Volumetric Hybrid Workspaces: Interactions with Objects in Remote and Co-located Telepresence

Conference Paper · May 2024

DOI: 10.1145/3613904.3642814

Andrew Irlitti
University of Melbourne

Thuong Hoang
Deakin University

Brandon Syiem
The University of Sydney

Volumetric Hybrid Workspaces: Interactions with Objects in Remote and Co-located Telepresence

Andrew Irlitti*
The University of Melbourne
Melbourne, Australia

Mesut Latifoglu
The University of Melbourne
Melbourne, Australia

Thuong Hoang
Deakin University
Geelong, Australia

Brandon Victor Syiem
Queensland University of Technology
Brisbane, Australia

Frank Vetere*
The University of Melbourne
Melbourne, Australia

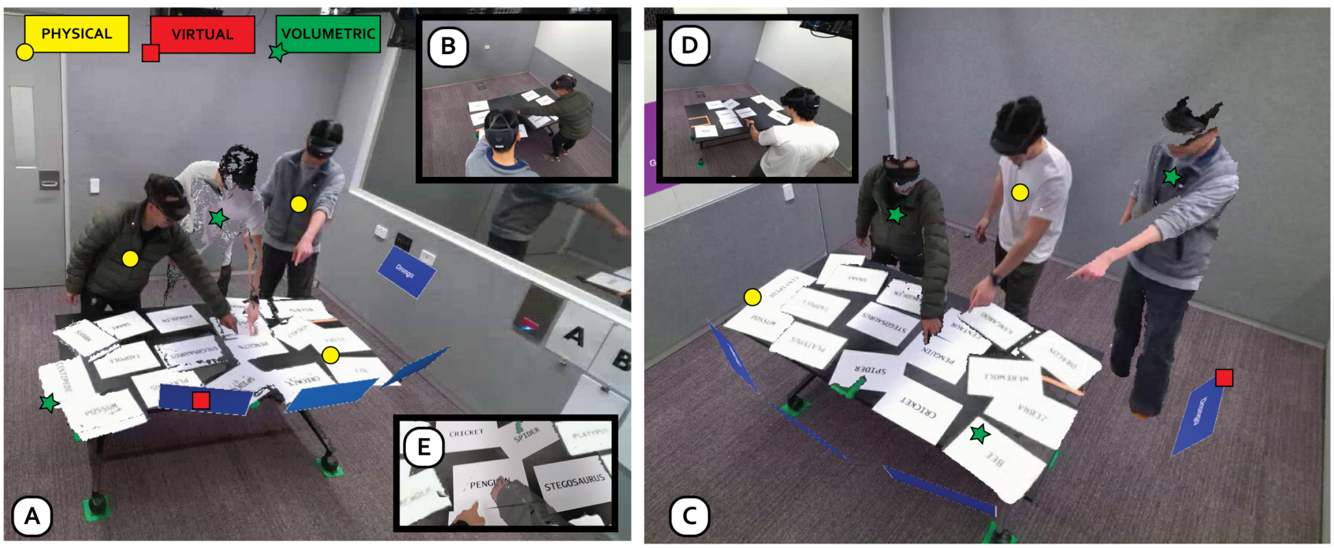


Figure 1: Hybrid Volumetric Telepresence across two rooms containing Physical, Volumetric and Virtual representations of objects and people. (A) Room 1: one remote volumetric participant (in center) working with two co-located collaborators. (B) Overhead view of Room 1 without the remote volumetric participant; (C) Room 2: one physical participant (in center) working with two remote volumetric collaborators (D) Overhead view of Room 2 without the remote volumetric participants; (E) First person view of cards from perspective of a participant wearing a Hololens 2.

ABSTRACT

Volumetric telepresence aims to create a shared space, allowing people in local and remote settings to collaborate seamlessly. Prior telepresence examples typically have asymmetrical designs, with volumetric capture in one location and objects in one format. In this paper, we present a volumetric telepresence mixed reality system that supports real-time, symmetrical, multi-user, partially distributed interactions, using objects in multiple formats, across multiple locations. We align two volumetric environments around a common spatial feature to create a shared workspace for remote and co-located people using objects in three formats: physical, virtual, and volumetric. We conducted a study with 18 participants over 6 sessions, evaluating how telepresence workspaces support spatial coordination and hybrid communication for co-located and remote users undertaking collaborative tasks. Our findings demonstrate the successful integration of remote spaces, effective use of proxemics and deixis to support negotiation, and strategies to manage interactivity in hybrid workspaces.

