significant role in guiding collaborative efforts. In group settings, leaders are often responsible for directing the group's attention, making key decisions, and resolving conflicts. Collaborative technologies can support leadership by enabling leaders to take control of shared virtual spaces, distribute tasks, or highlight important information for the group [11, 116]. For example, in co-located mobile AR collaborations, leaders can use augmented reality systems to direct participants' attention to specific virtual objects, guiding the group's focus and ensuring that everyone remains on task [60]. Additionally, technologies like the TrainAR system [20] offer frameworks that help leaders coordinate complex tasks by providing step-by-step instructions to team members within a shared AR environment.

However, leadership can be undermined if the technology fails to provide adequate mechanisms for managing the group. Conflicts can arise when multiple participants vie for control over the shared virtual space, leading to reduced cooperation. Systems not accounting for leadership structures may result in power struggles or disorganised workflows [34]. To address this, designers of collaborative systems must ensure that technological tools support, rather than hinder, effective leadership within group settings.

## Group Decision-Making

Decision-making is a critical aspect of collaborative work, requiring input from multiple participants to achieve consensus. In group settings, decisions are often made through a combination of discussion, negotiation, and compromise. CSCW technologies can facilitate decision-making by providing tools for information sharing, brainstorming, and voting [212]. For instance, shared displays in collaborative AR environments allow group members to visually explore different options and make decisions based on shared virtual content [97]. Systems like Studierstube [199] enable users to manipulate 3D models in real-time, fostering group discussions and collective decision-making.

However, decision-making can be hindered if the system limits communication or does not support group consensus-building. Research has shown that when technologies provide inadequate mechanisms for group discussions, such as poor turn-taking or information-sharing interfaces, group decision-making can become fragmented [19]. This highlights the need for CSCW systems to incorporate features that support open dialogue and equitable participation, ensuring that all group members have an opportunity to contribute to the decision-making process.

## Conflict Management in Collaborative Groups

Conflicts are inevitable in group collaboration, arising from differences in opinion, misunderstandings, or competition for resources. Effective conflict management is essential for maintaining group cohesion and ensuring that collaborative efforts stay on track. Technology can play a dual role in conflict management by either facilitating communication and conflict resolution or exacerbating tensions through poorly designed interaction mechanisms [150].

CSCW systems can aid in conflict management by providing communication tools that allow group members to express their views clearly and resolve disagreements amicably [10]. For example, in collaborative AR settings, systems that enable real-time annotations or virtual pointers can help clarify misunderstandings, as users can highlight specific aspects of the shared virtual environment to explain their perspectives [188]. In contrast, systems that lack these features may lead to further confusion as participants struggle to communicate effectively within the shared space.

Furthermore, the structure of the technology can either mitigate or exacerbate conflicts depending on how it manages access to shared resources. Systems that allow for equal access and control over shared virtual objects are more likely to

foster cooperation, while those that grant control to only a few users may lead to competition and frustration [130]. To address this, designers of collaborative technologies must consider how system affordances can either support or inhibit effective conflict management, ensuring that users have the tools they need to resolve conflicts productively.

### **The Role of Technology in Group Dynamics**

Technology plays a crucial role in shaping group dynamics in collaborative environments. CSCW systems, particularly in AR settings, provide affordances that can either enhance or inhibit group interactions. Effective collaborative technologies should support role distribution, leadership, decision-making, and conflict management by providing tools that foster communication, coordination, and cooperation [19, 165, 198]. However, if these systems are poorly designed or fail to account for the nuances of group dynamics, they may introduce barriers to collaboration, reducing the group's overall effectiveness [188].

In conclusion, group dynamics are a critical factor in the success of collaborative efforts. By understanding the roles, leadership structures, decision-making processes, and conflict management strategies within groups, designers can create technologies that support and enhance collaboration in CSCW settings. The challenge lies in developing systems that facilitate individual interactions with technology and nurture positive group dynamics, ensuring that all participants can contribute effectively to the collaborative process.

## **2.3 Cross-Device Computing in Collaborative Augmented Reality (AR)**

Cross-device computing has become a significant area of interest in the Human-Computer Interaction (HCI) domain, particularly for collaborative environments that require seamless interaction across multiple devices. As AR technologies continue to evolve, the integration of cross-device capabilities opens new possibilities for enhancing user experiences, especially in co-located collaboration scenarios. This section explores cross-device computing and its relevance to mobile, handheld AR systems, focusing on how interactions across multiple devices can inform the design of collaborative AR applications.

Cross-device computing refers to interactions where digital content or tasks span across more than one device, enabling users to leverage the affordances of multiple devices in concert. According to Brudy et al., cross-device computing presents unique opportunities to engage with content beyond the bounds of a single device screen, offering flexibility in how users interact with information [26]. Brudy et al. conducted a comprehensive survey of 510 papers in the cross-device computing domain and developed a taxonomy to unify terminology and facilitate future research. This taxonomy highlights historical trends, key interaction techniques, and under-explored application areas. As mobile AR matures, such taxonomies can be instrumental in designing systems that support seamless transitions and interactions across devices, enhancing the collaborative potential of AR.

The primary motivation behind cross-device systems is to extend the interaction capabilities of users by allowing digital content to flow naturally between devices, enabling more dynamic and flexible workflows. For instance, Chi and Li introduced Weave, a scripting framework for cross-device wearable interaction, which provides insights into how device synchronization and interaction can be managed across wearables and other handheld devices [35]. The framework focuses on enabling users to create fluid cross-device interactions, an essential feature in AR environments where different devices could serve as entry points to a shared augmented space. Such systems highlight the potential for collaboration between users who interact with virtual objects on different devices while maintaining real-time synchronization and mutual awareness.

A key challenge in cross-device computing, particularly in the context of mobile AR, is the design of interaction techniques that facilitate collaboration across various device types and sizes. Devices such as smartphones, tablets, and wearables offer different affordances that can either support or hinder collaboration depending on their configuration. As noted by Marquardt et al., cross-device interactions can benefit from the application of sociological constructs like micro-mobility and F-formations, which help users maintain spatial relationships and visual coherence when interacting with multiple devices [126]. These constructs are particularly relevant in AR settings where the spatial arrangement of devices and users can influence group dynamics and the overall collaborative experience.

Moreover, Schreiner et al. proposed the Connichiwa framework, which enables real-time synchronization and interaction across co-located devices using simple web-based technologies [182]. Such frameworks are valuable for AR systems, where



multiple users need to share a common virtual space, interact with the same set of digital objects, and maintain synchronized views of augmented content. The synchronization of content across devices is crucial in ensuring that all collaborators remain engaged and can contribute to the task without interruption or lag.

The importance of device configuration and its impact on collaboration has also been explored in various cross-device research studies. For instance, Fischer et al. examined how device size and positioning affect user interaction and group coordination in cross-device environments [49]. Their findings suggest that larger devices, such as tablets, tend to facilitate group work by providing more screen real estate for shared content, while smaller devices, like smartphones, are better suited for personal tasks within the collaborative workflow. In mobile AR, this distinction is essential for designing systems that balance individual and group activities, ensuring that both personal and shared content can be accessed seamlessly across devices.

As we move toward more integrated AR environments, the need for robust cross-device interaction models becomes more evident. Studies such as those by Goh et al. have shown that cross-device interaction can improve collaborative decision-making and sense-making activities by enabling shared understanding and mutual awareness among users [57]. This is particularly important in mobile AR systems where multiple users may be interacting with virtual objects from different perspectives and devices. By enabling fluid interactions across devices, cross-device computing enhances the collaborative experience, making it easier for users to coordinate their actions and share information.

Despite the promising advances in cross-device computing, there are still several open research challenges that need to be addressed to realise its potential in mobile AR fully. One of the main challenges is ensuring that the interaction techniques developed for cross-device systems are intuitive and scalable across different device types and contexts. Additionally, as noted by Brudy et al., there is a need for more research into the evaluation strategies used in cross-device computing to better understand how these systems can be optimized for collaborative use [26]. The development of standardized evaluation methods will be critical in advancing cross-device AR systems and ensuring that they meet the needs of users in diverse collaborative scenarios.

In conclusion, cross-device computing offers a promising framework for enhancing collaboration in mobile AR environments. By leveraging the unique affordances of multiple devices, cross-device systems can support more dynamic, flexible, and engaging interactions in augmented spaces. As researchers continue to explore the



opportunities and challenges of cross-device interaction, it is crucial to develop standardized frameworks, taxonomies, and evaluation methods that can guide the design of future AR systems. This will ensure that mobile AR applications are not only technologically feasible but also meaningful and effective for collaborative work. Furthermore, cross-device computing aligns with Milgram’s Reality-Virtuality Continuum by facilitating seamless transitions between real and virtual environments. By distributing interactions across multiple devices, cross-device systems enhance the spectrum of mixed reality experiences, allowing users to navigate fluidly between physical and digital spaces. This interplay between devices reinforces the hybrid nature of AR, where collaboration is not confined to a single reality but rather exists across a continuum of immersive and non-immersive interactions.

## 2.4 Conclusion

The background section of this thesis provides an exploration into Augmented Reality (AR) displays as well as their classifications, which reveals a nuanced landscape where technological advancements have made AR more accessible. However, there are still challenges that remain in its practical application. The historical evolution of AR – from early conceptual frameworks to modern handheld and head-mounted displays – shows how this technology has become embedded in various industries, including education, healthcare, and entertainment. Handheld AR, in particular, has emerged as a dominant platform due to the ubiquity of smartphones and tablets. This form of AR display offers portability and ease of use. However, the limitations of screen size, user interaction methods, and immersion highlight areas where further development is necessary to enhance the user’s experience.

In co-located collaborative settings, AR offers a potential for redefining group work by blending the virtual and physical environments. Yet, whilst it is assumed that this may be the case, there are large hurdles that need to be addressed. Research in the area of Computer-Supported Cooperative Work (CSCW) has provided insights into how AR could mediate such collaborations by supporting awareness, group dynamics, and affordances for interaction. But, despite the promise of AR in improving collaborative tasks, challenges such as managing group awareness, balancing individual and collaborative affordances, and designing intuitive interaction techniques that support interactions across both the virtual and physical realms remain key areas for improvement.

As we continue with this thesis, we build on the foundational understanding of AR technology and its integration into collaborative environments. Future research in this area must focus on addressing the limitations of handheld AR, exploring how this technology can – in its current state – better support group collaboration, and refining design principles that bridge the gap between AR’s potential and its current practical shortcomings. The exploration of these areas will directly inform the empirical investigations of this thesis, seeking to contribute to the development of more effective and immersive AR systems for co-located collaborative work.

# 3

## Exploring the Dynamics of Group Interaction in AR

### Contents

---

|     |                               |    |
|-----|-------------------------------|----|
| 3.1 | Introduction                  | 51 |
| 3.2 | Exploring the CollabAR System | 52 |
| 3.3 | Study Methodology             | 56 |
| 3.4 | Findings                      | 61 |
| 3.5 | Discussion                    | 69 |
| 3.6 | Limitations and Future Work   | 73 |
| 3.7 | Conclusion                    | 74 |

---

This thesis aims to investigate the role of Augmented Reality (AR) in enhancing co-located collaborative experiences. As an initial step, we seek to develop an understanding of how state-of-the-art collaborative handheld AR systems are used in real-world scenarios and identify the challenges that exist in this area, as defined in [Section 1.1](#). In this chapter, we explore how current mobile AR interfaces influence co-located group collaboration. By examining the cognitive load, group dynamics, and device handling associated with these systems, our aim is to answer [RQ1](#): *How do current group dynamics and practices affect co-located collaborations enhanced by handheld mobile AR applications?*

### 3.1 Introduction

While Augmented Reality (AR), particularly handheld mobile AR, is widely recognised for its potential to enhance co-located collaborative experiences, a fundamental gap persists in the absence of a unifying framework that addresses the challenges in understanding group dynamics, device configurations, and interaction techniques. This thesis aims to address this overarching problem by exploring how mobile handheld AR interfaces impact co-located collaborative work and inform the design of future AR systems.

Building on this, Chapter 3 delves explicitly into Research Question 1 (RQ1): “How do current group dynamics and practices affect co-located collaborations enhanced by handheld mobile AR applications?”. Despite the potential for AR interfaces to facilitate collaborative work [12], most mobile AR interfaces are currently designed primarily for individual use rather than explicit collaboration. Although some studies have explored co-located collaboration for mobile AR (e.g. [81, 146, 158]), our understanding of precisely how mobile AR influences collaboration and group interaction remains limited. This contrasts with the extensive research that demonstrates the importance of co-located group work for touch-based interfaces [25, 26, 74, 121, 213]. As Kim et al. concluded in their review of AR, more information is needed on how to create collaborative systems for AR [106]. This chapter, therefore, systematically investigates these current group dynamics and practices to provide fresh insights into the effect of contemporary mobile AR interfaces on co-located collaboration.

In an effort to develop a systematic understanding of how AR mediates co-located collaboration, we conducted an observational study exploring the impact of current mobile AR interfaces on co-located collaboration. Our primary aim is to shed light on the practices, challenges, and issues faced by groups and to provide a set of recommendations for the design and development of effective mobile AR interfaces for collaborative work. In our study, we investigate how groups of four individuals collaborate on basic tasks using standard mobile devices and a rudimentary mobile AR interface. The tasks involve the discovery, inspection, and comparison of various virtual models in a meeting setting (where virtual objects are projected onto a table). We designed six models that exemplify complex virtual data that are typically used in medical, educational, or knowledge work domains. We aim to build an understanding of existing group dynamics and practises in co-located ad-hoc situations. Our empirical study observes collaborative practises and classifies

them into three categories of group dynamics: (i) collaboration styles, (ii) context switching, and (iii) device handling and manipulation.

## 3.2 Exploring the CollabAR System

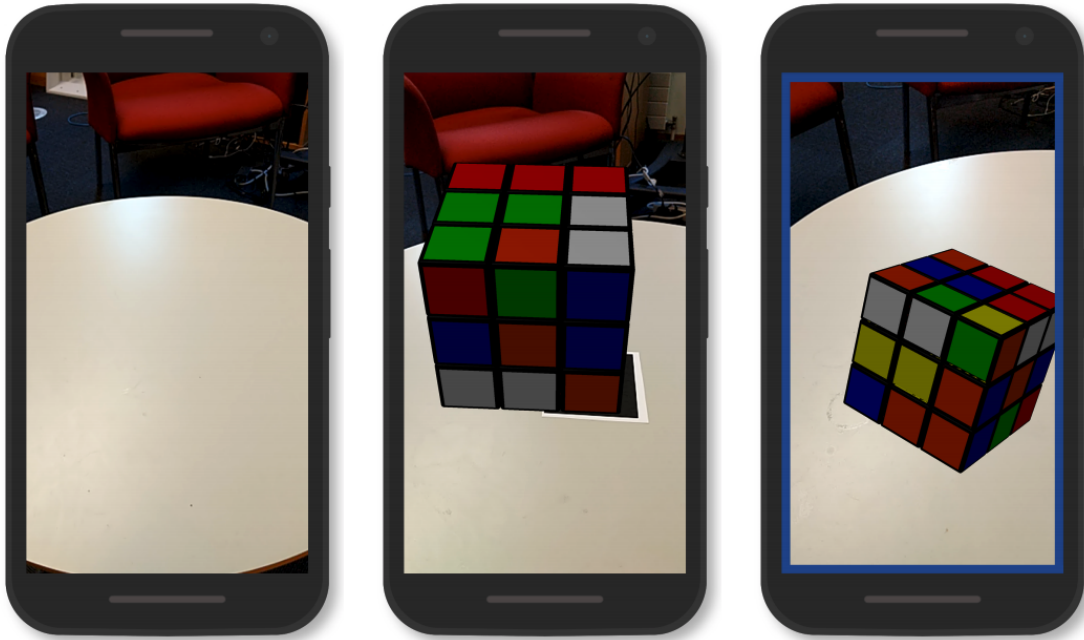
The goal of this chapter, and indeed the initial step within this thesis, is to elucidate the mediating role of handheld AR interfaces in the context of co-located collaborative work. Rather than immediately proposing novel interaction paradigms, the design of the *CollabAR* system was deliberately conceived as a simple exemplar AR interface for mobile devices to enable a foundational understanding of current group dynamics and practices when enhanced by handheld mobile AR applications.

This approach allowed for a systematic investigation into how existing, commercially prevalent mobile AR interfaces influence co-located group collaboration, providing a crucial baseline for identifying challenges and informing future design recommendations. Other more complex or experimental avenues for AR system design were intentionally set aside at this initial stage to focus on observing the inherent affordances and limitations of “out-of-the-box” AR experiences.

In essence, the design of CollabAR was a strategic decision to create a rudimentary mobile AR interface that served as a controlled environment to observe how groups naturally interact with AR in a collaborative context, particularly focusing on how existing designs influence group dynamics, cognitive load, and device handling. This foundational understanding was crucial before advancing to more complex design frameworks and interaction techniques in subsequent chapters.

### 3.2.1 User Interface

The user interface of CollabAR displays a full-screen camera feed. When AR markers are detected, virtual objects are overlaid on this feed ([Figure 3.1](#)). The application, designed to be an example of simplicity, incorporates features commonly found in commercially available AR technology. As a web-based interface, CollabAR can be accessed through a standard browser on most modern mobile devices, including smartphones and tablets. Multiple users can launch the web URL on their phone and point the device at the same AR marker to see and interact with the same 3D model. Users can physically move closer to or further from the marker to zoom in and out of the model or view it from a different angle. As the AR marker is stationary, altering the viewing perspective of the models requires physical movement around the marker.



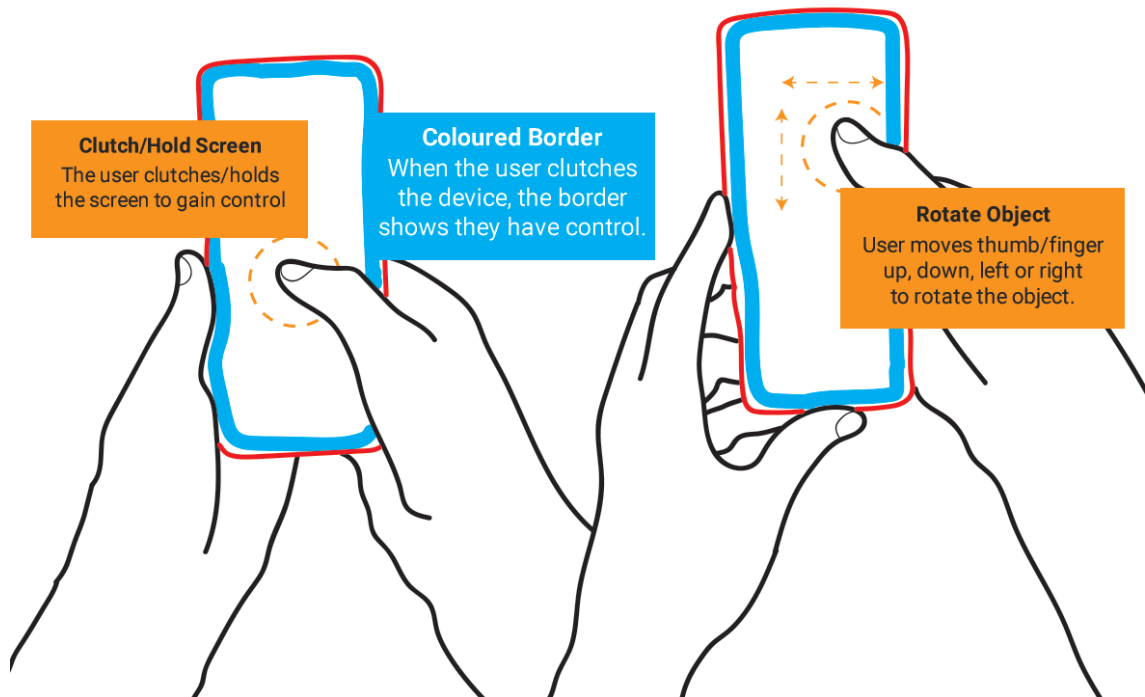
**Figure 3.1:** The three views of the CollabAR Interface: nothing in view (left), model in view (middle) and the locking border (right).

### 3.2.2 Gesture-Based Interaction

The gesture-based interaction in CollabAR was designed with a fundamental understanding of territoriality, conflict management, and sharing aspects in co-located collaborative work, particularly given the initial aim to understand how basic mobile AR interfaces influence group dynamics. While the gestures themselves were simple, the turn-based system for manipulation was a deliberate design choice that directly grounded these concepts.

Here is an expansion on how these aspects were incorporated and addressed:

- **Basic Gestures and Real-time Observation:** To interact with the 3D models, CollabAR implemented a set of basic gestures commonly found in standard AR applications, such as touch-and-drag for rotating the object and pinching gestures for scaling (Figure 3.2). Vertical finger movements tilted the model, while horizontal movements rotated it. Although only one user could actively manipulate the model at any given time, all other users could still view the model on their device and observe the modifications in real-time. This established a foundation of shared perception even when control was sequential.

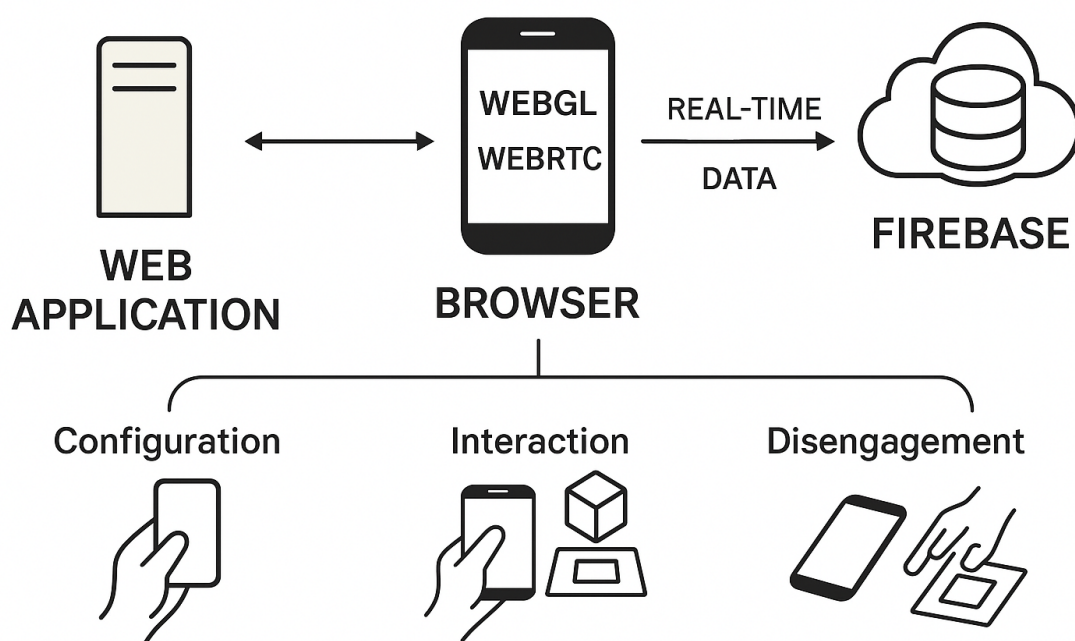


**Figure 3.2:** Basic interactions with CollabAR

- Territoriality and Control – The Clutching Mechanism:** A clutching mechanism was implemented to lock the model when a user initiated a modification. This design inherently created a temporary “territory” of control over the virtual object’s manipulation. By granting exclusive manipulation rights, albeit temporarily, the system implicitly defined who owned the active interaction space at that moment. To make this territoriality explicit and provide clear feedback, each user was assigned a specific colour, and when the user gained the lock, a coloured border appeared around the camera feed on all connected devices (Figure 3.1 – right). This visual cue immediately indicated the current proprietor of the manipulation control, a direct application of awareness cues identified as crucial for collaborative settings. This allowed researchers to observe how groups negotiated and managed this sequential ownership of the virtual object.
- Conflict Management through Turn-Based Control:** The primary motivation for this turn-based system was to simplify and prevent direct conflicts arising from simultaneous manipulation attempts by multiple users. In scenarios where multiple users could simultaneously try to rotate or scale

the same object, it could lead to confusing, conflicting inputs, unintended results, and user frustration. By enforcing that only one user at a time could modify the model, CollabAR avoided these concurrent input issues, allowing for a clearer observation of how groups coordinated their actions despite the sequential control. This design choice effectively externalised the negotiation of control into a turn-taking process rather than a simultaneous one.

### 3.2.3 Technical Implementation



**Figure 3.3:** Technical architecture of CollabAR. The system is implemented as a web application accessed through a WebGL- and WebRTC-enabled browser on a user’s device. 3D objects are rendered locally upon detecting a marker. Real-time collaboration is enabled via Firebase, which synchronises object data across clients. The three user interaction phases – Configuration, Interaction, and Disengagement – are illustrated along the bottom.

CollabAR was specifically engineered as a basic AR interface for mobile devices, with its technical implementation prioritising simplicity and broad accessibility (Figure 3.3). The system was developed as a web application, a deliberate choice made to maximise device compatibility and reduce limitations on the type of mobile device that could be used. This meant users could access the application via a standard web browser on most modern mobile devices, including smartphones and



tablets, provided they supported WebGL and WebRTC capabilities. To eliminate latency problems often associated with loading 3D objects over a network, each 3D object was rendered locally in the browser on the user’s device, ensuring a smoother and more responsive interaction with the virtual models. Google’s Firebase real-time database [58] facilitated the collaborative functionality of CollabAR. Firebase enabled the efficient updating of data, such as object coordinates, in the database and provided real-time notifications to clients about new events. This allowed us to send information, such as the coordinates of the object, to the database, which would then be updated in real-time on all clients. We recorded a device ID, enabling us to detect user interactions with the device and model, and subsequently lock these interactions to that user. This technical backbone ensured that all connected devices instantaneously displayed the same changes to the shared virtual model.

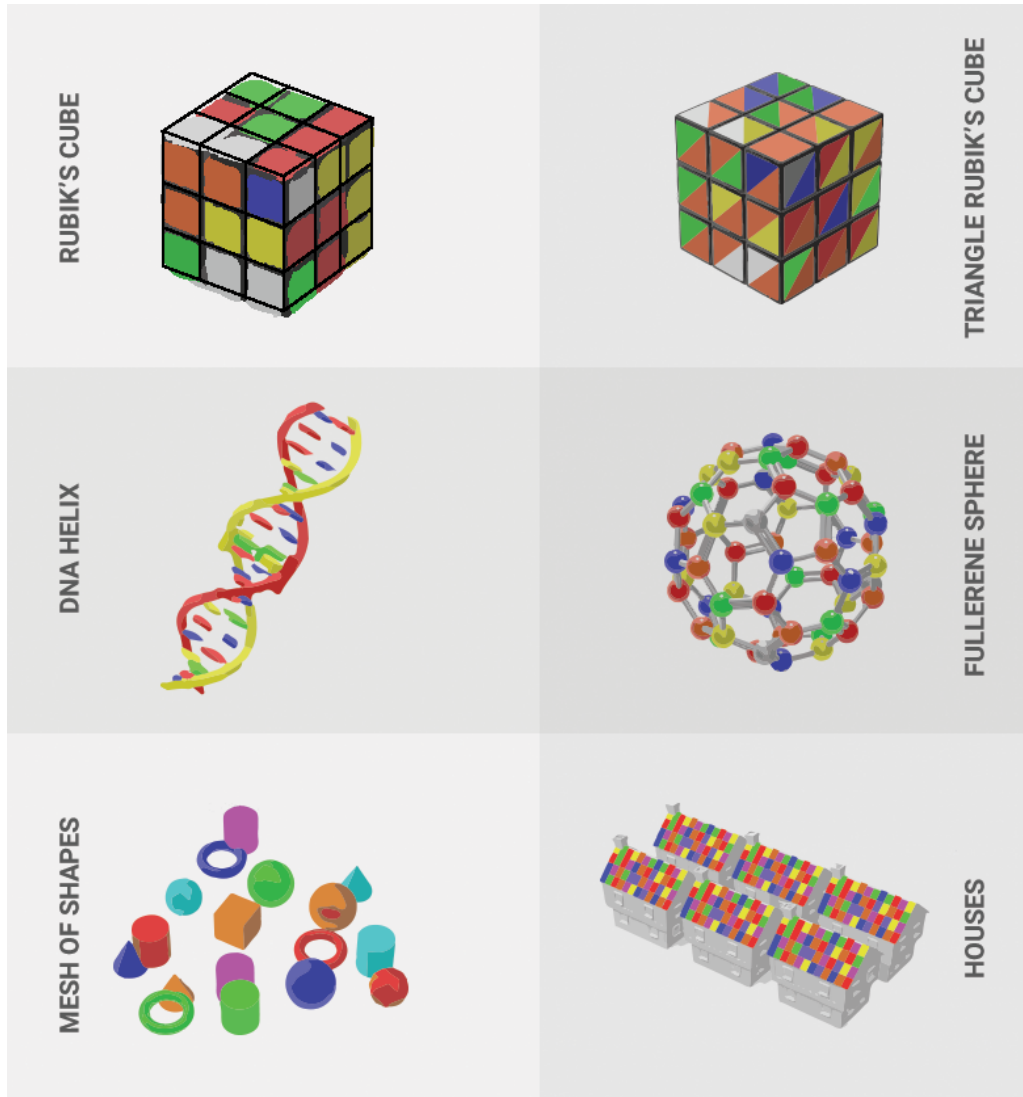
User interaction within the system was structured around three distinct interaction phases: a configuration phase where users would set up their devices and load the web page; an interaction phase that began when users placed an AR marker on the table and pointed their device’s camera at it, causing a 3D model to appear for manipulation; and a disengagement phase when users concluded their interaction by closing the browser and putting away their devices and the marker.

### 3.2.4 Virtual Models

We designed and created six models that were used during the study. These models serve as abstract examples of complex virtual data typically found in various AR domains such as education, healthcare, or other knowledge-based work settings. We designed the models to have varying levels of complexity by considering factors such as scale, geometric shape, colour density and sub-shapes, and intricacies (Figure 3.4). The models include a *Rubik’s Cube*, a *DNA helix*, a *mesh of shapes*, a *triangle Rubik’s Cube*, a *Fullerene Sphere*, and a *block of houses*.

## 3.3 Study Methodology

In our quest to understand how mobile AR influences and shapes collaborative group practices and dynamics, we utilised *CollabAR*, a basic AR interface for mobile devices, to examine collaborative activities in a co-located setting. Participants participated in collaborative tasks with virtual models of varying complexity. This section describes our study design and setup.



**Figure 3.4:** The six models of varying complexity used for the tasks.

### 3.3.1 Participants

We enrolled 20 participants (11 identified as female, 9 as male) through snowball sampling, who participated in groups of four. There were no strict prerequisites to participate in the study, barring the requirement that participants not be colour-blind, as some tasks were colour-specific. We did not disclose any task-related information prior to the study, as the tasks were abstract and not domain-specific.

The participants ranged in age from 18 to 24 years. Although prior experience with AR was not a criterion, seven participants had used AR before participating in the study. Examples of previously used AR applications included a Dulux application

for augmenting virtual furniture into a physical space, Pokemon GO, an AR-utilising game for capturing Pokemon, and the Google Pixel AR camera, which allows the user to overlay 3D objects into a room. All participants owned and frequently used at least one mobile device, such as a mobile phone or tablet.

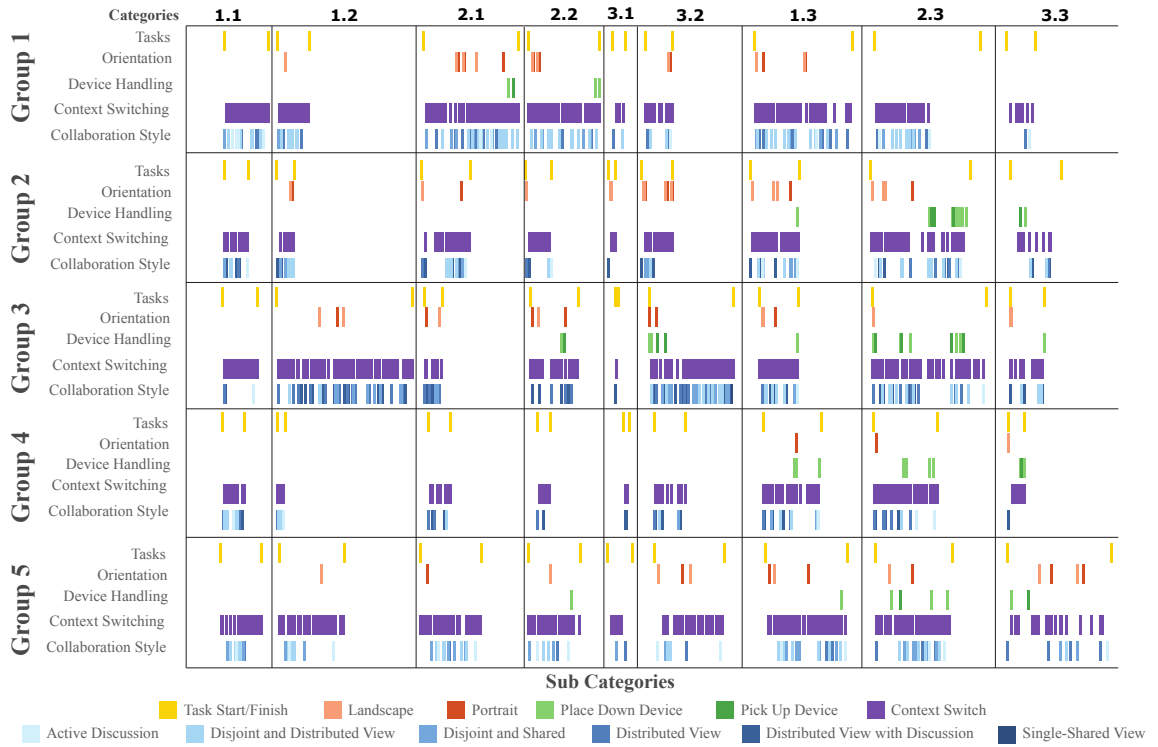
### 3.3.2 Apparatus

We invited the participants to a circular meeting room setup that facilitated free movement around a table (Figure 3.5). Chairs were provided, but participants were informed at the beginning of the study that they were not required to remain seated. After consenting to participate in the study, the participants completed a pre-study questionnaire that collected basic demographic data and information on previous experiences with AR technologies and commonly used technology. We provided each participant with a Moto G4 (XT1642) mobile device featuring a 5.5" display. The participants sat around the table, and the experimenter was opposite them. Pens and NASA-TLX forms were provided to complete the questionnaire after each task and in case the participants wished to take notes during the study.



**Figure 3.5:** The study set up in a meeting room

After briefing the participants about the task, we allowed them a 3-minute training session on using CollabAR. The participants were then asked to complete tasks based on the models given to them. Participants were not time-bound for



**Figure 3.6:** The timeline visualises the coding of the video recording per group according to observed categories. Tasks are labelled using a session-based scheme (e.g., 1.1, 1.2, 1.3) where the first digit indicates the model level and the second digit the task number. Tasks x.1 and x.2 correspond to counting tasks. Tasks x.3 represent comparison tasks, which introduced a second model.

each task: They simply had to complete the task and verbalise the final answer. After the task, participants were required to complete a NASA-TLX form to capture the perceived workload. Upon completion of all tasks, we conducted a semi-structured group interview.

### 3.3.3 Tasks

We designed a total of six models in three levels of varying difficulty for our study. Each difficulty level had three tasks. The first two tasks were inspection tasks, and the third task was a comparison task. During the first inspection task, participants were asked a question such as ‘count how many red tiles are on this cube’. The participants then inspected the cube as a group, counting the total number of red tiles. The second inspection task asked something specific about the model, such as ‘find how many shapes are touching each other’. Similarly to the first task, participants inspected the model, completed the given task, and reached a consensus

before providing the answer. For the third task in each level, we introduced a second model and asked participants to compare the original model with one they had not seen. An example question is ‘comparing the third and sixth models, listing the colours that do not appear in both models’.

We chose the questions for each task based on their simplicity, as we wanted to ensure that changes in collaboration were due solely to the complexity of the model rather than to the complexity of the tasks.

### 3.3.4 Data Collection and Analysis

We video-recorded each session to analyse the collaboration practices of the groups, including their collaboration styles and context switches. We define a context switch as an instance in which the participant looks away from the virtual content. Alongside the videos, the experimenter took notes to corroborate the video. We used the NASA-TLX questionnaire to record subjective workload, which was completed individually following each task. These questionnaires served as discussion points during a semi-structured group interview at the end of the session.

We analysed data collected from video recordings using ChronoViz [50]. We coded the video data for collaboration practises and participants’ interactions with each other or with their devices. Our analysis followed a grounded theory approach, utilising open coding of the video material in conjunction with the researcher’s notes. We adapted an existing coding framework previously developed by Tang et al. [200]. As in a study by Brudy et al. [25], we adapted this framework to include groups of four participants instead of pairs and to accommodate multiple devices. However, we also added an additional style, which was ‘Distributed View with Discussion’, as this happened frequently in each group. Our primary focus was on individuals within the groups and where their attention was focused. We differentiated this from a context switch based on duration. The collaboration styles we explored were as follows:

- C1 Active Discussion** – Active discussion indicated any face-to-face discussions that included all participants. As a result, there were limited interactions with their mobile devices.
- C2 Single-Shared View** – A single shared view indicated that all participants focused their attention on a single device.
- C3 Disjoint and Shared View** – Disjoint and shared view indicated 2-3 members focused on a single device, while others focused on their own device.

- C4 Disjoint and Distributed View** – Denoted 1-2 group members focusing on their device while the other two engaged in active discussion (not using their own devices).
- C5 Distributed View** – Denoted that the participants focused on their device with little to no discussion between each other. Complete focus on the task at hand.
- C6 Distributed View with Discussion** – Denoted that each participant was focused on their device while continuing the conversation with the others in the group.

## 3.4 Findings

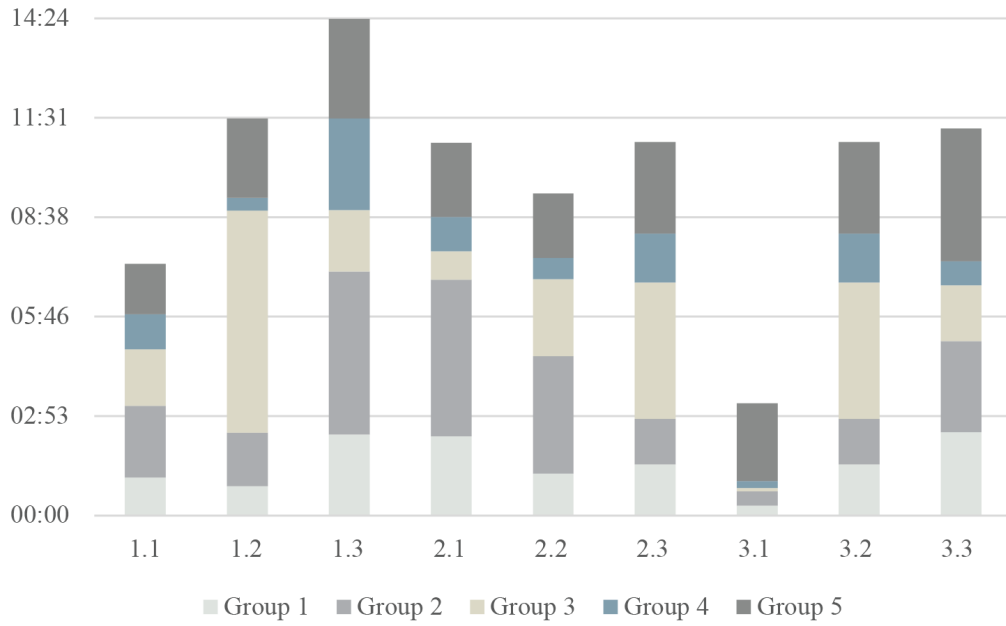
Our study explored how mobile AR interfaces influence group dynamics and practices. We noted how teams employ different collaboration styles and context switches to foster group conversation. This section presents our findings, categorised by the complexity of the tasks, which ranged from Level 1 (low complexity) to Level 3 (high complexity). Our study indicates significant differences between the difficulty levels in terms of mental demand and frustration. As the complexity of 3D objects increases, there is a decrease in the frequency of participant interactions while observing the model, with participants tending to concentrate on their individual devices instead.

### 3.4.1 Utilisation Patterns

[Figure 3.6](#) provides a detailed analysis of group collaboration, showing the entire timeline for each group with a breakdown of the different categories, including context switches, collaboration styles, device handling, and task start and stop times. The timelines reveal that there was a large number of context switches (all groups). More specifically, across all groups, there was a total of 1,589 context switches (mean: 317.8, min: 183, max: 422, SD: 24.4) over the 9 tasks that each group was given.

The duration of the tasks, shown in [Figure 3.7](#), was relatively short and quick to complete (mean: 128.19s, min: 6s, max: 386s, SD: 83.15s). Generally, comparison tasks took longer than inspection tasks, but there was clear variability between groups.

In terms of device mobility, we observed 44 instances of moving to landscape mode, 50 to portrait mode, 41 to place the device down, and 26 to pick it up. For overall collaboration styles, we observed 36 instances of Active Discussion, 2



**Figure 3.7:** Task completion time for each task.

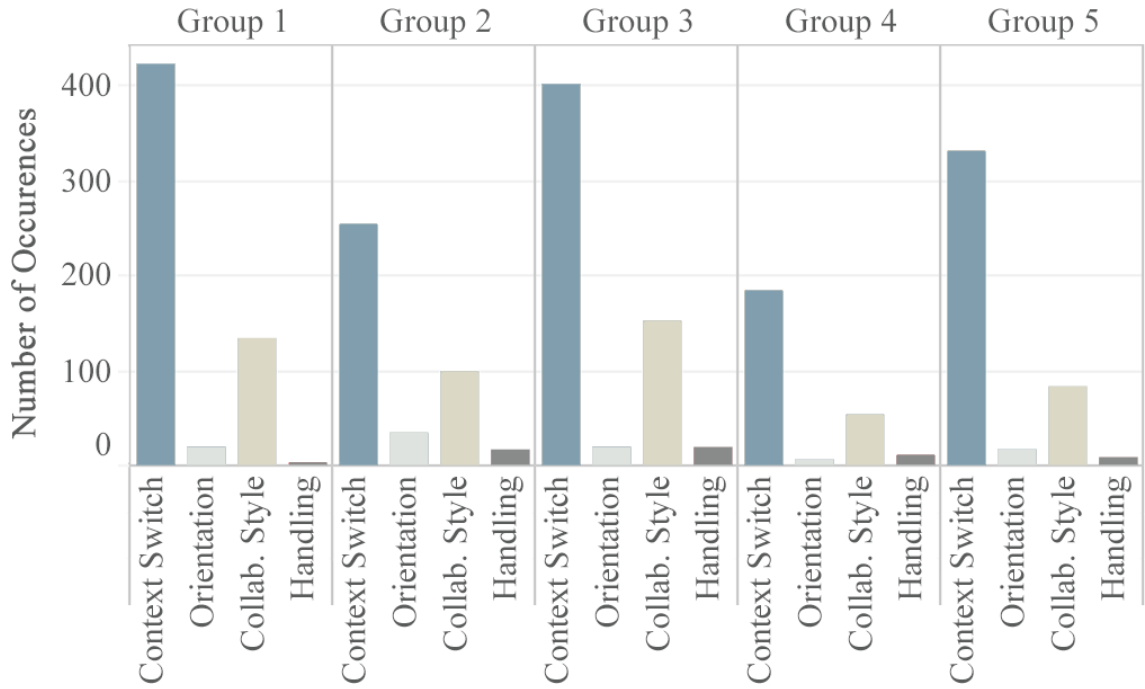
of Single Shared View, 57 of Disjoint and Shared, 128 of Disjoint and Distributed view, 104 of Distributed View, 167 of Distributed View with Discussion, and 27 of Disjoint Shared and Disjointed.

Figure 3.8 shows the breakdown of each group and the number of instances for each metric that was measured. The amount of context switching was extremely high compared to other metrics. A repeated measures ANOVA shows that on inspection tasks, there is no significant effect for the levels of difficulty ( $F(2, 8) = 1.690$ ,  $p > 0.1$ ) when it comes to context switches. However, further pairwise comparisons reveal that there is a significant difference between the third level of difficulty (mean: 17.8) and the other two levels (Level 2 mean: 34.8, Level 1 mean: 38.8).

### 3.4.2 Task Completion Patterns

Although the frequency of context switches has almost doubled between levels 2 and 3, the time it takes to complete each task is similar. Statistical analysis shows that during the categories of device handling and device orientation, there was no significant difference between each measure on the difficulty level. The handle measure specifically showed that participants would rarely put down or pick up their devices ( $F(2, 8) = 1.106$ ,  $p > 0.1$ ).





**Figure 3.8:** Overall occurrences for each group.

The maximum number of times a participant would put down their device was 1, which was toward the end of a task. Participants would generally hold their device for the entire time, even doing other tasks.

In general, there was no significant difference between the changes in collaboration style within inspection tasks ( $F(1.01, 8) = 0.659$ ,  $p > 0.1$ ). There was a significant difference between the first (mean: 12.2) and third (mean: 7.8) levels of difficulty, showing that the less difficult model had more collaboration style changes, particularly with participants looking through each other's devices and engaging in active discussion.

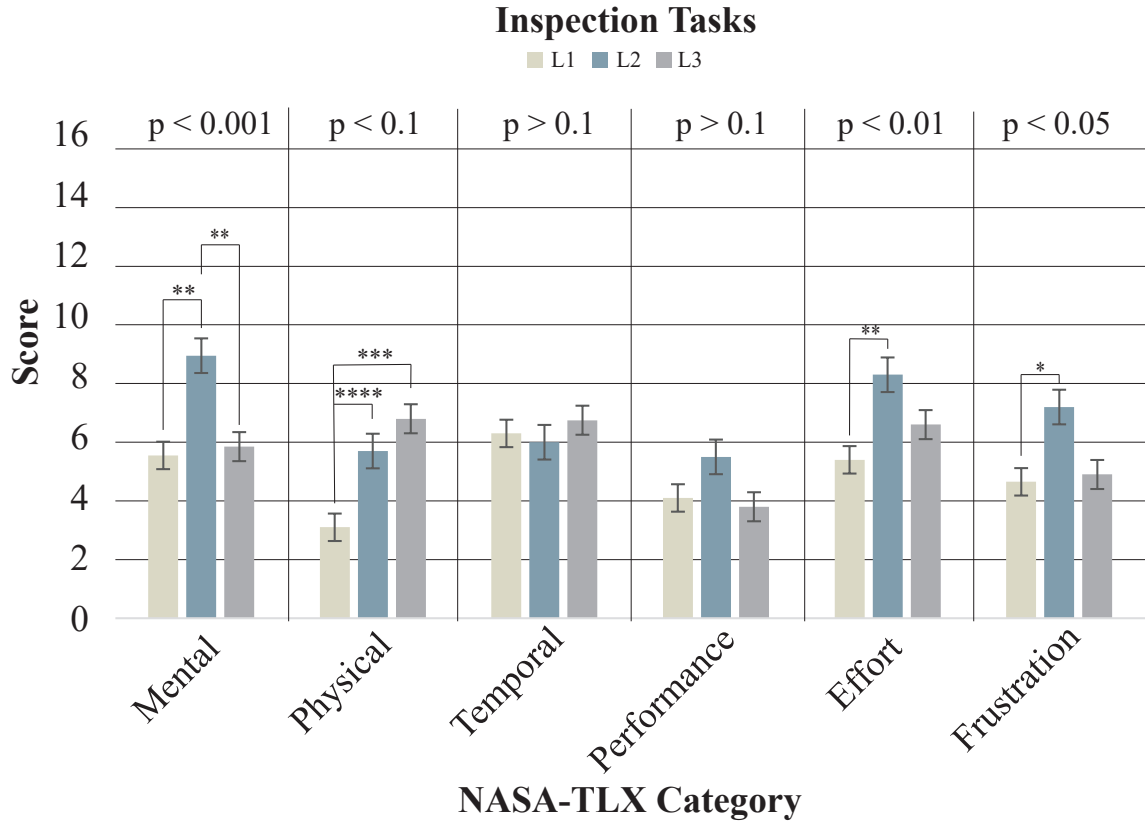
Similar results were obtained through the comparison tasks. We found a significant difference across difficulty levels ( $F(2, 8) = 9.003$ ,  $p < 0.05$ ). The pairwise comparisons revealed that there is a significant difference between level 1 (mean: 51.8) and level 3 (mean: 26.4), which is similar to what we found in the inspection tasks. However, there is also a significant difference between level 2 (mean: 57.2) and level 3.