* ✉ andrew.irlitti@unimelb.edu.au — ✉ f.vetere@unimelb.edu.au

CHI '24, May 11–16, 2024, Honolulu, HI, USA

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Mixed / augmented reality**; **Empirical studies in HCI**.

KEYWORDS

augmented reality, mixed reality, volumetric capture, telepresence, collaboration, partially distributed teams, workspace awareness

ACM Reference Format:
Andrew Irlitti, Mesut Latifoglu, Thuong Hoang, Brandon Victor Syiem, and Frank Vetere. 2024. Volumetric Hybrid Workspaces: Interactions with Objects in Remote and Co-located Telepresence. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3613904.3642814

1 INTRODUCTION

There has been growing interest in exploring the impact of hybrid collaborative interfaces for distributed teams [18, 49, 60, 61]. In such scenarios, teams involve both co-located and remote people undertaking synchronous and asynchronous tasks using different technologies. For these distributed hybrid activities, teams may transition between different *coupling strategies* [27, 66], such as working collectively as a group on a single problem, or breaking off to focus on tasks either individually or as sub-groups [48, 49], creating a real-time *workspace awareness* [22] of each other's actions. Co-located users achieve this awareness by conveying both verbal and non-verbal communication cues [47, 66]. However, for remote users, conveying non-verbal communication is a challenge [18, 48, 68]. Screen-mediated approaches have focused on presenting a consistent frame of reference between remote sites, using synchronized whiteboard displays [48], combining video feeds from each local site into the shared display as *mirrors* [18], or synchronizing remote avatars in front of the shared displays [12].

As an alternative approach to screen-based solutions, telepresence technologies enable collaboration between people who are geographically separated. Mixed reality [9, 10] has brought additional dimensions to telepresence with the inclusion of varying 3D representation of remote collaborators as avatars, allowing for awareness cues to be conveyed beyond the boundary of the interaction space. Approaches towards avatar representation are commonly either abstract or cartoon-like in design [2, 12, 63], reconstructions from photogrammetry scans or videos [36, 40, 45], or incorporate continuous volumetric reconstructions using arrays of camera sensors as point cloud or 3D mesh [7, 37, 39], referred to as *volumetric telepresence* [25]. Volumetric telepresence is commonly implemented unidirectionally; one user leaves their physical surrounds and joins their collaborators local space [15, 25, 34]. While bi-directional approaches have been presented [51, 55] the experience has been limited to only two users, with attention focused on remote interactions. For distributed teams using mixed reality, there are numerous examples demonstrating the potential of grounding communication around a shared environment to increase workspace awareness [12, 19, 24, 63]. However, prior approaches are limited by the collaboration predominately relying on virtual elements, with remote collaborators represented as virtual cartoon avatars, or focus only on virtual interactions inside the shared space. Alternatively, the volumetric capture process is capable of capturing *everything* inside the space, including physical objects, which can then be used as part of the underlying communication process [11].

In this paper, we present a volumetric telepresence mixed reality platform to support interactions between partially distributed teams separated across multiple sites. The physical environments of each site, including both objects and people, are captured and transmitted in real-time, and rendered onto collaborators' mixed reality headsets. This process allows remote elements to spatially align within each local environment. The shared environment is constructed around a common spatial feature inside each space—such as a table, desk, workbench or floor area—with the surrounding interaction area creating a *volumetric hybrid workspace*. With two remote physical locations aligned, as physical elements are produced in one environment, they are reconstructed in real-time in the same location at the remote site.

The volumetric telepresence platform therefore captures people, objects, and space within the capture area. From the perspective of each collaborator, the hybrid nature of the shared workspace affords three types of objects:

- *Physical* objects that exist in the same physical environment as the collaborator.
- *Volumetric* objects are volumetric renderings of physical objects that exist in a remote environment.
- *Virtual* objects that exist only in the virtual environment. They do not have a physical representation.

This categorization of objects also includes people, who can be physically present, volumetrically rendered, or virtually created (i.e., as 3D virtual avatars).

We conducted a study to evaluate how our volumetric telepresence platform supports partially distributed collaborations. Specifically, our study aims to understand two research questions with regards to volumetric telepresence:

(1) How do communication and negotiation techniques differ across remote and co-located users?
(2) What are the differences in perception, interaction, and engagement across three formats of objects used in collaboration tasks?

We recruited 18 participants over 6 sessions to collaborate on tasks involving physical, volumetric, and virtual objects. We created physical paper cards with written words and virtual cards to represent the three types of objects. Groups of three participants wore mixed reality headsets, completing seven sorting tasks with different distributions of physical, volumetric, and virtual cards across two rooms. We collected audio and video recordings of each session and conducted a semi-structured interview at their conclusion. Our findings show that a shared hybrid workspace in volumetric telepresence supports seamless interactions and natural proxemics among collaborators, with previously restricted co-located coupling styles being represented in a remote setting. Collaboration strategies differ across three types of tools and objects (physical, volumetric, and virtual). Physical and volumetric objects are often interchangeably mistaken despite the distinct level of control. The contributions of this paper are as follows:

- A multi-site, real-time, symmetrical, hybrid workspace volumetric telepresence system: The shared workspace supports the seamless grounding of communication between co-located and remote collaborators in partially distributed environments.
- A classification of hybrid representation of objects and people in volumetric telepresence: physical, volumetric, and virtual. The classification enables a new interpretation of collaborative and communication techniques.

- Empirical insights about the collaborative human experience involving co-located and remote participants using a real-time, symmetrical volumetric telepresence system, demonstrating the successful integration of remote spaces, effective use of proxemics and deixis to support negotiation, and strategies to manage interactivity in hybrid workspaces.

2 RELATED WORK

We divide the related work into two categories and discuss similar work on (1) volumetric telepresence, and (2) mixed reality mediated collaboration.

2.1 Volumetric Telepresence

Telepresence has a long history in science fiction, portraying a future where people are transported across large distances to remote environments instantly [17, 23]. Current approaches to telepresence however, take an alternative approach. Instead of transmitting users to remote locations, the remote environments are captured and digitally transmitted back to the user. This is achieved through the real-time reconstruction of remote environments. Seminal examples produced stereo reconstructions using large arrays of RGB cameras [30, 35]. More recently, the availability of RGB-D sensors has reduced the number of sensors required to produce high-quality reconstructions [28, 39, 73].

Researchers have now started to explore combining reconstructed environments with mixed reality devices, demonstrating the opportunities afforded by such interactive experiences [25, 33, 42, 43, 57]. Room2Room [55] enables the recreation of a life-sized virtual representation of a remote user in a local environment for collaborative experiences. Lindlbauer and Wilson [41] enabled a user in VR to spatially and temporally interact with a complete reconstruction of a physical environment that was captured using multiple depth sensors. Orts-Escolano et al. [51] produced *Holoportation*, a real-time, bi-directional telepresence system that provides remote users with communication affordances akin to co-located collaboration. These telepresence examples share the common characteristics of displaying the reconstructed environment and/or collaborator on head-mounted displays worn by users involved in the collaborative experience.

Telepresence experiences are not limited to head-mounted displays, and have been realized through other display configurations. For example, Fairchild et al. [10] developed the *withyou* telepresence system supporting collaboration across multiple European countries. Their multimodal solution used a combination of display technologies, including a stereo display, a wall display, a 2-sided CAVE, and a desktop computer system. Sites without the volumetric capture set-up used virtual avatars to represent users inside the experience. The Beaming telepresence system enabled a group of users to connect with a remote user that was presented as an avatar on a spherical display [64]. Similarly, Project Starline uses a head-tracked auto-stereoscopic display, along with a high-resolution 3D capture and rendering subsystems, to create a life-like face-to-face conversation between two remote users [37].

A key characteristic of collaborative telepresence examples is the underlying sense of social presence experienced by its users; known as the feeling of *being there* with collaborators in the same environment. Cho et al. [6] compared the effects of a live volumetric avatar to traditional 2D video and a photo-realistic prescanned avatar reporting that live volumetric avatars produced higher levels of social presence. In similar results, Kumaravel et al. [34] reported higher levels of co-presence with participants that performed instruction tasks with their peers using point cloud reconstructions. Most recently, Irlitti et al. [25] explored an asymmetrical cross-modal telepresence design with five users communicating across augmented reality, virtual reality, and an environment unencumbered from enabling technologies. While remote participants in VR reported feeling present inside the reconstructed physical environment, the lack of acknowledgment from their unencumbered peers had a negative impact on feeling socially connected.