The collaboration style metric showed that there was a significant difference between the levels of difficulty ( $F(2, 8) = 13.159$ ,  $p < 0.05$ ). The main difference found within the styles was that, at the highest difficulty level, there were significantly fewer changes to the collaboration style.



### 3.4.3 Usability Analysis

We applied the NASA-TLX metrics to gauge the perceived workload. We observed that the perceived load over all metrics were quite high, indicating that participants often had to multitask. This finding was consistent in both the Inspect and Compare categories.

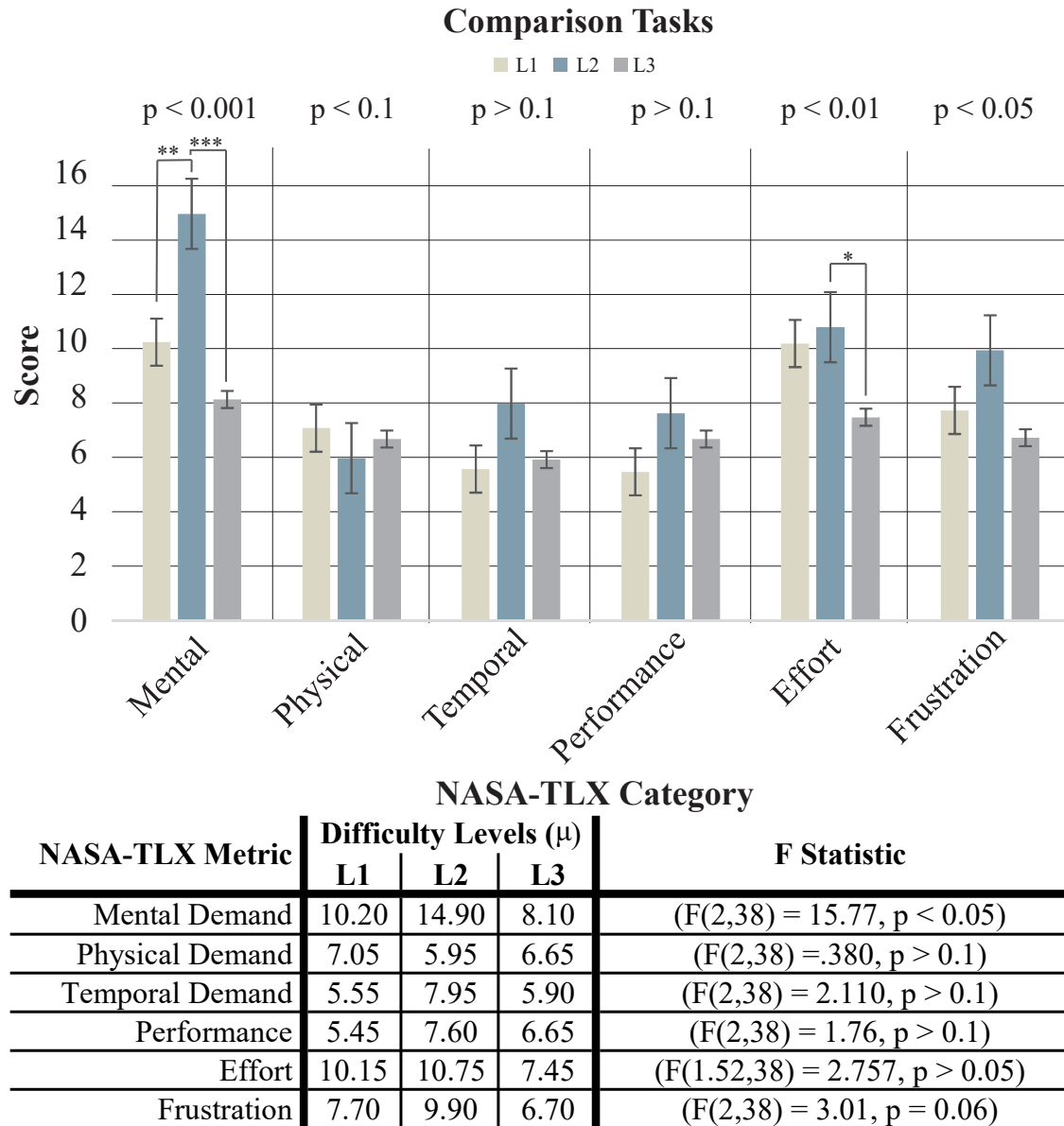


| NASA-TLX Metric | Difficulty Levels ( $\mu$ ) |      |      | F Statistic                          |
|-----------------|-----------------------------|------|------|--------------------------------------|
|                 | L1                          | L2   | L3   |                                      |
| Mental Demand   | 5.55                        | 8.95 | 5.85 | (F(141.73,233.6) = 11.528, p < 0.05) |
| Physical Demand | 3.10                        | 5.70 | 6.80 | (F(144.40,284.93) = 9.63, p < 0.05)  |
| Temporal Demand | 6.30                        | 6.00 | 6.75 | (F(2,38) = 0.421, p > 0.1)           |
| Performance     | 4.10                        | 5.50 | 3.80 | (F(2,38) = 2.39, p > 0.1)            |
| Effort          | 5.40                        | 8.30 | 6.60 | (F(2,38) = 6.87, p < 0.05)           |
| Frustration     | 4.65                        | 7.20 | 4.90 | (F(79.03,394.97) = 3.802, p < 0.05)  |

**Figure 3.9:** Statistics of the inspection tasks shown as figures. The top figure is a graph showing the pairwise comparisons and significance between each level. The bottom figure is the table of the mean values and the F-statistic for each metric.

The analysis of NASA-TLX shows significant differences in Mental Demand, Effort, and Frustration between both categories over the three different levels of

difficulty. This shows that our abstract exemplars were sufficiently different in difficulty, and when this level increases, it becomes more mentally challenging for the users to complete tasks. Within the Inspect category, a significant difference was found within the Physical Demand section.



**Figure 3.10:** Statistics of the comparison tasks shown as figures. The top figure is a graph showing the pairwise comparisons and significance between each level. The bottom figure is the table of the mean values and the F-statistic for each metric.

### 3.4.4 User Experience and Observations

We report our findings from our observations and interviews under three themes: (i) context switches, (ii) device handling, and (iii) collaboration styles.

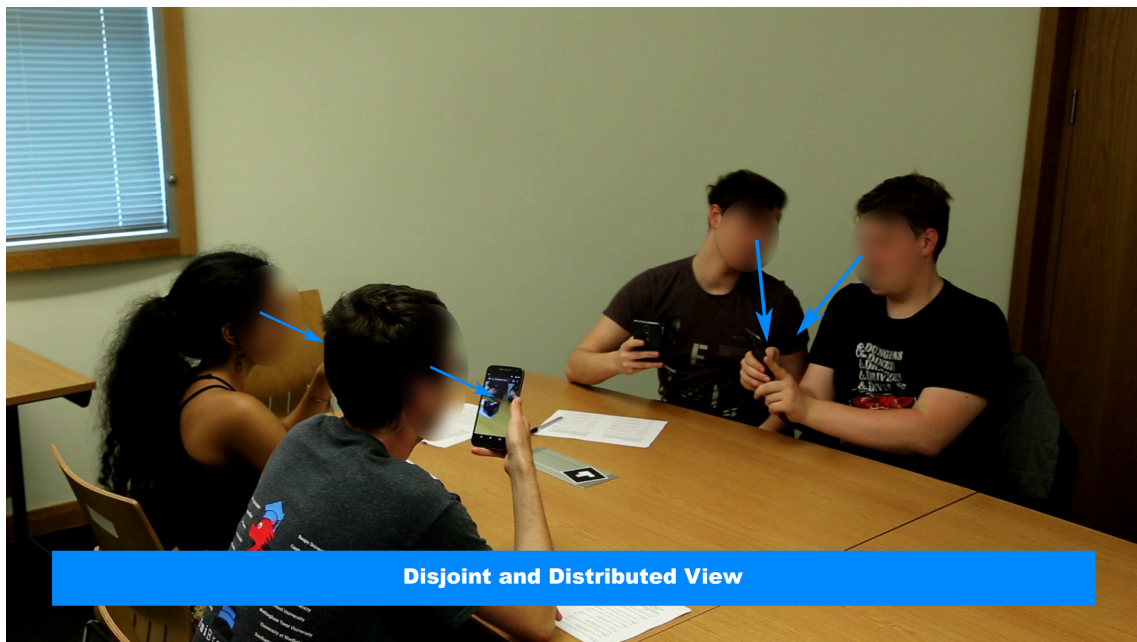
#### Theme 1 — Context Switches

AR allows collaborators to dynamically switch focus between a workspace and a communication space [14]. We observed that context switches occurred when referencing locations, for example, when a participant would point out “*the face closest to [the researcher]*” – P7, or “*the face closest to the camera*” – P1. These context switches also occurred when participants wanted confirmation of a particular statement they had said or even just in a natural conversation. If participants addressed a specific participant, they would generally glance in their direction, even if the participant did not look back.

Despite the high number of context switches (Figures 3.6 and 3.8), there was a lot less context switching during the inspection tasks in contrast to the comparison tasks. In particular, with the exception of group three (18%), the most difficult task had the lowest number of context switches (G1: 6%, G2: 8%, G4: 9%, G5: 12%). Even though the second level of difficulty was considered the most mentally demanding in our quantitative findings, the third level of difficulty received the least number of context switches, as the participants were really concentrating on the device. The same is also true within the comparison tasks, where the least amount of context switches occurred on the most difficult task for all groups (G1: 11%, G2: 14%, G3: 5%, G4: 4%, G5: 7%). Some participants mentioned that the sixth virtual model was the most difficult because of how much they had to focus. The participants mentioned that “*if you got closer, you could not see both sides [of the model] and it became more confusing*” – P4. This could explain why the context switching happened less frequently, as they did not want to lose where the side they were interacting with.

#### Theme 2 — Collaboration Styles

We were interested in what collaboration patterns would emerge between participants within the different levels of difficulty. The most used collaboration style in groups 2-5 was Distributed View with Discussion (34%, 36%, 36% and 30%). This means that each group would actively discuss the task at hand with each other while looking at their own device rather than each other. Group 1 tended to use a little more Disjoint and Distributed View (29% see Figure 3.11) when compared to Distributed



**Figure 3.11:** Disjoint and Distributed View was the second most common collaboration style.

View with Discussion (26%), in which we observed that they would often check each other’s devices to confirm that they were seeing the same object. The video data shows that there is a common trend with people looking through their device onto the model but using their words and hands to place spatial signals to other users when discussing. For example, a participant said “*it is the side closest to the camera*” – P14 when asked to find which face had an L shape made of red tiles. While *spatial referencing* contributes towards coordination, our observations show that it is commonly used in AR-mediated collaboration as a workaround and replacement for direct face-to-face interaction.

We observed that every group had a high percentage of their collaboration styles as Disjoint and Distributed View (29%, 28%, 19%, 23%, 26%), with two users sharing one device to view the model. A trigger generally initialised this to do a context switch from their own personal device to the group. Such triggers include a brief conversation about content that the other participant could view, a brief look at the participants sitting next to them, an explicit invitation to look at another device, or generally just walking to the other side of the table to view and confirm. In some cases, participants would walk around the table to the other side to look at the same view. But rather than looking at their group member’s device, they would look

through their own device instead. On one or two occasions, a participant glances at another group member’s device to confirm that they are seeing the same thing. Participants would point out a certain part of the model or bring attention to it to ensure that they were looking at the same thing through their own devices.

There was little to no Single Shared View usage, and only groups 2 and 3 used this shared view *once*. Common practice was that participants would sit opposite to each other so that they could view both sides of the model. They would only move around when they needed to be where all participants could view the same part of the model. Instead, they would trust each other when answers were given as part of the task. If participants had to confirm something and look at one part of the model, they would do so using their own devices instead of looking at the same device. This is likely caused by the form factor of the devices: because they were small, it was easier to focus on their own device and view.

### Theme 3 — Device Handling

We expected that the orientation of the device would change significantly during tasks to allow for more content to fit on the smaller form factor screens. However, our data show that only a few participants turned their device to create a different field of view. Our observations show that most of the participants decided to keep their phones in portrait instead of landscape mode. When asked about this, one participant said, *“I never really use my phone in landscape mode. It never crossed my mind to do that”* – P9. Another reason is that the device is easier to hold and more stable in portrait mode, thus increasing mobility and allowing participants to do other tasks more easily or perform context switches.

Most 3D objects were scaled in such a way that if the marker was in the middle of the table, the model would fit completely in the CollabAR viewport when standing next to the table. From there, the participant could get closer to inspecting the object in more detail. If the models were scaled larger, more participants may have felt the need to change their device landscape. This observation is supported by some of the groups, who turned and moved their device a lot more with the larger models, stating in the interview that *“sometimes it was difficult to concentrate on the larger objects because of the other things in the room”* – P7, such as a TV and cupboard. More specifically, it is supported by findings from the comparison tasks. Here, participants attempted to fit both models within their screen to be able to make the comparison. They would turn their phones to landscape, which meant we saw an increase in the orientation of the device during the comparison tasks when compared to the inspection tasks.

### 3.4.5 Summary of Findings

*Utilisation Patterns* – Video analysis shows diverse and changing collaboration styles, a high degree of context switching, and a decrease in collaboration for complex models.

*Usability* – Video analysis shows diverse and changing collaboration styles (Figure 3.6), high degree of context-switching (Figure 3.8), and decrease in collaboration for complex models.

*Observations* – We observed a range of group practices and appropriations, including hand signalling to coordinate work, approaches to switch context between the AR interface and the face-to-face collaboration, glancing through other peoples’ devices to confirm tasks, or physically moving around the table to inspect the model in order to ensure operation consistency for other users.

## 3.5 Discussion

Our study shows how mobile AR mediates collaboration. Although participants were able to complete the tasks, we identified recurring problems, observations, and workaround strategies to mitigate the absence of *collaborative features*. In the following, we discuss the findings and present design recommendations.

### 3.5.1 Context Switching and Device Handling

Our findings indicate extreme amounts of context switches for all the levels of difficulties for each model, especially considering the relatively short task duration times. Our data indicates that collaboration in M-AR specifically has a very high mental demand as users divide their attention between managing the device and collaborating with other users. In particular, the combination of configuring the device correctly in space to see the AR marker with actually looking through the device to see the model produces a significant amount of *configuration work* [88]. The number of context switches decreases slightly as models get more complex, which – if we class a switch of context as an ‘interruption’ – shows that there is more focus on the models but less explicit collaboration. This was mitigated by workarounds such as spatial signalling, waving, or conversing with others without face-to-face interaction. In general, our study shows that *as the intricacies and complexity of the models increase, there is less collaboration, and users report a high level of frustration in completing the tasks*.

Throughout the entirety of our study, we observed very little device handling - by which we mean the placing down and picking up of devices. Participants argued that it was important for the system to maintain the viewport and the model's angle so they could go back to the same point when analysing and discussing the tasks. As there are no tools available to help participants quickly recover from putting their device down and picking it back up again to continue where they left off, they would simply hold it in place while interacting with other participants or looking at other devices. While holding the device, participants constantly adjusted, moved and turned the device to ensure that the AR models were clearly visible. Especially in a collaborative setup, this constant re-alignment and fine-tuning of the device orientation and angle is a considerable distraction from the actual collaborative activities and imposes a high physical and cognitive demand on users. *The basic requirement to constantly hold up a device to be able to see the virtual model has a direct negative impact on collaboration.* Although this could be overcome by utilising HMDs to collaborate, it removes the portability and accessibility of this type of collaboration, which is what makes this type of AR important.

### 3.5.2 Impact of Model Complexity on Collaboration

For complex models, the number of changes in the collaboration style decreased as participants focused more on their own devices. In general, the more complex the model, the less dynamic collaboration occurs. Across all of the tasks, the most used type of collaboration style is 'Distributed View with Discussion'. In this style, each participant focused on their own device while actively discussing the task at hand. This type of collaboration style occurred 167 times in the study and was the dominant way for participants to collaborate. Participants, thus, collaborate almost literally through the device, and the many observed context switches are the shift from looking at the device to face-to-face discussions. The second most popular collaboration style occurred 128 times and was a Disjoint and Distributed View. Here, 2 or 3 participants focused their attention on one single mobile device (sharing one viewport), while others in the group focused on their own device. This happened more often during complex models, as participants wanted to confirm what they were seeing through another device even though what they were seeing was the same on both devices. Having a shared view allowed the participants to focus their attention on the same thing, and they felt more confident in the knowledge that they 'could see the same thing on the same device', so there was no confusion.

The least used collaboration style is the Single-Shared View - which is when all participants focus their attention on a single device. Over all tasks and groups, this type of collaboration style only happened twice, when the whole group wanted to confirm one part of a model. Upon observation, it appeared to be difficult for all participants to huddle around one mobile device due to the angle and position of the device but also because there were 4 participants. We observed that no more than two participants would usually share a single viewport for a short period of time. One of the report issues is that sharing one single viewport puts constraints on the angle and position of the device, as multiple participants needed to see the AR model. This implies that the participants could not get closer to the models and view intricacies due to physical limitations without a considerable amount of configuration work.

### **3.5.3 Mobility in Tabletop AR**

As AR relates digital models to a location, position, and anchor in the real world, we observed a considerable amount of mobility work. Zooming and inspecting different sides of the model or comparing multiple models all required participants to walk around the meeting table. Our results show that there is high physical demand throughout all tasks and that this increases when models are larger or more complex. Participants would view the model from different directions and thus move in all directions to obtain different vantage points over the model. Our observations even show that participants, in some cases, would have to stand on their tiptoes to view the top of the model. In addition, there was a lot of standing up and sitting down throughout the study. Participants also moved closer or farther away from the device to zoom in, causing them to lean far over tables to see the model closer in the physical space. Although CollabAR allowed participants to rotate and scale the model, most of the groups simply walked or moved around the model to see different sides. We believe this is to ensure that the model would remain ‘stable’ for the entire group and not make a modification while multiple participants are collaborating with the same model. This strong mobility requirement has implications for the accessibility and general usability of AR.

### **3.5.4 Design Recommendations**

We propose the following five concrete design principles to improve the co-located collaborative AR experience. These are directly grounded in our empirical findings and orientated towards building applications and systems for collaborative work.



- D1 View-Clutching** – Our observations show that a lot of effort went into aligning and configuring the AR view. A clutch mechanism could allow a user to hold the viewport in place whilst looking away from the device. This would mitigate the need for re-configuration and re-alignment steps when the user’s attention shifts between looking at the device and looking at other users during the collaboration. Clutching the view would also allow users to move the device out of the way without losing the model on the screen.
- D2 Spatial Awareness** – In our study, context switches were often caused by mobility (movement) around the table. This could be mitigated by supporting visual feedback on the location, distance and orientation of the other users in the room or by giving users accurate spatial references concerning the model. Particularly when multiple users are inspecting the virtual object, the system could provide indicators for where users are in relation to the physical marker or object representing the virtual model.
- D3 Visual Guidance** – To help users recover from context switches or change to another collaboration style, a system can provide a visual guidance mechanism which would help users ‘find’ an appropriate view when they pick up the device. To avoid longitudinal and manual configuration of the view, the system could suggest where, when, and how to look at the model by providing visual indicators for the right direction and distance. This particular design principle could also help provide various views for users, which allows for different collaboration styles to be used.
- D4 Awareness Cues** – While the coloured border helped users see if interactions were locked, more visual feedback can be provided to the user on where other users are looking and who are interacting/modifying the model in view. Similarly to screen-based groupware, collaborative M-AR interfaces could provide awareness cues for what content is being used by each person. The ‘turn-based’ coloured frame visualisation in CollabAR is one example, but more advanced visual cues could be included in the interface.
- D5 View Filters** – To handle increasing complexity in virtual models and objects, the system can remove the background clutter and noise from the physical space and allow users to toggle between the full-screen camera view with the AR model in view and a more annotated, collaborative view that is optimised

to see collaborative features, awareness cues, visual indicators, or even parts of the model which are visible to each user. These different filters could also support the articulation of certain viewing angles, enabling complementary information when users observe the model from different angles.

### 3.6 Limitations and Future Work

Understanding the impact of M-AR interfaces on collaboration is important to be able to push the field forward. There is a need for new design principles and studies that will pave the way for new frameworks for M-AR to consolidate different AR applications. Our work focuses on how current M-AR applications mediate ad hoc collaboration. We provide insights through analysis of collaboration metrics such as context switching, collaboration styles, and device handling. We present empirical findings of how users collaborate in a co-located setting and extrapolate design principles.

Our study was limited in size and scope, with only 20 participants with limited diversity in demographics. Although we attempted to create an environment similar to that of the real world (standard meeting room), the study itself was controlled without a specific focus on one application area. However, while there are limitations in the scope and size of the study, it sets a clear path for further research in this area. The results of this study provide general reflections and insights into how collaborative AR applications work in co-located settings using handheld, through-the-screen AR. Our work encourages further development in this area of mobile AR, and we believe that the design principles and concepts introduced in this chapter can inform future studies and AR technologies.

A more in-depth study of context switches is warranted to uncover important results. In addition to this, the mobility of the study (people walking around the table) is something that could also be taken into account, as well as how this affects the collaboration. A limitation of our study was the limited types of tasks for users. Future research could look at how a similar collaboration framework lends itself to more advanced tasks which take a longer amount of time. Through these studies, there could be further conversational analysis, as well as looking at how different age groups could play a role. It could also be beneficial for comparative research to be undertaken; for example, how might context switching differ between handheld vs. head-mounted AR displays? How could markerless AR solutions change the user experiences of relocating and manipulating objects? And how might different form factor displays affect shared-view experiences?

## 3.7 Conclusion

This chapter concludes that, while prior work argues that Augmented Reality (AR) facilitates collaboration, there is little empirical work into how AR affects co-located collaboration. We explored how current mobile AR interfaces mediate collaborative work. Our results show that while people can collaborate through a mobile AR interface, there is a high physical and cognitive load. We also observe extreme amounts of context switches between looking at the AR model through the device and face-to-face collaboration. As virtual models become more complex, the amount of collaboration decreases. Based on our empirical findings, we present design principles for implementing mobile AR interfaces that support co-located collaboration.

In the subsequent [Chapter 4](#), following lessons learned during this chapter, we look further into the completion of tasks using a simple handheld mobile AR interface, paying particular attention to how different device sizes and configurations impact key factors of collaboration such as collaboration strategy, behaviours, and efficacy.

# 4

## A Study into the Effect of Mobile Device Configurations on Co-Located Collaboration using AR

### Contents

---

|            |                                    |            |
|------------|------------------------------------|------------|
| <b>4.1</b> | <b>Introduction</b>                | <b>76</b>  |
| <b>4.2</b> | <b>Motivation</b>                  | <b>77</b>  |
| <b>4.3</b> | <b>Study Detail</b>                | <b>79</b>  |
| <b>4.4</b> | <b>Results</b>                     | <b>85</b>  |
| <b>4.5</b> | <b>Discussion</b>                  | <b>98</b>  |
| <b>4.6</b> | <b>Limitations and Future Work</b> | <b>101</b> |
| <b>4.7</b> | <b>Conclusion</b>                  | <b>102</b> |

---

Building upon our exploration of how Augmented Reality (AR) impacts co-located collaborative experiences, this chapter delves into the specific role that device configurations play in shaping these interactions. Following the findings of [Chapter 3](#), where we identified cognitive and physical demands imposed by current handheld mobile AR interfaces, we now turn our focus to the devices themselves and how these affect cognitive and physical demands. This chapter seeks to investigate how different handheld mobile AR devices influence user interactions and collaboration in real-world scenarios. This chapter and the research study within it were conducted during the COVID-19 pandemic, a period that was characterised by significant social and operational disruptions. To adhere to safety guidelines and to ensure

participant well-being, we implemented several adapted measures, including reduced group size and some social distancing protocols.

We aim to develop a comprehensive understanding of the interplay between device form factors, user engagement, and collaborative dynamics. By examining how various device configurations affect the efficiency and fluidity of AR-enhanced collaborative tasks, we address [RQ2](#): *How does device size and configuration affect the facilitation of co-located mobile AR collaboration?*

The increasing availability of portable handheld mobile Augmented Reality technology is revolutionising how digital information is embedded into the real world. As this data is embedded, it enables new forms of cross-device collaborative work. However, despite the widespread availability of handheld AR, little is known about the role that device configurations and size play in collaboration. This chapter presents a study that examines how completing tasks using a simple mobile AR interface on different device sizes and configurations impacts key factors of collaboration, such as collaboration strategy, behaviour, and efficacy. Our results show subtle differences between device size and configurations that have a direct influence on the way people approach tasks and interact with virtual models. We highlight key observations and strategies that people employ across different device sizes and configurations.

## 4.1 Introduction

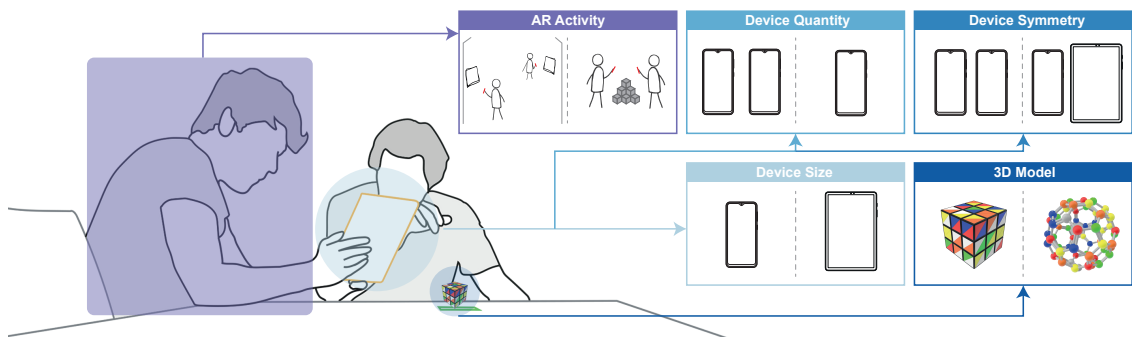
Currently, no studies or insights look at the role of device form factor and configurations in guiding group collaboration in mobile AR – meaning *we do not understand how using various device setups and sizes affects group work*. Inspired by work on device size in cross-device computing [26, 223], this chapter is an initial step in understanding the characteristics of mobile devices. As part of a study aimed at evaluating the impact of screen size and device configuration on AR-mediated co-located collaboration, we designed and conducted an experiment with dyads of participants to examine group strategies and usability metrics as they interacted with a test application. We investigated the effects of five different device configurations spanning two different device sizes (a mobile phone and a tablet) on collaboration with mobile AR. We examined (i) the user’s ability to complete tasks and the perceived workload, (ii) collaborative behaviours such as focus, communication, and device interaction, and (iii) general strategies of collaboration adopted by participants.

Through this study, we aim to understand the implications of device symmetry and sizes on how task labour is divided in AR for pairs of users in a collaborative

tabletop setting. We conducted an experimental study using the preexisting mobile AR system for model manipulation found in [Chapter 3](#). Participants were presented with two different models augmented onto the real world (hereby referred to as virtual objects) with different levels of occlusion. Participants were required to complete two types of collaborative tasks on these virtual objects in AR, namely identifying and locating specifics about them. We believe the results provided will help future researchers better understand how screen size affects their subjects, especially when designing and evaluating mobile applications. Our findings show preliminary evidence and insights into the nature of collaboration between device configurations, including device symmetry, size, and quantity, as well as the effect of model occlusion and type of AR activity on collaboration.

## 4.2 Motivation

Previous work illustrates that while AR supports new forms of collaboration [46, 188], it also induces problems and challenges related to device type, screen form factors and AR activity. In this chapter, we extend previous work related to hardware used for accessing augmented reality, along with AR components as well as 3D models that are augmented onto the real world, specifically exploring how these affect co-located AR collaborations. We dissect these challenges in co-located collaboration in AR into five individual research categories that we examine in our study: (i) *device symmetry*, (ii) *device size*, (iii) *device quantity*, (iv) *model occlusion* and (v) *AR activities* (see [Figure 4.1](#)).



**Figure 4.1:** The five research categories looked into throughout this chapter are demonstrated in the study set-up.

### 4.2.1 Device Symmetry (H1)

Device symmetry is a term used frequently in cross-device computing to describe how devices are arranged [127, 162]. In this chapter, we refer to device symmetry as when pairs of participants use the same device (i.e. two phones) or different devices (i.e. a tablet and a phone) in a task. When participants need to accomplish a task collaboratively but across separate devices, they may need to consider differences in asymmetric device configurations. As such, we hypothesise that device asymmetry will impede participant collaboration, with more focus on individual devices in multi-device setups and increased verbal communication as a workaround.

### 4.2.2 Device Size (H2)

Device size is a term commonly used to describe different form factors of device displays. Studies examining the effect of screen size exist (Section 2.2), but the effect of handheld devices on collaboration remains poorly explored [26]. Handheld devices can produce a gorilla arm effect [83, 96] depending on their weight, with designers reporting restrictions in mid-air gestures and handheld mobile devices as a result of high consumed endurance and fatigue [57, 83]. Larger devices, of course, allow more content to be viewed at once, but smaller devices are generally more portable and easier to handle. Our second hypothesis is that larger device sizes will afford more device sharing during collaboration with participants, giving more focus to each other's device(s).

### 4.2.3 Device Quantity (H3)

Device quantity is simply the number of devices available during collaborative tasks. As the number increases, so does the task of managing the devices in the scenario [76]. In pairs, with a maximum of two devices, the only real consideration is the claiming or sharing of ownership, along with the negotiation of control in multi-device setups. We hypothesise that in single-device configurations, participants with device ownership will interact with the virtual model more, as opposed to multi-device setups, as participants won't be concerned with breaking shared operational consistency.

#### 4.2.4 Model Occlusion (H4)

Model occlusion is a problem that occurs in 3D environments and is most commonly seen when a user navigates through them for 3D visualisation. Occlusion of objects is caused by the environment itself, the model, and its geometrical properties [45, 191]. In 3D environments and on 3D models, occlusion effects influence how users perceive concepts such as continuity, proximity and atomic orientation [173]. We hypothesise that a more occluded model will result in less autonomous work and promote more communication among participants. Increasing the amount of collaboration necessary to complete a task.

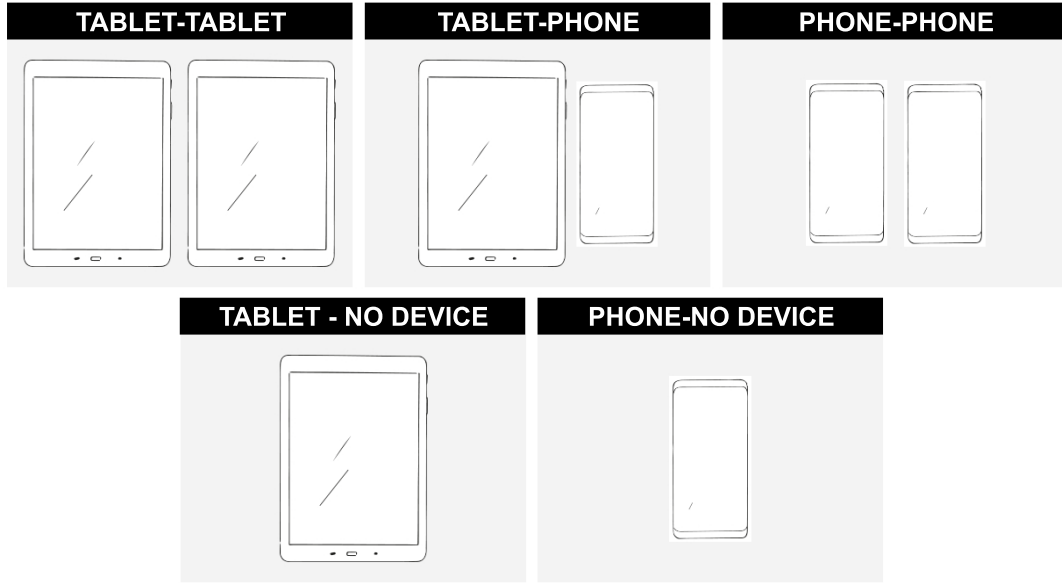
#### 4.2.5 AR Activity (H5)

An AR activity is a task or goal a participant needs to complete using Augmented Reality. This could be something as simple as different ways of experiencing a museum [103] or co-creating persistent structures in AR [71]. While the task may be the same, individual user goals may affect the overall AR activity and the overall collaboration to complete that task. We hypothesise that participants will need to adapt their collaboration strategies for different tasks, but the degree of adaptation will vary depending on the device configuration.

### 4.3 Study Detail

Extending on previous work in this area and within the defined research categories, this chapter aims to establish an initial understanding of the role of device size and configuration. We assume that in a world of heterogeneous devices, many real collaborations will involve a variety of device combinations and sizes. Similarly, we can assume that the 3D models being interacted with will vary in terms of complexity and occlusion. Additionally, users' activities can vary even when interaction techniques remain consistent when using handheld AR. Considering these assumptions, what implications do these factors have on the strategy and efficacy of user collaboration in AR?





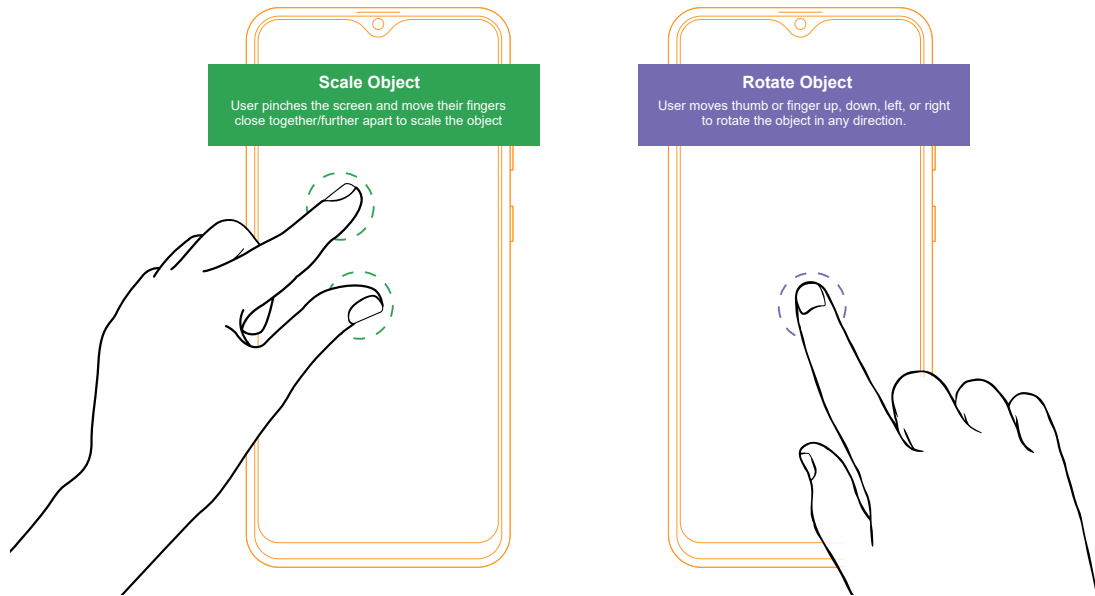
**Figure 4.2:** This figure illustrates the different combinations of device configurations between participants, including pairings of phones, tablets, and no device.

### 4.3.1 Device Ecology & Interaction Techniques

We study 5 different device configurations involving handheld mobile devices and tablets that vary in the number of devices and heterogeneity (Figure 4.2). These configurations include Phone + No Device, Tablet + No Device, Phone + Phone, Tablet + Tablet, and Tablet + Phone. We specifically focus on device configurations *designed for pair work*, as related work demonstrates that pairs are sufficient for evaluating general collaboration styles and strategies [95, 154, 223]. The device configuration, interactions, and tasks are designed specifically for tabletop collaborative settings with AR as a mediating technology. We used touch-based WYSIWYG interaction techniques for rotating and scaling the virtual model (see Figure 4.3). Inspired by the CollabAR system used in Chapter 3, we intentionally distribute the model interaction across all devices. Each user can use their device to manipulate the model whilst seeing the changes in real-time. This ensured a more organic development of rules and group strategies. We also utilise marker-based tracking to physically position the virtual model, which users can freely manipulate.

### 4.3.2 AR Tasks & Virtual Models

To evaluate group collaboration, we employ two types of virtual models and two types of tasks to be completed in every device configuration. The two task types were

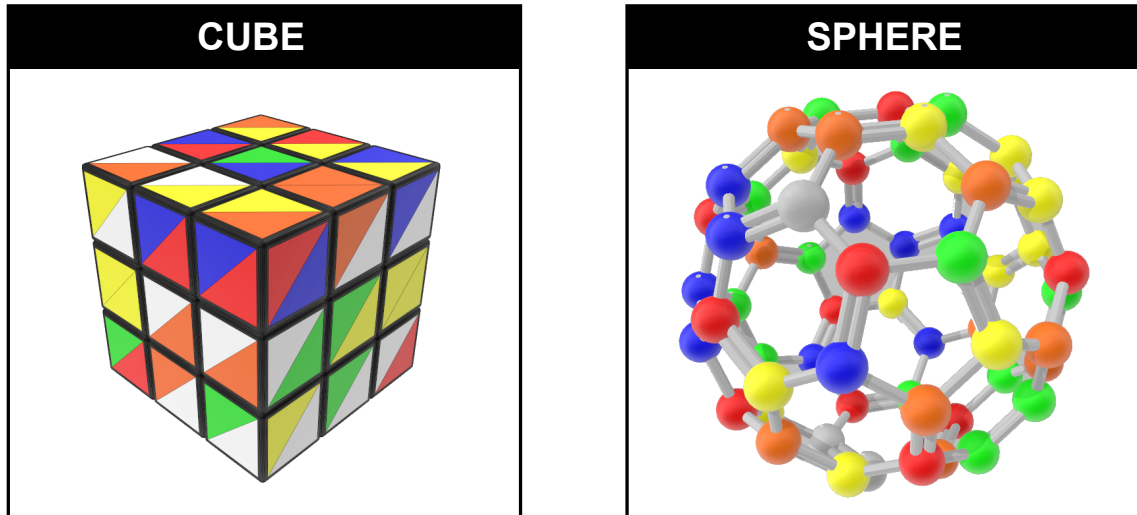


**Figure 4.3:** Basic interactions with the system. Rotation is controlled by swiping the object and scaling via pinch gestures, as is common in many commercial AR apps.

“identify”, describe something about the virtual model, and “locate”, find an object of interest on the virtual model. The tasks were designed for the study based on their simplicity and application independence, rather than being directly drawn from prior work. The aim was to ensure that any observed changes in collaboration were attributable solely to the complexity of the virtual models and device configurations, rather than to the inherent difficulty of the tasks themselves. The models were designed to be immediately recognisable by participants, incorporating assorted intricate details such as randomised colours. The two models, with varying levels of *occlusion*, were used for each device configuration and repeated for each AR task (see Figure 4.4). Utilising two different models, we can explore the effect of different levels of occlusion within a virtual model on how a device configuration is utilised and how collaboration strategies are formulated. Model 1, a variation on a Rubik’s cube, and Model 2, a variation of a fullerene sphere, were designed to be immediately recognisable by participants but with assorted intricate details, such as randomised colours.

For *identify* tasks, participants would be asked, for example, ‘count how many tiles on the cube that have the colour combination blue and yellow’. For *locate* tasks, participants would be asked to find something specific such as ‘locate the hexagon on the sphere which has the colour combination orange, red, blue, yellow, green, yellow’. Across both tasks, participants were encouraged to come to a

consensus through deliberation before finalising an answer for the task. Task time and task errors were recorded, but participants were informed to complete the tasks within their own time. Once a consensus was reached, participants would proceed with the next task and model.



**Figure 4.4:** The two different models used for each task. One with more occlusion (Cube, Left) and one with less occlusion (Sphere, Right).

### 4.3.3 Apparatus & Participants

We recruited 20 participants (9 identified as male, 11 identified as female) and divided them into pairs. Using a snowball sampling method, we recruited pairs of participants who were already acquainted. There were no special requirements for this study other than the participants who had no colour vision deficiency, as tasks required the identification of colours. The tasks were abstract and not domain-specific, so no information about the task content was provided prior to the study.

Most participants were between 18 and 24 years old, with two participants aged between 25 and 34. Whilst not a requirement, all participants had some AR experience prior to this study. Most cited social media filters (such as Snapchat and Instagram) as their use of AR. Other AR applications include Pokémon Go, Google AR, and Minecraft Earth. All participants owned and regularly used at least one handheld device (such as a mobile phone or tablet).

Participants were invited into a circular meeting room setting, which allowed for free movement around a table (see [Figure 4.5](#)). Chairs were provided, and participants were free to arrange themselves around the table as they pleased. Participants were

also informed that they did not have to remain seated. After completing a consent form, participants completed a brief prestudy questionnaire, which gathered basic demographic data, including their prior experience with AR technology.

The device configurations included either a Samsung A70 (mobile device), which has a display size of 6.7", or a Samsung S5 Tab (Tablet), which has a display size of 10.5". The order of device configuration presented to the participants was counterbalanced across each group using the Latin Square method to account for learning effects during the study. In configurations where there was only a single device, participants were free to decide how ownership and control of the device should work. Each participant was seated around the table, with the experimenter at the opposite end of the room. Pens and NASA-TLX forms were provided to complete after every task. Participants were also free to use pen and paper to note when completing the tasks.

After explaining the study, participants were given an introduction to the handheld AR system and time to become accustomed to the interaction. After this, they were given a brief sensemaking task, which allowed them to familiarise themselves with the models and also further solidify how to use the system. The introduction was completed using their initial device configuration. Participants were then informed of each task by the experimenter, who also provided the subsequent models by replacing the physical markers. After each task was completed, the participants would complete a post-task NASA-TLX questionnaire to capture the perceived workload for that task using that device configuration and a Likert scale survey regarding how they felt they communicated during this task. After all tasks and configurations were completed, we concluded the study with a semi-structured group interview, probing the different device configurations and the strategies adopted by the participants for the different tasks and models.

#### 4.3.4 Data Collection & Analysis

Each study session was video recorded, and the experimenter took observation notes to analyse general **collaborative strategies**, participant **focus**, **communication**, and **group interactions**. After all tasks were completed for a device configuration, the participants were asked to complete a NASA-TLX questionnaire to record their perceived workload for the entire device configuration. In addition, participants would also be asked to complete a Likert scale survey to record a self-reflection on how they felt they communicated during this configuration. On completion of



**Figure 4.5:** The study was set up in a circular meeting room portraying a pair of participants with the Phone + Phone configuration. Virtual models were projected using standard marker tracking.

all device configurations, the Likert scale answers, NASA-TLX questionnaires, and experimenter observations were used to elicit conversation during the semi-structured interview to conclude the study session.

The video footage was analysed by a colleague and me to define a set of collaboration strategies, and for each participant group, each task was labelled with a predefined strategy. We used and adapted an existing coding framework used by Zagermann et al. [223]. Similarly to this study, we recorded participants' interactions with their devices, i.e., how many rotations were performed on the virtual model, as well as their focus and communicative behaviour. The following codes were derived.

**0 - Focus** - To investigate participant's focus during the tasks, we distinguish between sub-codes:

0.1 *Focus on Own Device* - where the participant's focus is on their own device.

0.2 *Focus on Partner's Device* - where the participant's focus is on their partner's device.

0.3 *Focus on Partner* - where the participant's focus is on their partner.

0.4 *Focus on Other* - where the participant's focus is on anything else.

**1 - Communication** - To analyse the communication of participants during the study, we coded:

1.1 *Talking* - where the participant is talking.

1.2 Silence - where the participant is silent.

We also further coded communicative behaviour as:

1.4 Spatial Reference - where the participant makes a verbal reference to the virtual space.

1.5 Deictic Gesture - where the participant makes a physical gesture in the virtual space.

Participant interaction data was also recorded on the handheld devices during the study. Using this coding scheme, we can determine the level of participant focus and communication for a device configuration at certain intervals in a task or throughout. We can also further describe the nature of collaborative strategies adopted under different device configurations regarding focus, communication, and interaction. We analyse the NASA-TLX data, task time and error, as well as participant quotations to understand the impact of a device configuration and collaboration strategy on the perceived *workload* and *task efficacy*.

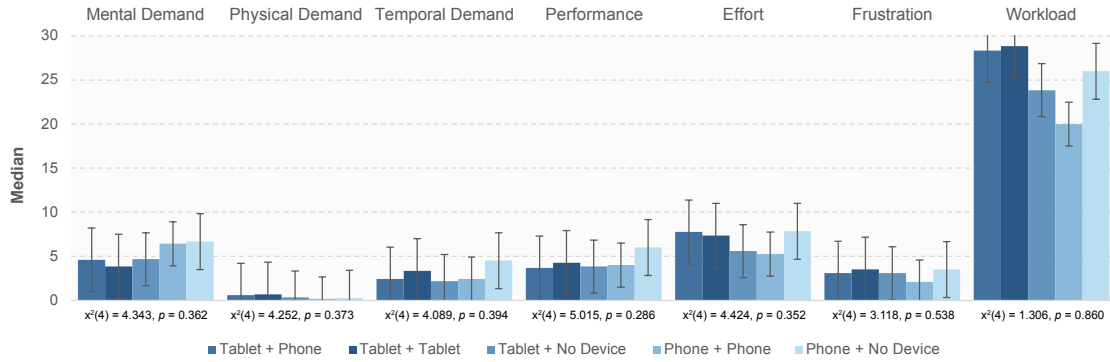
## 4.4 Results

The following section presents the results from the user study with dyads of participants. We categorise these results into findings associated with the efficacy, nature, and strategies of collaboration. We observe a balance between both positive and negative effects that relate to how dyads are affected in their focus, mental and task workload, and general communicative behaviours. We also observe how collaborative strategies are adopted, adjusted, and appropriated to meet these effects' challenges.

### 4.4.1 Efficacy of Collaboration

#### Individual NASA-TLX Metrics

A Friedman test was run on each individual TLX metric and overall Perceived Workload to determine differences exist in reported scores across device configurations. It was found that most metrics were not statistically significant (see Figure 4.6 for individual test results). We see a trend in the Median values that indicates a higher perceived workload across all metrics on configurations that contain a tablet, including Tablet + Tablet ( $Mdn = 28.85$ ) and Tablet + Phone ( $Mdn = 28.35$ ),



**Figure 4.6:** Median responses for the NASA-TLX questionnaire metrics and overall perceived workload, along with K-Related samples chi-square results, where the p-values report the significance between device configurations.

followed closely by Phone + No Device ( $Mdn = 26.00$ ). The lowest perceived workload was in Phone + Phone configurations ( $Mdn = 20.00$ ). Below, we break down each individual metric and note any observed differences across them.

### Mental Demand

Configurations that contained tablets had the lowest medians, with the symmetric configuration Tablet + Tablet having the lowest ( $Mdn = 3.84$ ), followed by Tablet + Phone ( $Mdn = 4.59$ ) and then Tablet + No Device ( $Mdn = 4.67$ ). Configurations containing only phones tended toward the highest Median, with Phone + No Device having the highest ( $Mdn = 6.67$ ) followed closely by Phone + Phone ( $Mdn = 6.42$ ). We see a large difference between the symmetric Tablet + Tablet configuration and the single Phone + No Device configuration regarding how mentally demanding they made collaboration.

### Physical Demand

We see that the medians for Physical Demand rest mostly the same across all configurations, which lends to the tasks not needing much physical exertion to complete. We do see that configurations containing a tablet tended to have slightly higher Physical Demand, such as Tablet + Tablet recorded as the highest ( $Mdn = .67$ ) and Tablet + Phone recorded as the second highest ( $Mdn = .58$ ). In group interviews, some participants noted that holding, manoeuvring, and interacting with a larger screen had an impact on their physical demand. This holds true with the median of Phone + Phone configurations ( $Mdn = .17$ ) reported as the lowest. However, when using the Phone + No Device configuration ( $Mdn = .25$ ), we observed that participants generally moved around the space more. Primarily to join their partner but also to look around the model.

### **Temporal Demand**

Participants tended to feel more time pressure in Tablet + Tablet ( $Mdn = 3.34$ ) and Phone + No Device ( $Mdn = 4.50$ ) configurations. Other configurations reported roughly the same, where participants did not feel rushed to complete the tasks (see [Figure 4.6](#)).

### **Performance**

Performance metrics with the highest ratings came from the single device configuration Phone + No Device ( $Mdn = 6.00$ ), and the dual device configuration Tablet + Tablet ( $Mdn = 4.25$ ). This shows that participants generally felt they performed better when they would either share a device or have a larger device screen available.

### **Effort**

There is a notable difference between the reported Medians for Tablet + No Device ( $Mdn = 5.58$ ) and Phone + Phone ( $Mdn = 5.25$ ) configurations being lower in Effort when compared with the other three configurations. The highest reported Effort was in the Phone + No Device configuration ( $Mdn = 7.83$ ) with Tablet + Phone ( $Mdn = 7.75$ ) not far behind. There seems to be no correlation between Effort and specific devices, only that sharing one larger device or having two devices of the same smaller size results in similar Effort.

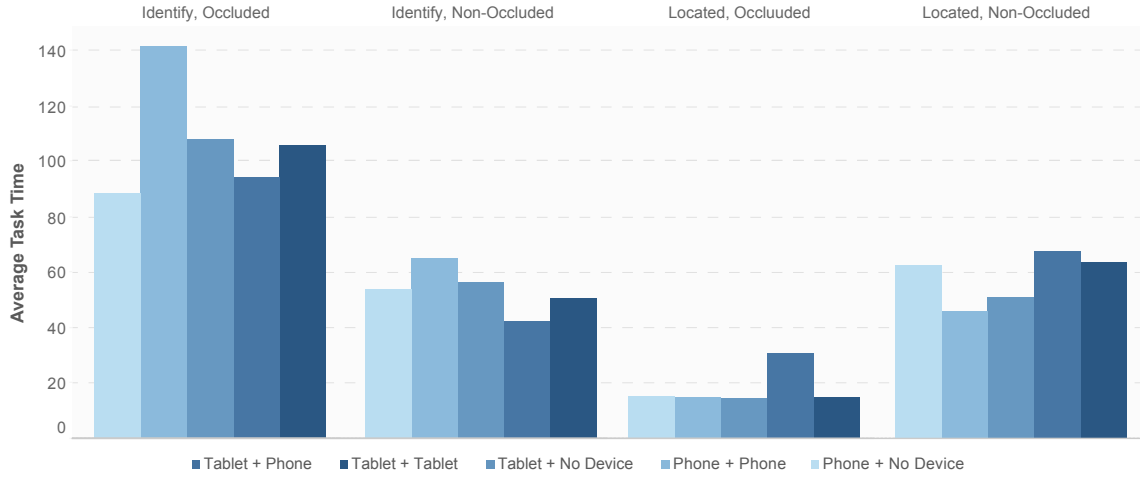
### **Frustration**

The level of frustration remains mostly the same across all device configurations. There seems to be a noticeable difference between the Phone + Phone ( $Mdn = 2.09$ ) configuration, which has slightly less Median frustration than the other configurations, and the opposite symmetric configuration Tablet + Tablet ( $Mdn = 3.50$ ) has the highest frustration. This shows that symmetrical configurations have a small difference in frustration between device sizes.

### **Task Completion Times**

We ran a two-way repeated measures ANOVA to determine the effect of different configurations and levels of model occlusion on task completion times. Analysis of the time data showed a non-normal distribution (right-skewed) when assessed by Shapiro-Wilk's test of normality. As such, we transformed the data using a log transformation and re-assessed it using the same test of normality and found it to





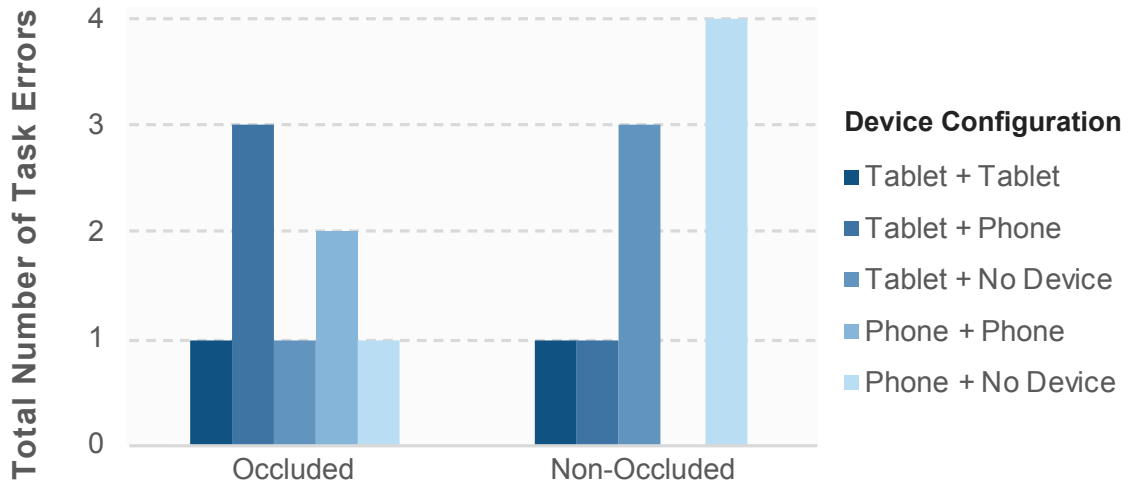
**Figure 4.7:** The mean task time for each device configuration to complete the tasks.

be normally distributed ( $p > .05$ ) with no outliers. Mauchly's test of sphericity indicated that the assumption of sphericity was met for the two-way interaction,  $X^2(9) = 15.50, p > .05$ . There was no statistically significant interaction between the model and configuration,  $F(4, 36) = .42, p > .05$ . However, the main effect of the model showed a statistically significant difference in task completion,  $F(1, 9) = 96.62, p < .001$ , showing that the observation of more occluded models having an increased average task completion time is correct.

Through further observation of Figure 4.7, we see that by far the shortest task to complete on average was the Locate task on the more occluded object. However, the occluded object also took the longest to complete Identify tasks. We see that over the Phone + No Device configuration, the Identify tasks generally took the shortest amount of time. Whilst in the dual-device Phone + Phone configuration, it took the most time. Locate tasks were, in fact, the shortest time on average in the Phone + Phone configuration but the most time-consuming in the Tablet + Phone configurations. We can infer that compared to asymmetrical configurations such as Tablet + Phone, symmetrical configurations such as Phone + Phone and Tablet + Tablet took longer on Identify tasks. We also see a trend that **more occluded** models increased the average task completion time in all cases except for the Phone + No Device configuration.

#### 4.4.2 Task Errors

A task error is defined as an incorrect answer given by a group of participants. For task error, we only report the error rate for identify tasks, since locate tasks were



**Figure 4.8:** A total of all errors across all device configurations and models occlusions on the ‘Identify’ task.

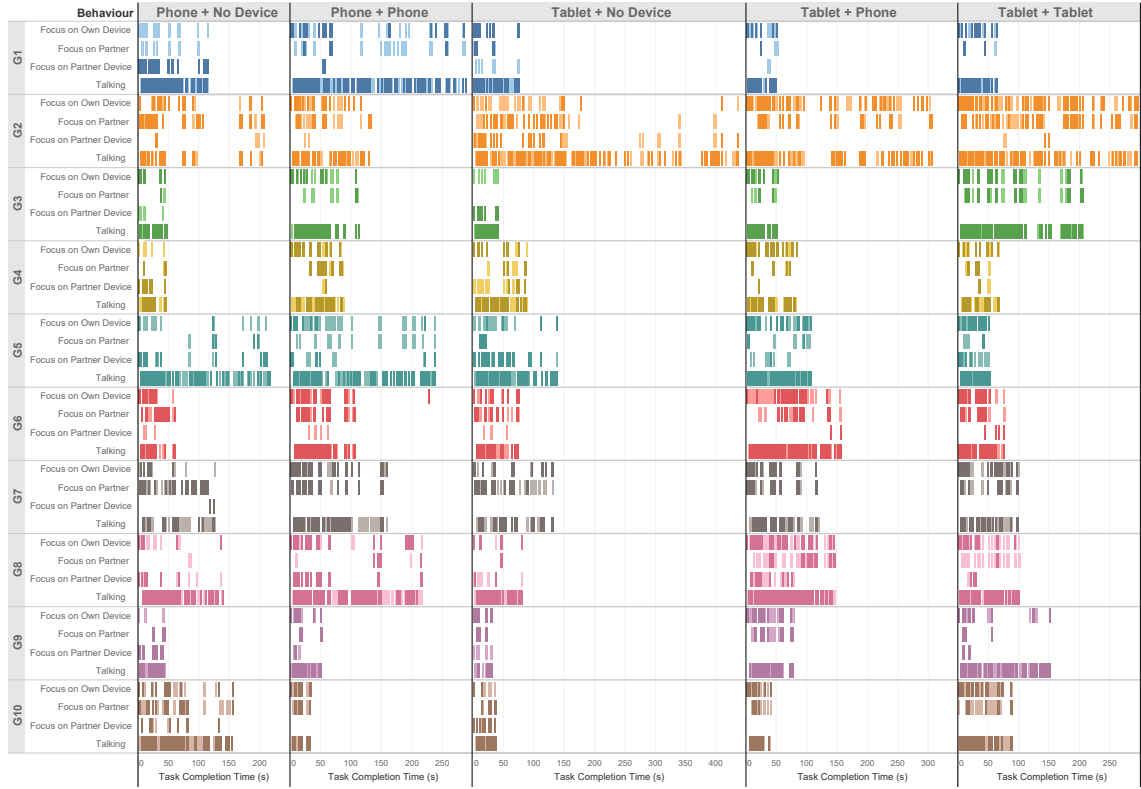
|                             | B      | S.E.  | Wald  | df | Sig.  | Exp(B) | 95% C.I.for EXP(B) |       |
|-----------------------------|--------|-------|-------|----|-------|--------|--------------------|-------|
|                             |        |       |       |    |       |        | Lower              | Upper |
| <b>Device Configuration</b> | -0.071 | 0.189 | 0.142 | 1  | 0.707 | 0.931  | 0.643              | 1.349 |
| <b>Model Occlusion</b>      | -0.142 | 0.534 | 0.071 | 1  | 0.790 | 0.868  | 0.305              | 2.469 |
| <b>Constant</b>             | -1.306 | 0.661 | 3.902 | 1  | 0.048 | 0.271  |                    |       |

**Table 4.1:** Logistic Regression Predicting Likelihood of Task Error based on Device Configuration and Model Occlusion.

not prone to error. We see that both occluded and non-occluded models have a similar number of errors, with the more occluded object having a total of 8 errors and the less occluded object having a total of 9 errors. A binomial logistic regression was performed to ascertain the effects of device configuration and model occlusion on the likelihood that participants would give an incorrect answer. The logistic regression model was not statistically significant,  $X^2(8) = 9.423, p > .05$ . Of the two predictor variables, none were statistically significant (as shown in Table 4.1). Through observation, and if we consider device configuration and task performance as a whole, it was found that while the Phone + No Device configuration had generally shorter task times, they were also the most error-prone (see Figure 4.8). We also see that symmetric configurations Tablet + Tablet and Phone + Phone were the least error-prone. Both tablets and phones were similarly prone to errors, but we see a trend in tablets being slightly faster to use.

### 4.4.3 Nature of Collaboration

Following our coding scheme (see Figure 4.9) outlined in the study detail, we observe how collaborative behaviours change between different device configurations,

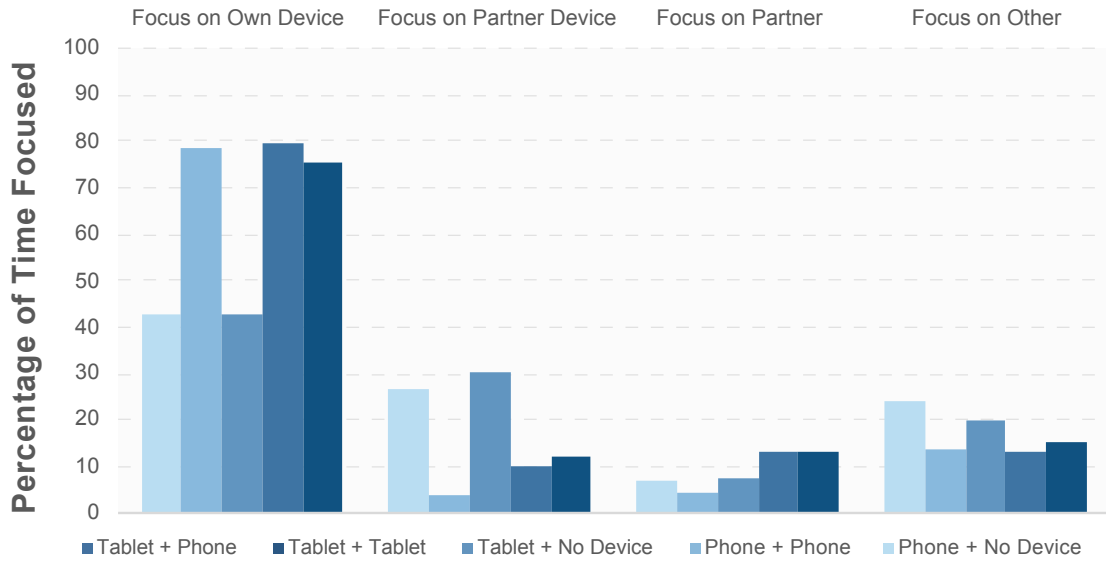


**Figure 4.9:** A timeline detailing how focus, communication and task completion time varied across groups. Different shades of the same colour represent the two different participants.

we analysed participant collaborative behaviours across all configurations, and we categorised our findings into three categories: (i) Focus, (ii) Communication, and (iii) Interaction.

#### 4.4.4 Focus

We ran four Friedman tests to determine if there were differences in Focus across the five different configurations during the tasks. Pairwise comparisons were performed with a Bonferroni correction for multiple comparisons. It was found that Focus on Own Device was statistically significantly different across different configurations,  $X^2(4) = 74.119, p < .001$ . Post hoc analysis revealed statistically significant differences between Tablet + No Device ( $Mdn = 12.275$ ) to Phone + Phone ( $Mdn = 30.667$ ) ( $p < .001$ ), Tablet + Tablet ( $Mdn = 38.576$ ) ( $p < .001$ ), and Tablet + Phone ( $Mdn = 37.264$ ) ( $p < .001$ ). Post hoc analysis also revealed statistically significant differences between Phone + No Device ( $Mdn = 10.343$ ) to Phone + Phone ( $Mdn = 30.667$ ) ( $p < .001$ ) and Tablet + Phone ( $Mdn = 37.264$ ) ( $p < .001$ ).

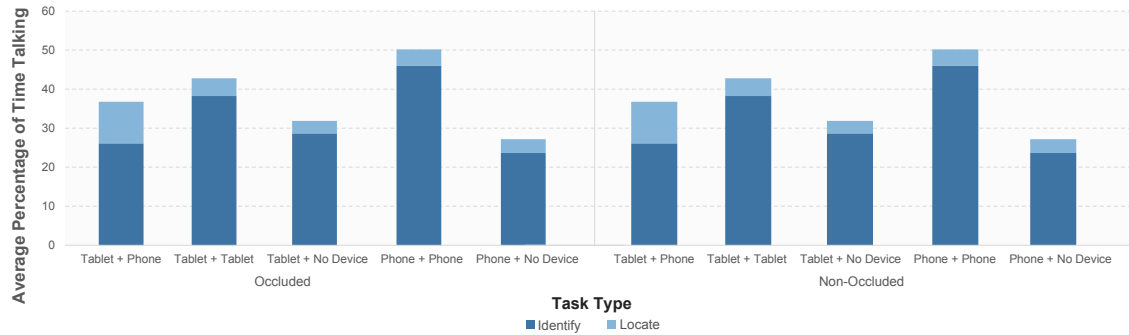


**Figure 4.10:** The average percentage of participant focus during all tasks. 'Focus on other' in this case refers to participants being idle, such as looking around the room.

It was also found that Focus on Partner Device was statistically significantly different across different configurations  $X^2(4) = 51.206, p < .001$ . Post hoc analysis revealed statistically significant differences between Tablet + No Device ( $Mdn = 0.000$ ) to Phone + Phone ( $Mdn = 0.000$ ) ( $p < .001$ ), Tablet + Tablet ( $Mdn = 0.000$ ) ( $p < .001$ ), and Tablet + Phone ( $Mdn = 0.000$ ) ( $p < .001$ ). Post hoc analysis also revealed statistically significant differences between Phone + No Device ( $Mdn = 0.000$ ) to Tablet + Tablet ( $Mdn = 0.000$ ) ( $p = 0.002$ ), Phone + Phone ( $Mdn = 0.000$ ) ( $p = 0.007$ ), and Tablet + Phone ( $Mdn = 0.000$ ) ( $p < .001$ ). No statistically significant differences were found between Focus on Partner and Focus on Other.

Across all device configurations, the participant's primary focus was on their own device without changing their gaze. It was especially prevalent in configurations with two devices, where people spent over 75% of their time looking at their own devices (Tablet + Phone, Tablet + Tablet, Phone + Phone). In these configurations, we observed participants looking out their viewports at the object during a discussion, even when they coordinated their actions. Sometimes, very briefly, they switch their focus to their partner.

In tasks in which only one device was present, 42% of the time was spent with one participant looking at their own device. Participants who were not in control of the device, however, spent less time concentrating on their partner's device when only a tablet was present and even less when only a phone was present. When



**Figure 4.11:** The average percentage of time participants spent talking in each task.

asked about this, participants mentioned that the tablet was “*easier to share*” (P12) because of the larger screen and that “*phones are more personal devices*” (P16), so in the real world “*you wouldn’t just look over someone’s shoulder to view their phone, so it was slightly more awkward*” (P19).

Although there is only a small difference in percentage (around 4%), we see pairs spent more time looking at their own device in configurations that included a phone (Tablet + Phone, Phone + Phone) than in those that included two tablets (Tablet + Tablet). With two phones (Phone + Phone), participants tended to spend most of their time looking at their own devices but very little time looking at their partners’ devices or their partners. This means they spend most of their time looking around the room, generally waiting for the task their partner is working on to be completed.

The results indicate that in both multi-device and single-phone configurations, participants’ attention is drawn more toward their own devices. This is also found in a configuration with a single tablet but to a lesser extent. Due to their focus on their own devices, there is less “face-to-face” interaction. Furthermore, we discovered that all configurations that include a phone seem to draw the user’s attention to their own device, lending to the idea that a phone has some innate privacy attached to it, thus making it feel “awkward” (P19) to look over someone’s shoulder at their device.

#### 4.4.5 Communication

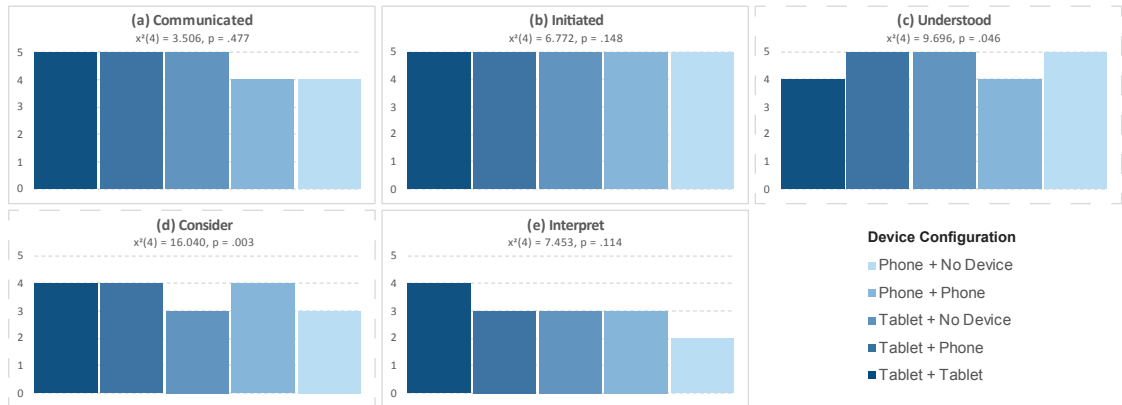
Based on our observations, we see how communication varies between configurations and tasks (Figure 4.11). Across the different configurations, we do not observe a large change in the amount of talking. Across all tasks, configurations with two devices had an average talking time of 26%, and configurations with one device had an average talking time of 24%. In the ‘Locate’ tasks, the configurations with a single device

experienced substantially less talk throughout. The participants generally operated from a shared device and adopted a *Independent Work* or a *Asynchronous Work* strategy. Participants felt they “didn’t need to discuss” (P6) the model as much as they were “seeing the same thing” (P6). In addition to communication across these configurations, we also considered the time participants spent counting out loud since sometimes this distracted other participants and would result in them losing their count and having to start over. Even with this in mind, the participants discussed very little in their groups, spending an average of 75% of the task time in silence.

We ran a three-way repeated measures ANOVA to determine the effect of different configurations, tasks, and levels of model occlusion on time spent talking during tasks. Analysis of the time data showed a non-normal distribution (right-skewed) when assessed by Shapiro-Wilk’s test of normality. As such, we transformed the data using a square root transformation and re-assessed them using the same normality test and found them to be normally distributed ( $p > .05$ ) with no outliers. There was no statistically significant difference between configuration, task, and level of occlusion,  $F(4, 76) = 1.72, p > .05$ .

However, there was a statistically significant simple two-way interaction between task and occlusion,  $F(1, 19) = 191.839, p < .001$ . Pairwise comparisons in this two-way interaction show that there was a statistically significant difference in time talking on the occluded model between the Identify task ( $Mean = 5.190$ ) and Locate task ( $Mean = 3.387$ )( $p < .001$ ). Pairwise comparisons in this two-way interaction also show that there was a statistically significant difference in time talking on the Identify task between the Occluded ( $Mean = 5.190$ ) and Non-Occluded Models ( $Mean = 3.387$ )( $p < .001$ ) and on the Locate task between Occluded ( $Mean = 1.958$ ) and Non-Occluded Models ( $Mean = 3.600$ )( $p < .001$ ).

There was also a statistically significant difference between in time spent talking between configuration and task,  $F(4, 76) = 3.788, p = .007$ . Pairwise comparisons show statistically significant differences in Phone + No Device to tasks Identify ( $Mean = 4.289$ )( $p = .002$ ) and Locate ( $Mean = 2.779$ )( $p = .002$ ), Phone + Phone to tasks Identify and Locate ( $p < .001$ ), Tablet + No Device and tasks Identify and Locate ( $p < .001$ ), and Tablet + Tablet to tasks Identify and Locate ( $p = .003$ ). Pairwise comparisons on time spent talking during Locate tasks found a statistically significant difference between Phone + No Device ( $Mean = 2.396$ ) to Tablet + Phone ( $Mean = 3.603$ )( $p = .030$ ), and Tablet + No Device ( $Mean = 2.311$ ) to Tablet + Phone ( $Mean = 3.603$ )( $p = .015$ ).

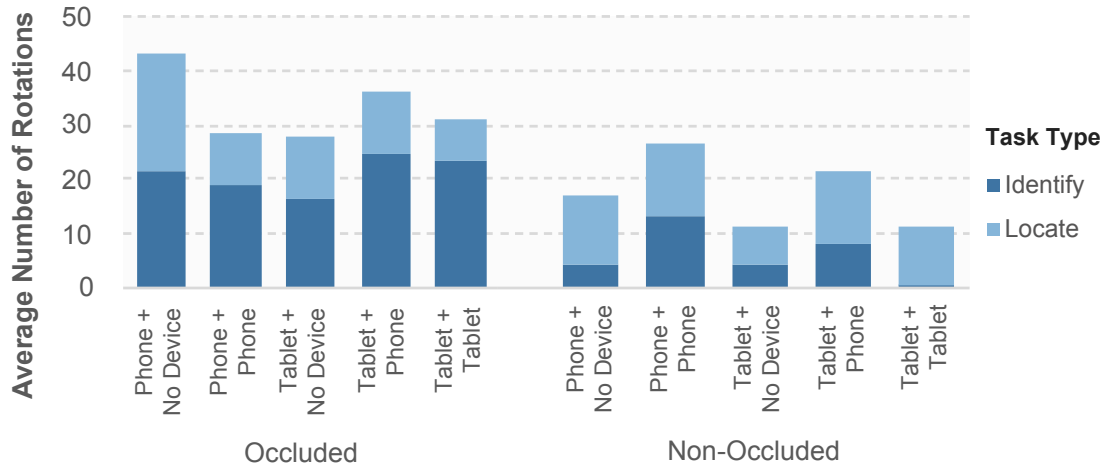


**Figure 4.12:** Median scores on a 5-point Likert scale of participants’ individual self-reflection on their communication. A dashed border is present to show significant differences in ‘Consider’ and ‘Understood’ questions.

When asked in a Likert-scale questionnaire how participants felt they communicated across all tasks, each participant generally Strongly Agreed that they communicated well with their partner and found it easy to initiate conversation (Figure 4.12a and b) on tablets, and mostly agreed that they communicated well on phone configurations (Phone + Phone and Phone + No Device). When asked if they felt their partner understood them at all times, we see that there is a trend in dual-device configurations (Tablet + Tablet and Phone + Phone) that they are slightly less agreeable.

A Friedman test was conducted to determine whether agreement in a self-reflection of communication differs between the different device configurations. The results show non-significant differences in how they felt they communicated, how easy they felt it was to initiate conversation, and how carefully they had to interpret their partner’s meanings, so therefore, retain the null hypothesis. However, we did find a significant difference when asked if the participants needed to carefully consider how to portray their thoughts (Figure 4.12d) ( $\chi^2(4) = 16.040, p = 0.003$ ) and how well they felt their partner understood them ( $\chi^2(4) = 9.696, p = 0.046$ ). The results indicate that on configurations that only contain a tablet (Tablet + Tablet and Tablet + No Device), participants had to more carefully consider what they would say to their partner to portray their thoughts. However, separately, on symmetric configurations Tablet + Tablet and Phone + Phone, participants felt less strongly that their partners understood what they were trying to communicate.

Participants used deictic gestures most often to point at their devices. This happened often when participants were working on Identify tasks, as they pointed



**Figure 4.13:** The average amount of rotations per task on occluded and non-occluded models.

at the model and counted individual objects. Participants displayed this behaviour primarily when both were focusing on a single device and when referring to a specific area on the virtual model. There were also instances when this happened naturally, even when participants did not share their views with one another.

#### 4.4.6 Interaction

The occluded model generally required more rotations to complete the task, as opposed to the less occluded model. According to our observation, on the less occluded model, using the symmetric configuration with tablets (Tablet + Tablet), there was a low amount of average rotations. We saw high average rotations on the more occluded object in the same configuration. We can see from our video observations that participants would typically scale the non-occluded model larger on most configurations and would be able to see through the object and identify the areas of interest more easily. This is in contrast to the more occluded object, where participants would be required to rotate the model to cover all areas for both types of tasks.

In general, configurations that included a phone had more interactions than configurations with a tablet. However, one interesting observation was that participants often preferred to use the tablet over the phone in the Tablet + Phone configuration when rotating more occluded models. The participants commented that



they preferred using the tablet because they could “see more” (P8) and they could “see rotations better” (P12) during the Tablet + Phone configuration. Regarding other configurations, participants commented that the phone was easier to manipulate the object with due to its size, mentioning “being able to hold it in one hand and rotate it using your thumb” (P1) and that the participants were also “used to using it” in every day life (P14).

#### 4.4.7 Strategies of Collaboration

A post hoc analysis of participant video data and an examination of collaboration metrics across all tasks were conducted to observe the effects of different device configurations on participants’ general collaboration strategies. Four distinct collaboration strategies emerged:

##### 1 – Independent Work

A task was completed by one participant within the pair *entirely independently*, with their primary **focus on their own device**, little to no **communication**, sharing of the device, or negotiation of control in multi-device configurations.

##### 2 – Asynchronous Work

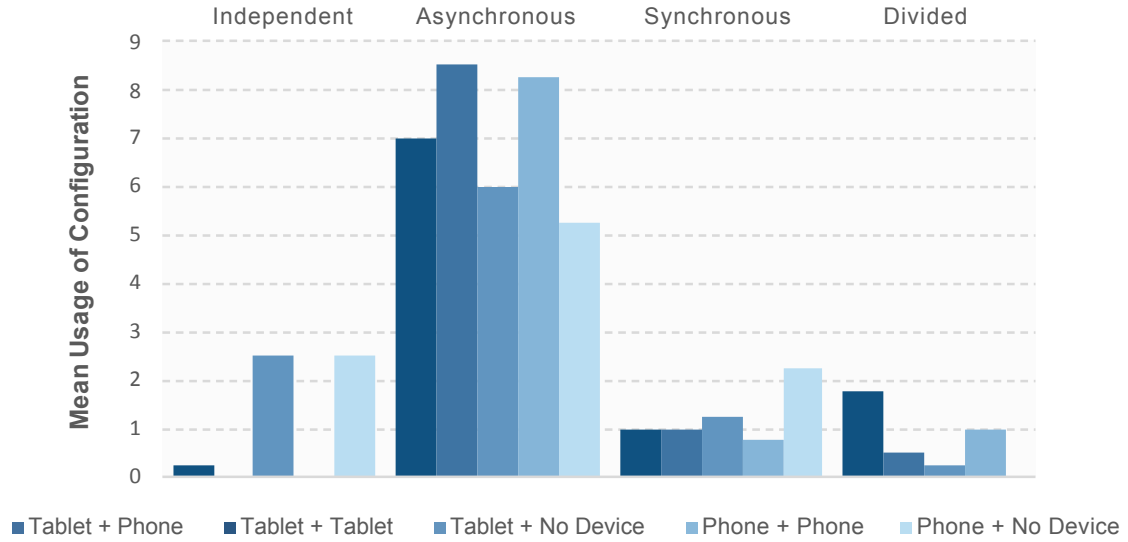
A task was completed independently by each participant *asynchronously*, with primary **focus on their own device**, by sharing device ownership in single-device configurations, or by negotiating control through turn turn-taking in multi-device configurations.

##### 3 – Synchronous Work

Participants completed the task *synchronously* either by sharing a device view in single-device configurations – demonstrated by higher **focus on partner’s device** – or **more communicating** with their partner to negotiate control of the virtual object in a multi-device configuration.

##### 4 – Divided Work

A task is explicitly divided into smaller components, often spatially, with participants taking ownership of a component and then working on this simultaneously with more **focus on own devices**. For example, spatially dividing a virtual model into separate areas for each participant to work on for either single or multi-device configurations.



**Figure 4.14:** The average amount of collaboration strategies adopted over each device configuration.

#### 4.4.8 Frequency of strategies across device configurations

These strategies were applied in varying frequencies depending on the device configuration, task type, and model type (Figure 4.14).

We see throughout our results that *Synchronous Work* was overwhelmingly the most popular strategy amongst participants across all device configurations, task, and model occlusions. We saw a reduction in the *Synchronous Work* strategy on the Tablet + Phone configuration, specifically when participants were working on *Identify* tasks on models that had *more occlusion*. We also see a reduction in this strategy when single-device configurations are in use (Tablet or Phone + No Device). While *Synchronous Work* was the most popular overall for single-device configurations, *Independent Work* and *Asynchronous Work* were adopted more than multi-device configurations and were mostly equal in popularity for both Phone + No Device and Tablet + No Device.

While *Divided Work* was the least popular, there were cases where this strategy increased in popularity. For instance, participants would more often use this approach in multi-device configurations, especially in cases where the devices were symmetrical, such as Phone + Phone and Tablet + Tablet. Furthermore, *Divided Work* was used more when working with **more occluded** models, in which case participants would always be positioned face-to-face [129].

## 4.5 Discussion

In this research, we have explored the impact of device configuration on collaboration efficacy, behaviour over time, and strategies formed. Based on five separate research categories, our findings provide insights into the nature of collaboration between different device configurations.

We first hypothesised (H1) that having asymmetric devices would yield more focus on individual devices and an increase in communication as a workaround. We found that participants tended to focus on one device during single-device configurations (one phone or one tablet) and share. However, the results show that there is still a slightly higher ‘Focus on Other’ during tasks, which indicates that in some groups, there is still a lot of time spent on individual work, with one participant doing their own thing, not in the AR space, which is in support of the hypothesis. This is further backed up by ‘Independent Work’ being higher on single-device configurations in general. We also see that single-device configurations had statistically significant differences to asymmetric multi-device configurations in communication on Locate tasks. However, communication was generally lower on single-device configurations.

Our second hypothesis (H2) was that a larger device size would afford more sharing during collaboration, giving more focus on own and partner devices. This is supported by a trend of somewhat more ‘Focus on Partner’ in tablet configurations, but no statistically significant differences were detected. While there was no statistically significant difference, we did see that configurations containing a tablet have a higher ‘Focus on Partner Device’ than symmetric phone configurations. However, there is insufficient statistical difference to conclude that device size facilitates more sharing during collaboration, therefore we reject the second hypothesis.

Our third hypothesis (H3) explains that single-device configurations would lend to increased interaction. From the results, we can observe that the single-device configuration containing a phone had more interactions than any other. In contrast, the single-device configurations containing a tablet had one of the lowest across both levels of model occlusion, meaning we reject this hypothesis. This can be related to the additional effort that is necessary to hold a tablet up while also interacting with the touch screen.

In our fourth hypothesis (H4), we wrote that more occlusion would mean more communication between participants. We observe that with more occluded models, participants would generally position themselves face-to-face. We also saw an increase in Divided Work adoption and less independent work. However, we do

not see a large change in the amount of talking when the occlusion changes across the different device configurations.

We hypothesised in our fifth and final hypothesis (H5) that collaborative strategies would need to be adapted depending on the type of task the participants were required to perform. We observe that the most common way participants would collaborate is through a *Synchronous Work* approach. Once adopted, this strategy would be carried throughout each configuration. The exception is in single-device configurations, in which we see an increase in the *Asynchronous Work* strategy. Therefore, we reject this hypothesis.

#### 4.5.1 Task Performance

Our results indicate a trend that task performance was better for symmetric configurations and single-device configurations, though single-device configurations were more prone to error. Observations indicate a reason for this could be due to the higher adoption of *Independent Work* approaches for the single-device configurations, which, in general, required less communication and negotiation of control. However, it has the disadvantage of tasks not being validated by another user. The increase in task performance in symmetric configurations could be the result of an equal perception of control from participants where *Divided Work* could be more readily applied to decrease the task space and workload of each participant, which in turn reduces the likelihood of errors within the task.

Further, the increased perceived workload and frustration seen in all phone-based configurations cannot be attributed to any noteworthy differences in collaboration strategy when compared to tablet-based configurations. However, we can speculate that the form factor and decreased screen ‘real estate’ of the phone may incur extra physical or cognitive steps for the user. Additionally, the much lower perceived workload of single-device configuration Tablet + No Device might be the result of better affordance of the tablet to the popular *Synchronous Work* collaboration strategy with the increased screen ‘real estate’ reducing the strain of viewing and perceiving the virtual model simultaneously. Interestingly, in configurations that were rated as requiring more effort and being more demanding, there was a small increase in participants’ perceiving their performance as better.

### 4.5.2 Impact of Device on Collaboration

Phones tended to facilitate slightly more focus on their own device, as opposed to tablets which afforded more screen sharing and shoulder-surfing, however these differences are relatively minor. We also see differences in interactions across device configurations. In asymmetric configurations, participants, depending on the device ownership and model's occlusion level, would adopt loose roles for completing the task and also manipulating the model, which is reflected in the average number of rotations in the Tablet + Phone configuration. For example, tablets were used more often to control more occluded models, whereas phones were used to control less occluded models.

### 4.5.3 Collaborative Strategies

Using the *Synchronous Work* strategy, negotiation of control was typically an ongoing and fluid process between participants based on verbal communication, implicit markers such as movement of the model or the other participant, and deictic gestures. This was also similar for *Divided Work*, as opposed to *Independent Work* and *Asynchronous Work* where control was often negotiated by device ownership. For both *Synchronous Work* and *Divided Work*, participants occasionally adopted supplementary tools to sketch, annotate, or spatially point on the device screen or in physical space, all in reference to the virtual model. The prevalence and popularity of *Synchronous Work* over a similar strategy such as *Divided Work* could be due to the added overhead of democratising the labour of a task. Due to the simplicity or low workload of the tasks, the ad-hoc nature of *Synchronous Work* may have been preferred as a 'quick-and-easy' approach as opposed to something more structured such as *Divided Work*. Despite *Synchronous Work* being the most popular collaborative strategy, participants would encounter issues not observed in other strategies. For example, in single-device configurations, participants sharing a view would be less strict about negotiating model control and would often cause unintended interactions.

For *Independent Work*, participants without device ownership were often comfortable letting the other participant complete the task and trusting the outcome. At the same time, they were mostly idle, with instances where they would even distract the other participant from completing the task. For *Asynchronous Work*, control of the model was mostly negotiated via device ownership, with the occasional instance of the device handler holding the device for the other participant and inviting

them to complete the task and interact. Regarding the type of task, **identify** tasks involved more *Asynchronous Work* and **locate** tasks had more *Independent Work* strategies. This could be due to the lower cognitive effort required for the locate tasks, meaning that *Independent Work* strategies are more desirable. For *Divided Work*, we can postulate that participants were more comfortable using this strategy in symmetric device configurations as there was more perceived equity in terms of the tools available to each participant despite the interactions being the same in asymmetric device configurations. We can also infer that participants were more encouraged to adopt this strategy if they were not able to perceive most of the information, i.e., when working with a more occluded model.

In summary, we observed and identified distinct collaborative strategies used by participants using handheld AR, and they adopted them at varying rates depending on device configuration. The effect of task type and model occlusion on collaboration strategies and user mobility was also observed. Generally, the results of this study do not make a strong claim that different categories influence co-located collaboration. However, the preliminary results provided insight and evidence about the nature of collaboration between devices of different configurations, including their symmetry, size, and quantity. In addition, there is evidence that model occlusion may affect collaboration, though further research is required.

## 4.6 Limitations and Future Work

Our work attempts to provide an initial investigation into the influence of different device ecologies and configurations on collaboration in co-located handheld AR. However, there are some limitations with our current work that show potential opportunities for future work to be built on. Firstly, our work primarily focused on tasks and interactions with a 3D virtual model, but AR can be used in different ways to support activities in the physical space, providing additional digital information and content. Naturally, it would be interesting to see how collaboration strategies, behaviour, and effectiveness are influenced when activities include physical and virtual interactions. The tasks and models used in the study were abstract and limited in scope but provided initial generalisable results for more application and context-specific work to build on.

Furthermore, the tasks were intentionally designed to be possible independently or collaboratively to keep the participant strategies open, sometimes resulting in less collaborative approaches from the participants. Future work could focus on tasks

that require more than one participant to complete, ensuring that in every instance, some manner of collaboration occurs. Regarding the analysis of collaboration, future work could better cross-reference strategy, behaviour, and task efficacy together to increase understanding of which device configurations and strategies work in different tasks and application contexts. Finally, the participant interaction could be analysed along with the mobility, spatial formation, and movement of the participant in similar handheld AR tasks to understand how collaboration strategies relate to on-screen interaction and group movement during a task.

## 4.7 Conclusion

This chapter shows that Augmented Reality has the potential to become an instrumental tool for collaboration. While many new technologies and devices are being introduced, currently, the dominant way to use and collaborate in AR is through mobile devices. Understanding the precise role that device size and configurations play in mobile handheld AR systems for collaboration is essential for developing new design approaches that enable and facilitate collaboration. In our study, we examine how mobile AR affects (i) the efficacy of collaborative work, (ii) changing collaboration behaviour and (iii) collaboration strategies. Our findings show a nuanced balance between positive and negative effects related to how participants were affected in their focus, mental/task load, and communication – but also how collaborative settings were adjusted, appropriated and adapted to adjust for such challenges.

In the next chapter, we consolidate the findings and insights gathered from the previous studies in [Chapter 3](#) and [Chapter 4](#), including the design recommendations, spatial configurations, and identified collaboration styles that emerged from our empirical investigations, and building on these foundational elements, we introduce an initial framework for a collaborative augmented reality interaction space. This framework aims to integrate the various elements discovered in our research, providing a cohesive structure that enhances co-located collaborative experiences through augmented reality. By leveraging the configurations and collaboration styles identified in the earlier chapters, this framework seeks to address the challenges and maximise the potential of handheld mobile AR systems in real-world applications.

# 5

## A Conceptual Lens on Handheld AR Collaboration

### Contents

---

|     |                                                                                     |     |
|-----|-------------------------------------------------------------------------------------|-----|
| 5.1 | Introduction . . . . .                                                              | 105 |
| 5.2 | A Conceptual Lens for Co-Located Collaborative Handheld Augmented Reality . . . . . | 107 |
| 5.3 | Research Through Design: Prototyping AR Interaction Techniques . . . . .            | 111 |
| 5.4 | Reality & Beyond: Designing Collaborative AR Interaction Techniques . . . . .       | 111 |
| 5.5 | Example Applications Scenarios . . . . .                                            | 119 |
| 5.6 | Discussion . . . . .                                                                | 123 |
| 5.7 | Conclusion . . . . .                                                                | 127 |

---

In the previous chapters, [Chapter 3](#) and [Chapter 4](#), this thesis examined the role of handheld mobile AR in co-located collaborative settings, each offering insights into how different aspects of AR systems and device configurations can enhance or hinder collaborative engagement. [Chapter 3](#) addressed Research Question 1 (RQ1) by investigating the “CollabAR” system, an exemplar prototype designed to support basic AR interactions on mobile devices in collaborative contexts. Through this system, the chapter explored how co-located users interact with virtual models of varying complexity, identifying key dynamics in group interactions and proposing initial design recommendations to enhance collaborative experiences. This foundational



work highlighted the potential of handheld AR to support meaningful collaboration by making virtual content accessible and engaging within shared physical spaces.

Chapter 4 then expanded on this foundation by focusing on Research Question 2 (RQ2), examining the influence of mobile device configurations on collaboration. Through controlled experimentation, the chapter explored how factors such as device size and orientation impact collaborative strategies, behaviours, and effectiveness in AR environments. The findings demonstrated that device characteristics play a crucial role in shaping how users engage with AR content and each other, underscoring the importance of designing AR systems with flexible configurations that cater to diverse collaborative needs. Together, Chapter 3 and Chapter 4 provide a comprehensive overview of how basic AR interactions and device configurations can enhance or limit collaborative efficacy in co-located AR settings.

Building on these foundational insights, we now shift the focus to the theoretical underpinnings that can further guide the design and understanding of co-located collaborative AR and begin to put these into practice. Researchers have shown that Augmented Reality (AR) effectively facilitates co-located collaboration [46], particularly with handheld devices due to their accessibility and versatility. Collaborative Handheld AR systems are becoming increasingly common in various domains such as education, social gaming, and design [8, 38, 172], supporting innovative forms of collaborative engagement and offering new ways for users to interact with virtual content in shared environments. Despite the practical advancements in Collaborative Handheld AR, there remains a gap in our understanding of how AR systems' design choices influence social dynamics and collaborative behaviours in co-located settings.

We propose *proxemics* as a conceptual lens to understand better and design co-located collaborative AR systems. Through this lens, we aim to support interaction designers in developing a socio-spatial literacy that informs their understanding of how AR systems shape user interactions in shared spaces [111]. Specifically, we expand upon Collaborative Coupling Styles found in previous chapters to explore how coupling styles and device configurations intersect in many-to-many relationships, providing more nuanced and adaptable collaboration options within AR environments.

To operationalise this approach, we conduct a design exploration of collaborative AR interaction techniques guided by proxemic concepts. Using a functional handheld multi-user AR prototype, we investigate the design space and identify ten interaction techniques that span a continuum from being “aligned with reality” to going “beyond reality” (illustrated in Figure 5.1).

The work in this chapter contributes the following:

- A conceptual lens grounded in proxemics, establishing a common vocabulary for articulating and comparing collaborative AR systems based on socio-spatial properties.
- An expansion of the design space for interaction techniques in co-located collaborative AR, promoting greater flexibility and adaptability.
- A discussion aimed at fostering a nuanced socio-spatial literacy, guiding designers in evaluating trade-offs when creating handheld AR interaction techniques for co-located settings.