2.2 Mixed Reality Mediated Collaboration

Gutwin and Greenberg define workspace awareness as *"the knowledge of others' locations, activities, and intentions relative to the task and to the space"* [21]. A fundamental challenge of workspace awareness in computer mediated collaboration is the notion of coordination, understanding the state of affairs and knowing when to, and how to, act accordingly [66]. This coordination involves a number of factors that allow the seamless integration of multiple people acting in a common interest, including *territoriality* [48], how users partition the activity space, *proxemics* [3, 70], how users position themselves inside the space, and *collaborative coupling* [27, 49, 66], how users transition back and forth between group and individual activities. In co-located settings, collaborators often incorporate non-verbal physical cues, along with verbal communication and existing social norms, to support this goal. Examples of such cues include head gaze, body gestures, object interactions, and the associated context of their actions [14]. However, cues that are highly dependent on physical actions or workspace context, such as gaze, are often lost in remote collaborative settings, creating additional challenges to achieve workspace awareness for remote environments involving interactive tabletop displays [27], shared wall displays [18, 48], or hybrid set-ups [12, 49].

Mixed reality is seen as a key enabler in supporting co-located and remote collaborative tasks due to its ability to provide each user with a combination of the real-world and an aligned virtual world, enhancing the underlying avenues for communication between collaborators [8, 16, 32, 46, 47, 54, 70]. Unlike traditional video conferencing systems, MR systems allow users to maintain *viewpoint independence*, enabling the exploration of the collaborative environment without being restricted to a static frame of reference. When combined with 3D reconstructions of collaborators and/or relevant workspace elements, users are capable of maintaining an up to date understanding of actions being executed by collaborators in the shared workspace [21]. This spatial knowledge is integral for collaborative teams to build and maintain a common understanding of each other's actions and behaviors [22]. There are numerous approaches to improving collaborative communication in MR including spatial referencing [50], gesturing [31, 67], gaze [20, 29], and virtual representations of objects and people [12, 26, 47, 57].

Recently, there has been a large body of research that has investigated solutions to enabling a shared perspective of a workspace with multiple collaborators [34, 44, 56, 57, 63]. In these scenarios, verbal and non-verbal cues leverage the collaborators' contextual awareness of the shared workspace. For example, deictic expressions are words or phrases that require both the individual conveying the message and the individual receiving the message to maintain shared perception of the current workspace context (such as '*this*', '*there*', '*here*', etc.,) are commonly used alongside gestures such as pointing. A limitation of these traditional approaches is the collaboration only occurs in a single workspace, with a remote user either joining a local user in their local environment, or joining a shared virtual environment.

Similar to our work are the concepts proposed in [12, 19, 24]. Herskovitz et al. [24] present a toolkit for facilitating distributed collaboration using portals, anchors, and world-in-miniature, however the provided interaction medium is only virtual and no discussion is provided on user presentation. Fink et al. [12] propose a dynamic mapping of interaction displays across incongruent remote spaces, however the collaborative interactions are mediated through traditional displays with users embodying featureless avatars. Likewise, Grønbæk et al. [19] also considers the mapping of multiple discrete physical work-spaces within the single environment at areas of importance such as a whiteboard or table, however the experience is only limited to two users, while users are represented by virtual avatars, limiting the opportunity for sharing physical interactions in the shared workspace.

2.3 Research Gap

Limited exploration of symmetrical telepresence has been achieved to date, with a majority of work focusing on single workspace approaches. Examples of symmetrical telepresence have also only explored dyad interactions [51, 55], focusing on the remote collaborative communication channel without considering the interplay with simultaneous co-located collaborators. Prior work that has been undertaken on multi-user telepresence collaboration introduced an asymmetry in the experience of its users [25]. To the best of our knowledge, this is the first work that has explored the collaborative interplay between co-located and remote users in a three-user, real-time, bi-directional telepresence system. While significant efforts towards collaborative mixed reality experiences have been previously undertaken, the unique nature of volumetric capture and reconstruction introduces new interactive experiences. The volumetric hybrid workspaces approach presented in this paper extends the traditional channels of physical and virtual interactions in collaborative settings with *volumetric* objects.