## 5.1 Introduction

Augmented Reality (AR) has demonstrated great potential to support co-located collaboration [46]. Due to their widespread availability, handheld devices have become the dominant platform for AR applications and use cases. They have been explored in domains such as education, social gaming, design, and architecture [8]. Collaborative Handheld AR (CHAR) systems allow for novel forms of co-located collaboration [38, 92, 222] and can be integrated into many different environments, such as the classroom [102, 172, 206, 208] to support new collaborative learning activities, as well as support new forms of collaborative design and architecture [71, 101, 189], and other domains such as sports [39].

Early research on Augmented Reality (AR) has focused on how technology supports this collaboration [15, 102, 164] and more recently, there has been an increase in research on collaborative AR systems [8, 46]. Although some of this work focuses on technical, usability, and human factor issues, less attention has been paid to the social aspects of collaborative AR [46]. In research on co-located AR, the concept of collaborative AR frequently emerges. However, its interpretation varies widely between different systems. For example, in HyperCubes [52], students collaborate around a singular AR-capable device, whereas in the Meta-AR-app [208], each student has their own device at their own table. Additionally, in Blocks [71], co-located users each have a device and collaborate to construct shared AR structures. Furthermore, there is a lack of clarity on how the system's design may affect co-located collaboration. As such, there is a need for shared terminology and richer notation styles for co-located AR systems [38, 46, 201].

Collaborative AR systems are often classified using traditional Computer-Supported Cooperative Work (CSCW) classifications, where they consider different dimensions

such as time, space, symmetry, scenario, and also input/output devices [46, 188]. While existing CSCW frameworks address certain terminology and notation challenges in collaborative AR [38, 46, 201], they do not fully capture the specific challenges of co-located collaborative AR [46, 188]. Although previous work has identified different categories of co-located collaboration in this thesis (Chapter 3) and beyond ([200]), current AR or spatial collaboration schemes are limited to a fixed one-to-one relationship between interpersonal distances and collaboration styles [200]. Therefore, there is a need to support a better and wider understanding of co-located AR collaboration and the designs of these systems and interactions.

Researchers commonly articulate Proxemics theory to explain how people use space to enact their intentions toward engaging with co-located others [75]. This research has also demonstrated that proxemic interaction [7, 63] serves as a useful analytical lens for designing interactive surfaces, spaces, and cross-device systems. However, there is yet to be an application of these theories and interactions to co-located collaborative handheld AR. By leveraging proxemics, designers can explore how to develop handheld AR systems that facilitate co-located collaborative interactions and examine how different design choices influence the socio-spatial dynamics of collaborating users.

This chapter presents proxemics as a conceptual lens to help researchers understand how handheld AR systems enable and constrain different socio-spatial configurations [75]. By using this lens, interaction designers can become more aware of the co-located collaboration dynamics they aim to support, ultimately strengthening their socio-spatial literacy [111]. In particular, we synthesise a series of proxemic concepts to expand on the notion of Collaborative Coupling Styles [156, 200] for co-located collaborative AR systems. Additionally, we extend the work discussed in Chapter 3 on collaboration styles for groups by demonstrating that coupling styles and device configurations share a many-to-many relationship, thereby increasing flexibility in collaborative AR.

We specifically apply the lens of contextualised proxemic concepts to systematically explore the design of collaborative AR interaction techniques. By designing and implementing a functional handheld multi-user AR prototype, we facilitate this exploration, allowing users to engage with and experience the design space [114]. Through this endeavour, we introduce *10 interaction techniques* that range from being *aligned with physical reality* to extending *beyond reality*. The findings in this chapter lead us to propose a continuum for designing AR systems, spanning those

that align with reality, those that go beyond it, and those that exist in between. To further clarify how these techniques apply in practice, we present two scenario-based examples using an Evaluation by Demonstration approach (Type 1) [114], illustrating how different interaction techniques integrate within key application domains such as design and education.

## 5.2 A Conceptual Lens for Co-Located Collaborative Handheld Augmented Reality

Throughout this chapter, we aim to develop a socio-spatial literacy that enhances our understanding of how AR system design influences co-located collaboration. To achieve this, we synthesise concepts from proxemics theory to create a vocabulary that articulates the socio-spatial dynamics of group collaboration among co-located users. This vocabulary enables us to systematically examine design rationales and explore alternatives for collaborative handheld AR systems based on their impact on the socio-spatial nature of co-located collaboration. In the following section, we define the individual dimensions of this framework and analyse how they interplay with each other. Additionally, we demonstrate how this lens can be used by applying it analytically and articulating how it relates to previously introduced systems.

### 5.2.1 The Interplay Between F-Formations, Collaboration Styles, and Physical Features

Since virtual content in AR applications is situated in the physical space, there is an argument to be made that AR interfaces should be understood by researchers as working in *interplay* with the spatial configuration of people, the devices, and the physical space – and the features within it – to shape an individual’s opportunities for collaborating with other co-located people.

In this section, we show how we can operationalise proxemics to understand collaborative handheld Augmented Reality (AR) by analysing its connection to existing mobile AR systems. By drawing on examples from previous research, we identify the different ways collaborative mobile AR can take different forms, helping to develop a vocabulary that strengthens socio-spatial literacy when considering design trade-offs in mobile AR systems. To illustrate how various systems influence group dynamics, we use the F-Formation system [104] to visually map spatial patterns. Alongside these illustrations, we explore three dimensions of Interaction

Proxemics [66, 131] – perceptual, deixis, and control proxemics – to explain how a group’s physical positioning affects their collaboration style and enable smooth transitions between different ways of working together.

More specifically, we use these three proxemic dimensions to articulate the following concerns of co-located communication needs in AR systems:

- **Configuring Perceptual Proxemics:** The support for achieving shared or private perspectives on AR objects and spaces.
- **Configuring Deixis Proxemics:** The support for expressively communicating through deictic referencing in the shared AR space.
- **Configuring Control Proxemics:** The support for configuring the access to control and manipulation of AR objects.

To contextualise these three proxemic dimensions for collaborative handheld AR systems, we can use prior handheld AR systems as examples of how the design of such systems can influence the social and spatial configurations of co-located collaborators. For instance, in research by Sarkar et al., ScholAR [172], two students work together on AR maths exercises using a single handheld mobile device. This type of collaboration requires users to be in close proximity to each other to share AR content perception, fostering a *closely coupled* collaboration. However, control of the application is limited to the user holding the AR device.

Similarly, in research by Fuste and Schmandt, HyperCubes [52], students collaborate around a single AR device but with the addition of multiple tangible cubes. As with ScholAR, the students have a closely coupled collaboration and must be close to each other to share AR content perception. However, the use of tangible cubes allows all users to interact with the system without having to hold the device. As the virtual content is anchored to the physical cubes, the users can directly share deictic gestures towards the physical objects.

In Blocks [71] by Guo et al., co-located users build AR structures, with the AR content anchored to the physical space around them. The users each have their own device with a networked connection synchronising their virtual content. When oriented towards virtual content, they both perceive and control the virtual content.

In Meta-AR-App [208] by Villanueva et al., co-located students sit at their own tables, each with their own device and tangible setup, collaborating by uploading, sharing, and downloading AR content. In this case, each student has control of their

own setup, and the sharing of perception and deictic gestures happens through the system itself. Due to this, the collaboration styles dynamically shift between loosely coupled, where each user works on their own tasks, and more tightly coupled, where participants help each other by sharing perception and deictic gestures.

Articulating these examples from the perspective of perception, deixis, and control proxemics illustrates how F-formations and collaboration styles can be related in different ways and how the physical features of space (such as tangible objects and permanent structures) also serve to configure possibilities for orientation and collaboration.

### 5.2.2 Cross-Device AR for Aligning With and Going Beyond Reality

Beyond understanding the proxemic relations of co-located collaborative AR systems, we also seek to develop a design lens for considering how cross-device interaction techniques [26] for multi-user handheld AR can enable new forms of proxemic relations for co-located handheld collaborative AR. For instance, when comparing the spatial cross-device configurations enabled by Meta-AR-App [208], ScholAR [172], and Blocks [71], we observe that ScholAR and Blocks enable a natural form of F-formation around shared virtual content as if the users organised around a real object. Whereas the co-located networking of devices for cross-device communication in Meta-AR-App shows how AR systems can also expand on the possibilities of using space for collaboration. Together, these systems exemplify each end of the *Reality & Beyond* that we propose in this chapter, emphasising how AR can either reinforce familiar spatial interactions that are aligned with reality or redefine them entirely by going beyond it (see Figure 5.1).

As demonstrated in the concept of Proxemic Interaction [7, 63, 126], designers can create spatial interactions that align with how people navigate a space. For example, as individuals move closer to a physical display, the UI can trigger implicit responses based on their proximity. Building on this, Marquardt et al. expand on how users engage in F-formations and introduce micro-mobility gestures as a means of sharing content across devices [126]. Some studies on co-located collaborative AR show how these spatial dynamics naturally emerge as users position themselves in F-formations during co-located collaboration around handheld mobile AR devices. Examples of this include Project IRL [38] and Blocks [71], along with observations discussed in Chapter 3. From a design perspective, research on collaborative immersive

analytics has explored how to integrate proxemic principles into physical reality when designing collaborative interactions with 3D visualisations [55, 91]. These studies highlight how spatial relationships between users and devices can shape engagement, demonstrating the potential for proxemic-aware design to enhance co-located collaborative AR experiences.

Even though Augmented Reality is anchored in the physical environment, the technology can still use the available cross-device possibilities to go beyond physical reality and overcome the constraints of the physical socio-spatial relations through wireless co-located sharing of views and controls. Grønbæk et al. have explored how this perspective can support interpersonal flexibility in cross-device systems [67].

Even though Augmented Reality (AR) is inherently anchored in the physical environment, it still has the ability to extend beyond the limitations that physical reality brings by leveraging cross-device interactions. Through the wireless, co-located sharing of views and controls, AR systems can transcend traditional socio-spatial constraints, enabling new forms of collaboration that are not restricted by physical proximity or direct line-of-sight. This approach allows users to interact fluidly across multiple devices, maintaining operational consistency and fostering greater adaptability in how information is accessed, shared, and manipulated within a shared AR experience. Grønbæk et al. has explored how these cross-device interactions enhance interpersonal flexibility, demonstrating how AR can dynamically support co-located collaboration without being bound strictly to physical spatial arrangements [67].

For the remainder of this chapter, we will focus on how a proxemic lens may generate new ideas for designing interaction techniques with handheld mobile AR to support co-located collaboration. In particular, we will use proxemics as a lens to understand how AR interactions can align with, and also go beyond, the physical reality. We methodically examine two complementary aspects. Initially, we look into how AR interactions might seamlessly broaden socio-spatial relations as they are experienced in terms of perception, deixis, and control within the physical realm among users. Then, we further analyse how AR interactions can go beyond the limitations of physical socio-spatial connections by enabling cross-device sharing of perception, deixis, and control.



### 5.3 Research Through Design: Prototyping AR Interaction Techniques

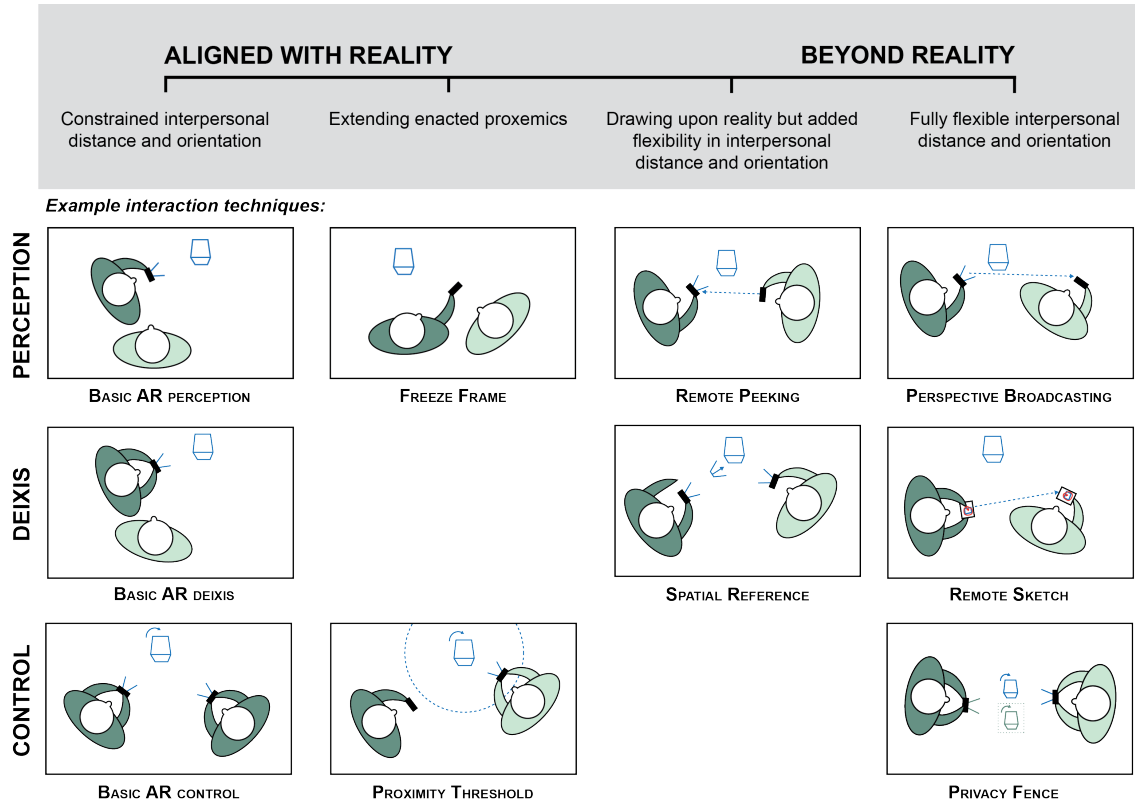
To guide our exploration of a broader design space for co-located AR, we draw inspiration from proxemics and the wider Computer-Supported Cooperative Work (CSCW) literature. Specifically, we operationalise key theoretical concepts from prior research, including *collaboration styles* [156, 200] and the three dimensions of interaction proxemics – *perceptual*, *control*, and *deixis proxemics* [66, 131]. These frameworks provide a structured foundation for the understanding and designing of co-located AR interactions. We apply these notions together as a conceptual lens to co-located collaborative handheld AR – a type of system to which they have not previously been applied. This leads to the construction of a design space, including exemplar interaction techniques, that are articulated in terms of their proxemic consequences.

We introduce terminology for articulating design intentions around collaboration within collaborative handheld AR systems, offering a foundation for future explorations of this emerging design space. This chapter presents the initial work in developing a theoretically grounded design framework, focusing on early-stage design exploration rather than direct user studies. As a result, we evaluate this work through demonstrations rather than usage-based assessments [114]. In terms of demonstration, we provide both an exploration of a design space as well as two scenarios of envisioned application usage [114]. This research does not validate that the proposed interaction techniques and enriched design space will be used in line with the proposed scenarios. Further, to limit the scope of the design space explorations, we focus on AR systems that are based solely on mobile, handheld technologies.

### 5.4 Reality & Beyond: Designing Collaborative AR Interaction Techniques

To explore proxemics as a lens for generating new ideas that support co-located collaboration in mobile AR systems, we systematically examine how to design interaction techniques that either align with reality – for example, by following the natural patterns of physical, interpersonal space – or extend beyond reality by intentionally breaking those same spatial conventions. By doing this, we focus on three dimensions of collaborative actions regarding proxemics: perception, deixis, and control. As a result of the design space exploration, we map out and position the





**Figure 5.1:** The design space of collaborative AR interaction techniques. The 10 interaction techniques explored range from aligning with reality to beyond reality.

interaction techniques along a continuum that ranges from techniques that align with reality to those that extend beyond the constraints of physical socio-spatial relations. This continuum, and also the design space exploration, can be seen in [Figure 5.1](#).

On the left side of this continuum ([Figure 5.1](#)), we have techniques that are aligned directly with reality. This means that these users' interactions with the AR content are constrained by interpersonal distance and orientation by the devices, the physical layout of the space, and also the position of the AR content. These techniques, grounded in reality, are often found in out-of-the-box multi-user AR systems, where the virtual content acts – as much as possible – as real objects are completely shared between the users.

On the right side of this continuum, we have techniques that explore how the users can collaborate with each other whilst still remaining completely flexible in their interpersonal distance and orientation. For example, the techniques that go beyond reality and make use of the networking between the AR devices.

In this chapter, we have techniques that mix these two different approaches between the two different ends of the continuum. Close to the alignment of reality,

we can design interaction techniques that are inspired by proxemic interactions [7, 63] that can extend the enacted proxemics. Learning further toward going beyond reality, we can design interaction techniques that add flexibility in interpersonal distance and orientation but still draw upon physical reality. In the following section, we present each interaction technique explored in this lens to support the three dimensions of collaborative actions: perception, deixis, and control. We also further illustrate how these align with, or go beyond, reality.

To further clarify the continuum, “aligned with reality” and “beyond reality” can be understood through their relationship with spatiality, the physical environment, and proxemic interactions within collaborative Augmented Reality (AR) systems.

1. **Aligned with Reality:** This side of the continuum focuses on designing AR interactions that **mirror or reinforce real-world socio-spatial conventions**. Interactions are typically constrained by physical factors such as interpersonal distance, user and device orientation, and the physical layout of the environment where the AR content is anchored. For example, to achieve a shared perception of a virtual object, users might need to physically move closer or stand side-by-side, much as they would when examining a tangible object. Basic AR control is often tied to immediate physical proximity and line of sight, respecting established social norms and the tangible environment. This approach aims for a natural, intuitive feel by leveraging existing human spatial understanding and behaviours.
2. **Beyond Reality:** In contrast, this end of the continuum explores how AR can **transcend the physical limitations and socio-spatial constraints** by leveraging the networking capabilities inherent in modern mobile handheld AR devices. This enables forms of collaboration that are not strictly bound by the physical proximity or line of sight. While it modifies perception and deixis through techniques like Remote Peeking and Remote Sketch – allowing users to share views or make deictic references without physical relocation – “beyond reality” also provides for the modification of control proxemics. An example of this is the Privacy Fence, which enables users to define a private virtual territory anywhere, irrespective of their physical standing or the existing physical boundaries. This flexibility extends to other aspects, allowing dynamic content access and manipulation that isn’t tied to a single physical viewpoint or a fixed F-formation, thereby supporting more fluid and adaptable collaborative

styles. These techniques allow AR systems to expand on the possibilities of using space for collaboration, moving beyond merely augmenting reality to actively reshaping the interactive experience.

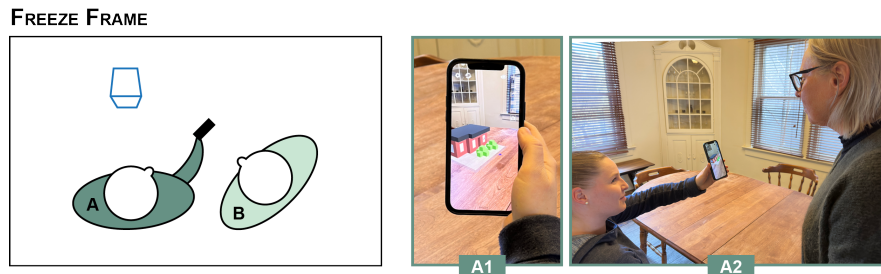
### 5.4.1 Configuring Perception Proxemics

#### Basic AR Perception

In Augmented Reality applications, developers tend to anchor virtual content to the physical world, which in turn defines how users perceive the content based on their distance and orientation. To share the same perception of virtual content in most basic AR applications, users must position themselves at the same distance and angle, either by standing closely together and looking through a single device or by each looking through their own device in close proximity. However, there may be some limiting factors that make this difficult. For example, the size of the mobile device and the field of view can restrict the space used, requiring users to arrange themselves in a specific F-Formation to share their perception.

#### Freeze Frame

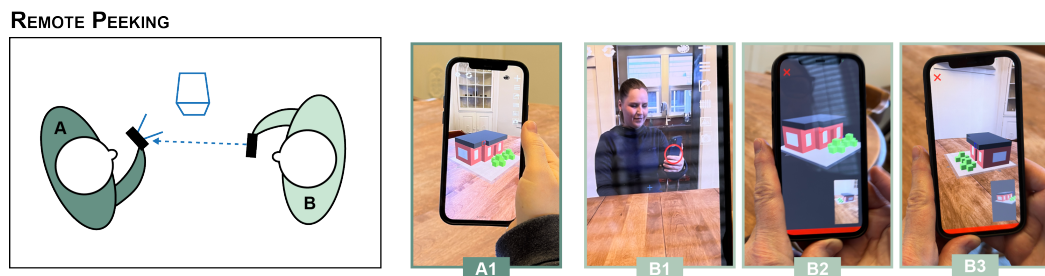
By utilising the possibilities of mobile devices, we can overcome the constrained interpersonal distance and orientation when sharing perception. In Freeze Frame, known from remote collaboration [31], users can take a snapshot of virtual content from the perspective they wish to share (see Figure 5.2). This allows them to attach the content directly to the device instead of having to perceive the content through the device. Using this technique in co-located collaboration enables the user to make micro-mobility gestures with the device, enabling more flexibility in their interpersonal distance and orientation.



**Figure 5.2:** In Freeze Frame Alice freezes a perspective on the AR content (A1). Alice can now use micro-mobility to share the perception of the AR object with Beth with flexibility in interpersonal distance and orientation (A2).

## Remote Peeking

While Freeze Frame allows for more flexibility in user interpersonal distance and orientation, users must still be relatively close to each other to perceive the attached content on the mobile device. In Remote Peeking (Figure 5.3), the networking possibilities of the devices allow users to peek at other users' perspectives. For example, if Alice wants to peek at Beth's perspective (see Figure 5.3 (A1)), Beth starts by selecting the location of Alice's device (Figure 5.3 (B1)). This enables the peeking view, allowing Beth to perceive the content from Alice's perspective (Figure 5.3 (B2)). Beth can switch between the peeking view and her own view (Figure 5.3 (B3)). Beth does not need to physically move to see Alice's perspective of the virtual content. This technique allows users to be flexible in their interpersonal distance and orientation. However, the use of this interaction technique still draws upon reality to understand which perspective they are seeing. This means that the user who wishes to peek at another user's device still needs to be able to view the device of the user they want to peek at.



**Figure 5.3:** In Remote Peeking, Beth can peek at Alice's perspective (A1) by selecting the position of Alice's device (B1). Beth can now swap between seeing Alice's perspective (B2) and her own perspective (B3) in focus.

## Perspective Broadcasting

The Perspective Broadcasting technique is similar to Remote Peeking, but in this technique, the initiative comes from the user sharing perspective instead of the peeking user. Users can broadcast their current perspective to all other users no matter where they are physically located in the space. This allows users to be fully flexible in their interpersonal distance and orientation while still being able to share their perception of the virtual content.

### 5.4.2 Configuring Deixis Proxemics

#### Basic AR Deixis

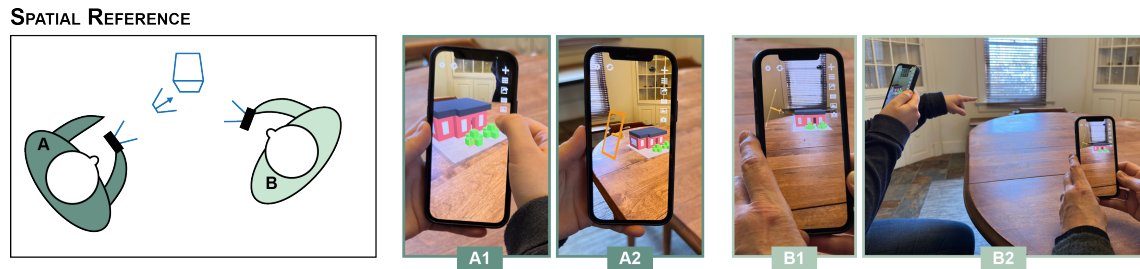
Even though AR content imitates physical objects in the way they are anchored in physical space, Poretski et al. [156] found that users use fewer deictic references when collaborating around virtual content in AR in comparison to physical objects. In basic AR, using deictic references to the AR content often means pointing to the content in the physical space as if it were real. These gestures are usually done either directly towards the objects in the physical space whilst maintaining a view of the screen or pointing directly to the screen. In handheld mobile AR, the devices have a limited field of view, making it challenging to see both the virtual content and the deictic references. If the deictic gesture was made in the physical space, it also has a non-specific location to most users other than the person pointing when they are not tightly coupled, as the users do not share the same perspective. As such, sharing deictic references is constrained by the users' interpersonal position, distance, and orientation.

#### Spatial Reference

Using the Spatial Reference feature, a user can anchor a spatial reference into the space to make deictic references later, as seen in [Figure 5.4](#). In this exemplar figure, Alice creates a spatial reference on a perspective she wishes to save for later ([Figure 5.4 \(A1\)](#)). She sees the reference as an AR object pointing ([Figure 5.4 \(A2\)](#)). Beth also sees the spatial reference on her device ([Figure 5.4 \(B1\)](#)). When Alice makes a deictic gesture roughly in the direction of this spatial reference, Beth easily follows her deictic reference ([Figure 5.4 \(B2\)](#)). This technique draws upon reality by augmenting a position in space, but it enables users to share more precise deictic references in more flexible f-formations.

#### Remote Sketch

Whilst Spatial References can create more interpersonal flexibility, users still need to be oriented towards the virtual object and reference point to be able to understand the deictic references made by other users. However, with Remote Sketch ([Figure 5.5](#)), the users can be even more flexible with their interpersonal positions and orientation while sharing deictic references. Remote Sketch builds on the Freeze Frame interaction technique, allowing users to broadcast the frozen frame to other users and annotate



**Figure 5.4:** Alice can create a Spatial Reference at a desired perspective (A1). An AR object appears pointing from the previously selected perspective (A2). Beth also sees the spatial reference (B2), and when Alice points towards the AR object, Beth knows what she is referencing with help from the spatial reference (B2).

the frame for everyone, as seen in Figure 5.5. In this figure, Alice freezes a frame and broadcasts it to the other users no matter where they are (Figure 5.5 (A1)). She sketches on the frame to reference some specific detail (Figure 5.5 (A2)). Beth receives the frame and sees Alice's live sketch on top (Figure 5.5 (B1)). By freezing the frame, the user that makes the deictic gesture is suddenly released from the need to hold their position, making them flexible in moving to a more comfortable position before making the references. Other users are able to see the deictic references from anywhere, enabling flexible F-formations.



**Figure 5.5:** In Remote Sketch, Alice can freeze a frame and share it with everybody (A1). She sketches the frame in order to explain a detail (A2). Beth sees the frame and follows Alice's sketch in real-time.

### 5.4.3 Configuring Control Proxemics

#### Basic AR Control

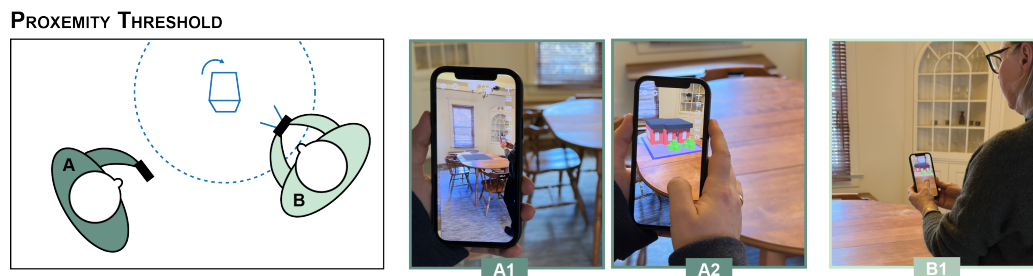
As the input device that is used to view and manipulate virtual objects in handheld mobile AR is in the hands of the users, in basic AR applications, users can realistically control objects whenever they are within the field of view of their device. However,



in practice, the control that is demonstrated is contained in users' personal space, fixed and semi-fixed features of the physical environment, and social norms. For example, you would rarely manipulate an object that is right in front of another user, an object that is behind a table or at another table other than the one you are sitting at, or an object in the process of being manipulated by others.

### Proximity Threshold

Although basic control of AR follows physical and social rules, there is also the opportunity to enhance these within the system. With the Proximity Threshold, users can only control objects that are within close proximity (Figure 5.6). In this figure, we see that Beth is manipulating the AR object in front of her (Figure 5.6 (B1)). Alice sees that the object is there, but she cannot see any details of the object (Figure 5.6 (A1)). When approaching the object and when in close enough proximity, Alice now sees the object and can start manipulating it along with Beth (Figure 5.6 (A2)). Since Alice has to be close to the object to see and manipulate it, she also moves closer to Beth, who is already interacting with the object, entering a closer coupled collaboration. Beth can easily see if other users have control of the object along with her. In this way, the AR system extends the enacted proxemics.

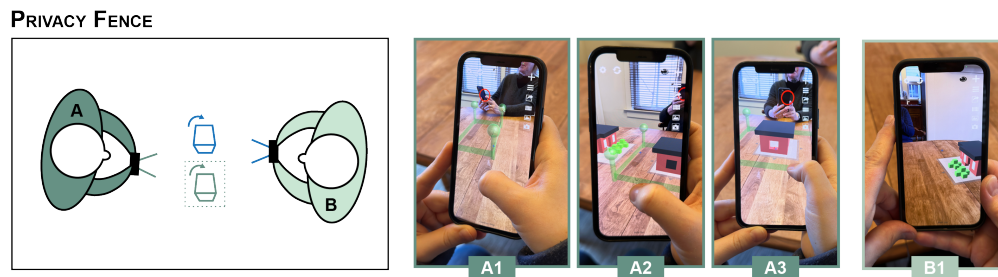


**Figure 5.6:** In Proximity Threshold, users are able to control AR objects when being close by. Beth is manipulating an object in front of her (B1). Alice only sees that some object is there from a distance, but she does not see details and has no access to manipulate the object (A1). When she moves closer, the object appears, and she can now manipulate the object (A2).

### Privacy Fence

By using Proximity Threshold, users have more privacy around virtual objects within their personal space, but the control of the virtual environment still has socio-spatial meaning. With Privacy Fence, users can create their own private territory anywhere.

The user defines a spatial fence, where everything they put inside the fence will be their private objects completely hidden from other users, giving them full control as seen in Figure 5.7. If Alice wants to work on something private along with the public content, she creates a Privacy Fence next to the public object (Figure 5.7 (A1)), adds the content she wants to work on inside (Figure 5.7 (A2)), and starts manipulating her private content (Figure 5.7 (A3)). Beth still only sees the shared content (Figure 5.7 (B1)). This interaction technique allows users to control directly what they want to share with other users in the group and also what they want to keep completely private by moving virtual objects inside and outside of the Privacy Fence.



**Figure 5.7:** In Privacy Fence, users can create their own private territory anywhere. Alice creates a Privacy Fence next to a shared object (A1-A2). Alice works on her private object (A3), while Beth still only sees the shared object (B1).

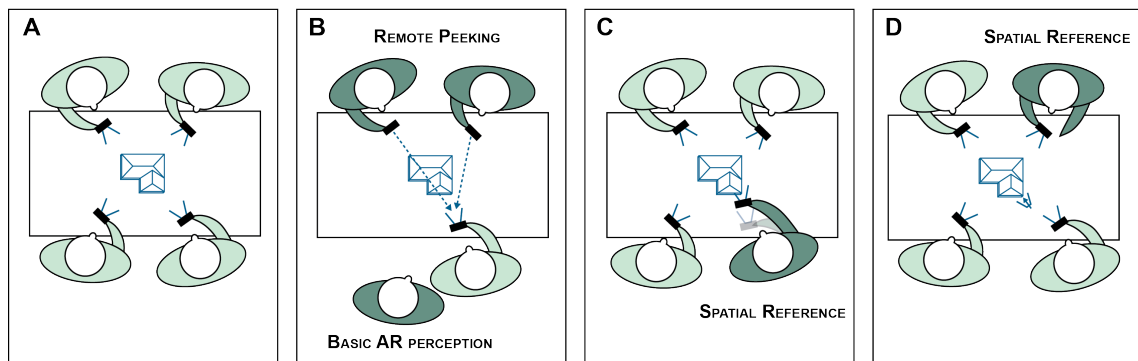
## 5.5 Example Applications Scenarios

To demonstrate how the conceptual lens can be used as a tool for developing new AR applications, we consider two different domains and lean upon application areas that have already been explored within co-located AR [8]. This includes an AR application for architectural design and an AR application for education. With these example scenarios, we want to illustrate how we can use the proxemics lens to design for flexible social interactions. The described scenarios illustrate how the interaction techniques can be complementary to and also support the dynamics of the collaborations over time. In addition, they illustrate the potential value of interaction techniques that are not only grounded in the physical space but go beyond reality.

### 5.5.1 Architectural Design Meetings

In this scenario, an architect company wants to have an AR application that their architects can use to collaborate around 3 models of building designs in internal





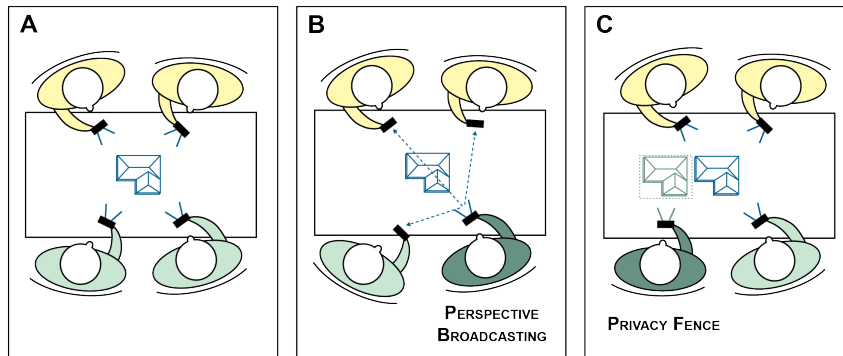
**Figure 5.8:** A closely coupled collaboration between architects in an internal design meeting. A) The architects view and discuss the overall design of the building. B) They shift into discussing a detail, one peeking over the shoulder (Basic AR Perception), while the others use Remote Peeking to see the detail. C) One creates a Spatial Reference highlighting the detail. D) Later, one of the architects points towards the Spatial Reference to reference the detail again.

design meetings and external client meetings. In these internal design meetings, the architects meet within a meeting room to discuss different specific designs, and they use this time to agree on the details of the designs. Most of the time, architects should have a tightly coupled collaboration while they discuss the design of a building (see [Figure 5.8](#)).

Sharing perception should have a low proxemic threshold to rapidly shift between the overview and the specific details. If the architects are side by side, sharing perception is made easy by looking at each of their own devices from almost the same angle, basically placing their phones almost side by side. However, if one is sitting on the opposite side of the table, the proxemic threshold is much higher. In this case, being able to provide *multiple alternatives*, such as also enabling the Perspective Peeking technique, could create a low proxemic threshold for all the users and support these rapid shifts between contexts, as illustrated in [Figure 5.8](#).

When the architects find problems with the designs, they need to be able to make detailed deictic references to the problem areas and then remember them later. Here, the Spatial References interaction technique comes in handy, supporting both referencing the problems when they are found and the spatiality of the reference point, making it easy to remember and revisit the point later.

In these internal design meetings, the architects have symmetric roles. However, in external client meetings, the roles between architects and clients are asymmetric, which creates other needs for collaboration. At the external client meetings, the architects and clients discuss the design prepared by the architects (see [Figure 5.9A](#)).

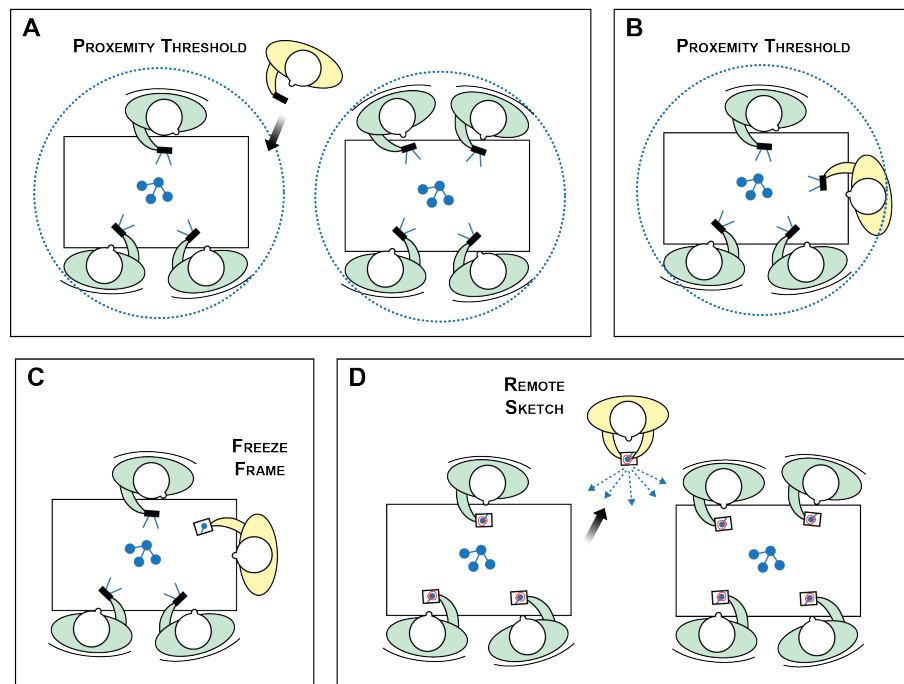


**Figure 5.9:** A meeting between architects (green) and clients (yellow). A) The architects and clients discuss the overall design of a building. B) One of the architects shows details to everyone using Perspective Broadcasting. One of the architects breaks out of the closely coupled collaboration and creates a Privacy Fence to work on a new design without interrupting the others’ conversation.

In this case, the architects want to show a detailed design to the clients. Using the Perspective Broadcast interaction technique, the architects are able to share their perspective with everyone without needing to change the current F-formation, as is illustrated in (see [Figure 5.9B](#)). In this meeting, the client may not like some of the designs that are being proposed to them. In this case, one of the architects may want to create a quick design variation for the client. By creating a Privacy Fence alongside the shared model, it is now possible for the architect to privately create a variation of the design without breaking out of the current F-formation, retaining operational consistency, and not disrupting the discussion that is already taking place, as illustrated in (see [Figure 5.9C](#))

### 5.5.2 Classroom Group Work

In a classroom setting, group work is an activity that commonly happens in which students work together on an exercise. At the same time, the teacher moves around the room to assist the different groups of students. Designing an Augmented Reality (AR) application for this usage case requires the consideration of the collaborative work among students, as well as the teacher’s role in the collaboration. In group work scenarios, students have similar roles, and ideally, they adopt a closely coupled collaboration style. With regard to the role of the teaching, the teacher moves around the classroom to discuss the exercises with each of the student groups and is only interested in helping one group at a time. In this situation, the use of a Proximity Threshold ensures that the AR content is only perceived and controlled once the



**Figure 5.10:** Classroom scenario where a teacher (yellow) walks around to groups of students to help with their AR exercises. A-B) The teacher walks within the Proximity Threshold to a group to see their AR content. C) The teacher uses Freeze Frame to show a problem to a student. D) The teacher shows the problem to the entire class by broadcasting a live sketch (Remote Sketch) of the problem to the class while explaining.

teacher is close to it. For example, when the teacher approaches a group to help them, the content appears once within close proximity, as illustrated in [Figure 5.10A-B](#). In addition to the use of the threshold, the teacher could also use the Freeze Frame feature to allow them to show details of the virtual content to one of the students by attaching the view to the device viewport and turning the device away from the AR collaboration to display the information, as can be seen in [Figure 5.10C](#).

In the case that the problem being shown is relevant to all the students in the class, the teacher can keep the frame on the device, move away from that group and use Remote Sketch to talk to the whole class, using Perspective Broadcast to display their content to the entire class, and using the live sketch to highlight parts of the object, as shown in [Figure 5.10D](#). These features enable teachers to better manage group work in a classroom setting using AR technology.

## 5.6 Discussion

In this section, we return to the initial goals of this chapter, where we aimed to explore how proxemics theory could provide a socio-spatial conceptual lens that would help to understand and design co-located collaborative handheld AR systems. More specifically, these goals were threefold. Firstly, we aimed to establish a conceptual vocabulary for AR systems and design that focuses on the socio-spatial dynamics. Secondly, we looked to investigate how proxemic concepts – such as perception, deixis, and control proxemics – can help to operationalise flexible and collaborative experiences. Finally, we aimed to offer a continuum of AR interaction techniques that not only align with the physical reality but extend beyond this into virtuality. This then provides AR interaction designers with versatile tools to navigate and shape social configurations in AR.

By looking at AR through the conceptual proxemics lens, we have been able to identify how interaction techniques can either constrain or extend spatial relationships, directly influencing the way AR systems can support and shape collaboration in co-located collaborative situations. By grounding the exploration of this chapter in practical and applicable application scenarios – ranging from educational to architectural design – we demonstrated how these interaction techniques can facilitate dynamic collaborations, allowing users to more fluidly switch between closely and loosely coupled collaboration. The findings of this chapter offer interaction designers a structured design space to draw upon, fostering an awareness of how proxemic relationships may be leveraged to enhance social and spatial adaptability in co-located collaborative handheld AR interactions.

Below, we summarise the key insights from this exploration of proxemics as a lens for co-located collaborative handheld AR.

### 5.6.1 AR for Aligning With and Going Beyond Reality

In this chapter, we have explored the potential of designing AR applications that not only align with the physical reality but also leverage the networking capabilities that inherently come with modern mobile handheld AR devices to go beyond reality. The focus was not to compare the different systems or judge them as better or worse, but rather to showcase the different possibilities of designing collaborative AR systems, taking a first step from shifting the attention from viewing AR as a linear interaction to one grounded in reality. The exploration of the design space

for different interaction techniques that are based on the dimensions of Interaction Proxemics [131] – perception, deixis, and control – highlighted how AR interaction techniques can be based on reality, whilst also extending beyond it into virtuality [84]. This exploration indicates how we can utilise the medium of AR to transcend our physical limitations. The contribution of this chapter is not only to demonstrate individual interaction techniques, which have been explored before in the context of AR, but also to present a broader design space for interaction techniques that can be used in co-located collaborative handheld AR.

### 5.6.2 A Many-To-Many Relationship Between Collaboration Styles and F-Formations

The model of group dynamics in this chapter builds on prior research of collaborative coupling styles [156, 200], which are based on empirical studies. We presented interaction techniques that are derived from the use of the proxemics lens, which expands the possible forms of collaboration styles from a different perspective. This presents a novel opportunity for AR to transform the one-to-one relationships between coupling styles and interpersonal distance, as well as between coupling styles and device configurations that have been identified in empirical studies [200].

Using the conceptual lens of proxemics, we show that group dynamics within AR can be understood as a many-to-many relationship between F-formations and collaboration styles. Co-located collaborative handheld AR systems can enable users to engage in the same collaboration style while enacting it in various F-formations. Conversely, these systems can also allow the users to switch between multiple collaboration styles while maintaining the same F-formation. For example, both the Perspective Broadcast and Remote Sketch interaction techniques enable users to be in and also transition between different F-formations, all while still keeping the same collaboration style, such as discussing the same details of the virtual content. Another example can be seen in the Privacy Fence interaction technique within the client meeting. This technique enables users to transition between collaboration styles while keeping the same F-formation, such as an architect shifting from closely coupled collaboration with the clients to working on something private for a brief period of time before then returning to active discussion.

### 5.6.3 Designing for Awareness and Control in Flexible Proxemic Interactions

Designing for going beyond the physical reality not only enhances flexibility in socio-spatial configurations but also adds complexity to the application in terms of the user's understanding. Remotely controlled interactions, which provide flexible options, may feel less natural to the user's collaborations because they do not immediately extend our physical spatial relations, which we rely on when we are interpreting non-verbal cues. As such, future investigations in this area should focus on the design of appropriate awareness cues to address questions such as how users can understand the application's functionalities, what the different ways to share user perception and deictic gestures could be, and also who controls the content. For example, with the Remote Peeking interaction technique, users can now flexibly peek into another user's perspective – but how can they determine who is seeing what? With the Privacy Fence interaction technique, a user can work on something privately, even while standing directly in front of another user, who remains unable to see the content. Although awareness cues have been extensively explored in remote collaboration, more work is required in this area to understand how these cues can be used to support co-located handheld mobile collaboration. This also highlights the trade-off between adding flexibility and understanding the application.

### 5.6.4 Conceptual Framework for Real-World Application Scenarios

The scenarios in architectural design and classroom group work illustrate how the proxemic lens can be applied practically to meet the unique needs of real-world co-located handheld collaborative settings. In particular, the ability to transition fluidly between aligning with the physical reality and also going beyond it can support diverse group configurations and collaborative demands. For example, in the architectural application scenario, the Privacy Fence interaction technique demonstrates how a designer might support individual ideation within a collaborative setting, a capability that is particularly beneficial for client-focused work in design industries. The classroom application scenario highlights how proxemics could enable more adaptable group configurations, helping educators to balance individual, small-group, and whole-class interactions seamlessly. These examples showcase how this conceptual framework can be applied and extended to cater to the evolving needs of diverse domains.

### 5.6.5 Insights on Trade-Offs in Flexibility and Intuitiveness

An important insight from this exploration is the inherent trade-off between enabling flexibility in AR interactions and maintaining intuitive, contextually relevant cues for user understanding. Systems that go beyond reality offer enriched possibilities for remote control and flexibly available views but risk introducing cognitive challenges by departing from natural spatial cues. For example, the Remote Sketch and Perspective Broadcast interaction techniques both empower users to access shared perspectives from a distance but may require users to interpret non-physical representations of proximity and control. These findings underscore the need for careful calibration in designing interaction cues, especially in settings where real-time group coordination and seamless flow of interactions are crucial, as with architectural teams and classroom dynamics. Future work within this area should investigate the optimal balance between flexibility and ease of use, considering how intuitive feedback mechanisms might help users understand shared virtual space without losing track of co-located cues and intentions.

### 5.6.6 Potential for Extending Proxemics Theory to AR

The work in this chapter also highlights a gap in the research and paves the way for further theoretical development by demonstrating how proxemics, traditionally applied to physical interactions, can inform AR design. This extension of proxemics to AR reveals new ways of understanding spatial dynamics in digital settings and suggests potential refinements to the proxemic theory itself. For example, the concepts of deixis proxemics and control proxemics, introduced here, reflect how AR allows users to engage in complex, multi-layered socio-spatial relationships that go beyond physical constraints. By articulating these dimensions, this research encourages future studies to expand and refine proxemic concepts to accommodate the unique affordances and interaction patterns of AR, potentially influencing more expansive fields such as Human-Computer Interaction (HCI) and Computer-Supported Cooperative Work (CSCW).

### 5.6.7 Limitations and Future Work

Based on the design exploration within this chapter, we have mapped out the different interaction techniques on a continuum, ranging from aligning with the physical reality to going beyond it. However, using a design space exploration and evaluation by

demonstration [114] has certain limitations. As the interaction techniques are based on an exploration and not an extensive review, the presented techniques should not be seen as an exhaustive list. Rather, they should be treated as a starting point for the further exploration of this design space. Furthermore, the interaction techniques have not been evaluated in the system usage section of this chapter. Further to this, the interaction techniques have not been deeply evaluated in usage, so more work is needed to validate the techniques further and study their interplay and how they are useful in specific contexts and domains.