3 VOLUMETRIC HYBRID WORKSPACES: DESIGN AND IMPLEMENTATION

In response to the prior research in Telepresence and Mixed Reality Mediated Collaboration, we created a system of *Volumetric Hybrid Workspaces*. Volumetric Hybrid Workspaces is the merging of two remote physical spaces, creating a feeling of being co-located in the same shared space. A volumetric capture system at each site captures and transmits in real-time, the local environment to connected remote sites, with the resulting bi-directional reconstruction generating positive user experiences of co-location. The volumetric capture system captures all people inside each local capture area, thereby supporting multi-user, group-to-group telepresence. The following sections provide a brief description of the technical details of volumetric hybrid workspaces.

3.1 Hybrid Workspace: Users

Inspired from prior work demonstrating the opportunities for sharing full environmental reconstruction with multiple users [25], the implemented volumetric capture platform is a modular system that supports co-located and remote user interactions using mixed reality devices. Bi-directional telepresence has been challenging [34, 37, 51], due to a combination of performance, latency, bandwidth, and coverage requirements. Volumetric Hybrid Workspaces addresses these concerns, by delivering a real-time, low-cost, capture and rendering system, deployed across multiple locations.

The modular capture design is built around the concept of decentralizing camera placement, allowing for larger capture areas. Kinect Network Streaming Servers (KN-SS) are lightweight PC's that have 1-2x Azure Kinect cameras attached via USB connection. The workstations are connected to an enterprise network over a gigabit Ethernet connection. On each KN-SS, a standalone .NET application processes arriving camera frames through connected RGB-D sensors. Frames are requested from Kinect Network Rendering Servers (KN-RS) at each site, where payloads of n frames are requested at a time. Upon receipt, the KN-SS will initiate the processing of each current frame from requested cameras through a compression pipeline, comprising of H.264 encoding for color and an implementation of RVL [71] for depth. The depth compression [71] utilizes a bit shifting algorithm reducing the depth payload by 60%. Color compression is delivered through the application of H.264 encoding, where image forecasting and dynamic bit-rate allow for the reduction in overall network payload. Comparing the required bandwidth for streaming a raw RGB-D stream with JPEG color compression, the implemented compression techniques reduce necessary bandwidth for a single camera from 255mbps to 30mbps.

The rendering module implements the approaches described by [25], undertaking color and depth processing, masking, and segmentation. Once data has been processed, frames are sent onto Kinect Network Rendering Clients (KN-RC), where remote frames are combined with an underlying map of the local environment. Extending on the approach described by Irlitti et al. [25], the point cloud rendering for volumetric users is improved by incorporating the generation of mesh derived from incoming depth images. The points on the depth image are projected into 3D space, using a nearest neighbor approach applying a configurable tolerance to determine if triangulation should occur. The points are also assigned a UV index to allow the color image to be correctly interpolated against. The entire process is undertaken on the GPU without any CPU involvement, rendering the volumetric avatar directly into a scene. To further address issues relating to noisy and missing information from RGB-D sensors, a depth averaging function is applied to interpolate missing depth information during the data processing phase.

A Microsoft Hololens 2 optical see-through head mounted display (HMD) renders the resulting image through a Microsoft Holographic Remoting[1] connection over a WiFi hotspot. The use of

[1]https://learn.microsoft.com/en-us/windows/mixed-reality/develop/native/holographic-remoting-overview

(a) (b)

Figure 2: Examples of the volumetric components presented in *Volumetric Hybrid Workspaces*. (a) Scene viewed from behind a remote collaborator *(white shirt)*. They are rendered as a volumetric avatar while holding a physical sheet of paper which is also rendered as a volumetric object. Virtual instructions *(in a dark rectangle)* can be seen registered to the wall to the left of the physical co-located collaborator. (b) The same scene as (a), viewed from in front of a remote collaborator *(white shirt)*. The physical card being held up contains printed information. Virtual cards can also be seen registered in front of the other co-located user. — Images are captured through a Hololens 2 HMD worn by each of the co-located collaborators.

an optical see-through display allows users to maintain a natural view of their local environment, affording the ability to maintain peripheral vision of co-located colleagues to support workspace awareness. Users remotely captured by the telepresence system are reconstructed as a real-time volumetric avatar, representing their exact physical state in the connected remote location. The entire technical processing and rendering environment was developed using Unity 2022.3 LTS on the Windows Platform.

3.2 Hybrid Workspace: Environment & Objects

The implementation of *Volumetric Hybrid Workspaces* is facilitated inside two individual meeting rooms on our university campus. Each room contains 2x Kinect Network Streaming Servers (KN-SS) and 3x Azure Kinect cameras, mounted in the ceiling. Each camera is mounted $2.4m$ above floor level, directed at 45° downwards towards the alternate wall resulting in a 360° capture of the enclosed environment covering an area of $3m^2$. The hybrid workspace is created surrounding the middle of each meeting room, using a table as the spatial feature to calculate the inter-room calibration.

To calibrate each environment, we adopted the approach implemented by [25], using an extrinsic alignment of synchronized key frames using OpenCV and an optical rigid marker for each camera pair. The resulting calibration creates a geometric relationship between the physical environment and its volumetric representation. This calibration is then used by each HMD to calibrate its view into its local environment, creating a spatial anchor to render registered content. To align both co-located and remote HMD's into the one hybrid workspace, we infer a relationship between known local spatial locations at each site, calculating an internal transformation which is shared across all users. The sharing of transformations ensures that co-located users see registered content in the same location inside their own physical environment, while also ensuring the same calibration is used for every remote user to achieve the same outcome across multiple sites.

A key characteristic of the hybrid workspace experience is the merging of remote locations into a perceived single environment. Like similar work that blends two remote locations together [12, 19], our approach facilitates the sharing of virtual content in a shared physical space. A Mirror network[2] is deployed alongside the volumetric capture system to synchronize the underlying virtual environment across all connected KN-RC users. The network also facilitates voice over IP connections between all connected clients. In the hybrid environment, virtual objects can be instantiated, shared, and manipulated by all connected parties. Through the capturing process, our approach extends on other works by also introducing *volumetric* objects, alongside physical and virtual objects (Figure 2). In local environments, physical objects can be introduced into the workspace, volumetrically captured alongside users using the process described in Section 3.1. The resulting object is transmitted and registered into the remote hybrid workspace in the same geometric position as represented in its own local environment. This new form of interactive quality introduces another level of interaction in group-to-group mixed reality telepresence experiences.

4 STUDY DESIGN

We conducted a user-study to evaluate how the volumetric telepresence mixed reality system supports collaboration with mixed-format objects among partially distributed users around a shared hybrid workspace. The study evaluated the impact of volumetric telepresence on communication, negotiation, organization, and content creation across the three formats of objects within the shared workspace: *physical, virtual, and volumetric*. The study aims to address two research questions in the context of a volumetric telepresence mixed reality system:

(1) How do communication and negotiation techniques differ across remote and co-located users?

[2] https://mirror-networking.com/

(2) What are the differences in perception, interaction, and engagement across three formats of objects used in collaboration tasks?

The study adopted a between-subject design across two types of collaborations (*co-located and remote*) and a within-subject design across three object formats (*physical*, *volumetric*, and *virtual*), through a combination of collaboration tasks.

4.1 Procedure

We set up two soundproofed physical meeting rooms (4.2m x 3.5m) with volumetric real-time capture. Both rooms contained a one-way mirror along one side allowing for observations to be made throughout the study. In the middle of both rooms, a table measuring 1600mm x 800mm x 720mm was positioned inside the space. On the mirror side of the table, a stack of cards were placed to be used throughout the study. Each room contained a single chair per user, positioned along the same wall. The chairs were only provided as a support mechanism throughout the study, no restrictions were placed on how or what the user did inside the space throughout the study. The hybrid workspace was constructed using the table as a spatial feature, using the approach described in Section 3.2.

We conducted six sessions, each with groups of three users: two co-located in Room 1, and the other located in Room 2. Co-located users interacted with each other through *physical* objects and users across two rooms interacted with one another through *volumetric* representations of the physical objects. Users from one room could see the user(s) from the other room as a volumetric rendering through the headset. Participants were asked to complete 7 tasks (see Section 4.3) involving sorting and grouping cards placed on the table within the shared hybrid workspace.