In this chapter, we have focused on AR systems based on handheld mobile technologies. In handheld mobile AR, the device acts as a *portal* to perceive the AR content anchored in the physical space. The limited Field of View (FoV) of the handheld mobile device creates some challenges in the co-located collaborative setting. However, the mobility of the device also creates opportunities, as shown in the design exploration. In other AR technologies, AR devices are in the same way a portal to the AR virtual world, but devices have different roles in the physical environment.

For example, with AR glasses, the device is worn directly on the user's head. At the same time, in spatial Augmented Reality, virtual content is seamlessly integrated with the physical environment by being anchored directly onto physical space. This direct connection – whether linking the viewing device to the user or embedding content within the environment – simplifies the relationships between users, devices, the physical space, and AR content when compared to standard handheld mobile AR. However, we argue that proxemics still hold value in understanding and informing the design of these display types, offering deeper insights into how users interact within augmented spaces. Future research is needed to broaden the space for the use of proxemics in designing for collaborative handheld AR.

## 5.7 Conclusion

In this work, we have demonstrated how proxemics theory can be used as a conceptual lens for articulating the socio-spatial dynamics of collaborative AR systems and informing their design. Through design space exploration, we have explored how interaction techniques for handheld multi-user AR applications can either align with physical reality or extend beyond it, offering new possibilities for flexible collaboration. Our findings reveal a many-to-many relationship between collaboration styles and F-formations, highlighting the dynamic ways users can engage with both virtual content and each other in co-located settings. Building on this, we introduced the Reality and



Beyond continuum, illustrating how AR systems can extend enacted proxemics by reinforcing natural spatial relations or supporting greater flexibility by transcending physical constraints. We also emphasise the importance of understanding trade-offs in AR design – particularly in terms of perceptual, deixis, and control proxemics – to balance flexibility with intuitive and socially grounded interactions. This framework not only provides valuable guidance for designing co-located collaborative handheld AR systems but also underscores the broader potential of proxemic theory to drive future innovations in augmented reality experiences.

In the next chapter, we build on the theoretical foundation of proxemics to introduce the Collaborative Augmented Reality Interaction Space (CARIS) framework, which translates socio-spatial insights into a practical model for designing co-located collaborative AR systems. CARIS integrates key proxemic dimensions into a flexible, component-based architecture that facilitates intuitive interactions across users, devices, and physical space. Grounded in research on co-located AR, proxemics, and Computer-Supported Cooperative Work (CSCW), the framework supports dynamic collaboration by accommodating diverse spatial configurations and interaction styles. The chapter outlines CARIS’s core elements – People, Interaction Portals, and Physical Features – and demonstrates its application through an exemplar system and a small-scale architectural design study. Together, these contributions offer a foundation for designing AR systems that foster meaningful collaboration and align with natural human interactions.

# 6

## Designing Collaborative Spaces in Mobile Augmented Reality: Integrating Proxemics and Cross-Device Interactions

### Contents

---

|            |                                                          |            |
|------------|----------------------------------------------------------|------------|
| <b>6.1</b> | <b>Introduction</b>                                      | <b>131</b> |
| <b>6.2</b> | <b>CARIS Framework: Operationalising Proxemics in AR</b> | <b>132</b> |
| <b>6.3</b> | <b>System Implementation</b>                             | <b>139</b> |
| <b>6.4</b> | <b>Study Design</b>                                      | <b>146</b> |
| <b>6.5</b> | <b>Results</b>                                           | <b>148</b> |
| <b>6.6</b> | <b>Discussion</b>                                        | <b>156</b> |
| <b>6.7</b> | <b>Conclusion</b>                                        | <b>160</b> |

---

In the previous chapter, we introduced proxemics as a conceptual lens for the understanding of socio-spatial dynamics in co-located, collaborative handheld Augmented Reality AR. Through dimensions of perception, deixis, and control, the chapter explored how spatial configurations influence user interactions and collaborative styles within co-located handheld AR environments. These insights show the critical role of AR systems design in shaping not only the individual user experiences but also the broader socio-spatial dynamics among collaborators.

In this chapter, we extend these theoretical insights by presenting a Collaborative Augmented Reality Interaction Space (CARIS) framework. CARIS moves from abstract proxemic principles to a structured, practical model that addresses the

challenges of designing co-located collaborative AR systems. By integrating proxemics theory with interaction design components, CARIS provides a flexible, component-based architecture allowing intuitive, proxemic-aware interactions across people, devices, and physical space. This framework serves as both a conceptual and operational space, enabling AR designers to create socio-spatially aware systems that foster genuine collaborative engagement within a shared AR space without breaking operational consistency.

CARIS is shaped by empirical research on co-located AR and theoretical concepts from proxemics and Computer-Supported Cooperative Work (CSCW). This research reveals the importance of fluid, flexible interactions across users, devices, and physical space, underscoring the need for a shared, cohesive AR experience that supports collaborative engagement. By synthesising these insights, CARIS offers a structured approach to designing AR systems that honour proxemic principles in a shared interaction space. This chapter unfolds the CARIS framework journey and its three components: People, Interaction Portals, and Physical Features. We provide:

1. An overview of the CARIS framework's core elements – People, Interaction Portals, and Physical Features – and how each plays a critical role in structuring collaborative AR environments.
2. An exploration of how these components operationalise proxemic dimensions (perception, deixis, and control) to create a dynamic AR collaboration space that accommodates various collaborative styles and spatial configurations.
3. An exemplar system implementation demonstrating proxemic- and cross-device-aware interaction techniques, as discussed in [Chapter 5](#).
4. A small-scale study based on an architectural application scenario, aligned with the examples presented in [Chapter 5](#). This study evaluates collaborative behaviours, context-switching dynamics, and user engagement within the CARIS framework.

Together, these sections provide a first illustration of how we can integrate proxemics principles into a cohesive AR interaction space, guiding the design of socio-spatially dynamic, proxemic-aware collaborative AR systems. With this framework and study, CARIS not only extends proxemic principles into actionable AR components but also validates these through an empirical examination of collaborative dynamics. By designing for socio-spatial awareness, CARIS sets a foundation for AR systems that align closely with natural human interactions, redefining the collaborative potential of handheld AR environments.

## 6.1 Introduction

Augmented Reality (AR) has come a long way since its initial conceptualisations, evolving into a transformative medium that seamlessly blends the digital and physical realms [5]. Handheld mobile AR is currently the dominant way people can consume AR [18], and recent advances in mobile technology have accelerated this evolution, ushering in an era where handheld AR devices are increasingly accessible and play a key role in shaping digital interactions [1, 46, 59]. These frameworks have led to the development of a variety of AR applications in different domains, for example, IKEA Place [93], which allows users to place furniture in their home, Pokémon Go [143], which combines gaming with the real world by overlaying virtual creatures, and QuiverVision [159], which brings colouring pages to life to improve engagement in classroom environments. However, while the ubiquity of these devices offers unparalleled opportunities, they can disrupt face-to-face social interactions [135, 149, 204], and their prevalent design paradigms lean toward individualistic consumption, often at the cost of genuine collaborative experiences [123, 204, 217].

Historically, AR's primary vision was not solely individualistic. From its inception, there was a promise of a shared and collaborative space where users could interact with digital content and each other in meaningful ways [9, 15, 177]. However, the current state of AR, particularly in the context of handheld devices, seems to have drifted away from this vision. This drift is evident in the overwhelming emphasis on device-centric on-screen interactions, which often impede the fluidity and dynamism inherent to collaborative experiences [68, 106]. Common handheld device applications use device-centric on-screen interaction metaphors designed for devices with touch-based screens [57]. Previous research has explored general interaction techniques with handheld devices and also gestures that occur between [33, 68, 178] and around [220, 224] devices. These techniques of cross-device interaction transcend the typical linear view of device-centric AR, making use of multi-device tracking and external cameras in one system [26, 89] to enable expressive 3D motions for techniques. However, these approaches typically rely on the modification of existing devices using advanced sensing techniques, making them difficult to build, test, and deploy to ad hoc mass audiences [26]. In addition to cross-device interactions, research on proxemic interaction shows an improved semantic relationship between devices, which is a "*central precondition for any application or technique to work across devices*" [26, 66], and can facilitate further control of digital devices when interpersonal distance is introduced [7, 63]. While the benefits of these works for

interaction and collaboration have been well demonstrated, they have yet to be introduced to co-located AR collaboration.

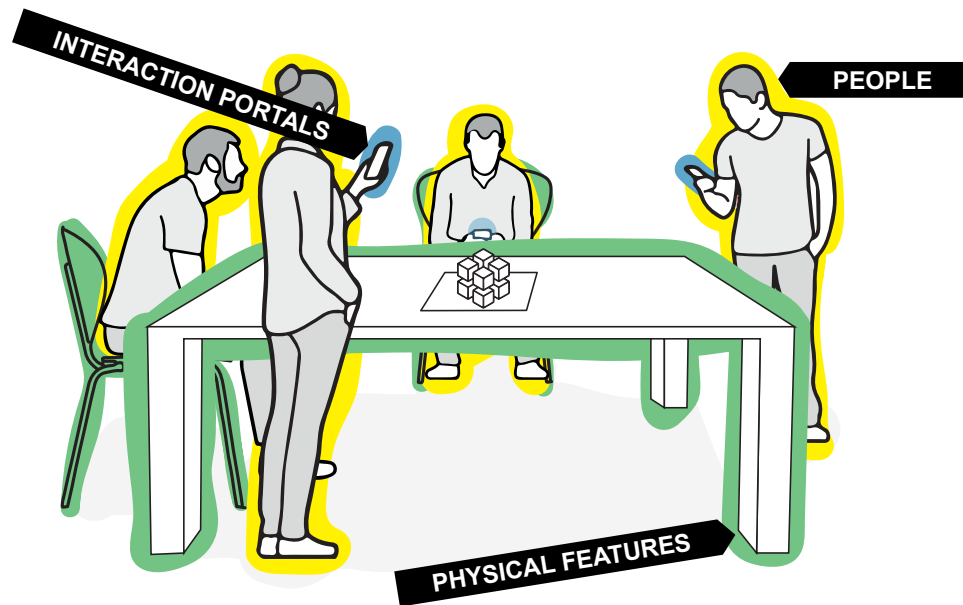
Recognising this gap, CARIS revisits AR's collaborative roots by drawing from diverse collaborative domains such as tabletop interactions, cross-device computing, and proxemic interactions. CARIS aims to rekindle the collaborative spirit of AR [7, 26, 63, 89]. The framework not only conceptualises a holistic AR space but also offers tangible guidelines for its realisation. Building on CARIS's principles, we developed an exemplar component-based system that encapsulates its principles. This system introduces interaction techniques that, although familiar in other domains, present a fresh perspective in the co-located AR domain. These techniques are designed to foster various interaction modalities, increasing user immersion and offering a more nuanced AR experience [33, 178, 220, 224].

Drawing on the conceptual insights of proxemics, CARIS responds to the limitation of an individualised device-centric AR space by reintroducing the collaborative potential of AR. By designing AR spaces that integrate both physical and digital elements through proxemic principles, CARIS enables a dynamic interaction space that supports diverse collaborative styles and spatial configurations. This transition from individualised device-centric AR to a flexible, interactive AR space establishes CARIS as a framework for designing AR systems that are responsive to socio-spatial dynamics.

## 6.2 CARIS Framework: Operationalising Proxemics in AR

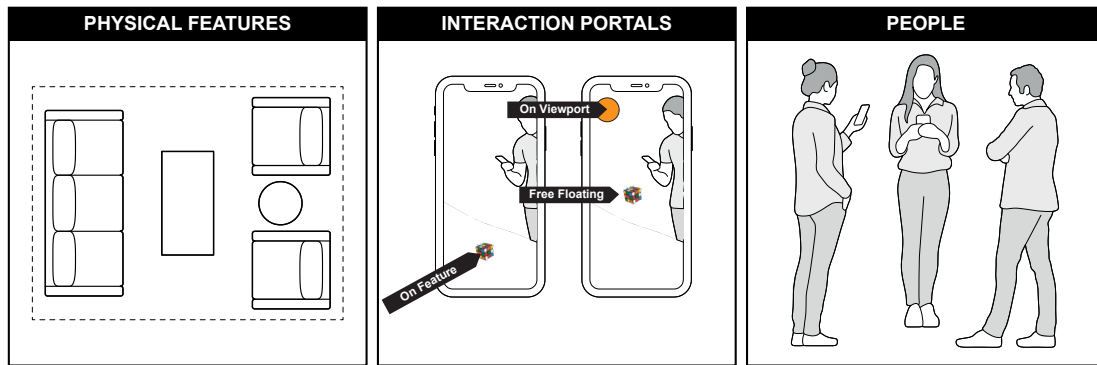
CARIS builds on proxemic insights to establish a collaborative AR framework centred on three core components: People, Interaction Portals, and Physical Features. Each component embodies a specific proxemic function – perception, deixis, or control – enabling a fluid, shared AR experience. In contrast to the typical linear view from user to device, CARIS treats the AR space as an interconnected environment where users, devices, and spatial features interact fluidly and meaningfully.

Mobile AR applications typically focus on device-centric, on-screen interactions, thereby providing an individualistic linear view between the components of co-located AR collaboration. Related work illustrates that co-located collaboration can, therefore, adversely affect group collaborations by, for example, causing high levels of physical and cognitive load, as well as large amounts of attention switching [123, 217,



**Figure 6.1:** A sketch of the conceptual Collaborative Augmented Reality Interaction Space (CARIS) with each of the components highlighted and labelled.

218]. To address this problem and move beyond an individual AR experience, CARIS aims to interconnect the individual components of co-located collaborative AR into one seamless space, allowing free interaction across devices, people, and 3D models. Specifically, we define CARIS as being composed of a commonly shared ground plane and a 3D volumetric space above and around users and handheld devices (see Figure 6.1). As such, rather than the typical linear view from the user through the handheld device into the AR space, we propose that everything is connected and aware of everything else within the single interaction space. As a result of this approach, users are enabled to bring their own device into the space and interact with the virtual content, being able to interact fluently in this 3D area – either with other users, their own devices, or the fixed/semi-fixed features around the room. We believe that interactions should not be limited to those that happen in a device-centric, on-the-screen interaction below the user’s fingertips but also across users and devices; gestures should flow freely through each component. At the same time, the user is directly in control of one device, which does not preclude gestural acts from extending across multiple users and devices beyond the physical limits of one person.



**Figure 6.2:** The three foundational elements of CARIS: Physical Features (left), Interaction Portals (middle), and People (right).

### 6.2.1 Foundational Elements of CARIS

In the realm of co-located AR collaborations, three fundamental elements consistently emerge as pivotal for fostering collaboration [4, 5, 18, 30, 46, 106, 163]: *Physical Features*, serving as the focal point or periphery around which collaboration revolves; *Interaction Portals*, offering a window into the virtual realm; and *People*, the driving force behind collaboration. The CARIS framework is intricately woven around these components, ensuring that AR transcends a solitary experience bound to a device and thrives as a collaborative venture seamlessly integrated with the ambient environment. The foundational elements of CARIS – People, Interaction Portals, and Physical Features – translate the dimensions of proxemics into practical design components that structure collaborative AR interactions. These elements address key socio-spatial needs in AR design, each serving as a point of interplay where perception, deixis, and control proxemics are expressed and refined in the collaborative space. Together, these components facilitate a seamless, proxemic-aware AR environment that fosters flexible, engaging interactions.

#### C1 - People

In CARIS, People are the primary agents driving collaborative interactions and configuring perception proxemics. Each user's position, orientation, and proximity to others shape how AR content is perceived and shared within the group, enabling seamless transitions between tightly coupled group tasks and loosely coupled individual activities based on the proxemic requirements of each interaction [75]. This flexibility empowers users to arrange themselves as needed, adapting their engagement levels dynamically. Additionally, CARIS provides

options for users to share or maintain privacy over their Interaction Portals, facilitating both private views and group-based public perspectives within the AR environment [169].

## **C2 - Interaction Portals**

Interaction Portals are handheld devices that function as adaptable windows into the AR space, anchoring deixis proxemics by enabling users to perform and observe deictic gestures across devices [169]. These portals facilitate clear, proxemic-aware communication, allowing users to navigate public and private views of the AR space, which fosters flexible, real-time collaboration. Serving as both input and output interfaces, Interaction Portals can transition between states: for instance, shifting from standard AR views to “Perspective Broadcasting” or “Freeze Frame” modes, where users can share live or static perspectives, dynamically adapting to the collaborative context [26, 77].

## **C3 - Physical Features**

Physical Features encompass the fixed and semi-fixed elements in the AR space, such as furniture, tables, walls, or doors [75], that help to structure and anchor collaborative activities within the AR environment. These features support control proxemics by providing stable anchor points for positioning AR content, serving as territorial zones that distinguish between public and private areas [25, 77]. For instance, semi-fixed elements like tables can serve as Overview Devices [25], facilitating collaboration by providing a shared focal point for AR content, thereby enhancing decision-making and group interactions [95, 213]. Additionally, Physical Features enable users to set privacy boundaries, such as the “Privacy Fence,” which demarcates zones for private and collaborative work, balancing individual and shared control within the AR space.

While these foundational elements define the core components of CARIS, their dynamic interrelationships and interactions truly bring the framework to life and enable versatile collaborative experiences.



### 6.2.2 Interplay of CARIS Components

Having established the foundational elements of the CARIS framework, it is essential to understand how these components interact within the collaborative AR space. The strength of CARIS lies not just in how these foundational components – People, Interaction Portals, and Physical Features – interact within the AR space but in their interplay, fostering dynamic and meaningful interactions. Each component’s dynamic relationships support the continuous, proxemic-aware flow of interactions across users, devices, and space, facilitating a seamless transition between collaborative modes.

#### **People → People**

Within the CARIS framework, interpersonal interactions are as dynamic as the AR space itself. People can shift from private, individual engagement to more public, collaborative modes, reflecting the fluid nature of proxemics in co-located AR. This flexibility supports various forms of perception sharing, ranging from individual views to group-wide broadcasts, thus enhancing collaborative immersion.

#### **People → Interaction Portal**

People’s interaction with portals goes beyond the device-centric model, allowing them to share perspectives and engage in deixis with ease. CARIS’s structure accommodates both private and shared interactions across portals, empowering users to manage their visibility and deictic references effectively. This shift facilitates proxemic-informed deixis that promotes clearer, more intentional collaborative exchanges.

#### **People → Physical Feature**

Physical Features within CARIS act as spatial anchors that support both shared and private collaborative zones. By positioning devices and AR content in relation to physical elements like tables or walls, users can control the visibility and accessibility of AR content. For instance, a user might place an AR object on a shared table for public viewing or create a “Privacy Fence” around specific content, balancing control proxemics with the need for shared interaction.

CARIS is shaped by empirical research on co-located AR and theoretical concepts from proxemics and Computer-Supported Cooperative Work (CSCW). This research reveals the importance of fluid, flexible interactions across users, devices, and physical

space, underscoring the need for a shared, cohesive AR experience that supports collaborative engagement. By synthesising these insights, CARIS offers a structured approach to designing AR systems that honour proxemic principles in a shared interaction space.

The CARIS framework builds on the theoretical insights of proxemics, transforming these into an operational model for co-located collaborative AR. By defining three core components and exploring their dynamic interplay, CARIS establishes a proxemic-aware AR space where users, devices, and physical features interact seamlessly. This framework invites AR designers to craft flexible, socio-spatially responsive AR systems that foster genuine collaboration, bridging the gap between individual agency and collective engagement.

This approach aligns CARIS more explicitly with proxemic insights, creating a natural transition from the theoretical to the practical. Each component is clearly related to proxemic functions (perception, deixis, control), grounding the framework in the socio-spatial principles discussed in the prior chapter. This ensures that CARIS is not only a continuation but also a realisation of the proxemic lens tailored for collaborative AR.

### 6.2.3 Inspiring Design Guidelines

Using the components that established the foundational elements of the CARIS framework, we examine the guiding principles that inform its design. Drawing inspiration from empirical studies in co-located collaborative domains [7, 67, 126, 200], as well as findings throughout this thesis, we discern important considerations for the shaping of collaborative AR systems. These studies not only highlight the challenges inherent in AR-mediated collaborations but also suggest potential pathways to mitigate them.

The CARIS framework captures these insights and advocates for a paradigm shift in how we perceive the capabilities of AR handheld devices. Instead of a siloed, individual-centric perspective, CARIS envisions a cohesive space where all components are interconnected and contextually aware. This vision led us to formulate four central design guidelines for co-located AR collaboration, each deeply rooted in tried-and-true co-located collaborative practices. Here, we elucidate the core design guidelines that underpin the CARIS framework:

### **D1 - Enable Proximity and Orientation**

To enable natural and fluent interactions, it is imperative to consider the spatial relationships between the CARIS components. The way we physically interact and communicate with other people, as well as objects in our daily lives, is profoundly influenced by spatial relationships [7, 63, 75]. These proxemic considerations offer invaluable insights into group processes and dynamics. Furthermore, the gradual engagement between digital devices as a function of proximity – ranging from mere awareness to progressive reveal and eventual information transfer – underscores the importance of these spatial relationships in the design of collaborative systems [125]. As such, systems that adhere to the CARIS framework should prioritise the proximity and orientation of components within the system to facilitate and enhance interactions with the surrounding components.

### **D2 - Support Device Mobility**

Mobile devices are not inherently confined to a specific interaction space. This allows users to select their desired viewing angles by orienting the device in a manner that suits them. This freedom of movement and orientation of the device has paved the way for cross-device sharing techniques [33, 126]. However, challenges arise with static content on devices that are meant to be dynamic. The resultant mobility can sometimes hamper the user experience, especially when users are expected to change their viewing perspectives frequently [217]. Current AR applications generally do not factor in the flexibility of roaming within the virtual space, restricting user interactions. Adherence to the CARIS framework implies that systems should accommodate and even encourage micro-mobility during content interactions. They should also ensure that larger-scale interactions are seamless, facilitating user movement within the virtual environment [200].

### **D3 - Support Collaborative Transitions**

Collaborative transitions, from a user's primary focus to secondary areas of interest, are crucial in co-located AR collaborations. The fluidity between the collaboration styles, from loose interactions to tightly coupled interactions, is paramount for fostering seamless collaboration [89, 200]. However, existing AR applications often fail, lacking the agility to allow users to quickly transition between focal points and various collaboration styles without affecting operational

flow [46, 217]. For AR systems that adhere to the CARIS framework, there is an implicit need to incorporate techniques that allow rapid switching between primary and secondary collaboration styles or interaction spaces, ensuring uninterrupted and harmonious collaboration [126].

#### **D4 - Provide Multiple Alternatives**

When considering flexible social interactions in group collaborations, it is vital to offer users flexibility in their interactions. Research has highlighted the importance of “supporting multiple alternative manipulation opportunities to achieve the same effect” [67]. Current AR applications often have a single, prescribed way to complete an interaction, which can constrain the fluidity of co-located collaborations. Systems designed with the CARIS framework in mind should entertain the idea of multiple avenues of interaction. For example, a user could physically approach another participant to share a view or receive a method to see this view by proxy without the need for physical relocation.

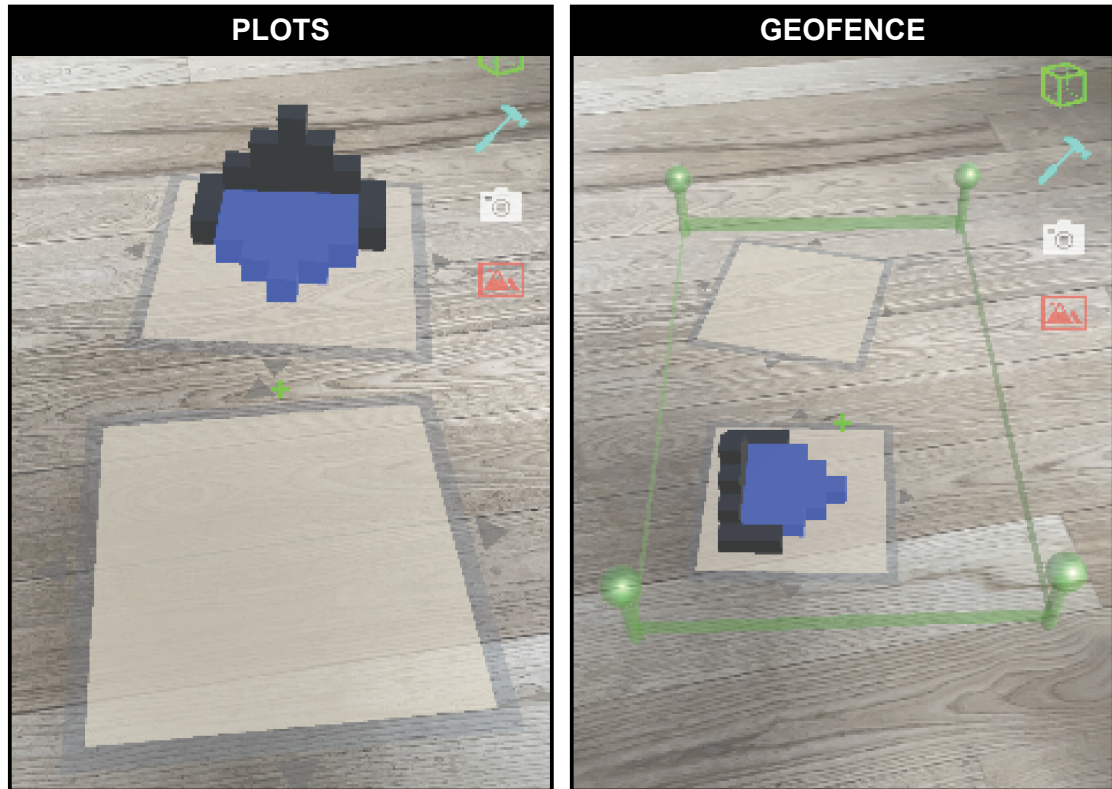
## **6.3 System Implementation**

Building upon the established foundational elements and inspired by the design guidelines, we present the system implementation of the Collaborative Augmented Reality Interaction Space. Pioneering interaction techniques as realised in [Chapter 5](#), our system upholds recognised principles of co-located collaboration, such as proxemic and cross-device interactions. Developed as an iOS mobile application, the CARIS system champions the co-creation of structures in synchronous, co-located environments. Drawing cues from Minecraft [134] and Blocks [71], it empowers users to prototype AR scenes using coloured blocks through an intuitive interface.

### **6.3.1 System Architecture and Component Structure**

The CARIS system follows a modular component-based architecture where each component is responsible for handling user input, managing tools, processing interactions, and synchronising across devices. The architecture consists of the following core components:

1. **Collaborative Features:** Drawing inspiration from common tabletop collaboration strategies such as territoriality [184], we developed a solution in which the building process is managed through collaborative plots. Users can **claim**



**Figure 6.3:** Two of the core features of the CARIS system implementation. The left figure shows a full and empty plot, and the right figure shows the Geofence placement.

**plots**, build on them, or work collaboratively. This feature allows users to manipulate entire structures, including **moving, rotating, scaling, resizing, copying, saving, deleting, and replacing plots** (see [Figure 6.3](#) (left)).

An additional feature, Geofence (see [Figure 6.3](#) (right)), allows users to define spatial boundaries around specific areas or physical objects. The Geofence creates a **rectangular public space** within the collaborative system, while anything outside remains private. Users can dynamically scale and rotate the Geofence, affecting all enclosed plots accordingly.

## 2. Input Handling and Tool Management:

- **InputHandler:** Captures user interactions, such as **screen taps, gestures, and device movement**.
- **ToolManager:** Determines the currently active **tool** (e.g., BuildTool for placing objects).

- **Selected Tool (e.g., BuildTool):** Executes user actions based on the tool's state.

### 3. Plot Management and Object Placement:

- **PlotGlobal:** Manages global modifications to plots, ensuring consistency across multiple users.
- **PlotLocal:** Handles local object placement before broadcasting changes to the network.
- **Unity Object Integration:** Converts user interactions into AR-rendered objects in the virtual scene.

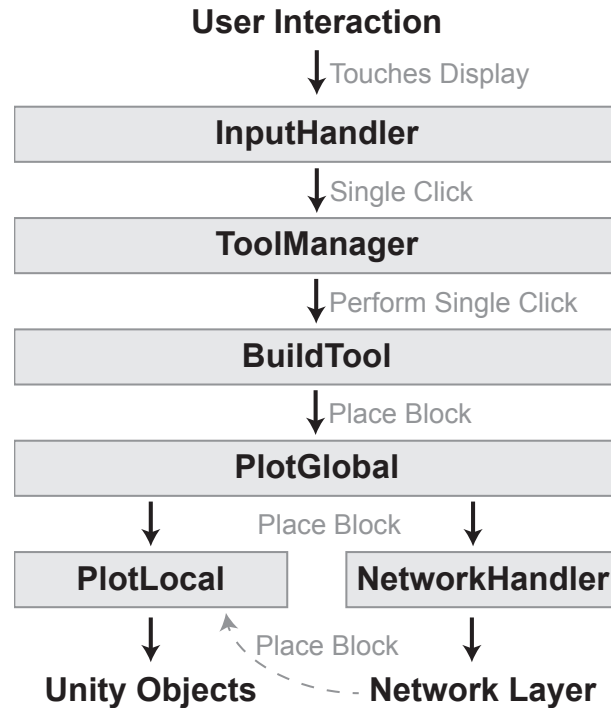
### 4. Network Communication and Synchronisation:

- **NetworkHandler:** Passes updates to the NetworkLayer, ensuring real-time collaboration.
- **NetworkLayer:** Utilises Apple's Multi-Peer Connectivity Framework to synchronise interactions across devices.
- **Anchor Points & Shared Dot Clouds:** Ensures that all devices align 3D spatial coordinates in the same AR environment.

## 6.3.2 Flow of Information in CARIS

The CARIS system processes interactions through a structured flow of information. User actions propagate through multiple system components before being rendered in the AR environment and synchronised across devices. The interaction flow begins when a user engages with the system (e.g., tapping the screen) and passes through input handling, tool selection, spatial processing, and network synchronisation.

At the core of this flow, the ToolManager dynamically switches states based on the system's current state, the object the phone is focused on, and the touch interaction. This ensures the system interprets user inputs contextually and selects the appropriate action.



**Figure 6.4:** Example flow of information during block placement in CARIS

**General Flow of Information** While multiple interaction flows exist in CARIS, a generalised pipeline follows these steps:

1. The **user initiates an interaction** (e.g., tapping the screen, dragging an object), which is first processed by the **InputHandler**.
2. The **ToolManager** determines the active tool (e.g., **BuildTool**, **DeleteTool**), switching states accordingly.
3. The **selected tool executes the action**, modifying a targeted AR object (e.g., an AR plot handled by **PlotGlobal** or **PlotLocal**).
4. If the action affects a local AR object, **PlotLocal updates the Unity environment**, rendering the changes visually.
5. If collaboration is required, the **NetworkHandler transmits updates** to the **NetworkLayer**.
6. The **NetworkLayer synchronises changes** across all connected devices, ensuring consistency.

7. **Other devices receive and apply the update**, ensuring a **shared and synchronous AR experience**.

**Example: Block Placement Flow** An example of this flow can be seen in [Figure 6.4](#), which illustrates how information moves through the system when placing a block in AR. In this case:

- The ToolManager determines that the user is using the BuildTool.
- Both the PlotManager and the Plot are notified of the interaction.
- The Plot or PlotManager then modifies the local scene, placing the block within AR.
- This updated information is shared with other peers via the network layer to ensure synchronisation.

This structured flow enables CARIS to support real-time co-creation, ensuring that local interactions are efficiently processed and consistently reflected across all devices in the collaborative AR space.

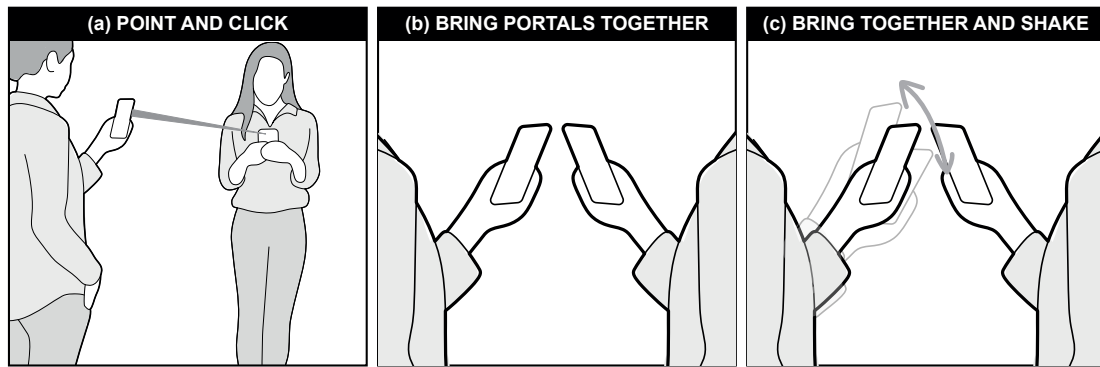
### 6.3.3 Interaction Space

CARIS defines the input and output of interactions as separate units, providing a foundation for flexible software interactions. Unity's in-built event system makes it possible to create valuable classes that define an input or an output. The event system ensures that the output components, which subscribe to certain input components, receive information as user input information. This gives us a set of high-level input and output components that facilitate the configuration of individual parts, allowing explorative and flexible studies.

#### Input Events

A set of input interactions has been implemented in the system. Primary input includes orienting the phone and clicking on the display. To determine which objects users wish to manipulate, they simply point a crosshair in the middle of the screen at the given object. This gives users the means to focus on objects by changing the orientation of the phone, which engages more muscle memory and frees up screen space for other interactions. For now, screen interactions include left-click, right-click, continuously pressing, long-click, erase gesture, scale gesture, and rotate gesture. Other means of input into the system include the following.





**Figure 6.5:** Illustration of input actions in the CARIS system. (a) **Point and Click:** Users interact by aligning their device’s crosshair with an object or another portal and tapping, allowing for selection without direct touch input. (b) **Bring Portals Together:** Users initiate interaction by moving two AR portals into close proximity, triggering a spatial event. (c) **Bring Together and Shake:** In addition to proximity-based interaction, shaking the device while portals are near each other activates a distinct input event. These techniques leverage spatial positioning, device movement, and proxemic interactions to enable more natural and intuitive user experiences.

**Point and Click** As devices are aware of each other’s coordinates and angles, looking at other portals also gives valuable input. In the system, each portal has a ring attached, which you can click through your own interaction portal. This fires an event, and the output may involve one or both of the portals. This is illustrated in (a) on [Figure 6.5](#)

**Distance to Devices** Utilising distance to other phones provides a foundation for interesting interaction abilities. For example, shaking or clutching one’s phone when close to others triggers two different events. These are illustrated in (b) and (c) on [Figure 6.5](#)

**Distance to Plots** Distance to plots also serves as meaningful information for implicit interaction with the environment. This concept draws on proxemics and how distance from objects in the physical world changes, complicates or prevents interaction.

## Output Events

Output events can be configured to listen to any type of input event, provided there is a sensible connection between the two. The following output events include both the manipulation of objects and the change in viewpoint to increase flexibility and aid communication.



(a) A user decides to peek through another user's portal by aiming and pressing the screen. (b) The user's view is split into two, with the other user's portal being the new main view. (c) The user can flip between views by clicking the bottom right corner.

**Figure 6.6:** Demonstration of Perspective Peeking.

**Save, Load, or Duplicate Objects** House manipulation allows much more than the placement and removal of blocks. This includes placing, moving, erasing, duplicating, scaling, rotating, saving, loading, and resizing the grid size of the plots.

**Perspective Freeze** To increase communication flexibility, users can freeze the current frame on the display. Freeze frames break the continuous timeline of augmented reality and can provide a focal point for discussion.

**Perspective Peeking** To aid visual communication between users, a peek-interaction allows FaceTime-similar communication, where a user sees both their own and can be shown simultaneously (see [Figure 6.6](#)). For two or more peers to share, an identical view gives way to well-informed conversations and discussions.

**Visual Guidance** Expanding on the ideas of peaking, peers can create guidance objects whose purpose is twofold. First, they spawn camera points through which users can peek in the same manner as peeking through other portals. This is interesting. Second, they provide a guide in the air from which users can obtain the same view of the world as when the frame was created. If a specific

task involves creating something with a specific point of view in mind, this may be useful.

**Geofence** Building and marking the Geofence requires selecting three points in the environment from which a rectangle can be defined. The act of choosing these points, as well as rotating and scaling the fence, are acts of output and require coordinates.

## 6.4 Study Design

In our research, we used the CARIS system to explore collaborative activities within a synchronous, co-located setting. Our study participants were engaged in co-creation tasks that encompassed a range of collaboration levels and roles. The objective of this study was to gain insight into the dynamics of collaboration, the impact of spatial configurations, and the effectiveness of various interaction techniques in a CHAR environment. This involved a comprehensive examination of the behaviour of the participants and the tasks performed. The data collected was then analysed to identify patterns, challenges, and opportunities that could inform CHAR systems' future design and implementation.

### 6.4.1 Participants

A total of 20 participants were recruited for this study using a snowball sampling method. The participant group was diverse, with 10 individuals identified as female, 8 as male, and 2 as non-binary. The study did not impose special requirements on participants. The tasks assigned during the study were abstract co-creation tasks framed within the context of architectural design. This theme was introduced as part of a narrative during the study to engage the participants and provide a context for the tasks. Most of the participants (19 of 20) were between 18 and 24 years old, with one participant falling in the age bracket of 25 to 34 years. No previous experience with AR was required for participation, although 11 participants had used AR technologies before the study. Examples of their previous AR experiences included applications like Pokémon Go, Harry Potter: Wizards Unite, Apple's product visualisation feature, and various social media filters. All participants owned and regularly used at least one mobile device, such as a mobile phone or tablet.

### 6.4.2 Apparatus

The study was conducted in a circular meeting room, which was chosen for its ability to facilitate free movement around semi-fixed and fixed furniture. Chairs were provided, but participants were explicitly informed before and during the study that they were not required to remain seated. On arrival, participants were asked to complete a consent form, followed by a prestudy questionnaire. This questionnaire was designed to gather basic demographic data and information on previous participants' experiences with AR technologies. Each participant was provided with an iPhone 11, equipped with a 6.1" display, with the CARIS application preloaded onto it for the duration of the study. At the start of the study, the participants were seated across from each other on one side of the table, with the researcher positioned at the opposite end. For the purpose of data collection and note-taking, pens, NASA-TLX forms, and an SUS form were made available to participants. These materials were to be used at specific intervals during the study, as well as whenever participants wished to write down any observations or thoughts.

### 6.4.3 Tasks

Our study focused on examining collaborative dynamics in the context of Augmented Reality (AR) through a task termed the 'City Co-creation Task'. This task aimed to understand the natural progression and nuances of collaboration without strict guidelines. To accomplish this, the task was divided into various subtasks, which included directives like "add more greenery to your houses" and "in pairs, build a fire station", among others. Although participants were provided with some direction, such as the need to transition to dyad work at specific moments, they were largely given creative freedom. This approach allowed us to observe organic collaborative behaviours as they emerged in the AR setting without imposing specific spatial dynamics or collaboration strategies.

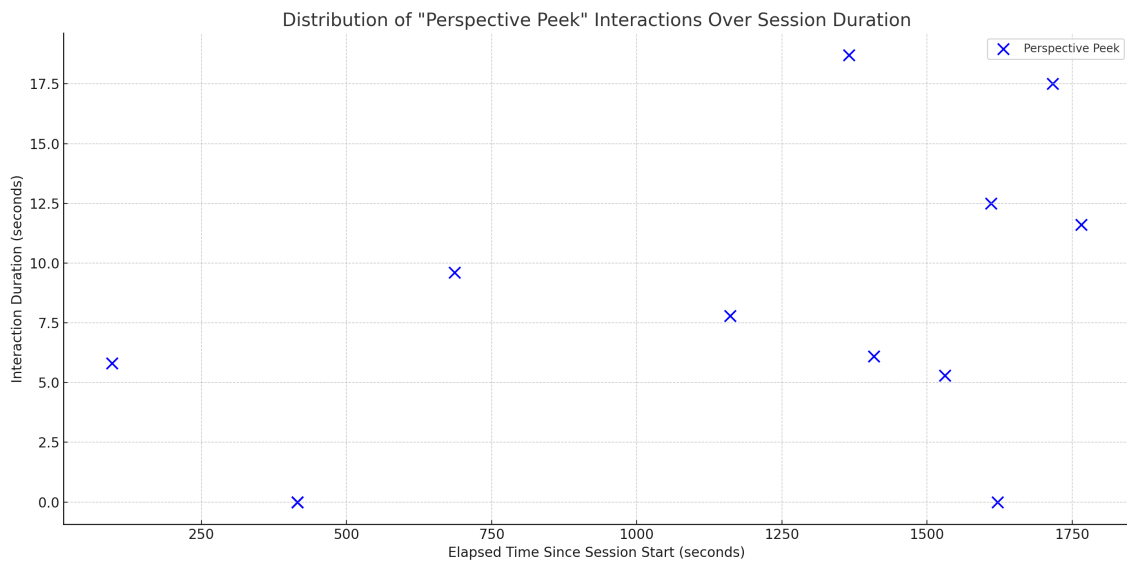
### 6.4.4 Data Collection and Analysis

Each session was recorded on video to analyse the group's collaborative behaviours, including collaborative styles, F-formations, and context switches. Context switches were identified as instances where users disengaged from the system's immersion, i.e. looked away from their device. Alongside the video, the investigator recorded observations that corresponded to the video events. After the session, participants

completed a NASA-TLX questionnaire to gauge their subjective workload and a System Usability Scale questionnaire to evaluate the system’s performance. A semi-structured group interview concluded the data collection. During this, reflections on tasks and feedback from the questionnaires facilitated deeper discussions. We analysed the data collected using ChronoViz [50]. The video and investigator notes were coded for collaborative practises and participant interactions. We followed a grounded theory approach and used open coding. We used a framework of Tang et al. [200] and modified it according to Chapter 3 to include groups of four and multiple devices. Our focus was on individual participants for system immersion and on group and subgroup dynamics when coding F-formations.

## 6.5 Results

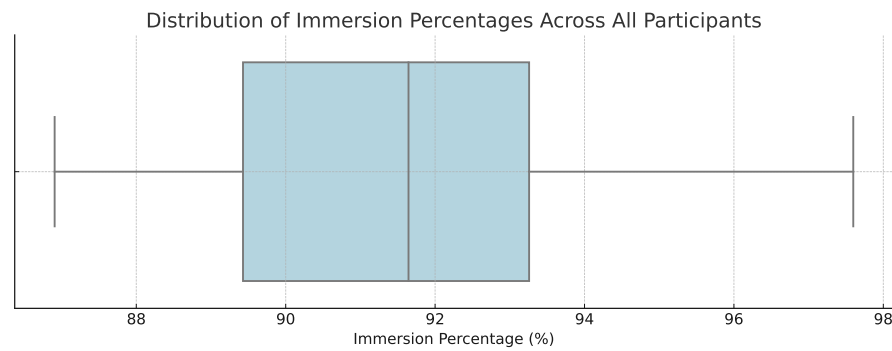
### 6.5.1 Interactions



**Figure 6.7:** Distribution of “Perspective Peek” Interactions Over Time. Each dot represents a unique instance of the interaction, with its vertical position indicating the duration.

In the analysed session, the “Perspective Peek” interaction technique was employed 12 times by the participants. The average duration of these interactions was approximately 7.91 seconds. However, the duration exhibited variability, with some instances being brief and others extending longer. When examining the temporal distribution of these interactions, they were found to be spread throughout the session,

suggesting sporadic usage based on need rather than regular intervals. Additionally, when contextualised within a 1-minute window around each “Perspective Peek” occurrence, the technique was frequently associated with context-switching events, especially for participants labelled P2 and P3. Furthermore, the technique was often used during collaborative styles characterised by individual focus, both with and without verbal communication. This suggests that users might resort to the “Perspective Peek” technique during moments of individual concentration, task transitions, or when adjusting to new contexts.



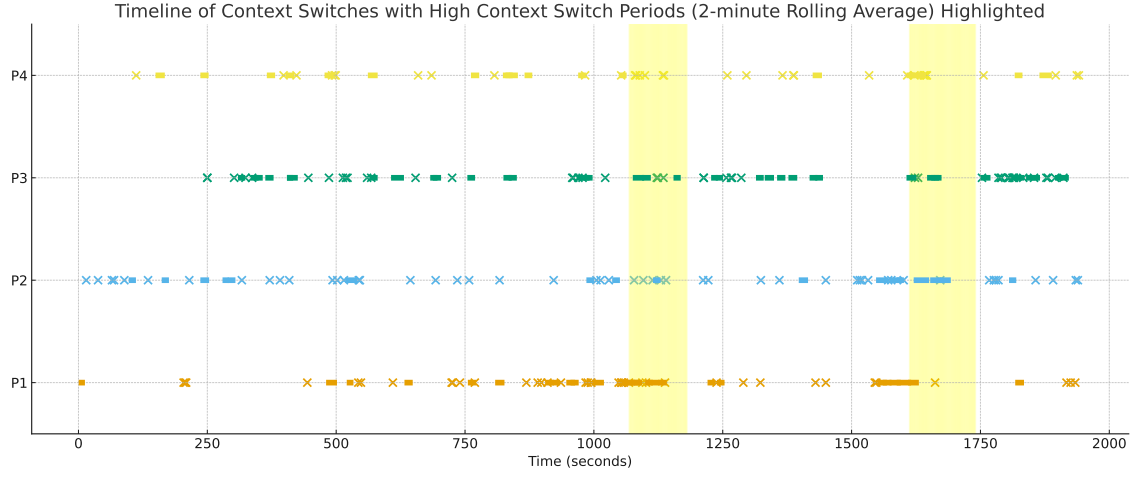
**Figure 6.8:** Boxplot visualising the distribution of immersion percentages across all participants. The central line within the box represents the median immersion level. The interquartile range (IQR), indicated by the width of the box, demonstrates the range within which the middle 50% of participants’ immersion levels fall. The whiskers of the boxplot extend to the minimum and maximum values, providing a comprehensive overview of immersion variations.

## 6.5.2 Context Switching

To analyse the immersion of the user within our system, we examined the context switches during the task. Context switches were defined as moments when users removed themselves from the immersive experience of the system. Data was extracted from the video coding of these switches, and each entry captured both the start and end times of the interruption. To quantify the level of immersion of each participant, we calculated:

$$\text{Percentage Immersed}_{P_n} = 100 - \left( \frac{\text{Total Non-Immersed Time}_{P_n}}{\text{Task Duration}} \times 100 \right)$$

Where  $P_n$  represents a participant (e.g., P1, P2, ...), and  $\text{Total Non-Immersed Time}_{P_n}$  is the sum of all durations of context switches for that participant. In cases where



**Figure 6.9:** Timeline illustrating the context switches of four participants (P1 to P4). Each horizontal bar indicates a period during which a participant disengaged from the immersive experience, while the crosses represent momentary switches. The highlighted regions, shown in a light yellow overlay, indicate periods with a significantly high number of context switches based on a 2-minute rolling average.

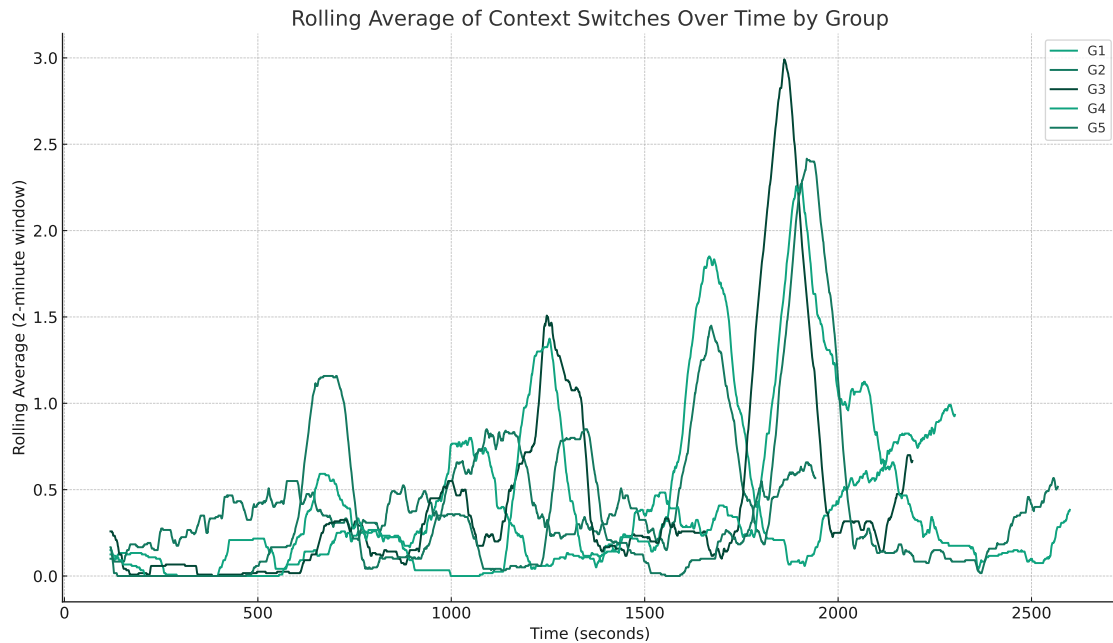
the start and end times of a context switch were identical, a default duration of one second was assumed, representing a brief momentary lapse in immersion. We also calculated the total immersion time using:

$$\text{Total Immersion Time}_{P_n} = \text{Task Duration} - \text{Total Non-Immersed Time}_{P_n}$$

For visual analysis, we illustrate the rolling average of context switches over time by group (see Figure 6.10). The graph presents the rolling average of context switches over a 2-minute window for each group (G1 to G5) during the AR task. The rolling average, calculated at every second, provides a smoothed representation of how often participants transitioned between tasks or contexts. Each group is delineated by a distinct line, allowing for comparisons between the different groups.

$$\text{Rolling Average}(t) = \frac{1}{\text{Window Size}} \sum_{i=t-\text{Window Size}+1}^t \text{Context Switches}(i)$$

This rolling average was calculated over a 2-minute window for each group to further elucidate patterns in the context-switching behaviour across participants. This analytical approach facilitated the identification of periods characterised by



**Figure 6.10:** Rolling average of context switches over a 2-minute window for each group (G1 to G5) during the AR task. The graph illustrates the frequency of context switches, with each group represented by a distinct line. The rolling average offers insights into the periodicity and consistency of participants' engagement and transitions during the task.

increased or diminished context-switching activity. By plotting these rolling averages, distinct trends emerged, revealing intervals during the task where participants, collectively, were more prone to disengage from the AR system. Reference to the rolling average graph (see [Figure 6.10](#)) underscores these pivotal moments. In particular, the highlighted regions in the graph mark periods where the frequency of context switches notably exceeded the rolling average. Such intervals may correspond to parts of the task that were either inherently challenging or where the AR interface possibly introduced ambiguities. Conversely, periods with a low frequency of context switches, lying below the rolling average, signify moments of sustained immersion, potentially indicating smoother user interactions or more intuitive task segments.

The analysis of immersion levels centred on the time durations in which the participants were not immersed, which served as an indicator of distraction or loss of engagement. A summary of the non-immersion times and corresponding immersion percentages for each participant can be found in [Table 6.1](#). On average, participants in G1 were not immersed for about 187.3 seconds. Those in G2 experienced an average non-immersion duration of 186.5 seconds. G3 participants, on the other hand, had the highest average non-immersion time across all groups, with a mean duration



**Table 6.1:** Non-Immersion Times and Immersion Percentages for Each Participant

| Participant | Non-Immersion Time (s) | Immersion (%) |
|-------------|------------------------|---------------|
| P1          | 240.0                  | 89.57         |
| P2          | 134.0                  | 94.18         |
| P3          | 219.0                  | 90.49         |
| P4          | 156.0                  | 93.22         |
| P5          | 199.5                  | 89.73         |
| P6          | 169.0                  | 91.30         |
| P7          | 239.0                  | 87.70         |
| P8          | 138.4                  | 92.88         |
| P9          | 287.0                  | 86.91         |
| P10         | 283.0                  | 87.09         |
| P11         | 153.0                  | 93.02         |
| P12         | 145.0                  | 93.39         |
| P13         | 173.0                  | 93.36         |
| P14         | 119.0                  | 95.44         |
| P15         | 287.0                  | 88.99         |
| P16         | 270.0                  | 89.64         |
| P17         | 203.5                  | 92.08         |
| P18         | 61.7                   | 97.60         |
| P19         | 205.8                  | 91.99         |
| P20         | 321.7                  | 87.47         |

of 217.0 seconds. Participants in G4 and G5 were not immersed for an average of 212.3 and 198.2 seconds, respectively. A deeper dive into individual performances revealed that P20 had the most prolonged non-immersion time of 321.7 seconds.

For a complete view of all participants' immersion levels, we refer to [Figure 6.8](#), which provides a visual summary of the data distribution. The median immersion level, represented by the central line within the box, was approximately 91.64%. The interquartile range (IQR), indicated by the width of the box, ranged from 89.43% to 93.26%, which indicates that the middle 50% of the participants had immersion levels within this range. The whiskers of the boxplot extended from the minimum value of 86.91% to the maximum value of 97.60%. No outliers were detected, suggesting consistent engagement across all participants. This uniform

engagement underscores the effectiveness and appeal of the collaborative co-creation AR tasks presented to the participants.

### 6.5.3 System Usability

Participants evaluated the CARIS system using the System Usability Scale (SUS), responding to ten statements on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). The average scores and standard deviations for each statement are summarised in [Table 6.2](#).

**Table 6.2:** Summary of System Usability Scale (SUS) Scores, showing the average score and standard deviation for each statement.

| Statement | Average Score | Standard Deviation |
|-----------|---------------|--------------------|
| S1        | 2.50          | 1.000              |
| S2        | 2.95          | 0.759              |
| S3        | 2.65          | 1.040              |
| S4        | 2.15          | 1.268              |
| S5        | 3.20          | 0.834              |
| S6        | 3.25          | 1.020              |
| S7        | 2.90          | 0.718              |
| S8        | 2.45          | 1.050              |
| S9        | 2.45          | 1.146              |
| S10       | 2.35          | 0.988              |

**S1** – I think that I would like to use this system frequently.

**S2** – I found the system unnecessarily complex.

**S3** – I thought the system was easy to use.

**S4** – I think that I would need the support of a technical person to be able to use this system.

**S5** – I found the various functions in this system were well integrated.

**S6** – I thought there was too much inconsistency in this system.

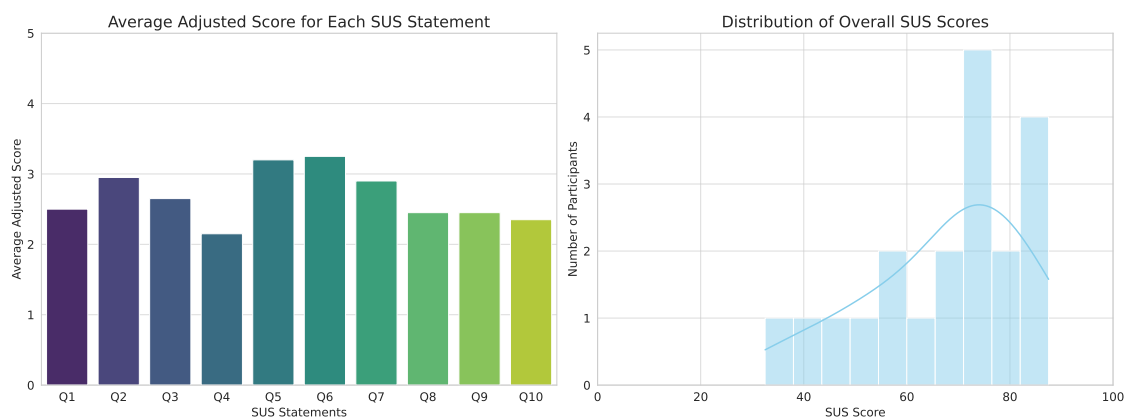
**S7** – I would imagine that most people would learn to use this system very quickly.

**S8** – I found the system very cumbersome to use.

**S9** – I felt very confident using the system.

**S10** – I needed to learn a lot of things before I could get going with this system.

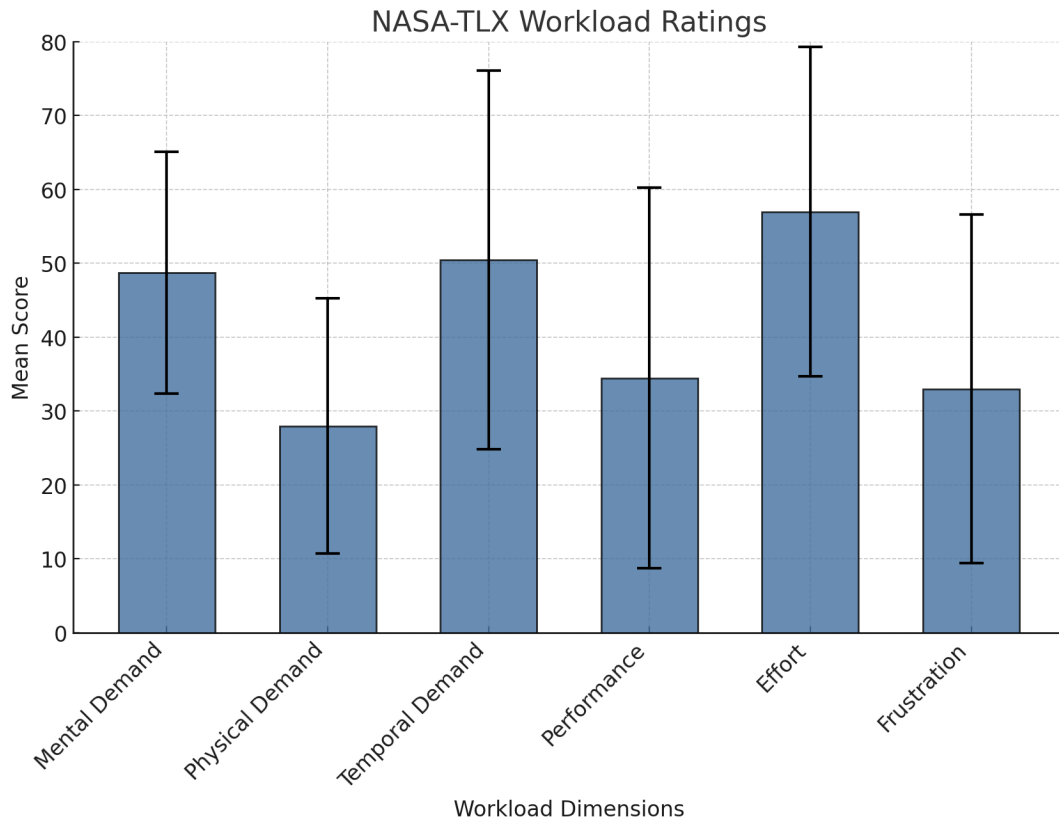
The highest-rated statement was S6 ( $M = 3.25$ ,  $SD = 1.020$ ), indicating that participants found the system somewhat inconsistent. Similarly, S5 ( $M = 3.20$ ,  $SD = 0.834$ ) suggested moderate agreement that the system's functions were well integrated. The lowest-rated statement, S4 ( $M = 2.15$ ,  $SD = 1.268$ ), reflected a perception that users did not require significant technical support to use the system. The overall SUS score – calculated based on adjusted participant responses – ranged between 32.5 and 87.5, with an average of 67.5 ( $SD = 18.7$ ). These results suggest a mixed user experience, with some participants finding the system intuitive while others encountered usability challenges. Notably, the lowest SUS score (32.5) was observed for one participant (P1), whereas the highest scores (87.5) were recorded for two participants (P2 and P20), illustrating the variation in perceptions.



**Figure 6.11:** System Usability Scale (SUS) Analysis. (a) Average adjusted scores for each of the ten SUS statements, highlighting the variability in participant responses to different usability aspects of the system. S5, related to the integration of functions, stands out with a notably high score. (b) Distribution of overall SUS scores across participants, showcasing a varied user experience with notable peaks around scores of 32.5 and 87.5.

Visualisations were constructed to offer a clearer representation of these findings (see Figure 6.11). A bar graph detailed the average adjusted score for each SUS statement, with S5 related to the integration of functions, which is particularly highlighted. A histogram further illustrated the distribution of overall SUS scores, indicating a varied user experience with peaks around scores of 32.5 and 87.5.

### 6.5.4 NASA-TLX Analysis



**Figure 6.12:** Distribution of NASA-TLX scores across six dimensions of perceived workload during an AR-based co-creation task. Each bar represents the mean score, with error bars indicating the standard deviation. This visualisation highlights variations in participants' experiences across different workload dimensions.

To assess the perceived workload of the participants during the task, a NASA Task Load Index (NASA-TLX) was used. This subjective workload assessment tool was used to gauge participants' experiences across six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Each participant rated their experience on a scale from 0 (very low) to 100 (very high) for each dimension.

#### Mental Demand

Among the 20 participants, the mean score for mental demand was found to be 48.75 (SD = 16.37), ranging from a minimum of 20 to a maximum of 75. It captures the cognitive and perceptual challenges encountered during the task.

### **Physical Demand**

For this metric, participants generally found the task to be not physically demanding. The mean score was 28.00 ( $SD = 17.27$ ), with scores ranging from 5 to 65. Given the AR nature of the task, it is noteworthy that the physical demand was relatively low, suggesting that the task was not overly taxing on the physical faculties of the participants.

### **Temporal Demand**

Temporal Demand had a mean score of 50.50 ( $SD = 25.64$ ). The scores for this dimension, which gauges perceived time pressure, ranged from 10 to 90. This wide range provides information on participants' varied experiences, some of whom may feel more rushed than others.

### **Performance**

The self-assessment of the performance of the participants produced a mean score of 34.50 ( $SD = 25.75$ ), with scores ranging from 5 to 85. This dimension evaluates how successful participants felt they were in completing the co-creation task.

### **Effort**

The level of effort required to achieve their level of performance produced the highest mean score of 57.00 ( $SD = 22.27$ ). Scores for this dimension ranged from 20 to 90, indicating a notable degree of variability in perceived effort among participants.

### **Frustration**

The levels of frustration among the participants had a mean score of 33.00 ( $SD = 23.59$ ), with individual scores ranging from 5 to 85. This dimension captures feelings of insecurity, discouragement, irritation, and stress versus feelings of contentment and relaxation during the task.

## **6.6 Discussion**

In examining the influence of providing multiple alternatives for the same task within co-located AR on collaboration styles such as F-formations, a critical aspect to examine is the interaction cost associated with each alternative. The interaction cost, which encapsulates the cognitive, physical, and temporal efforts necessary to engage

with these alternatives, could influence user preferences and, consequently, shape collaborative dynamics. For example, if an alternative presents a high cognitive load, such as complicated procedures or a steep learning curve, users might face challenges in collaboration, resulting in potential disruptions in collaboration dynamics. On the other hand, alternatives that reduce cognitive demands but perhaps increase physical or temporal costs could trigger different collaboration patterns as users negotiate the balance between mental and physical effort. Furthermore, the ability and speed with which users can switch between alternatives might be influenced by these interaction costs. If an alternative switch entails high interaction costs, users might be disinclined to change their method of task execution, thereby influencing the versatility and adaptability of collaborative processes.

### 6.6.1 Perceived Workload of CARIS

Our reference system was used in a co-creation task, where participants collaborated as architects to build a city. This presented varied workload experiences as captured by the NASA-TLX scores. This multidimensional assessment of subjective workload metrics offers valuable insights into the task's demands and the participants' perceived experiences.

Mental Demand scores - with a mean of 48.75 - suggest a moderate cognitive challenge for users and a usable system. With the SUS average score located at around 67.5, we see an above-average usability for the system. Given the collaborative nature of the task, as well as the AR interface, participants had to engage in both spatial reasoning and coordination with others in both the real world and the virtual plane. Although the scores did not reach the extreme high end of the scale, it indicates that there is moderate cognitive involvement. The low mean score for Physical Demand (28.00) shows that the system allows users to interact seamlessly without imposing undue physical strain. Although we observed that participants would often put themselves in uncomfortable positions to reach places on their buildings (for example, standing on a chair, reaching high on tiptoes, and crouching on the floor), the low physical demand is indicative of an interface that does not require excessive manipulation. We observe a wide range in Temporal Demand scores, with a mean of 50.50. Participants were given 10 minutes for each subtask, and some participants might have felt time pressure more acutely than others. This could be attributed to variations in individual working pace or perhaps specific subtasks being more time-consuming than anticipated, particularly toward the end of the study. Performance,

with its mean score of 34.50, highlights that participants generally felt that they were somewhat successful, but the wide score range suggests variability in perceived achievements. This could be a result of varying levels of familiarity with AR systems but also indicative of the learning curve for new interaction techniques provided or certain subtasks being more challenging than others. Effort was the most notable metric, with the highest mean score of 57.00. This suggests that the participants felt they had to exert themselves strongly. This could be due to the multifaceted nature of the tasks. Not only were participants engaged in a co-creation task within an AR environment, but they also had different tasks asked of them, requiring negotiation, planning, and execution. The Frustration level, with a mean of 33.00, was moderate, yet we observe its range is quite broad. This indicates certain pain points in the system or the task that caused frustration for some participants but not for others.

The NASA-TLX scores provide a nuanced understanding of the perceived workload during the co-creation task. While certain dimensions like Physical Demand were favourably low, areas such as Effort and the variability in Temporal Demand and Performance suggest avenues for further refinement in the system interface. Future iterations should focus on providing clearer guidance to improve usability and reduce perceived effort.

### 6.6.2 System Usability

The CARIS framework and, in turn, the reference system were conceived as a response to the challenges encountered in co-located AR collaborations. Its primary aim is to showcase the evolution of interaction techniques that buttress well-established co-located collaborative principles. Given this backdrop, the results from the SUS questionnaire offer a lens into *how participants perceived the system's usability*.

The statement “*I think that I would like to use this system frequently*” receiving an average score of 2.50 indicates that while some participants might find frequent use appealing, some participants harbour reservations. However, all participants currently do not work or have regular use for AR in their everyday lives, and as such, the moderate score here could be attributed to this, considering CARIS is designed for collaborative use in synchronous AR environments. One participant mentioned during the discussion that they “do not have a need for the system” (P19) so they wouldn’t use it, however, they think “some of the techniques would be useful for other tasks, like if I was presenting I could broadcast” (P5). Feedback from statements like “*I found the system unnecessarily complex*” and “*I thought the system was easy to*

*use*” provides insight into the system’s user-friendliness and intuitiveness. The study aimed to delve into collaborative dynamics in the AR realm, particularly through a co-creation task. When asked about this, the participants generally agreed that the system was “nice and fun to use” (P6), but the techniques “need getting used to as [they] forgot about them after a while” (P14) and that the “interface was easy to use, but there were times [they] forgot what techniques [they] could use” (P15). As just a reference system and exemplar implementation, the system is just used as a proof of concept, and one participant mentioned, “it reminds me a lot of Minecraft in some ways, but I think there’s a learning curve before it becomes second nature” (P12).

The combination of quantitative results and qualitative insights from the discussion suggests that there is an appreciation for the techniques. Participants’ comparisons to established platforms like Minecraft underscore its potential. However, the learning curve and the occasional forgetfulness about available techniques highlight areas for refinement. As AR continues to mature and find more commonplace applications, systems like our reference could pave the way for more intuitive and efficient collaborative experiences.

### 6.6.3 Context Switches & Immersion

The delicate balance between immersion and real-world engagement is central to the design and usability of AR systems. Our analysis of context switches offers unique insights into this equilibrium. A context switch, as defined within our study, represents a momentary lapse in the immersive experience, a moment when the user temporarily disengages from the AR environment. Such disruptions serve as markers, highlighting the points of friction between the AR experience and the demands of the physical environment or the user’s cognitive load.

Our results indicate variable immersion rates among participants. While some maintained a high level of engagement, occasionally breaking away, others exhibited more frequent interruptions. The reasons for such context switches could be manifold. External stimuli, interface complexities, cognitive overload, or even simple curiosity could draw users away from the immersive experience. Momentary context switches, those lasting just a second, may hint at rapid distractions or brief uncertainties in the user’s interaction with the system. On the other hand, extended interruptions might be indicative of more profound challenges, possibly related to task comprehension, usability of the interface, or real-world distractions. The patterns of these changes among participants offer another layer of understanding. Periods with a high context



switching frequency, exceeding the 2-minute rolling average, might denote universally challenging or ambiguous sections of the task. Such collective disruptions could be areas where the AR interface or task design needs refining. Alternatively, there could also be moments when the real-world environment presented challenges, necessitating a break from AR immersion. It is crucial to note that while uninterrupted immersion is often lauded as an ideal state for AR experiences, consistent engagement without breaks might not always be desirable. Extended, unbroken immersion could lead to cognitive fatigue or detachment from necessary real-world interactions. Context switches, in moderate amounts, might serve as essential cognitive “breathers”, allowing users to recalibrate, reorient, or even just rest momentarily.

Understanding the dynamics of context switches and their implications for immersion can offer actionable insights to AR system designers. A nuanced balance between immersive engagement and real-world reconnection is crucial to creating AR experiences that are not only captivating but also cognitively sustainable and contextually aware.

## 6.7 Conclusion

In this work, we have demonstrated how proxemics theory can be used as a conceptual lens articulating collaborative AR systems and their properties. Through a design space exploration, we explored how interaction techniques for handheld multi-user AR applications can be designed to either align with or go beyond reality. This work has highlighted how there can be a many-to-many relationship between different collaboration styles and F-formations. Based on this exploration, we present reality and beyond as a continuum for designing co-located collaborative AR systems, illustrating how we can design AR systems to extend enacted proxemics or support flexibility in interpersonal distance and orientation. Finally, we highlight how proxemics can be used to understand the tradeoffs in the design of AR in relation to perceptual, deixis, and control proxemics.

# 7

## Discussion

### Contents

---

|            |                                                                 |            |
|------------|-----------------------------------------------------------------|------------|
| <b>7.1</b> | <b>Revisiting Research Questions . . . . .</b>                  | <b>163</b> |
| <b>7.2</b> | <b>Broader Implications and Relation to Prior Work . . . .</b>  | <b>166</b> |
| <b>7.3</b> | <b>Implications for Future Design in Collaborative Settings</b> | <b>175</b> |
| <b>7.4</b> | <b>Summary and Key Takeaways . . . . .</b>                      | <b>184</b> |
| <b>7.5</b> | <b>Future Work . . . . .</b>                                    | <b>186</b> |

---

In this chapter, the contributions of this thesis to the broader field of HCI are summarised, including progress toward answering the three research questions introduced in [Chapter 1](#). The chapter also reviews the key contributions and implications, discusses directions for future work in [Section 7.5](#), and concludes in [Section 2.4](#).

This thesis explored how Augmented Reality (AR) has emerged to enhance co-located collaborative experiences by integrating innovative interaction techniques and applying proxemics theory as a new conceptual lens to cross-device collaborative handheld mobile AR. The previous chapters have examined both the theoretical and practical aspects of the effects of handheld AR on co-located collaborations. The work contributes to the field of HCI, and in particular to the field of CSCW, in the following ways:

1. In [Chapter 3](#), the current dynamics of handheld mobile augmented reality technology in co-located collaborative environments were explored, providing a comprehensive analysis of how these interfaces presently influence group

interactions. By conducting an empirical study focused on the interaction patterns and attention within group tasks, the chapter maps out a nuanced understanding of collaborative challenges posed by current AR interfaces but also points to under-explored areas, such as the effect AR complexity has on co-located group dynamics. The observed high frequency of context switches in mobile AR, which contribute to mental load and reduced collaboration, suggests that further in-depth analysis is warranted. The insights gained from this chapter contribute substantially to the design of interaction techniques and lay the foundations for a focused research agenda in augmenting co-located collaborations through handheld mobile AR.

2. In [Chapter 4](#), the impact of device configurations and sizes on mobile augmented reality (AR)-mediated collaboration is thoroughly explored, providing valuable insights into the subtleties of cross-device collaborative work. Through a structured study, this chapter examines how different mobile device sizes and setups influence collaboration strategies, behaviours, and effectiveness in AR settings. The findings highlight the differences in how participants approach and interact with AR tasks depending on their device's form factor and setup. This exploration contributes to a deeper understanding of the physical and cognitive demands of handheld mobile AR. It highlights several under-explored areas, such as the role of device symmetry and screen size in collaborative efficiency. This chapter reveals that mobile device configurations, including size and symmetry, influence collaboration strategies, behaviours, and efficacy in co-located AR settings. These insights advance the design and application of interaction techniques in mobile AR, setting a focus on device attributes within AR for enhancing user experience in cross-device computing within co-located collaborative environments.
3. In [Chapter 5](#), the application of proxemics theory is introduced as a conceptual lens for understanding and designing co-located collaborative AR systems. Proxemics, in this case, provides a socio-spatial perspective, allowing designers to consider how AR interactions can align with or go beyond real-world spatial configurations. Through this lens, the chapter explores how AR systems shape interpersonal dynamics by supporting a range of collaborative coupling styles and physical proximities. By detailing design principles and interaction techniques, this chapter establishes a socio-spatial vocabulary that aids in

understanding and creating AR systems tailored for enhanced co-located collaboration. This conceptual framework encourages AR applications to move from isolated user interactions to cohesive, context-aware multi-user experiences, strengthening the field's approach to co-located collaborative AR.

4. In [Chapter 6](#), we present a novel conceptual framework that integrates AR with collaborative work environments. Focusing on the seamless merging of virtual information and the physical world, the framework addresses spatial and contextual dynamics often overlooked in traditional AR systems, thereby advancing collaborative dynamics in co-located settings. By extending interaction paradigms beyond individual usage to support complex, intuitive group interactions, this chapter not only broadens the application of AR in collaborative contexts but also sets new standards for immersive, user-centric collaborative technologies. This approach not only enriches HCI and CSCW research by providing a practical example of innovative system architecture but also pushes the boundaries of how digital and physical environments can interact to enhance collective engagement and productivity.

In the next section, we use the results of these discussions to revisit the research questions set at the beginning of this thesis.

## 7.1 Revisiting Research Questions

This thesis explores how current mobile AR applications affect co-located collaborations and introduces ways to better these collaborations. In this section, I revisit the research questions from [Chapter 1](#) and summarise the findings from the corresponding studies for each question. The superordinate research question addressed in this thesis is as follows.

*How do mobile handheld Augmented Reality interfaces impact co-located collaborative work and inform the design of future Augmented Reality systems?*

To help address this question, we formed three sub-research questions (RQ1–RQ3). As an initial step in setting the scene for the rest of the thesis work, in [Chapter 2](#), we looked at the transformative capabilities of AR in enhancing collaborative environments, mainly focusing on its integration within Human-Computer Interaction. AR's ability to merge digital information with the physical world

enhances the user's sensory perception and transforms interaction dynamics in shared spaces. Despite its advanced integration, current AR applications often lack a comprehensive understanding of how AR can dynamically facilitate and enhance co-located collaborative interactions, mainly through the nuances of spatial and social contexts. As such, the need has emerged to conceptually and practically explore how AR technologies can be designed to support and enrich collaborative experiences in various real-world settings effectively. However, we first needed an understanding of how groups interact with a standard, basic AR system, leading to the formulation of my first sub-research question:

**RQ1: How do current group dynamics and practices affect co-located collaborations enhanced by handheld mobile AR applications?**

In [Chapter 3](#), the investigation centred on the influence of mobile handheld AR technology on co-located group interactions. We mainly focused on collaborative practices mediated by these mobile handheld devices. We conducted a detailed study examining the dynamics of collaboration between groups using mobile handheld AR to interact with virtual models in a real-world setting. This research found that mobile AR could induce significant context switching and influence group collaboration styles. Participants often switched between interacting with the AR environment and engaging directly with other group members. These switches were influenced by the complexity of the tasks and the AR models used. We also found that mobile device handling often distracted participants from smoother collaborative interactions. The necessity of constantly managing the device interface while participating in group tasks added a layer of complexity to the collaborative process. We also saw that, as the complexity of AR models increased, there was a noticeable change in how groups collaborated. Simpler models facilitated more dynamic and engaged collaboration, while complex models tended to lead participants to focus more on their individual interactions with the AR system, reducing overall group discussion and interaction. Based on these observations, we proposed five design recommendations to improve collaborative experiences in AR. These included ways of improving the user interface to minimise the introduction of new interaction techniques that minimise the need to context switch and increase immersion.

With a better understanding of the People aspect of co-located mobile handheld AR collaborations, I shifted my attention to the devices. With the changes in model size and participants' comments mentioning a larger device size may help, we wanted to look into how this might affect user interactions, leading to the second research question:

**RQ2: How does device size and configuration affect the facilitation of co-located mobile AR collaboration?**

In [Chapter 4](#), we investigated the effects of mobile device size and configuration on collaborative strategies and behaviours in a co-located AR setting. The study examines explicitly the interaction dynamics between dyads using different combinations of mobile phones and tablets. We explore how these variables influence task performance, communication patterns, and the overall effectiveness of collaboration. We saw that different device configurations affected how participants interacted with the AR environment and with each other. Larger screens facilitated better shared viewing and interaction, while smaller screens tended to promote more personal engagement with the device. This difference in screen size and the resulting visibility of the information affected how the participants coordinated and shared tasks. The configuration of the devices also influenced the collaboration strategies adopted by the participants. When both participants used devices of similar size, there was a more balanced interaction and task sharing. However, asymmetrical setups often led to one participant taking a lead role, especially when manipulating complex 3D models in the AR space. The study also revealed that device configurations impacted task performance, with different configurations favouring certain types of tasks over others. For example, tasks that required detailed examination of 3D models were easier to perform on larger tablets, providing better visibility and less occlusion. Communication patterns varied with device configuration, and participants using the same type of device communicated more fluidly and coordinated their actions more effectively than mixed-device pairs. This was attributed to a more consistent user experience and the ease of sharing visual content.

Finally, having investigated two major components of co-located handheld AR collaborations – People and Interaction Devices – taking principles identified in both studies, we formed the final sub-research question:

**RQ3: How can we operationalise design principles into a design space that enhances co-located mobile AR collaborations?**

In [Chapter 5](#) and [Chapter 6](#), this thesis delves into operationalising design principles into a structured design space aimed at enhancing co-located mobile AR collaborations. Building on the foundational insights from CollabAR and DeviceAR chapters, this section introduces a comprehensive framework that integrates the dynamic and proxemic interactions essential to collaborative handheld AR systems.

By applying the conceptual lens of proxemics, [Chapter 5](#) articulates how socio-spatial considerations, such as perceptual, deixis, and control proxemics, shape user interactions in co-located settings. This lens frames how collaboration styles intersect with device configurations, extending the coupling styles framework to a broader, many-to-many relationship between F-formations and AR-mediated collaboration styles.

Expanding on these theoretical insights, [Chapter 6](#) illustrates the practical application of these design principles through an exemplar system that showcases how interaction techniques can span from *aligning with reality* to *going beyond reality*. This continuum enables a flexible, proxemic-aware AR design space that can adapt to the shifting spatial needs of users engaged in collaborative tasks. For instance, Remote Sketch and Perspective Broadcasting allow for fluid shifts between closely and loosely coupled interactions, supporting synchronous and asynchronous communication among co-located users.

The chapters address how specific design principles can transform into a tangible, flexible design space that prioritises user engagement and interaction fluidity in shared augmented spaces through this dual approach. By offering structured guidelines that leverage socio-spatial awareness, the framework supports the creation of handheld AR systems that actively enhance collaborative tasks in diverse settings, setting a foundation for more intuitive, effective, and adaptable AR experiences.

## 7.2 Broader Implications and Relation to Prior Work

This section explores the broader implications of the study's findings on co-located collaborative handheld AR and situates them in the context of prior research. The discussion spans proxemics, mobile handheld AR, cross-device collaboration, group collaboration strategies, and methodologies, highlighting both the theoretical contributions and the practical applications of my work.

### 7.2.1 Proxemics in Collaborative AR

Applying proxemics in collaborative AR introduces a nuanced understanding of socio-spatial configurations and interpersonal dynamics within shared digital environments. This study builds on foundational insights from Hall's proxemic theory [75], positioning proxemics as a core analytical lens for designing and evaluating handheld

AR systems. By contextualising proxemic interaction through perceptual, deixis, and control dimensions (as introduced in [Chapter 5](#)), we identify practical implications for enhancing collaborative user experiences in AR settings.

### **Perceptual Proxemics and Shared AR Spaces**

Perceptual proxemics, as defined in this study, address the spatial relationships that govern user access to shared AR content. Prior research by Mentis et al. [131] highlighted that shared perceptual access is critical for effective collaboration, particularly when users interact with complex, information-dense AR objects. This study extends these insights by demonstrating that flexible perceptual configurations, such as the *Freeze Frame* and *Perspective Broadcasting* techniques, allow users to adaptively manage their viewpoints relative to collaborative content. By enabling users to “freeze” their perspective or broadcast views across devices, these techniques support fluid transitions between individual and shared focus within an AR environment. This flexibility in viewpoint configuration not only enhances user awareness of collaborative content but also reduces physical constraints on interpersonal positioning, a key consideration identified by Sereno et al. for co-located collaborative AR [188].

### **Deixis Proxemics for Effective Spatial Referencing**

The concept of deixis proxemics, where users communicate through spatial referencing, is particularly relevant to collaborative handheld AR settings. Prior findings, such as those by Billinghurst et al. on deictic referencing in virtual environments [13], emphasised the value of natural gesturing in AR. Techniques such as *Spatial Reference* and *Remote Sketch* expand this notion by allowing users to share and annotate perspectives across devices, thereby facilitating precise communication without necessitating physical co-location. This approach aligns with recommendations from [Chapter 3](#), where enhancing deictic flexibility was identified as essential for bridging the gap between real and virtual interaction spaces. Incorporating remote sketching capabilities offers an innovative solution for maintaining alignment across diverse F-formation configurations, thereby supporting both tightly and loosely coupled collaborative modes as described by Tang et al. [200].



### Control Proxemics and User Autonomy

Control proxemics within collaborative AR settings address the allocation and regulation of user control over shared and individual AR objects. Including a *Privacy Fence* technique exemplifies how proxemic-informed design can afford users autonomous spaces within a shared AR environment. This concept is rooted in the idea of “territories of interaction,” as initially discussed by Kendon in relation to spatial interaction boundaries [104]. By allowing users to establish personal territories within a collaborative space, this feature acknowledges and adapts to the socio-spatial nuances of co-located work, providing both shared and individual engagement options. This implementation mirrors findings from Grønbæk et al. on proxemic transitions in shared interfaces [65], enabling more seamless transitions between individual and group interactions while respecting personal space boundaries within a collective environment.

### Proxemics Beyond Reality and Cross-Device Flexibility

Going beyond reality through cross-device interactions adds an additional layer of flexibility to proxemic configurations in AR. As this study explored, techniques such as *Remote Peeking* and *Perspective Broadcasting* enable users to seamlessly transition between private and shared views across multiple devices, irrespective of their physical positioning. These techniques resonate with the cross-device proximity concepts presented by Marquardt et al. [126], whose research emphasises the importance of flexible device-sharing mechanisms in collaborative spaces. By leveraging networking capabilities and adaptive perspective-sharing mechanisms, proxemics can support interaction designs that go “beyond reality,” enabling users to navigate and interact with AR content in ways that are not spatially or physically constrained. This concept challenges the one-to-one correspondence between interpersonal distance and collaborative coupling styles outlined by Tang et al. [200], proposing instead a many-to-many relationship where various proxemic and collaborative configurations coexist.

#### 7.2.2 Implications for Mobile Handheld AR

The findings from this thesis accentuate the unique role of handheld AR in shaping co-located collaborative experiences, providing insights into how mobile AR’s form factor, interaction styles, and technological affordances influence group interaction and task performance. Mobile handheld AR enables new forms of spatial and contextually

aware collaboration, facilitating interaction with digital objects in shared physical spaces. However, findings from studies within both [Chapter 3](#) and [Chapter 4](#) reveal that this form of AR can also impose substantial cognitive and physical demands on users, potentially limiting its collaborative effectiveness. Here, we discuss the broader implications of these findings for the development and implementation of handheld AR systems in collaborative settings, focusing on three main areas: cognitive load, group dynamics, and device handling.

### **Cognitive Load and Task Engagement**

A recurring theme throughout the chapters was the high cognitive demand associated with handheld AR interactions, especially when users were required to engage in continuous context-switching between physical and digital realms and encountered more complex virtual models. As indicated in the CollabAR study ([Chapter 3](#)), tasks involving complex virtual models frequently resulted in high mental load, with users expressing frustration and needing to redirect focus between the AR interface and other group members. This finding highlights a key challenge in handheld AR. While AR can support collaborative tasks by overlaying digital information within physical environments, the cognitive overhead associated with maintaining spatial and object-based awareness can detract from user engagement and collaboration quality. Previous research in collaborative AR [[12](#), [81](#), [106](#)] has similarly pointed to the need for AR interfaces that better manage cognitive demands, as cognitive load influences both task efficacy and group interaction [[46](#)]. As handheld AR applications are increasingly integrated into educational, professional, and social environments, reducing cognitive load through improved stability, visual clarity, and interaction flow may help to maintain collaborative momentum and reduce fatigue.

### **Impact on Group Dynamics**

The thesis findings suggest that handheld AR, by requiring individuals to hold and manage their devices, often leads to more individualised interaction patterns, which can influence group cohesion. For example, in the DeviceAR study ([Chapter 4](#)), participants using asymmetric device configurations (such as phone + tablet) exhibited behaviours that prioritised individual device interaction over collaborative engagement, aligning with findings in cross-device systems, where asymmetry can create task division [[127](#)]. Additionally, complex models that necessitated high levels of attention to AR content tended to limit verbal and face-to-face interaction among

group members. This aligns with observations in tabletop interfaces, where personal device usage in collaborative settings can dilute shared focus and increase individual tasking [25, 146]. For future handheld AR systems, designers may consider collaborative features that support simultaneous, multi-user interactions while minimising device-related barriers, such as dynamic view-sharing or simplified gesture-based controls that facilitate fluid interaction between physical and digital environments.

### Device Handling and Physical Demand

Physical demand was a critical factor in shaping user experience across the *CollabAR* and *DeviceAR* studies (Chapter 3 and Chapter 4), as participants frequently reported the need to hold, move, and adjust their devices to maintain visual clarity of virtual models. This aspect was particularly noticeable with larger, heavier devices, which, while offering improved visibility, also increased fatigue and impeded task fluidity. The challenge of continuously holding up handheld devices aligns with prior research highlighting the “gorilla arm” effect [83] and suggests a need for device-handling support in future AR designs. The inherent physical demands of handheld AR could be mitigated by incorporating ergonomically optimised grips or implementing augmented stability features. This would likely increase comfort, reduce cognitive interference due to physical discomfort, and thus enhance collaborative effectiveness. Further, the findings suggest that future research could explore alternative forms of AR display, such as lighter or head-mounted options, to reduce the continuous handling burden, particularly for longer or more complex collaborative tasks.

Overall, this thesis illustrates the potential of handheld AR to foster innovative, collaborative experiences in co-located settings. At the same time, the findings emphasise the need for user-centred improvements that address cognitive load, social dynamics, and physical demands. These broader implications contribute to a growing body of research aiming to optimise AR interfaces for collaborative work, ultimately guiding the development of handheld AR systems that are more intuitive, engaging, and effective for co-located group interaction.

### 7.2.3 Implications for Cross-Device Collaboration

The findings from this thesis highlight implications for the development of cross-device collaboration within handheld Augmented Reality (AR). The CARIS framework, building on prior conceptualisations from the *CollabAR* and *DeviceAR* chapters, illustrates the potential for integrating multiple devices in a unified AR collaboration

space [26, 68]. By incorporating proxemic principles, such as perception, deixis, and control, CARIS aligns with previous research showing that spatial dynamics between devices and users can optimise collaboration within co-located environments [7, 63]. This approach emphasises the value of cross-device interactions that allow AR content to extend seamlessly across devices, facilitating collaborative engagement while minimising the interruption of immersion.

A key implication of cross-device collaboration in handheld AR, as supported by the findings in the DeviceAR chapter, is the necessity of fluid, adaptable interfaces that respond to users' proxemic shifts. Cross-device applications in handheld AR often require users to navigate complex digital-physical interactions, such as transitioning between individual tasks on personal devices and group tasks involving shared AR content. The findings in CollabAR suggest that by enabling proxemic-aware transitions, such as moving from individual to shared perspectives, AR systems can create a more cohesive and immersive experience [67, 200]. However, these transitions also present challenges in managing attention and context switching, as participants engage with both personal and shared spaces, an issue observed in the study of context switching within collaborative AR tasks.

Furthermore, CARIS's inclusion of Interaction Portals as dynamic, proxemic-aware entry points into the AR environment suggests that the flexibility of AR spaces can be enhanced by allowing for diverse interaction modalities across devices [7]. As seen in the architectural design study, Interaction Portals facilitate seamless collaborative transitions by adapting to proxemic changes, thus enhancing deixis and shared control over digital content. This adaptability is particularly relevant for cross-device AR systems that may support dynamic interactions across different device types, such as handhelds, head-mounted displays, or tabletop devices [68, 89]. These findings extend the theoretical implications of proxemics by showing that spatially aware interaction techniques can foster collaborative engagement and accommodate individual user preferences, which has the potential to reduce cognitive load in multi-device AR scenarios.

Additionally, integrating proxemic principles into cross-device AR collaboration highlights the importance of supporting multiple interaction alternatives, as suggested by CARIS's component-based structure. This flexibility allows AR systems to meet varied collaborative needs and styles, aligning with previous research that advocates for adaptable interaction frameworks in AR [26]. The implications of these findings suggest that AR systems designed with cross-device collaboration in mind must go

beyond static interfaces, offering dynamic, proxemically aware options that enable users to navigate between devices and modes of interaction effortlessly. Such designs not only support more cohesive collaboration but also align with the overarching goal of reducing device-centric limitations within AR spaces [33, 68].

The CARIS framework’s application of proxemics to cross-device AR systems presents actionable insights for future AR design. By facilitating proxemic-aware transitions and supporting multi-device interactions, CARIS offers a pathway towards more flexible, collaborative AR environments. These findings show the potential of proxemic integration to address challenges of immersion and cognitive load, setting the stage for a new wave of AR systems that fully leverage the collaborative potential of cross-device interactions.

#### 7.2.4 Implications for Group Collaboration Strategies

The research conducted through the *CollabAR* and *DeviceAR* studies illuminate how handheld AR interfaces shape group collaboration strategies in co-located settings, with findings indicating both enabling and constraining effects on collaboration. Specifically, the analysis reveals how device configurations, model complexity, and the inherent limitations of handheld AR drive users to adopt distinct collaboration strategies, such as *Independent Work*, *Turn-Taking*, *Synchronous Work*, and *Divided Work*. This section discusses the broader implications of these observed strategies, particularly in relation to group dynamics, task management, and the evolving expectations of collaborative AR systems.

##### Influence of Device Configuration on Strategy Selection

Device configuration significantly impacts the strategies that groups adopt when collaborating in AR. As demonstrated in the *DeviceAR* study, symmetric device configurations (e.g., phone + phone or tablet + tablet) tended to facilitate more evenly distributed participation, enabling groups to adopt synchronous work strategies that allowed both users to interact with the AR content simultaneously. In contrast, asymmetric configurations, such as tablet + phone, often led to a “leader-follower” dynamic where one participant took on a more active role in interacting with the AR content, while the other participant observed or supported. This pattern aligns with findings in cross-device collaboration [26, 127], where device asymmetry can create an unintentional hierarchy, leading participants to divide tasks or take turns in managing interactions. This thesis suggests that future AR systems should consider how device symmetry or asymmetry can affect group dynamics, especially in contexts where equitable collaboration is essential.

### Task Complexity and Adaptive Collaboration Strategies

The research also emphasises how the complexity of virtual models influences collaboration strategies, with groups frequently adapting their approach based on task demands. In the *CollabAR* study, simpler models with lower cognitive demands encouraged more synchronous and discussion-oriented strategies, allowing participants to adopt synchronous or shared-view approaches that involved simultaneous inspection and rotation of virtual objects. However, as task complexity increased, participants commonly switched to independent or turn-based strategies, especially in multi-device configurations, to manage the cognitive load associated with complex model features. This finding aligns with previous work showing that, in collaborative settings, high task complexity often necessitates more individualised approaches to minimise cognitive overload [46, 184]. The implications here suggest that AR interfaces could benefit from adaptive interaction modes that respond to task complexity, enabling users to fluidly transition between independent, turn-based, and synchronous modes as task demands fluctuate.

### Group Dynamics and Collaboration Equity

The thesis findings highlight a critical implication for collaboration equity in handheld AR: the format and accessibility of interaction with virtual content can impact how engaged each group member feels. For instance, the *DeviceAR* findings indicate that larger devices, such as tablets, afforded a more equitable view and facilitated synchronous engagement. However, smaller devices or those with limited visibility created instances where participants had to rely on turn-taking, often reducing the overall sense of collaboration equity within the group. This supports existing research on co-located collaborative devices that suggests device size and interaction mode influence group balance and communication richness [106, 120]. Future AR applications could integrate features that promote shared control, such as dynamic resizing, screen sharing, or multi-view capabilities, to ensure that all participants have equal access to AR content regardless of device size or configuration.

#### 7.2.5 Implications for Methodologies

The exploration of collaborative Augmented Reality (AR) in this thesis highlights key methodological implications for studying multi-device, proxemic-aware systems. Drawing on approaches utilised in the *CollabAR*, *DeviceAR*, *Conceptual Lens*,

and Framework chapters, this research adopts a diverse methodological foundation, encompassing empirical observational studies, empirical design evaluation studies, and research-through-design methods. This methodological plurality reflects the nuanced requirements of collaborative AR studies. It highlights the need for adaptable, mixed-method approaches to capture the complex interplay between users, devices, and spatial contexts.

A primary methodological insight from this research is the importance of *empirical observational studies* in establishing a baseline understanding of collaborative dynamics. For example, the observational analysis conducted in [Chapter 3](#) provided foundational insights into user behaviours and interaction patterns as they engage with AR in a co-located, device-centric setting without control groups. By observing participants under uniform conditions, the empirical approach established patterns that could inform future comparative studies. In a similar vein, the empirical analyses in [Chapter 4](#) further highlighted how device configurations within collaborative AR can influence interaction outcomes. These empirical studies collectively reinforce the value of initial observational methods, providing a baseline that contextualises the impact of subsequent interventions and design adjustments on user experience.

The *research-through-design* approach undertaken in [Chapter 5](#) represents another methodological insight, especially for creating and testing principles-based frameworks in collaborative AR. The conceptual lens developed here grounded the CARIS framework in a proxemic-driven understanding of socio-spatial dynamics in collaborative settings. By iteratively refining the CARIS framework based on design principles grounded in this thesis, the research-through-design approach allowed for the synthesis of empirical findings and theoretical principles into a structured, adaptable toolkit for cross-device AR interactions. This iterative design process not only established CARIS as a conceptual model but validated it as a practical tool for enhancing user engagement in AR environments. This demonstrates the capacity of research-through-design to bridge theory and practice effectively in multi-device AR systems by integrating proxemics directly into the design space.

Building upon the principles and insights from research-through-design, the *empirical design evaluation study* in [Chapter 6](#) provided practical demonstrations of the CARIS framework through the implementation of an exemplar system. This small-scale, empirical study allowed the framework's principles to be tested under real-world collaborative conditions, capturing data on user immersion, context switching, and interaction with AR features. This design evaluation demonstrated



the operational effectiveness of CARIS and highlights the importance of iterative, empirically-grounded testing in assessing the framework's ability to support proxemic-aware, cross-device AR collaboration. This methodology emphasises how empirical design evaluation is essential for validating design frameworks, enabling researchers to explore how theoretical constructs perform in practice.

The diversity of methods adopted in this thesis highlights the broader implication that AR studies benefit from a *mixed-method approach* that captures both quantitative and qualitative aspects of collaboration. Given the multi-dimensional nature of collaborative AR, quantitative measures alone may overlook the nuanced ways users adapt to new interaction techniques or the subjective factors influencing user satisfaction and engagement. For example, quantitative data on context switching provided insights into user immersion, while qualitative feedback on interface usability offered a deeper understanding of user experience aspects that metrics alone might miss [135]. This mixed-method approach allows for a richer, more holistic view of user experience, crucial for the development of functional, user-centred AR systems.

In summary, this thesis demonstrates that studying collaborative AR requires a flexible, iterative methodology that accommodates the complexity of multi-device, proxemic-aware interactions. By integrating empirical observational and design evaluation studies alongside research-through-design, this research offers a comprehensive methodological framework that advances our understanding of collaborative AR while establishing a foundation for future studies. This multi-method approach highlights the importance of capturing diverse data types and balancing theoretical development with practical experimentation, laying the initial foundations for more robust and user-responsive AR design methodologies.

### 7.3 Implications for Future Design in Collaborative Settings

The following section details implications for the future design of AR systems in collaborative settings, broken down into technical, empirical, and design-specific recommendations. These considerations stem from the study's findings. They are informed by prior research, emphasising technical constraints, empirical observations, and innovative design ideas that can facilitate more effective and engaging collaborative AR experiences.



### 7.3.1 Technical Implications

The technical aspects of AR system design are critical in shaping user experience and effectiveness in collaborative contexts. Our study identified several technical challenges, such as device handling, context-switching, and latency, that significantly impact the fluidity of collaboration. These findings reinforce prior work by Henrysson et al., who identified the importance of stable and synchronised views for collaborative AR tasks, especially on handheld devices [81].

#### Device Handling and Stability

Our results highlighted the necessity for enhanced device handling solutions in handheld AR. Users frequently encountered interruptions due to the need to continually adjust their devices to maintain stable views, which aligns with observations by Norman regarding the cognitive strain imposed by device manipulation [147]. Future AR systems should implement stabilisation features, such as “view-clutching” mechanisms, that allow users to lock their view temporarily to mitigate these interruptions.

#### Low-Latency Cross-Device Synchronisation

Low-latency communication is crucial for maintaining synchronised views across devices during collaborative tasks. Similar to findings by Gutwin and Greenberg, who observed the importance of workspace awareness in collaborative systems [72], our study found that delays in content synchronisation disrupted collaboration. Future AR designs should employ robust communication architectures that support real-time data exchange even in low-bandwidth environments, ensuring that content remains synchronised across devices to minimise disruptions.

#### Scalable Network Infrastructure

To support the potential scalability of AR systems in real-world deployments, systems must be designed to handle variable group sizes and network conditions. Building on recommendations by Sereno et al. regarding the flexibility of AR in group settings [188], future systems should incorporate network architectures that can dynamically adapt to larger user groups without compromising performance. This would involve modular design principles and adjustable data handling protocols to optimise resource allocation based on group size and network quality.

### 7.3.2 Empirical Implications

Our findings also present empirical implications for understanding user behaviour, collaborative strategies, and interaction patterns in AR settings. By analysing these factors, future research can provide deeper insights into user needs and inform the design of AR systems that better support co-located collaboration.

#### User Behavioural Patterns

As shown across this thesis, user behaviour in collaborative AR settings is shaped by a combination of proxemic dynamics and cognitive demands, which arise from the complexities of collaborative tasks. Proxemic awareness emerged as a central factor, influencing how users positioned themselves relative to AR content and one another. Consistent with Mentis et al.'s findings on proxemics as a facilitator of natural collaboration [131], this research observed that users frequently adjusted their spatial orientation and distance to better interact with shared AR content. For instance, the Freeze Frame interaction technique enabled users to capture and share specific AR perspectives, allowing other collaborators to view frozen snapshots without needing to be physically close. This technique proved especially helpful in maintaining focus in complex tasks, as users could shift between real-time and static views, thereby enhancing group awareness without disrupting their proxemic arrangement.

Cognitive load was also a key consideration, with context-switching metrics providing insights into engagement and the demands of AR tasks. These metrics allowed us to quantify the impact of various interaction techniques, offering a practical way to assess how collaboration styles, device handling, and immersion levels influenced user engagement. For example, the Remote Sketch technique supported context-switching by allowing users to annotate a frozen frame and broadcast it to the group, enabling clear, deictic references without requiring all collaborators to be in the same viewing angle or physical space. This reduced the need for physical repositioning, helping users manage their cognitive load by minimising unnecessary context switches and focusing on relevant task information.

Future AR systems can benefit from refining these proxemic-responsive metrics to provide a deeper understanding of how proximity-based cues and task-switching interactions jointly influence cognitive load, interaction quality, and overall user experience. By developing frameworks that integrate proxemic cues with context-switch measures, designers can better support adaptable, user-centric collaborative AR environments that seamlessly align with the spatial and cognitive needs of co-located users.

### User-Centric Evaluation Metrics

This thesis identifies the diversity of user roles, collaboration strategies, and spatial interactions in collaborative AR settings, highlighting the importance of user-centric evaluation metrics that capture the nuances of individual experiences, task demands, and collaborative contributions. Consistent with Billingham and Henrysson's recommendation for tailored metrics in collaborative AR [19], this research reinforces the ongoing need for metrics that reflect subjective workload assessments, role-based task performance, and engagement levels to provide a holistic overview of user experience and collaboration quality.

These user-centric metrics should also consider role-based interaction patterns, proxemic preferences, and task complexity, enabling designers to examine how individuals in primary or supportive roles navigate collaborative AR environments. For instance, extending proxemic theory in AR design can clarify how users adjust spatial configurations and interpersonal distances to optimise engagement, as observed in applications like Blocks [71], where users' relative positioning affects shared understanding and ease of collaboration. Furthermore, incorporating context-switching measures can help quantify cognitive load shifts that arise during complex or overlapping tasks, allowing for insights into how proxemic and cognitive dynamics intersect to impact user performance and satisfaction.

Ultimately, these user-centric metrics empower designers and developers of collaborative AR spaces to create adaptive, responsive AR systems that meet individual needs in collaborative environments. By focusing on role-specific engagement, spatial orientation, and cognitive demands, future evaluations can guide AR designs that flexibly cater to both collective and individual user experiences, thereby enhancing collaborative efficacy and user satisfaction across varied AR contexts.

#### 7.3.3 Design Ideas

The design recommendations from this study suggest several innovative ideas that could improve the usability and collaborative potential of future AR systems. These design ideas draw from empirical observations and align with the principles and theories discussed in prior research.

### **Adaptable Collaboration Styles**

Our findings indicate that users benefit from the ability to switch fluidly between different collaboration styles, ranging from tightly coupled to loosely coupled interactions. This supports the perspective of Billingham et al., who highlighted the need for flexible collaboration modes in AR [13]. To accommodate this, future AR designs could incorporate dynamic collaboration modes that adjust based on user activity and context. For instance, features like synchronised views for tightly coupled tasks and isolated perspectives for individual work could allow users to engage in varied collaboration styles as needed.

### **Intuitive, Multi-Modal Interfaces**

The user interface plays a crucial role in determining how easily users can perform collaborative tasks in AR environments. In line with Norman's concept of affordances [147], our study found that intuitive visual cues and straightforward interaction methods reduced cognitive load and improved task performance. Future systems should leverage multi-modal interfaces that integrate visual, auditory, and tactile feedback to accommodate diverse user preferences and enhance task efficiency. Incorporating gestures, voice commands, and haptic feedback could further enrich the interaction experience and reduce reliance on manual device handling.

### **Role-Specific Control Mechanisms**

Our study observed that users naturally adopted roles, such as leader or navigator, during collaborative tasks. This aligns with findings by Schmidt regarding the importance of role distribution in collaborative work [180]. Future AR systems could support these natural roles by providing role-specific controls, where each user has access to features tailored to their function within the group. For example, a "lead navigator" could control navigation and viewpoint settings, while other users might access annotation tools or model manipulation features, enhancing group cohesion and reducing task redundancy.

### **Adaptive Spatial Interaction Proxemics**

Our findings emphasise the need for AR systems to respond dynamically to user positions and proximity in shared spaces. Echoing the spatial interaction principles noted by Kirk et al. in mixed reality systems [107], future designs could integrate proximity-based interaction zones that enable seamless transitions between private

and shared spaces. Adaptive spatial boundaries could be defined to allow specific users exclusive access to particular content while enabling others to view or interact within designated zones, thus accommodating both shared and individual engagement.

### **Equitable Participation and Inclusivity**

The importance of equitable access to shared content emerged as a crucial design consideration. Reflecting findings by Li et al. on balancing control and visibility in collaborative AR [116], future designs should include shared control mechanisms and flexible visibility options that promote inclusivity and equitable participation. For instance, enabling all users to contribute and view content from different perspectives—without any single participant dominating the interaction—could enhance engagement and make collaborative AR more accessible across varying skill levels and abilities.

### **7.3.4 Design Guidelines**

Building on the findings, ideas, and theoretical insights presented in this thesis, we outline a set of design guidelines aimed at shaping future AR systems to enhance collaborative experiences. These guidelines distil key observations on user behaviour, spatial dynamics, and interaction affordances into actionable recommendations, bridging empirical findings with design practice. Each guideline addresses a specific aspect of collaborative AR, from supporting fluid transitions between collaboration styles to fostering inclusivity and accessibility. By embedding these principles into AR design, future systems can better support the diverse needs of users in co-located, multi-device environments.

#### **D1 - Fostering Seamless Collaboration**

This study observed that user collaboration in handheld AR often shifts between tightly and loosely coupled interactions, as users alternated between individual tasks and group discussions. Billinghurst and Kato and Rekimoto and Nagao emphasise that effective collaboration systems should support seamless transitions between these modes to enhance overall interaction quality [14, 166]. Therefore, we recommend:

- **Design for Varying Collaboration Styles.** AR systems should provide flexible interaction modes that adapt to various collaboration styles, enabling users to alternate smoothly between shared and individual views. For example,

features like shared workspaces and synchronised views can support a tightly coupled approach when needed while providing independent views for tasks that require individual focus.

- **Enable Fluid Interaction Transitions.** Integrating techniques such as *Perspective Sharing* (for shared understanding) and *Freeze Frame* (to stabilise the view when pausing) allows users to move seamlessly between collaborative and individual tasks. This fluidity can help users maintain engagement without cognitive disruption, aligning with Schmidt and Bannon’s articulation work model that emphasises the importance of minimising context-switching friction [180].

## D2 - Enhancing User Experience

The study found that constant device reconfiguration and handling detract from user focus and introduce cognitive strain, especially in complex tasks. In line with Norman’s affordances concept, it’s critical for handheld AR systems to offer intuitive, streamlined interactions that minimise user effort [147]. To achieve this:

- **Prioritise Intuitive Interfaces.** User interfaces should be designed for ease of use, particularly to support first-time or casual users. Visual cues and clear feedback mechanisms (e.g., prompts for successful marker recognition or realignment assistance) are essential to reduce configuration work, making complex tasks easier to navigate.
- **Incorporate Adaptive Feedback.** Real-time, context-aware feedback can enhance the user experience by guiding users through tasks and adapting to different environmental conditions. For example, dynamic visualisations and context-sensitive prompts (e.g., distance markers or rotation aids) provide users with the support they need to focus on task completion rather than constant device adjustments.

## D3 - Supporting Spatial and Contextual Awareness

Collaborative systems depend on spatial awareness for users to engage effectively. Building on Gutwin and Greenberg’s findings on workspace awareness, this study observed that mobility within physical spaces and spatial interactions were essential for AR users to explore models collaboratively [72]. However, physical limitations often interrupted engagement, indicating a need for spatially aware systems that support intuitive, context-rich interactions:

- **Leverage Environmental Cues.** Integrating environmental mapping, such as spatial anchors or dynamic markers that guide users' attention, can enhance AR experiences by acknowledging the physical space in which collaboration occurs. These cues can reduce physical demands by providing stable, shared reference points that reduce the need for constant physical movement.
- **Facilitate Spatial Interaction Proxemics.** Interaction designs should consider proxemic interactions, as outlined by Mentis et al. in healthcare and Schmidt in collaborative work [131, 180]. Spatial boundaries and cues can help manage private and public spaces while allowing participants to observe or interact with different parts of the model based on their position. Proximity-based interaction zones, for instance, could automatically trigger view adjustments or highlight shared content when users are near one another, supporting seamless transitions in group dynamics.

#### D4 - Ensuring Accessibility and Inclusivity

Handheld AR has the potential to engage a diverse user base across skill levels and abilities. The limitations identified in this study around device handling and visibility further show the need for accessible design, in line with the inclusivity goals of modern HCI and CSCW studies [13]. Recommendations include:

- **Design for Diverse Abilities.** To accommodate users with varying physical capabilities or technical expertise, AR interfaces should offer customisable options. For example, providing alternative interaction methods, such as voice or gesture control, can make AR more accessible to users who may struggle with device handling.
- **Promote Equitable Participation.** Ensuring that all users have equal opportunities to contribute is essential for collaborative success. To this end, AR systems should feature shared control mechanisms (e.g., token-based access to tools or rotational control of shared views) and equitable visibility options that prevent one user from monopolising the shared AR space.

## D5 - Fostering Long-Term Engagement

Given that handheld AR is often deployed as a short-term solution, this study highlights the need for long-term usability to ensure sustained user engagement. Drawing on insights from gaming and educational applications of AR [102, 110], we recommend:

- **Support for Longitudinal Use.** For AR to integrate effectively into daily work or learning contexts, systems must evolve to support sustained usage. Frequent updates, user-driven customisation options, and iterative design improvements based on user feedback will be essential to meet the demands of ongoing engagement and adaptability.
- **Encourage Continuous Learning.** To keep users engaged and improve their proficiency with AR over time, gradual exposure to advanced features and options should be supported. This can be achieved through in-system tutorials, staged learning paths, and community support tools, enabling users to build confidence in navigating complex collaborative tasks.

## D6 - Facilitating Real-World Deployment

Finally, achieving the collaborative potential of AR in real-world settings requires addressing infrastructure and scalability concerns. Recognising the challenges of networked AR, as observed in mobile applications and discussed by Billinghurst and Kato and Szalavári et al. [15, 199], this study recommends:

- **Develop Robust Communication Architectures.** Reliable, low-latency connections are critical for maintaining synchronicity in collaborative AR. Systems should be designed with fault-tolerant communication architectures that handle fluctuations in network quality and recover smoothly from disconnections to ensure real-time collaboration remains uninterrupted.
- **Ensure Scalability and Flexibility.** To support a variety of group sizes and use cases, AR systems must be designed to scale effectively. Modular, flexible designs can enable tailored configurations for different applications, whether for small educational groups or large-scale industrial teams. These systems should also be adaptable to evolving requirements, making them suitable for deployment across diverse environments and user needs.



These recommendations synthesise the study's findings with insights from prior CSCW and HCI research, focusing on enhancing the accessibility, usability, and collaborative dynamics of handheld AR. By prioritising adaptability, inclusivity, and context-sensitive interaction, future AR technologies can better support meaningful collaboration and accommodate the diverse needs of co-located users across varied applications.

## 7.4 Summary and Key Takeaways

This thesis advances the understanding and design of co-located collaborative Augmented Reality (AR), with a particular focus on handheld mobile devices. Through a series of empirical studies, research-through-design, and a structured framework, this research explores the complex dynamics of group collaboration in AR environments, providing insights that guide the design of adaptive, user-centred AR systems. By investigating device configurations, spatial awareness, and proxemic interactions, this work contributes a flexible design space for creating socio-spatially aware AR systems that effectively support both individual and group collaboration. These insights inform future AR designs, enabling developers to create immersive, adaptable systems that enhance co-located collaboration and enrich user engagement across a variety of applications.

- **Cognitive Load and Context Switching:** The studies reveal that handheld AR interfaces impose considerable cognitive load, primarily due to frequent context switching between digital interaction and face-to-face communication. This observation aligns with previous findings on the cognitive strain of managing dual interaction spaces [88]. By examining these transitions in real-world tasks, this thesis highlights the need for AR systems to reduce cognitive demands, particularly through design features that minimise context switches or support smoother transitions between physical and digital engagement.
- **Device Configuration and Collaborative Dynamics:** Device configurations influence user interaction styles and group dynamics in collaborative AR. Larger devices, like tablets, were found to foster shared viewing and synchronous engagement, while smaller devices tended to promote individual focus, corroborating insights from cross-device collaboration studies [26]. This thesis demonstrates that the physical attributes of handheld devices shape not only cognitive load but also collaborative strategies, informing future AR design to consider device symmetry and screen size for optimal user engagement.

- **Application of Proxemics Theory:** This work applies proxemics as a socio-spatial framework to analyse and enhance co-located collaborative AR. Proxemics theory, which describes how spatial relationships affect interaction, was used to define three critical dimensions in AR: perceptual proxemics, deixis proxemics, and control proxemics. By using these dimensions, this thesis provides a structured approach to designing AR systems that are responsive to users' spatial arrangements, enabling adaptive and fluid collaboration [131, 200]. This conceptual lens expands AR interaction design, allowing developers to embed socio-spatial cues that facilitate shared and individualised AR experiences.
- **CARIS Framework for Proxemic-Aware Collaborative AR:** The thesis introduces the Collaborative Augmented Reality Interaction Space (CARIS) framework, which structures co-located AR collaboration around three components: People, Interaction Portals, and Physical Features. CARIS leverages proxemics to support spatially responsive AR interactions, such as shared views and personal interaction zones. This framework offers practical design guidelines, including *Spatial Awareness Cues* and *View-Clutching*, to reduce cognitive load and support dynamic collaboration, echoing principles seen in collaborative work research [13, 188].
- **Design Recommendations for Handheld AR Systems:** The research outlines key design recommendations for handheld AR that address device handling, collaborative flexibility, and user-centred feedback mechanisms. Suggested features include mechanisms for stabilising user views (to minimise physical strain), proxemic cues for spatial awareness, and task-specific toggles for handling complex models. These guidelines are based on empirical findings and reveal the importance of adaptive, context-aware AR systems that facilitate both tightly and loosely coupled collaboration [147, 180].

The thesis concludes by identifying areas for future research, such as scaling proxemic-aware AR frameworks to larger groups and investigating role-based control in collaborative environments. It also suggests that future studies explore privacy considerations, particularly when users' personal and shared views converge in close proximity [131]. These directions highlight the potential for CARIS and proxemic-aware AR to support complex, large-scale interactions, setting a foundation for more inclusive, immersive, and user-centred AR experiences.

## 7.5 Future Work

The findings of this thesis offer a strong foundation for collaborative Augmented Reality (AR) research, particularly in the domains of cross-device interaction, proxemic-aware systems, and co-located, multi-user AR. While the work within this thesis makes strong contributions, it also opens up several avenues for further investigation. This section reflects on the limitations encountered, outlines open questions that remain unaddressed, and presents new directions for research inspired both by the individual studies and the synthesis of ideas across the thesis.

### 7.5.1 Limitations

One limitation of this research lies in the scope and scale of empirical studies. Both *CollabAR* and *DeviceAR* chapters relied on small-scale, co-located group studies to analyse collaboration and interaction patterns. Although these sessions offered valuable insights into the collaborative dynamics in handheld AR, they may not be strongly generalisable and fully capture the diverse range of contexts and environments in which collaborative AR could be deployed. For example, the relatively controlled lab setting may have mitigated factors such as environmental noise or network variability, which could have an effect on collaboration in real-world scenarios, such as network dropouts.

Additionally, the *Conceptual Lens* and *Framework* chapters, while advancing our understanding of proxemics in AR design, were limited by the lack of long-term, longitudinal studies. Short-term testing provided insights into proxemic interaction patterns, but it did not capture how users might adapt to these systems over time or how their behaviours might evolve with repeated exposure. It relied heavily on the initial demonstration of interaction techniques and the ability of participants to remember them. Moreover, the CARIS framework, although implemented as an exemplar system, was not extensively tested across diverse user demographics, which could have yielded richer data on inclusivity and accessibility in AR.

### 7.5.2 Open Questions and Concerns

Several open questions and methodological concerns arose throughout the research, particularly regarding the integration and optimisation of proxemics within AR systems. One key area is the scalability of proxemic-aware AR in larger groups or

industrial applications. While CARIS introduced mechanisms to enable spatially-aware, flexible interactions, it remains unclear how such systems would function in larger, multi-room environments or with a broader range of devices beyond handheld AR. Questions about the trade-offs in computational cost, device synchronisation, and interface complexity in such settings also need further exploration.

Another open question pertains to user privacy in collaborative AR spaces. Although proxemics was shown to be effective in delineating private and public spaces, maintaining these distinctions becomes challenging as user density and proximity increase. Future research could explore privacy-preserving AR design strategies, examining how private and shared content can be dynamically managed while minimising cognitive load and maintaining collaboration fluidity.

Further, the design of role-specific controls in collaborative AR, as discussed in *DeviceAR*, introduces a conceptual challenge around role dynamics and control distribution. A more detailed understanding of how roles emerge and shift in collaborative AR could inform guidelines for adaptive role-based controls that respond to user behaviours in real-time.

### 7.5.3 Emerging Ideas and New Directions

In this thesis, I propose several promising ideas for future research that were not fully explored within this scope. One important area of potential is the integration of adaptive learning features in AR systems, particularly to aid new users in mastering collaborative interactions. Building on insights from the *CollabAR* chapter, which demonstrated how varying levels of AR familiarity impacted interaction styles, a learning-based feature could tailor prompts or tutorials based on real-time user proficiency, providing smoother onboarding experiences for collaborative AR.

Future work could also focus on developing multi-modal, cross-device interaction techniques that build on the CARIS framework's proxemic insights. For instance, incorporating audio cues, tactile feedback, or even holographic and spatial avatar representations could expand interaction modalities, making collaborative AR systems more immersive and responsive to user intent. Additionally, the potential for integrating wearable AR devices with handheld devices in collaborative settings warrants further investigation, especially in light of recent advances in wearable computing.

Another promising direction lies in the integration of Artificial Intelligence (AI) to enhance the immersive and distraction-free nature of collaborative AR environments. By leveraging AI-driven scene processing, future AR systems could dynamically

filter out visual noise, such as paintings or posters on walls, to provide a cleaner and more focused workspace. Such functionality could increase user immersion by removing non-essential background elements, allowing users to concentrate on the collaborative task at hand. Adaptive AI filters could also recognise and differentiate between essential and non-essential objects in real-time, dynamically adjusting the environment to maintain a clear view of AR content. These AI capabilities would be particularly useful in educational or corporate settings, where a clutter-free environment may help teams focus and work more effectively.

Lastly, this thesis opens new directions for longitudinal research that examines collaborative AR's impact on group dynamics, productivity, and long-term engagement. Observing users in prolonged tasks or across multiple sessions could provide a richer understanding of user adaptation and retention of collaborative behaviours over time. Such studies would not only validate and refine proxemic-aware AR frameworks like CARIS but also offer actionable insights for integrating AR into everyday collaborative work, from education to industrial training.

# 8

## Conclusion

This thesis explored the transformative potential of handheld mobile Augmented Reality (AR) in enhancing co-located collaborative experiences, focusing on the socio-spatial dynamics that influence group interactions and device configurations in AR settings. By employing proxemics theory as a guiding conceptual lens, this work extends the understanding of how AR systems can be designed to support flexible, adaptive, and socio-spatially aware collaboration. Throughout this thesis, I have aimed to identify design principles and interaction techniques that foster more seamless and effective AR-based collaboration, addressing core research questions and demonstrating a nuanced approach to AR system development.

In earlier chapters, we examined basic AR interactions and device configurations, highlighting the cognitive and physical demands associated with mobile AR interfaces in collaborative environments ([Chapter 3](#) and [Chapter 4](#)). The findings revealed that larger devices generally facilitate shared focus and interaction, while smaller devices often lead to individualised interactions. The insights provided the foundation for a flexible, proxemically grounded framework for designing AR experiences that respect personal space while enhancing collaboration. Through this, we highlighted the importance of developing interaction techniques that transition smoothly between individual and group tasks, reinforcing user engagement in shared AR spaces.

In [Chapter 5](#), we introduced proxemics as a valuable perspective for understanding and designing co-located AR. This approach yielded a conceptual framework that articulated the socio-spatial dynamics of collaborative handheld AR, addressing previously unexplored factors like deixis, perceptual, and control proxemics in

shared AR environments. The interaction techniques derived from this framework ranged from aligning closely with physical reality to more flexible configurations that transcend real-world constraints, illustrating a continuum of design choices that balance physical proximity with digital flexibility. These techniques demonstrated how AR systems could support both tightly coupled collaborations (e.g., shared views and gestures in close proximity) and loosely coupled interactions (e.g., remote sketching and broadcasted perspectives).

Reflecting on the findings, this thesis proposes that AR systems designed with proxemics in mind can better accommodate diverse collaboration styles and foster more intuitive social interactions. Techniques such as Remote Peeking, Perspective Broadcasting, and Privacy Fence were shown to offer users various options for personalising their interaction space, thus supporting a richer and more adaptable collaboration style. Furthermore, these techniques illuminate the trade-offs between providing flexibility and maintaining the intuitive, contextually relevant cues that users rely on in real-world interactions. I suggest that future research should explore these dynamics further, focusing on enhancing user understanding of shared virtual spaces without sacrificing co-located cues and intentions.

In conclusion, this thesis contributes to the field of collaborative AR by offering empirically grounded design principles, a structured framework for AR interaction, and practical interaction techniques that guide the development of future AR applications. By examining the unique socio-spatial interactions fostered by handheld AR, I have highlighted the accessibility and versatility of mobile devices as a platform for AR, which holds promise for broad adoption across collaborative settings. With continued research and innovation, AR has the potential to reshape how people work, learn, and connect in augmented spaces, transcending physical limitations to create more engaging, flexible, and socially connected environments.

## References

- [1] Apple. *ARKit / Apple Developer Documentation*.
- [2] Dedy Ariansyah et al. “Augmented Reality Training for Improved Learnability”. In: *CIRP Journal of Manufacturing Science and Technology* 48 (Feb. 1, 2024), pp. 19–27. DOI: [10.1016/j.cirpj.2023.11.003](https://doi.org/10.1016/j.cirpj.2023.11.003).
- [3] I. Aslan and E. André. “Pre-Touch Proxemics: Moving the Design Space of Touch Targets from Still Graphics towards Proxemic Behaviors”. In: pp. 101–109. DOI: [10.1145/3136755.3136808](https://doi.org/10.1145/3136755.3136808).
- [4] R. Azuma et al. “Recent Advances in Augmented Reality”. In: *IEEE Computer Graphics and Applications* 21.6 (Nov. 2001), pp. 34–47. DOI: [10.1109/38.963459](https://doi.org/10.1109/38.963459).
- [5] Ronald T. Azuma. “A Survey of Augmented Reality”. In: *Presence: Teleoperators and Virtual Environments* 6.4 (Aug. 1, 1997), pp. 355–385. DOI: [10.1162/pres.1997.6.4.355](https://doi.org/10.1162/pres.1997.6.4.355).
- [6] Robert F. Bales. “A Set of Categories for the Analysis of Small Group Interaction”. In: *American Sociological Review* 15.2 (1950), pp. 257–263. DOI: [10.2307/2086790](https://doi.org/10.2307/2086790). JSTOR: [2086790](https://www.jstor.org/stable/2086790).
- [7] Till Ballendat et al. “Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment”. In: *ACM International Conference on Interactive Tabletops and Surfaces - ITS '10*. ACM International Conference. Saarbrücken, Germany: ACM Press, 2010, p. 121. DOI: [10.1145/1936652.1936676](https://doi.org/10.1145/1936652.1936676).
- [8] Ryan Anthony J. de Belen et al. “A Systematic Review of the Current State of Collaborative Mixed Reality Technologies: 2013–2018”. In: *AIMS Electronics and Electrical Engineering* 3.2 (2019), pp. 181–223. DOI: [10.3934/ElectrEng.2019.2.181](https://doi.org/10.3934/ElectrEng.2019.2.181).
- [9] Steve Benford et al. “VR-VIBE: A Virtual Environment for Co-operative Information Retrieval”. In: *Computer Graphics Forum* 14.3 (Aug. 1995), pp. 349–360. DOI: [10.1111/j.1467-8659.1995.cgf143\\_0349.x](https://doi.org/10.1111/j.1467-8659.1995.cgf143_0349.x).
- [10] Steve Benford et al. “Understanding and Constructing Shared Spaces with Mixed-Reality Boundaries”. In: *ACM Transactions on Computer-Human Interaction* 5.3 (Sept. 1, 1998), pp. 185–223. DOI: [10.1145/292834.292836](https://doi.org/10.1145/292834.292836).
- [11] H. Benko et al. “Cross-Dimensional Gestural Interaction Techniques for Hybrid Immersive Environments”. In: *IEEE Proceedings. VR 2005. Virtual Reality, 2005*. IEEE Proceedings. VR 2005. Virtual Reality, 2005. Mar. 2005, pp. 209–216. DOI: [10.1109/VR.2005.1492776](https://doi.org/10.1109/VR.2005.1492776).



- [12] M. Billinghurst et al. “Shared Space: An Augmented Reality Approach for Computer Supported Collaborative Work”. In: *Virtual Reality 3.1* (Mar. 1, 1998), pp. 25–36. DOI: [10.1007/BF01409795](https://doi.org/10.1007/BF01409795).
- [13] M. Billinghurst et al. “Mixing Realities in Shared Space: An Augmented Reality Interface for Collaborative Computing”. In: *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No.00TH8532)*. International Conference on Multimedia and Expo. Vol. 3. New York, NY, USA: IEEE, 2000, pp. 1641–1644. DOI: [10.1109/ICME.2000.871085](https://doi.org/10.1109/ICME.2000.871085).
- [14] Mark Billinghurst and Hirokazu Kato. “Collaborative Mixed Reality”. In: *Mixed Reality*. Ed. by Yuichi Ohta and Hideyuki Tamura. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999, pp. 261–284. DOI: [10.1007/978-3-642-87512-0\\_15](https://doi.org/10.1007/978-3-642-87512-0_15).
- [15] Mark Billinghurst and Hirokazu Kato. “Collaborative Augmented Reality”. In: *Communications of the ACM* 45.7 (July 1, 2002), pp. 64–70. DOI: [10.1145/514236.514265](https://doi.org/10.1145/514236.514265).
- [16] Mark Billinghurst et al. “The MagicBook: A Transitional AR Interface”. In: *Computers & Graphics. Mixed Realities - beyond Conventions* 25.5 (Oct. 1, 2001), pp. 745–753. DOI: [10.1016/S0097-8493\(01\)00117-0](https://doi.org/10.1016/S0097-8493(01)00117-0).
- [17] Mark Billinghurst et al. “Advanced Interaction Techniques for Augmented Reality Applications”. In: *Virtual and Mixed Reality: Third International Conference, VMR 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. Proceedings 3*. Springer, 2009, pp. 13–22. DOI: [10.1007/978-3-642-02771-0\\_2](https://doi.org/10.1007/978-3-642-02771-0_2).
- [18] Mark Billinghurst et al. “A Survey of Augmented Reality”. In: *Foundations and Trends® in Human-Computer Interaction* 8.2-3 (Mar. 30, 2015), pp. 73–272. DOI: [10.1561/11000000049](https://doi.org/10.1561/11000000049).
- [19] Mark N. Billinghurst and Anders Henrysson. “Research Directions in Handheld AR”. In: *International Journal of Virtual Reality* 5.2 (2 Jan. 1, 2006), pp. 51–58. DOI: [10.20870/IJVR.2006.5.2.2690](https://doi.org/10.20870/IJVR.2006.5.2.2690).
- [20] Jonas Blattgerste et al. “TrainAR: A Scalable Interaction Concept and Didactic Framework for Procedural Trainings Using Handheld Augmented Reality”. In: *Multimodal Technologies and Interaction* 5.7 (7 July 2021), p. 30. DOI: [10.3390/mti5070030](https://doi.org/10.3390/mti5070030).
- [21] Eleonora Bottani and Giuseppe Vignali. “Augmented Reality Technology in the Manufacturing Industry: A Review of the Last Decade”. In: *IISE Transactions* 51.3 (Mar. 4, 2019), pp. 284–310. DOI: [10.1080/24725854.2018.1493244](https://doi.org/10.1080/24725854.2018.1493244).
- [22] Andrew C Boud et al. “Virtual Reality and Augmented Reality as a Training Tool for Assembly Tasks”. In: *1999 IEEE International Conference on Information Visualization (Cat. No. PR00210)*. IEEE, 1999, pp. 32–36. DOI: [10.1109/IV.1999.781532](https://doi.org/10.1109/IV.1999.781532).
- [23] D.m. Bressler and A.m. Bodzin. “A Mixed Methods Assessment of Students’ Flow Experiences during a Mobile Augmented Reality Science Game”. In: *Journal of Computer Assisted Learning* 29.6 (2013), pp. 505–517. DOI: [10.1111/jcal.12008](https://doi.org/10.1111/jcal.12008).

- [24] Frederick P. Brooks. “The Computer Scientist as Toolsmith II”. In: *Commun. ACM* 39.3 (Mar. 1, 1996), pp. 61–68. DOI: [10.1145/227234.227243](https://doi.org/10.1145/227234.227243).
- [25] Frederik Brudy et al. “Investigating the Role of an Overview Device in Multi-Device Collaboration”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18: CHI Conference on Human Factors in Computing Systems. Montreal QC Canada: ACM, Apr. 21, 2018, pp. 1–13. DOI: [10.1145/3173574.3173874](https://doi.org/10.1145/3173574.3173874).
- [26] Frederik Brudy et al. “Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices”. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI ’19: CHI Conference on Human Factors in Computing Systems. Glasgow Scotland Uk: ACM, May 2, 2019, pp. 1–28. DOI: [10.1145/3290605.3300792](https://doi.org/10.1145/3290605.3300792).
- [27] A. Butz et al. “Enveloping Users and Computers in a Collaborative 3D Augmented Reality”. In: *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR’99)*. Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR’99). Oct. 1999, pp. 35–44. DOI: [10.1109/IWAR.1999.803804](https://doi.org/10.1109/IWAR.1999.803804).
- [28] Julie Carmigniani and Borko Furht. “Augmented Reality: An Overview”. In: *Handbook of augmented reality* (2011), pp. 3–46. DOI: [10.1007/978-1-4614-0064-6\\_1](https://doi.org/10.1007/978-1-4614-0064-6_1).
- [29] Peter Carstensen and Kjeld Schmidt. “Computer Supported Cooperative Work: New Challenges to Systems Design”. In: *Handbook in Human Factors/Ergonomics* (in Japanese). Asakura Publishing, 1999.
- [30] Dimitris Chatzopoulos et al. “Mobile Augmented Reality Survey: From Where We Are to Where We Go”. In: *IEEE Access* 5 (2017), pp. 6917–6950. DOI: [10.1109/ACCESS.2017.2698164](https://doi.org/10.1109/ACCESS.2017.2698164).
- [31] Henry Chen et al. “3D Collaboration Method over HoloLens™ and Skype™ End Points”. In: *Proceedings of the 3rd International Workshop on Immersive Media Experiences*. ImmersiveME ’15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 27–30. DOI: [10.1145/2814347.2814350](https://doi.org/10.1145/2814347.2814350).
- [32] Weiqin Chen. “Historical Oslo on a Handheld Device—a Mobile Augmented Reality Application”. In: *Procedia Computer Science* 35 (2014), pp. 979–985. DOI: [10.1016/j.procs.2014.08.180](https://doi.org/10.1016/j.procs.2014.08.180).
- [33] Xiang ’Anthony’ Chen et al. “Duet: Exploring Joint Interactions on a Smart Phone and a Smart Watch”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14: CHI Conference on Human Factors in Computing Systems. Toronto Ontario Canada: ACM, Apr. 26, 2014, pp. 159–168. DOI: [10.1145/2556288.2556955](https://doi.org/10.1145/2556288.2556955).
- [34] Kun-Hung Cheng and Chin-Chung Tsai. “Affordances of Augmented Reality in Science Learning: Suggestions for Future Research”. In: *Journal of Science Education and Technology* 22.4 (Aug. 1, 2013), pp. 449–462. DOI: [10.1007/s10956-012-9405-9](https://doi.org/10.1007/s10956-012-9405-9).

- [35] Pei-Yu (Peggy) Chi and Yang Li. “Weave: Scripting Cross-Device Wearable Interaction”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, Apr. 18, 2015, pp. 3923–3932.
- [36] John W Creswell and Vicki L Plano Clark. *Designing and Conducting Mixed Methods Research*. Sage publications, 2017.
- [37] Scott G. Dacko. “Enabling Smart Retail Settings via Mobile Augmented Reality Shopping Apps”. In: *Technological Forecasting and Social Change* 124 (Nov. 1, 2017), pp. 243–256. DOI: [10.1016/j.techfore.2016.09.032](https://doi.org/10.1016/j.techfore.2016.09.032).
- [38] Ella Dagan et al. “Project IRL: Playful Co-Located Interactions with Mobile Augmented Reality”. In: *Proceedings of the ACM on Human-Computer Interaction* 6 (CSCW1 Apr. 7, 2022), 62:1–62:27. DOI: [10.1145/3512909](https://doi.org/10.1145/3512909).
- [39] Florian Daiber et al. “BouldAR: Using Augmented Reality to Support Collaborative Boulder Training”. In: *CHI '13 Extended Abstracts on Human Factors in Computing Systems*. Chi Ea '13. New York, NY, USA: Association for Computing Machinery, 2013, pp. 949–954. DOI: [10.1145/2468356.2468526](https://doi.org/10.1145/2468356.2468526).
- [40] M. Daza et al. “An Approach of Social Navigation Based on Proxemics for Crowded Environments of Humans and Robots”. In: *Micromachines* 12.2 (Feb. 13, 2021), p. 193. DOI: [10.3390/mi12020193](https://doi.org/10.3390/mi12020193).
- [41] Arindam Dey et al. “A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014”. In: *Frontiers in Robotics and AI* 5 (2018).
- [42] Joan Morris DiMicco et al. “Influencing Group Participation with a Shared Display”. In: *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. CSCW '04. New York, NY, USA: Association for Computing Machinery, Nov. 6, 2004, pp. 614–623. DOI: [10.1145/1031607.1031713](https://doi.org/10.1145/1031607.1031713).
- [43] Gino Dini and Michela Dalle Mura. “Application of Augmented Reality Techniques in Through-Life Engineering Services”. In: *Procedia Cirp* 38 (2015), pp. 14–23. DOI: [10.1016/j.procir.2015.07.044](https://doi.org/10.1016/j.procir.2015.07.044).
- [44] echo3D. *15 Industries Using Augmented Reality and Virtual Reality*. echo3D. May 18, 2022. URL: <https://medium.com/echo3d/15-industries-using-augmented-reality-and-virtual-reality-3ef4577b9df4> (visited on 07/18/2024).
- [45] Niklas Elmqvist and Philippos Tsigas. “A Taxonomy of 3D Occlusion Management for Visualization”. In: *IEEE Transactions on Visualization and Computer Graphics* 14.5 (Sept. 2008), pp. 1095–1109. DOI: [10.1109/TVCG.2008.59](https://doi.org/10.1109/TVCG.2008.59).
- [46] Barrett Ens et al. “Revisiting Collaboration through Mixed Reality: The Evolution of Groupware”. In: *International Journal of Human-Computer Studies*. 50 Years of the International Journal of Human-Computer Studies. Reflections on the Past, Present and Future of Human-Centred Technologies 131 (Nov. 1, 2019), pp. 81–98. DOI: [10.1016/j.ijhcs.2019.05.011](https://doi.org/10.1016/j.ijhcs.2019.05.011).

- [47] Jørgen Ernstsens et al. “Incidental Memory Recall in Virtual Reality: An Empirical Investigation”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 63.1 (Nov. 1, 2019), pp. 2277–2281. DOI: [10.1177/1071181319631411](https://doi.org/10.1177/1071181319631411).
- [48] F. E. Fadzli et al. “ARGarden: 3D Outdoor Landscape Design Using Handheld Augmented Reality with Multi-User Interaction”. In: *IOP Conference Series: Materials Science and Engineering* 979.1 (Nov. 2020), p. 012001. DOI: [10.1088/1757-899X/979/1/012001](https://doi.org/10.1088/1757-899X/979/1/012001).
- [49] Joel E. Fischer et al. “Beyond “Same Time, Same Place”: Introduction to the Special Issue on Collocated Interaction”. In: *Human-Computer Interaction* 33.5-6 (Sept. 4, 2018), pp. 305–310. DOI: [10.1080/07370024.2018.1440556](https://doi.org/10.1080/07370024.2018.1440556).
- [50] Adam Fouse et al. “ChronoViz: A System for Supporting Navigation of Time-Coded Data”. In: *CHI '11 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '11. New York, NY, USA: Association for Computing Machinery, May 7, 2011, pp. 299–304. DOI: [10.1145/1979742.1979706](https://doi.org/10.1145/1979742.1979706).
- [51] Markus Funk et al. “A Benchmark for Interactive Augmented Reality Instructions for Assembly Tasks”. In: *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. 2015, pp. 253–257.
- [52] Anna Fuste and Chris Schmandt. “HyperCubes: A Playful Introduction to Computational Thinking in Augmented Reality”. In: *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. CHI PLAY '19 Extended Abstracts. New York, NY, USA: Association for Computing Machinery, Oct. 17, 2019, pp. 379–387. DOI: [10.1145/3341215.3356264](https://doi.org/10.1145/3341215.3356264).
- [53] Riva G et al. “Interreality: The Experiential Use of Technology in the Treatment of Obesity”. In: *Clinical Practice and Epidemiology in Mental Health* 7.1 (Mar. 4, 2011). DOI: [10.2174/1745017901107010051](https://doi.org/10.2174/1745017901107010051).
- [54] Joseph L. Gabbard et al. “Effects of AR Display Context Switching and Focal Distance Switching on Human Performance”. In: *IEEE Transactions on Visualization and Computer Graphics* 25.6 (June 2019), pp. 2228–2241. DOI: [10.1109/TVCG.2018.2832633](https://doi.org/10.1109/TVCG.2018.2832633).
- [55] Zeinab Ghaemi et al. “Proxemic Maps for Immersive Visualization”. In: *Cartography and Geographic Information Science* 49.3 (2022), pp. 205–219. DOI: [10.1080/15230406.2021.2013946](https://doi.org/10.1080/15230406.2021.2013946). eprint: <https://doi.org/10.1080/15230406.2021.2013946>.
- [56] James J Gibson. “The Theory of Affordances”. In: *Hilldale, USA* 1.2 (1977), pp. 67–82.
- [57] Eg Su Goh et al. “3D Object Manipulation Techniques in Handheld Mobile Augmented Reality Interface: A Review”. In: *IEEE Access* 7 (2019), pp. 40581–40601. DOI: [10.1109/ACCESS.2019.2906394](https://doi.org/10.1109/ACCESS.2019.2906394).
- [58] Google. *Firebase Realtime Database*. 2019.
- [59] Google. *ARCore*.

- [60] Jamie Gower et al. “Augmented Reality Support for Performance and Decision-Making in Collaborative Time-Critical Tasks”. In: *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). Oct. 2022, pp. 958–963. DOI: [10.1109/ISMAR-Adjunct57072.2022.00215](https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00215).
- [61] Scott A. Green et al. “Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design”. In: *International Journal of Advanced Robotic Systems* 5.1 (Mar. 1, 2008), p. 1. DOI: [10.5772/5664](https://doi.org/10.5772/5664).
- [62] Saul Greenberg and Ernest Chang. “Computer Support for Real Time Collaborative Work”. In: *Proceedings of the Conference on Numerical Mathematics and Computing*. Vol. 74. Citeseer, 1989.
- [63] Saul Greenberg et al. “Proxemic Interactions: The New Ubicomp?” In: *Interactions* 18.1 (Jan. 2011), pp. 42–50. DOI: [10.1145/1897239.1897250](https://doi.org/10.1145/1897239.1897250).
- [64] Irene Greif, ed. *Computer-Supported Cooperative Work: A Book of Readings*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1988.
- [65] Jens Emil Grønbæk et al. “Proxemic Transitions: Designing Shape-Changing Furniture for Informal Meetings”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. New York, NY, USA: Association for Computing Machinery, May 2, 2017, pp. 7029–7041. DOI: [10.1145/3025453.3025487](https://doi.org/10.1145/3025453.3025487).
- [66] Jens Emil Grønbæk et al. “Proxemics Play: Exploring the Interplay between Mobile Devices and Interiors”. In: *Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion*. DIS ’19: Designing Interactive Systems Conference 2019. San Diego CA USA: ACM, June 18, 2019, pp. 177–181. DOI: [10.1145/3301019.3323886](https://doi.org/10.1145/3301019.3323886).
- [67] Jens Emil Grønbæk et al. “Proxemics beyond Proximity: Designing for Flexible Social Interaction through Cross-Device Interaction”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Chi ’20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–14. DOI: [10.1145/3313831.3376379](https://doi.org/10.1145/3313831.3376379).
- [68] Jens Grubert and Matthias Kränz. “Towards Ad Hoc Mobile Multi-Display Environments on Commodity Mobile Devices”. In: *2017 IEEE Virtual Reality (VR)*. 2017 IEEE Virtual Reality (VR). Mar. 2017, pp. 461–462. DOI: [10.1109/VR.2017.7892379](https://doi.org/10.1109/VR.2017.7892379).
- [69] J. Grudin. “Computer-Supported Cooperative Work: History and Focus”. In: *Computer* 27.5 (May 1994), pp. 19–26. DOI: [10.1109/2.291294](https://doi.org/10.1109/2.291294).
- [70] Luis A. Guerrero et al. “Selecting Computing Devices to Support Mobile Collaboration”. In: *Group Decision and Negotiation* 15.3 (May 1, 2006), pp. 243–271. DOI: [10.1007/s10726-006-9020-3](https://doi.org/10.1007/s10726-006-9020-3).
- [71] Anhong Guo et al. “Blocks: Collaborative and Persistent Augmented Reality Experiences”. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3.3 (Sept. 9, 2019), 83:1–83:24. DOI: [10.1145/3351241](https://doi.org/10.1145/3351241).



- [72] Carl Gutwin and Saul Greenberg. “Design for Individuals, Design for Groups: Tradeoffs between Power and Workspace Awareness”. In: *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*. CSCW '98. New York, NY, USA: Association for Computing Machinery, Nov. 1, 1998, pp. 207–216. DOI: [10.1145/289444.289495](https://doi.org/10.1145/289444.289495).
- [73] Carl Gutwin and Saul Greenberg. “The Importance of Awareness for Team Cognition in Distributed Collaboration.” In: *Team Cognition: Understanding the Factors That Drive Process and Performance*. Ed. by Eduardo Salas and Stephen M. Fiore. Washington: American Psychological Association, 2004, pp. 177–201. DOI: [10.1037/10690-009](https://doi.org/10.1037/10690-009).
- [74] Jonathan Haber et al. “Paper vs. Tablets: The Effect of Document Media in Co-Located Collaborative Work”. In: *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*. AVI '14. New York, NY, USA: Association for Computing Machinery, May 27, 2014, pp. 89–96. DOI: [10.1145/2598153.2598170](https://doi.org/10.1145/2598153.2598170).
- [75] Edmund T. Hall. *The Hidden Dimension*. Doubleday, 1966. 1082 pp. Google Books: [rvs\\_DtXv47EC](https://books.google.com/books?id=rvs_DtXv47EC).
- [76] Peter Hamilton and Daniel J. Wigdor. “Conductor: Enabling and Understanding Cross-Device Interaction”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. New York, NY, USA: Association for Computing Machinery, Apr. 26, 2014, pp. 2773–2782. DOI: [10.1145/2556288.2557170](https://doi.org/10.1145/2556288.2557170).
- [77] Preben Hansen and Kalervo Järvelin. “Collaborative Information Retrieval in an Information-Intensive Domain”. In: *Information Processing & Management* 41.5 (Sept. 1, 2005), pp. 1101–1119. DOI: [10.1016/j.ipm.2004.04.016](https://doi.org/10.1016/j.ipm.2004.04.016).
- [78] Alexa M. Harris et al. “Joining Together Online: The Trajectory of CSCW Scholarship on Group Formation”. In: *Proceedings of the ACM on Human-Computer Interaction* 3 (CSCW Nov. 7, 2019), 148:1–148:27. DOI: [10.1145/3359250](https://doi.org/10.1145/3359250).
- [79] Hillevi Hedberg et al. “A Systematic Review of Learning Through Mobile Augmented Reality”. In: *International Journal of Interactive Mobile Technologies (iJIM)* 12.3 (3 July 20, 2018), pp. 75–85. DOI: [10.3991/ijim.v12i3.8404](https://doi.org/10.3991/ijim.v12i3.8404).
- [80] Steven Henderson and Steven Feiner. “Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair”. In: *IEEE transactions on visualization and computer graphics* 17.10 (Oct. 2011), pp. 1355–1368. DOI: [10.1109/TVCG.2010.245](https://doi.org/10.1109/TVCG.2010.245). pmid: [21041888](https://pubmed.ncbi.nlm.nih.gov/21041888/).
- [81] Anders Henrysson et al. “Face to Face Collaborative AR on Mobile Phones”. In: *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*. ISMAR '05. USA: IEEE Computer Society, Oct. 5, 2005, pp. 80–89. DOI: [10.1109/ISMAR.2005.32](https://doi.org/10.1109/ISMAR.2005.32).

- [82] Hillary and Angela Scott-Briggs TechBullion. *The Rise of Augmented Reality in Cell Phone Applications*. TechBullion. Jan. 12, 2024. URL: <https://techbullion.com/the-rise-of-augmented-reality-in-cell-phone-applications/> (visited on 03/06/2025).
- [83] Juan David Hincapié-Ramos et al. “Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14: CHI Conference on Human Factors in Computing Systems. Toronto Ontario Canada: ACM, Apr. 26, 2014, pp. 1063–1072. DOI: [10.1145/2556288.2557130](https://doi.org/10.1145/2556288.2557130).
- [84] Jim Hollan and Scott Stornetta. “Beyond Being There”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Chi ’92. New York, NY, USA: Association for Computing Machinery, 1992, pp. 119–125. DOI: [10.1145/142750.142769](https://doi.org/10.1145/142750.142769).
- [85] Patrick Holland et al. *The 29 Best ARKit Apps to Try on iPhone and iPad*. CNET. Dec. 15, 2018. URL: <https://www.cnet.com/pictures/best-ar-apps-for-ios-that-you-need-to-try> (visited on 07/18/2024).
- [86] Svend Hollensen et al. “Metaverse – the New Marketing Universe”. In: *Journal of Business Strategy* 44.3 (Jan. 1, 2022), pp. 119–125. DOI: [10.1108/JBS-01-2022-0014](https://doi.org/10.1108/JBS-01-2022-0014).
- [87] Tobias Höllerer and Steve Feiner. “Mobile Augmented Reality”. In: *Telegeoinformatics: Location-based computing and services* 21 (2004).
- [88] Steven Houben et al. “Activityspace: Managing Device Ecologies in an Activity-Centric Configuration Space”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. ACM, 2014, pp. 119–128. DOI: [10.1145/2669485.2669493](https://doi.org/10.1145/2669485.2669493).
- [89] Steven Houben et al. “Opportunities and Challenges for Cross-Device Interactions in the Wild”. In: *Interactions* 24.5 (Aug. 25, 2017), pp. 58–63. DOI: [10.1145/3121348](https://doi.org/10.1145/3121348).
- [90] Kuo-Ting Huang et al. “Augmented Versus Virtual Reality in Education: An Exploratory Study Examining Science Knowledge Retention When Using Augmented Reality/Virtual Reality Mobile Applications”. In: *Cyberpsychology, Behavior, and Social Networking* 22.2 (Feb. 2019), pp. 105–110. DOI: [10.1089/cyber.2018.0150](https://doi.org/10.1089/cyber.2018.0150).
- [91] Sebastian Hubenschmid et al. “STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Chi ’21. New York, NY, USA: Association for Computing Machinery, 2021. DOI: [10.1145/3411764.3445298](https://doi.org/10.1145/3411764.3445298).
- [92] Duy-Nguyen Ta Huynh et al. “Art of Defense: A Collaborative Handheld Augmented Reality Board Game”. In: *Proceedings of the 2009 ACM SIGGRAPH Symposium on Video Games*. Sandbox ’09. New York, NY, USA: Association for Computing Machinery, 2009, pp. 135–142. DOI: [10.1145/1581073.1581095](https://doi.org/10.1145/1581073.1581095).
- [93] IKEA. *IKEA Place*.

- [94] P. Isenberg et al. “Collaborative Visualization: Definition, Challenges, and Research Agenda”. In: *Information Visualization* 10.4 (2011), pp. 310–326. DOI: [10.1177/1473871611412817](https://doi.org/10.1177/1473871611412817).
- [95] Petra Isenberg et al. “Co-Located Collaborative Visual Analytics around a Tabletop Display”. In: *IEEE Transactions on Visualization and Computer Graphics* 18.5 (May 2012), pp. 689–702. DOI: [10.1109/TVCG.2011.287](https://doi.org/10.1109/TVCG.2011.287).
- [96] Sujin Jang et al. “Modeling Cumulative Arm Fatigue in Mid-Air Interaction Based on Perceived Exertion and Kinetics of Arm Motion”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17: CHI Conference on Human Factors in Computing Systems. Denver Colorado USA: ACM, May 2, 2017, pp. 3328–3339. DOI: [10.1145/3025453.3025523](https://doi.org/10.1145/3025453.3025523).
- [97] Jérôme Jetter et al. “Augmented Reality Tools for Industrial Applications: What Are Potential Key Performance Indicators and Who Benefits?” In: *Computers in Human Behavior* 87 (Oct. 1, 2018), pp. 18–33. DOI: [10.1016/j.chb.2018.04.054](https://doi.org/10.1016/j.chb.2018.04.054).
- [98] Robert Johansen et al. *Groupware: Computer Support for Business Teams*. Free Press, 1988. 232 pp. Google Books: [TicTAQAAMAJ](https://books.google.com/books?id=TIcTAQAAMAJ).
- [99] Peter Johnson-Lenz and Trudy Johnson-Lenz. “Consider the Groupware: Design and Group Process Impacts on Communication in the Electronic Medium”. In: *Studies of Computer-Mediated Communications Systems: A Synthesis of the Findings* 16 (1981).
- [100] Alexander Kan et al. “Mixed Reality Stories: How Creative Writers Integrate Virtual and Physical Realities”. In: *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*. OzCHI ’14. New York, NY, USA: Association for Computing Machinery, Dec. 2, 2014, pp. 115–118. DOI: [10.1145/2686612.2686628](https://doi.org/10.1145/2686612.2686628).
- [101] Shunichi Kasahara et al. “Second Surface: Multi-user Spatial Collaboration System Based on Augmented Reality”. In: *SIGGRAPH Asia 2012 Emerging Technologies*. Sa ’12. New York, NY, USA: Association for Computing Machinery, 2012, pp. 1–4. DOI: [10.1145/2407707.2407727](https://doi.org/10.1145/2407707.2407727).
- [102] Hannes Kaufmann and Dieter Schmalstieg. “Mathematics and Geometry Education with Collaborative Augmented Reality”. In: *ACM SIGGRAPH 2002 Conference Abstracts and Applications*. SIGGRAPH ’02. New York, NY, USA: Association for Computing Machinery, July 21, 2002, pp. 37–41. DOI: [10.1145/1242073.1242086](https://doi.org/10.1145/1242073.1242086).
- [103] Jens Keil et al. “A Digital Look at Physical Museum Exhibits: Designing Personalized Stories with Handheld Augmented Reality in Museums”. In: *2013 Digital Heritage International Congress (DigitalHeritage)*. 2013 Digital Heritage International Congress (DigitalHeritage). Vol. 2. Oct. 2013, pp. 685–688. DOI: [10.1109/DigitalHeritage.2013.6744836](https://doi.org/10.1109/DigitalHeritage.2013.6744836).
- [104] Adam Kendon. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. CUP Archive, Nov. 30, 1990. 308 pp.
- [105] Thomas Kilgus et al. “Mobile Markerless Augmented Reality and Its Application in Forensic Medicine”. In: *International journal of computer assisted radiology and surgery* 10 (2015), pp. 573–586. DOI: [10.1007/s11548-014-1106-9](https://doi.org/10.1007/s11548-014-1106-9).



- [106] Kangsoo Kim et al. “Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008–2017)”. In: *IEEE Transactions on Visualization and Computer Graphics* 24.11 (Nov. 2018), pp. 2947–2962. DOI: [10.1109/TVCG.2018.2868591](https://doi.org/10.1109/TVCG.2018.2868591).
- [107] David Kirk et al. “Ways of the Hands”. In: *ECSCW 2005*. Ed. by Hans Gellersen et al. Dordrecht: Springer Netherlands, 2005, pp. 1–21. DOI: [10.1007/1-4020-4023-7\\_1](https://doi.org/10.1007/1-4020-4023-7_1).
- [108] K. Kiyokawa et al. “A Collaboration Support Technique by Integrating a Shared Virtual Reality and a Shared Augmented Reality”. In: *IEEE SMC’99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.99CH37028)*. IEEE SMC’99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics. Vol. 6. Tokyo, Japan: IEEE, 1999, pp. 48–53. DOI: [10.1109/ICSMC.1999.816444](https://doi.org/10.1109/ICSMC.1999.816444).
- [109] Hoshang Kolivand et al. “LivePhantom: Retrieving Virtual World Light Data to Real Environments”. In: *PLOS ONE* 11.12 (Dec. 8, 2016), e0166424. DOI: [10.1371/journal.pone.0166424](https://doi.org/10.1371/journal.pone.0166424).
- [110] George Koutromanos et al. “The Use of Augmented Reality Games in Education: A Review of the Literature”. In: *Educational Media International* 52.4 (Oct. 2, 2015), pp. 253–271. DOI: [10.1080/09523987.2015.1125988](https://doi.org/10.1080/09523987.2015.1125988).
- [111] Peter Gall Krogh et al. “Sensitizing Concepts for Socio-spatial Literacy in HCI”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI ’17. New York, NY, USA: Association for Computing Machinery, May 2, 2017, pp. 6449–6460. DOI: [10.1145/3025453.3025756](https://doi.org/10.1145/3025453.3025756).
- [112] Teemu H. Laine. “Mobile Educational Augmented Reality Games: A Systematic Literature Review and Two Case Studies”. In: *Computers* 7.1 (1 Mar. 2018), p. 19. DOI: [10.3390/computers7010019](https://doi.org/10.3390/computers7010019).
- [113] Magic Leap. *Magic Leap*. 2018. URL: <https://www.magicleap.com>.
- [114] David Ledo et al. “Evaluation Strategies for HCI Toolkit Research”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. New York, NY, USA: Association for Computing Machinery, Apr. 19, 2018, pp. 1–17. DOI: [10.1145/3173574.3173610](https://doi.org/10.1145/3173574.3173610).
- [115] Dumi Lee. *Ikea Place Is an AR App That Lets You Put Furniture on the Street*. 2017.
- [116] Wenkai Li et al. “A State-of-the-Art Review of Augmented Reality in Engineering Analysis and Simulation”. In: *Multimodal Technologies and Interaction* 1.3 (3 Sept. 2017), p. 17. DOI: [10.3390/mti1030017](https://doi.org/10.3390/mti1030017).
- [117] Yuan Li et al. “The Effects of Incorrect Occlusion Cues on the Understanding of Barehanded Referencing in Collaborative Augmented Reality”. In: *Frontiers in Virtual Reality* 2 (July 1, 2021). DOI: [10.3389/frvir.2021.681585](https://doi.org/10.3389/frvir.2021.681585).

- [118] Andrés Lucero et al. “Collaborative Use of Mobile Phones for Brainstorming”. In: *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*. MobileHCI '10. New York, NY, USA: Association for Computing Machinery, Sept. 7, 2010, pp. 337–340. DOI: [10.1145/1851600.1851659](https://doi.org/10.1145/1851600.1851659).
- [119] Andrés Lucero et al. “Pass-Them-around: Collaborative Use of Mobile Phones for Photo Sharing”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, May 7, 2011, pp. 1787–1796.
- [120] Stephan Lukosch et al. “Collaboration in Augmented Reality”. In: *Computer Supported Cooperative Work (CSCW)* 24.6 (Dec. 1, 2015), pp. 515–525. DOI: [10.1007/s10606-015-9239-0](https://doi.org/10.1007/s10606-015-9239-0).
- [121] Sus Lundgren et al. “Designing Mobile Experiences for Collocated Interaction”. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW '15. New York, NY, USA: Association for Computing Machinery, Feb. 28, 2015, pp. 496–507. DOI: [10.1145/2675133.2675171](https://doi.org/10.1145/2675133.2675171).
- [122] Mille Skovhus Lunding et al. “Reality and Beyond: Proxemics as a Lens for Designing Handheld Collaborative Augmented Reality”. In: *Proc. ACM Hum.-Comput. Interact.* 7 (ISS Nov. 1, 2023), 427:21–427:40. DOI: [10.1145/3626463](https://doi.org/10.1145/3626463).
- [123] Eva Babette Mackamul and Augusto Esteves. “A Look at the Effects of Handheld and Projected Augmented-reality on a Collaborative Task”. In: *Proceedings of the Symposium on Spatial User Interaction*. SUI '18. New York, NY, USA: Association for Computing Machinery, Oct. 13, 2018, pp. 74–78. DOI: [10.1145/3267782.3267793](https://doi.org/10.1145/3267782.3267793).
- [124] K. Mahadevan et al. ““Grip-that-there”: An Investigation of Explicit and Implicit Task Allocation Techniques for Human-Robot Collaboration”. In: DOI: [10.1145/3411764.3445355](https://doi.org/10.1145/3411764.3445355).
- [125] Nicolai Marquardt et al. “The Continuous Interaction Space: Interaction Techniques Unifying Touch and Gesture on and above a Digital Surface”. In: *Human-Computer Interaction – INTERACT 2011*. Ed. by Pedro Campos et al. Red. by David Hutchison et al. Vol. 6948. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 461–476. DOI: [10.1007/978-3-642-23765-2\\_32](https://doi.org/10.1007/978-3-642-23765-2_32).
- [126] Nicolai Marquardt et al. “Cross-Device Interaction via Micro-Mobility and f-Formations”. In: *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*. UIST '12. New York, NY, USA: Association for Computing Machinery, Oct. 7, 2012, pp. 13–22. DOI: [10.1145/2380116.2380121](https://doi.org/10.1145/2380116.2380121).
- [127] Nicolai Marquardt et al. “SurfaceConstellations: A Modular Hardware Platform for Ad-Hoc Reconfigurable Cross-Device Workspaces”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18: CHI Conference on Human Factors in Computing Systems. Montreal QC Canada: ACM, Apr. 21, 2018, pp. 1–14. DOI: [10.1145/3173574.3173928](https://doi.org/10.1145/3173574.3173928).

- [128] Bernard Marr. *9 Powerful Real-World Applications Of Augmented Reality (AR) Today*. Forbes. URL: <https://www.forbes.com/sites/bernardmarr/2018/07/30/9-powerful-real-world-applications-of-augmented-reality-ar-today/> (visited on 05/10/2022).
- [129] Paul Marshall et al. "Using F-formations to Analyse Spatial Patterns of Interaction in Physical Environments". In: *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work - CSCW '11*. The ACM 2011 Conference. Hangzhou, China: ACM Press, 2011, p. 445. DOI: [10.1145/1958824.1958893](https://doi.org/10.1145/1958824.1958893).
- [130] Will McGrath et al. "Branch-Explore-Merge: Facilitating Real-Time Revision Control in Collaborative Visual Exploration". In: *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*. ITS '12. New York, NY, USA: Association for Computing Machinery, Nov. 11, 2012, pp. 235–244. DOI: [10.1145/2396636.2396673](https://doi.org/10.1145/2396636.2396673).
- [131] Helena M. Mentis et al. "Interaction Proxemics and Image Use in Neurosurgery". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. New York, NY, USA: Association for Computing Machinery, May 5, 2012, pp. 927–936. DOI: [10.1145/2207676.2208536](https://doi.org/10.1145/2207676.2208536).
- [132] Microsoft. *HoloLens*. 2016.
- [133] Paul Milgram and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays". In: *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994), pp. 1321–1329.
- [134] Minecraft. *Minecraft*. Minecraft.net. URL: <https://www.minecraft.net/en-us> (visited on 02/25/2025).
- [135] Shalini Misra et al. "The iPhone Effect: The Quality of In-Person Social Interactions in the Presence of Mobile Devices". In: *Environment and Behavior* 48.2 (Feb. 1, 2016), pp. 275–298. DOI: [10.1177/0013916514539755](https://doi.org/10.1177/0013916514539755).
- [136] M. Modell and C. Gray. "Searching for Personal Territory in a Human-Computer Interaction Design Studio". In: *Journal for Education in the Built Environment* 6.2 (), pp. 54–78. DOI: [10.11120/jebe.2011.06020054](https://doi.org/10.11120/jebe.2011.06020054).
- [137] D. Mogilev et al. "AR Pad: An Interface for Face-to-Face AR Collaboration". In: *CHI '02 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '02. New York, NY, USA: Association for Computing Machinery, Apr. 20, 2002, pp. 654–655. DOI: [10.1145/506443.506530](https://doi.org/10.1145/506443.506530).
- [138] Dimitris Mourtzis et al. "Augmented Reality Application to Support Remote Maintenance as a Service in the Robotics Industry". In: *Procedia Cirp* 63 (2017), pp. 46–51. DOI: [10.1016/j.procir.2017.03.154](https://doi.org/10.1016/j.procir.2017.03.154).
- [139] Epson Moverio. *Moverio BT-40 / See-Through Mobile Viewer / Smart Glasses / Products / Epson United Kingdom*. URL: [https://www.epson.co.uk/en\\_GB/products/smart-glasses/see-through-mobile-viewer/moverio-bt-40/p/31095](https://www.epson.co.uk/en_GB/products/smart-glasses/see-through-mobile-viewer/moverio-bt-40/p/31095) (visited on 02/25/2025).
- [140] Tek-Jin Nam and Kyung Sakong. "Collaborative 3D Workspace and Interaction Techniques for Synchronous Distributed Product Design Reviews". In: (2009).

- [141] Wolfgang Narzt et al. “Augmented Reality Navigation Systems”. In: *Universal Access in the Information Society* 4.3 (Mar. 1, 2006), pp. 177–187. DOI: [10.1007/s10209-005-0017-5](https://doi.org/10.1007/s10209-005-0017-5).
- [142] U. Neumann and A. Majoros. “Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance”. In: *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180)*. Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180). Mar. 1998, pp. 4–11. DOI: [10.1109/VRAIS.1998.658416](https://doi.org/10.1109/VRAIS.1998.658416).
- [143] Niantic. *Pokémon GO*.
- [144] M. Nigro et al. “An Interactive Augmented Reality Interface for Personalized Proxemics Modeling: Comfort and Human–Robot Interactions”. In: *IEEE Robotics & Automation Magazine* (), pp. 2–11. DOI: [10.1109/mra.2024.3415108](https://doi.org/10.1109/mra.2024.3415108).
- [145] A. Nijholt. “Social Augmented Reality: A Multiperspective Survey”. In: (), pp. 1–8. DOI: [10.1109/icievicivpr52578.2021.9564182](https://doi.org/10.1109/icievicivpr52578.2021.9564182).
- [146] Susanna Nilsson et al. “Using AR to Support Cross-Organisational Collaboration in Dynamic Tasks”. In: *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. 2009 8th IEEE International Symposium on Mixed and Augmented Reality. Oct. 2009, pp. 3–12. DOI: [10.1109/ISMAR.2009.5336522](https://doi.org/10.1109/ISMAR.2009.5336522).
- [147] Donald A. Norman. *The Psychology of Everyday Things*. Basic Books, 1988. 257 pp. Google Books: [0lNSRAAACAAJ](https://books.google.com/books?id=0lNSRAAACAAJ).
- [148] Max M. North and Sarah M. North. “A Comparative Study of Sense of Presence of Traditional Virtual Reality and Immersive Environments”. In: *Australasian Journal of Information Systems* 20 (Feb. 24, 2016). DOI: [10.3127/ajis.v20i0.1168](https://doi.org/10.3127/ajis.v20i0.1168).
- [149] Erick Oduor et al. “The Frustrations and Benefits of Mobile Device Usage in the Home When Co-Present with Family Members”. In: *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. DIS ’16: Designing Interactive Systems Conference 2016. Brisbane QLD Australia: ACM, June 4, 2016, pp. 1315–1327. DOI: [10.1145/2901790.2901809](https://doi.org/10.1145/2901790.2901809).
- [150] Gary M. Olson and Judith S. Olson. “Distance Matters”. In: *Human–Computer Interaction* 15.2-3 (Sept. 1, 2000), pp. 139–178. DOI: [10.1207/S15327051HCI1523\\_4](https://doi.org/10.1207/S15327051HCI1523_4).
- [151] Paulo Pérez et al. “Proxemic Environments: A Framework for Developing Mobile Applications Based on Proxemic Interactions”. In: *Annals of Computer Science and Information Systems* 21 (2020). DOI: [10.15439/2020F29](https://doi.org/10.15439/2020F29).
- [152] B. Petrak et al. “Let Me Show You Your New Home: Studying the Effect of Proxemic-Awareness of Robots on Users’ First Impressions”. In: DOI: [10.1109/ro-man46459.2019.8956463](https://doi.org/10.1109/ro-man46459.2019.8956463).
- [153] David Pinelle et al. “Task Analysis for Groupware Usability Evaluation: Modeling Shared-Workspace Tasks with the Mechanics of Collaboration”. In: *ACM Transactions on Computer-Human Interaction* 10.4 (Dec. 1, 2003), pp. 281–311. DOI: [10.1145/966930.966932](https://doi.org/10.1145/966930.966932).

- [154] Thomas Plank et al. “Is Two Enough? ! Studying Benefits, Barriers, and Biases of Multi-Tablet Use for Collaborative Visualization”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, May 2, 2017, pp. 4548–4560.
- [155] Jarkko Polvi et al. “Handheld Guides in Inspection Tasks: Augmented Reality versus Picture”. In: *IEEE transactions on visualization and computer graphics* 24.7 (2017), pp. 2118–2128. DOI: [10.1109/TVCG.2017.2709746](https://doi.org/10.1109/TVCG.2017.2709746).
- [156] Lev Poretski Poretski et al. “Physicality As an Anchor for Coordination: Examining Collocated Collaboration in Physical and Mobile Augmented Reality Settings”. In: *Proceedings of the ACM on Human-Computer Interaction* 5 (CSCW2 Oct. 13, 2021), pp. 1–29. DOI: [10.1145/3479857](https://doi.org/10.1145/3479857).
- [157] Giambattista della Porta. *Natural Magick*. London: Printed for T. Young and S. Speed, 1584. 3 pp.
- [158] Erik Prytz et al. “The Importance of Eye-Contact for Collaboration in AR Systems”. In: *2010 IEEE International Symposium on Mixed and Augmented Reality*. 2010 IEEE International Symposium on Mixed and Augmented Reality. Oct. 2010, pp. 119–126. DOI: [10.1109/ISMAR.2010.5643559](https://doi.org/10.1109/ISMAR.2010.5643559).
- [159] QuiverVision. *QuiverVision*.
- [160] Rafael Radkowski et al. “Augmented Reality-Based Manual Assembly Support with Visual Features for Different Degrees of Difficulty”. In: *International Journal of Human-Computer Interaction* 31.5 (2015), pp. 337–349. DOI: [10.1080/10447318.2014.994194](https://doi.org/10.1080/10447318.2014.994194).
- [161] Roman Rädle et al. “HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration”. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. ITS ’14. New York, NY, USA: Association for Computing Machinery, Nov. 16, 2014, pp. 45–54. DOI: [10.1145/2669485.2669500](https://doi.org/10.1145/2669485.2669500).
- [162] Roman Rädle et al. “Spatially-Aware or Spatially-agnostic? Elicitation and Evaluation of User-Defined Cross-Device Interactions”. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, Apr. 18, 2015, pp. 3913–3922.
- [163] Iulian Radu et al. “A Survey of Needs and Features for Augmented Reality Collaborations in Collocated Spaces”. In: *Proceedings of the ACM on Human-Computer Interaction* 5 (CSCW1 Apr. 13, 2021), pp. 1–21. DOI: [10.1145/3449243](https://doi.org/10.1145/3449243).
- [164] G. Reitmayr and D. Schmalstieg. “Mobile Collaborative Augmented Reality”. In: *Proceedings IEEE and ACM International Symposium on Augmented Reality*. Proceedings IEEE and ACM International Symposium on Augmented Reality. Oct. 2001, pp. 114–123. DOI: [10.1109/ISAR.2001.970521](https://doi.org/10.1109/ISAR.2001.970521).
- [165] Jun Rekimoto. “Transvision: A Hand-Held Augmented Reality System For Collaborative Design”. In: *Proceeding of Virtual Systems and Multimedia*. Vol. 96. 1996, pp. 18–20.



- [166] Jun Rekimoto and Katashi Nagao. “The World through the Computer: Computer Augmented Interaction with Real World Environments”. In: *Proceedings of the 8th Annual ACM Symposium on User Interface and Software Technology*. UIST '95. New York, NY, USA: Association for Computing Machinery, Dec. 1, 1995, pp. 29–36. DOI: [10.1145/215585.215639](https://doi.org/10.1145/215585.215639).
- [167] Walter J Rezende et al. “Use of Augmented Reality to Support Education”. In: *Creating a Mobile E-learning Tool and using it with an Inquiry-based Approach 1* (2017), pp. 100–107.
- [168] Giuseppe Riva et al. “Interreality in Practice: Bridging Virtual and Real Worlds in the Treatment of Posttraumatic Stress Disorders”. In: *Cyberpsychology, Behavior, and Social Networking* 13.1 (Feb. 2010), pp. 55–65. DOI: [10.1089/cyber.2009.0320](https://doi.org/10.1089/cyber.2009.0320).
- [169] Yvonne Rogers and Siân Lindley. “Collaborating around Vertical and Horizontal Large Interactive Displays: Which Way Is Best?” In: *Interacting with Computers* 16.6 (Dec. 2004), pp. 1133–1152. DOI: [10.1016/j.intcom.2004.07.008](https://doi.org/10.1016/j.intcom.2004.07.008).
- [170] S Rutherford. *Meet Endgame: Proving Ground - Google's AR Mobile Game / Tom's Guide*. 2015.
- [171] Pratiti Sarkar et al. “ScholAR: A Collaborative Learning Experience for Rural Schools Using Augmented Reality Application”. In: *2018 IEEE Tenth International Conference on Technology for Education (T4E)*. 2018 IEEE Tenth International Conference on Technology for Education (T4E). Dec. 2018, pp. 8–15. DOI: [10.1109/T4E.2018.00010](https://doi.org/10.1109/T4E.2018.00010).
- [172] Pratiti Sarkar et al. “Learners’ Approaches, Motivation and Patterns of Problem-Solving on Lines and Angles in Geometry Using Augmented Reality”. In: *Smart Learning Environments* 7.1 (2020), pp. 1–23. DOI: [10.1186/s40561-020-00124-9](https://doi.org/10.1186/s40561-020-00124-9).
- [173] Kim Sauvé et al. “A Change of Perspective: How User Orientation Influences the Perception of Physicalizations”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20: CHI Conference on Human Factors in Computing Systems. Honolulu HI USA: ACM, Apr. 21, 2020, pp. 1–12. DOI: [10.1145/3313831.3376312](https://doi.org/10.1145/3313831.3376312).
- [174] Nina Savela et al. “Does Augmented Reality Affect Sociability, Entertainment, and Learning? A Field Experiment”. In: *Applied Sciences* 10.4 (4 Jan. 2020), p. 1392. DOI: [10.3390/app10041392](https://doi.org/10.3390/app10041392).
- [175] Dieter Schmalstieg and Tobias Höllerer. “Augmented Reality: Principles and Practice”. In: *2017 IEEE Virtual Reality (VR)*. 2017 IEEE Virtual Reality (VR). Mar. 2017, pp. 425–426. DOI: [10.1109/VR.2017.7892358](https://doi.org/10.1109/VR.2017.7892358).
- [176] Dieter Schmalstieg and Daniel Wagner. “Experiences with Handheld Augmented Reality”. In: *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. Nov. 2007, pp. 3–18. DOI: [10.1109/ISMAR.2007.4538819](https://doi.org/10.1109/ISMAR.2007.4538819).

- [177] Dieter Schmalstieg et al. “The Studierstube Augmented Reality Project”. In: *Presence: Teleoperators and Virtual Environments* 11.1 (Feb. 1, 2002), pp. 33–54. DOI: [10.1162/105474602317343640](https://doi.org/10.1162/105474602317343640).
- [178] Dominik Schmidt et al. “A Cross-Device Interaction Style for Mobiles and Surfaces”. In: *Proceedings of the Designing Interactive Systems Conference on - DIS '12*. The Designing Interactive Systems Conference. Newcastle Upon Tyne, United Kingdom: ACM Press, 2012, p. 318. DOI: [10.1145/2317956.2318005](https://doi.org/10.1145/2317956.2318005).
- [179] Kjeld Schmidt. “The Problem Withawareness’: Introductory Remarks Onawareness in CSCW”. In: *Computer Supported Cooperative Work (CSCW)* 11.3 (2002), pp. 285–298. DOI: [10.1023/A:1021272909573](https://doi.org/10.1023/A:1021272909573).
- [180] Kjeld Schmidt. “Cooperative Work and Coordinative Practices”. In: *Cooperative Work and Coordinative Practices: Contributions to the Conceptual Foundations of Computer-Supported Cooperative Work (CSCW)*. Ed. by Kjeld Schmidt. Computer Supported Cooperative Work. London: Springer, 2011, pp. 3–27. DOI: [10.1007/978-1-84800-068-1\\_1](https://doi.org/10.1007/978-1-84800-068-1_1).
- [181] Kjeld Schmidt and Carla Simonee. “Coordination Mechanisms: Towards a Conceptual Foundation of CSCW Systems Design”. In: *Computer Supported Cooperative Work (CSCW)* 5.2 (June 1, 1996), pp. 155–200. DOI: [10.1007/BF00133655](https://doi.org/10.1007/BF00133655).
- [182] Mario Schreiner et al. “Connichiwa: A Framework for Cross-Device Web Applications”. In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA ’15. New York, NY, USA: Association for Computing Machinery, Apr. 18, 2015, pp. 2163–2168. DOI: [10.1145/2702613.2732909](https://doi.org/10.1145/2702613.2732909).
- [183] Jan-Henrik Schröder et al. “Collaborating Across Realities: Analytical Lenses for Understanding Dyadic Collaboration in Transitional Interfaces”. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI ’23. New York, NY, USA: Association for Computing Machinery, Apr. 19, 2023, pp. 1–16. DOI: [10.1145/3544548.3580879](https://doi.org/10.1145/3544548.3580879).
- [184] Stacey D Scott. “Territory-Based Interaction Techniques for Tabletop Collaboration”. In: (2003).
- [185] Stacey D. Scott and Sheelagh Carpendale. “Theory of Tabletop Territoriality”. In: *Tabletops - Horizontal Interactive Displays*. Ed. by Christian Müller-Tomfelde. London: Springer, 2010, pp. 357–385. DOI: [10.1007/978-1-84996-113-4\\_15](https://doi.org/10.1007/978-1-84996-113-4_15).
- [186] Stacey D. Scott et al. “Territoriality in Collaborative Tabletop Workspaces”. In: *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. CSCW ’04. New York, NY, USA: Association for Computing Machinery, Nov. 6, 2004, pp. 294–303. DOI: [10.1145/1031607.1031655](https://doi.org/10.1145/1031607.1031655).
- [187] Julian Seifert et al. “MobiSurf: Improving Co-Located Collaboration through Integrating Mobile Devices and Interactive Surfaces”. In: *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*. ITS ’12. New York, NY, USA: Association for Computing Machinery, Nov. 11, 2012, pp. 51–60. DOI: [10.1145/2396636.2396644](https://doi.org/10.1145/2396636.2396644).

- [188] Mickael Sereno et al. “Collaborative Work in Augmented Reality: A Survey”. In: *IEEE Transactions on Visualization and Computer Graphics* 28.6 (June 2022), pp. 2530–2549. DOI: [10.1109/TVCG.2020.3032761](https://doi.org/10.1109/TVCG.2020.3032761).
- [189] Joon Gi Shin et al. “Couples Designing Their Living Room Together: A Study with Collaborative Handheld Augmented Reality”. In: *Proceedings of the 9th Augmented Human International Conference*. Ah '18. New York, NY, USA: Association for Computing Machinery, 2018. DOI: [10.1145/3174910.3174930](https://doi.org/10.1145/3174910.3174930).
- [190] Jeffrey H Shuhaiber. “Augmented Reality in Surgery”. In: *Archives of surgery* 139.2 (2004), pp. 170–174. DOI: [10.1001/archsurg.139.2.170](https://doi.org/10.1001/archsurg.139.2.170).
- [191] Ludwig Sidenmark et al. “Outline Pursuits: Gaze-assisted Selection of Occluded Objects in Virtual Reality”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20: CHI Conference on Human Factors in Computing Systems. Honolulu HI USA: ACM, Apr. 21, 2020, pp. 1–13. DOI: [10.1145/3313831.3376438](https://doi.org/10.1145/3313831.3376438).
- [192] Gaurav Sinha et al. “Human Computer Interaction”. In: *2010 3rd International Conference on Emerging Trends in Engineering and Technology*. 2010 3rd International Conference on Emerging Trends in Engineering and Technology. Nov. 2010, pp. 1–4. DOI: [10.1109/ICETET.2010.85](https://doi.org/10.1109/ICETET.2010.85).
- [193] Thad Starner. “Project Glass: An Extension of the Self”. In: *IEEE Pervasive Computing* 12.2 (2013), pp. 14–16.
- [194] Anthony Steed et al. “Immersive Competence and Immersive Literacy: Exploring How Users Learn about Immersive Experiences”. In: *Frontiers in Virtual Reality* 4 (2023).
- [195] Anselm Strauss. “Work and the Division of Labor”. In: *The Sociological Quarterly* 26.1 (Mar. 1, 1985), pp. 1–19. DOI: [10.1111/j.1533-8525.1985.tb00212.x](https://doi.org/10.1111/j.1533-8525.1985.tb00212.x).
- [196] Ivan E. Sutherland. “A Head-Mounted Three Dimensional Display”. In: *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*. AFIPS '68 (Fall, Part I). New York, NY, USA: Association for Computing Machinery, Dec. 9, 1968, pp. 757–764. DOI: [10.1145/1476589.1476686](https://doi.org/10.1145/1476589.1476686).
- [197] Toqeer Ali Syed et al. “In-Depth Review of Augmented Reality: Tracking Technologies, Development Tools, AR Displays, Collaborative AR, and Security Concerns”. In: *Sensors* 23.1 (1 Jan. 2023), p. 146. DOI: [10.3390/s23010146](https://doi.org/10.3390/s23010146).
- [198] Brandon Victor Syiem et al. “Impact of Task on Attentional Tunneling in Handheld Augmented Reality”. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21: CHI Conference on Human Factors in Computing Systems. Yokohama Japan: ACM, May 6, 2021, pp. 1–14. DOI: [10.1145/3411764.3445580](https://doi.org/10.1145/3411764.3445580).
- [199] Z. Szalavári et al. ““Studierstube”: An Environment for Collaboration in Augmented Reality”. In: *Virtual Reality* 3.1 (Mar. 1, 1998), pp. 37–48. DOI: [10.1007/BF01409796](https://doi.org/10.1007/BF01409796).



- [200] Anthony Tang et al. “Collaborative Coupling over Tabletop Displays”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '06. New York, NY, USA: Association for Computing Machinery, Apr. 22, 2006, pp. 1181–1190. DOI: [10.1145/1124772.1124950](https://doi.org/10.1145/1124772.1124950).
- [201] Lili Tong et al. “It’s Not How You Stand, It’s How You Move: F-formations and Collaboration Dynamics in a Mobile Learning Game”. In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. MobileHCI '16. New York, NY, USA: Association for Computing Machinery, Sept. 6, 2016, pp. 318–329. DOI: [10.1145/2935334.2935343](https://doi.org/10.1145/2935334.2935343).
- [202] Marcus Tönnis et al. “Representing Information – Classifying the Augmented Reality Presentation Space”. In: *Computers & Graphics* 37.8 (Dec. 1, 2013), pp. 997–1011. DOI: [10.1016/j.cag.2013.09.002](https://doi.org/10.1016/j.cag.2013.09.002).
- [203] Crispino Tosto et al. “The Potential of AR Solutions for Behavioral Learning: A Scoping Review”. In: *Computers* 11.6 (6 June 2022), p. 87. DOI: [10.3390/computers11060087](https://doi.org/10.3390/computers11060087).
- [204] Sherry Turkle. *Alone Together: Why We Expect More from Technology and Less from Each Other*. 3rd ed. edition. New York: Basic Books, Nov. 30, 2017. 400 pp.
- [205] I. Tussyadiah. “Territoriality and Consumption Behaviour with Location-Based Media”. In: pp. 249–259. DOI: [10.1007/978-3-7091-1142-0\\_22](https://doi.org/10.1007/978-3-7091-1142-0_22).
- [206] Almar van der Stappen et al. “MathBuilder: A Collaborative AR Math Game for Elementary School Students”. In: *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. CHI PLAY '19 Extended Abstracts. New York, NY, USA: Association for Computing Machinery, 2019, pp. 731–738. DOI: [10.1145/3341215.3356295](https://doi.org/10.1145/3341215.3356295).
- [207] Petr Vávra et al. “Recent Development of Augmented Reality in Surgery: A Review”. In: *Journal of healthcare engineering* 2017.1 (2017), p. 4574172. DOI: [10.1155/2017/4574172](https://doi.org/10.1155/2017/4574172).
- [208] Ana Villanueva et al. “Meta-AR-App: An Authoring Platform for Collaborative Augmented Reality in STEM Classrooms”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: Association for Computing Machinery, Apr. 23, 2020, pp. 1–14. DOI: [10.1145/3313831.3376146](https://doi.org/10.1145/3313831.3376146).
- [209] Vuforia. *Vuforia Enterprise Augmented Reality (AR) Software / PTC*. URL: <https://www.ptc.com/en/products/vuforia> (visited on 02/25/2025).
- [210] Vuzix. *Vuzix / Heads-Up, Hands-Free AR Smart Glasses*. Vuzix Corporation. URL: <https://www.vuzix.com/> (visited on 02/25/2025).
- [211] Zachary Walker et al. “Beyond Pokémon: Augmented Reality Is a Universal Design for Learning Tool”. In: *Sage Open* 7.4 (Oct. 1, 2017), p. 2158244017737815. DOI: [10.1177/2158244017737815](https://doi.org/10.1177/2158244017737815).
- [212] James R. Wallace et al. “Investigating the Role of a Large, Shared Display in Multi-Display Environments”. In: *Computer Supported Cooperative Work (CSCW)* 20.6 (Oct. 4, 2011), p. 529. DOI: [10.1007/s10606-011-9149-8](https://doi.org/10.1007/s10606-011-9149-8).

- [213] James R. Wallace et al. “Collaborative Sensemaking on a Digital Tabletop and Personal Tablets: Prioritization, Comparisons, and Tableaux”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. New York, NY, USA: Association for Computing Machinery, Apr. 27, 2013, pp. 3345–3354. DOI: [10.1145/2470654.2466458](https://doi.org/10.1145/2470654.2466458).
- [214] Miaosen Wang et al. “Proxemic Peddler: A Public Advertising Display That Captures and Preserves the Attention of a Passerby”. In: *Proceedings of the 2012 International Symposium on Pervasive Displays*. PerDis '12. New York, NY, USA: Association for Computing Machinery, June 4, 2012, pp. 1–6. DOI: [10.1145/2307798.2307801](https://doi.org/10.1145/2307798.2307801).
- [215] Yitong Wang and Jun Zhao. *Mobile Edge Computing, Metaverse, 6G Wireless Communications, Artificial Intelligence, and Blockchain: Survey and Their Convergence*. Sept. 28, 2022. DOI: [10.48550/arXiv.2209.14147](https://doi.org/10.48550/arXiv.2209.14147). arXiv: [2209.14147 \[cs\]](https://arxiv.org/abs/2209.14147). URL: <http://arxiv.org/abs/2209.14147> (visited on 03/04/2024). Pre-published.
- [216] Sabine Webel et al. “Augmented Reality Training for Assembly and Maintenance Skills”. In: *BIO Web of Conferences*. Vol. 1. EDP Sciences, 2011, p. 00097. DOI: [10.1051/bioconf/20110100097](https://doi.org/10.1051/bioconf/20110100097).
- [217] Thomas Wells and Steven Houben. “CollabAR - Investigating the Mediating Role of Mobile AR Interfaces on Co-Located Group Collaboration”. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–13. DOI: [10.1145/3313831.3376541](https://doi.org/10.1145/3313831.3376541).
- [218] Thomas Wells et al. “A Study into the Effect of Mobile Device Configurations on Co-Located Collaboration Using AR”. In: *Proceedings of the ACM on Human-Computer Interaction* 6 (MHCI Sept. 20, 2022), 200:1–200:23. DOI: [10.1145/3546735](https://doi.org/10.1145/3546735).
- [219] E. Wong et al. “Practice-Informed Patterns for Organising Large Groups in Distributed Mixed Reality Collaboration”. In: pp. 1–18. DOI: [10.1145/3613904.3642502](https://doi.org/10.1145/3613904.3642502).
- [220] Robert Xiao et al. “Toffee: Enabling Ad Hoc, around-Device Interaction with Acoustic Time-of-Arrival Correlation”. In: *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services - MobileHCI '14*. The 16th International Conference. Toronto, ON, Canada: ACM Press, 2014, pp. 67–76. DOI: [10.1145/2628363.2628383](https://doi.org/10.1145/2628363.2628383).
- [221] A. Yim et al. ““I Don’t Feel like It Is ‘Mine’ at All”: Assessing Wikipedia Editors’ Sense of Individual and Community Ownership”. In: *Written Communication* 41.3 (), pp. 419–448. DOI: [10.1177/07410883241242103](https://doi.org/10.1177/07410883241242103).
- [222] Cik Suhaimi Yusof et al. “Collaborative Augmented Reality for Chess Game in Handheld Devices”. In: *2019 IEEE Conference on Graphics and Media (GAME)*. IEEE, 2019, pp. 32–37. DOI: [10.1109/GAME47560.2019.8980979](https://doi.org/10.1109/GAME47560.2019.8980979).

- [223] Johannes Zagermann et al. “When Tablets Meet Tabletops: The Effect of Tabletop Size on Around-the-Table Collaboration with Personal Tablets”. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI’16: CHI Conference on Human Factors in Computing Systems. San Jose California USA: ACM, May 7, 2016, pp. 5470–5481. DOI: [10.1145/2858036.2858224](https://doi.org/10.1145/2858036.2858224).
- [224] Chen Zhao et al. “SideSwipe: Detecting in-Air Gestures around Mobile Devices Using Actual GSM Signal”. In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*. UIST ’14. New York, NY, USA: Association for Computing Machinery, Oct. 5, 2014, pp. 527–534. DOI: [10.1145/2642918.2647380](https://doi.org/10.1145/2642918.2647380).
- [225] Yuting Zhou et al. “A Meta-Analytic Review on Incorporating Virtual and Augmented Reality in Museum Learning”. In: *Educational Research Review* 36 (June 1, 2022), p. 100454. DOI: [10.1016/j.edurev.2022.100454](https://doi.org/10.1016/j.edurev.2022.100454).