Each user was provided a Microsoft Hololens 2 with an accompanied lavalier microphone to support voice communication. All participants were presented with a tutorial on the use of the Hololens 2 regardless of their familiarity with the device. The introduction described the gestures for showing and hiding the home screen, and near- and far-interactivity. Each participant undertook an eye calibration, then demonstrated their understanding of point and click gestures with menu options, their ability to show and hide the start menu, and finally initiating the holographic remoting application. Once complete, the three participants were closed into their rooms, maintaining voice communication through a Voice over IP through the HMD.

At the conclusion of the study, each participant participated in a group, semi-structured interview discussing their experience. The questions focused on participants' perception of technology, self-rated task success, and communication and interactions approaches with collaborators. We also asked participants to reflect and compare their experience with prior experience with telecommunication systems. The entire study and interview lasted approximately 90 minutes for each group. Participants were provided with a $30 gift card for their involvement. The study was granted approval from the University's Office of Research and Ethics.

4.2 Participants

We recruited 18 participants to undertake the study across 6 sessions (Table 1). Each participant was given an ID, consisting of the Session (A-F), Room number (room 1 with co-located participants or room 2 with remote participant), and Participant number (1-18). The participants were aged between 22 and 45 years old, with a mean age of 28.89 (SD = 5.15). Ten participants identified as male and eight as female. Gender distribution was maintained across the sessions, such that: participants co-located as pairs in Room 1 (MM = 1, FF = 1, MF = 4) and participants solely located in Room 2 (M = 4, F = 2). The groups were composed of participants who could converse in English, and all had some degree of familiarity with each other to remove any conversational difficulties. The six groups were composed of users that had predominately high video telephony experience in their roles, and had some familiarity with mixed reality devices.

Table 1: Self-Reported Participant Information with solo participants highlighted in gray.

Session	Room	Paricipant ID	Gender	Age
A	Room 1	A_{1-P1}	M	31
		A_{1-P2}	F	27
	Room 2	A_{2-P3}	M	26
B	Room 1	B_{1-P4}	M	29
		B_{1-P5}	M	26
	Room 2	B_{2-P6}	F	34
C	Room 1	C_{1-P8}	M	26
		C_{1-P9}	F	27
	Room 2	C_{2-P7}	M	27
D	Room 1	D_{1-P11}	F	28
		D_{1-P12}	M	45
	Room 2	D_{2-P10}	F	37
E	Room 1	E_{1-P14}	F	28
		E_{1-P15}	F	27
	Room 2	E_{2-P13}	M	26
F	Room 1	F_{1-P17}	M	22
		F_{1-P18}	F	24
	Room 2	F_{2-P16}	M	30

4.3 Tasks

Participants were asked to group and sort cards printed with animal names. Table 2 describes the task order, the distribution of cards across both rooms and virtual cards, as well as the instructions given to participants to sort and group animals. Table 3 presents a breakdown of the animal names for each task type, and their physical or virtual state. Physical cards in one room were rendered as volumetric cards in the other room, which were viewed through headsets. The tasks were conducted in the same order across all groups, however tasks 4 & 5 were counterbalanced across sessions. Printed cards prepared for each task were placed faced down in separate piles. Blank cards were provided for the creation activities in task 6 and 7.

Table 2: Tasks and Card types. The physical cards in one room are available as a volumetric cards in the other room. The virtual cards are available to both rooms.

Tasks #	Details	Physical Cards Room 1	Physical Cards Room 2	Virtual Cards
	TRAINING			
1	Arrange the cards into groups according to card colour	0	0	8
	HYBRID (small set)			
2	Arrange the cards in alphabetical order	3	3	2
	HYBRID (large set)			
3	Arrange the cards into groups according to animal habitat	7	7	3
	CONTROL (Room 1) ⋆			
4	Arrange the cards into groups according to animals with same number of limbs	18	0	4
	CONTROL (Room 2) ⋆			
5	Arrange the cards into groups according to animals from the same geographical region	0	18	4
	NEGOTIATION (card creation) ‡			
6	Write the name of animals that are similar in some way to the printed cards	3 + 4†	3 + 2†	0
	NEGOTIATION (category creation) §			
7	Arrange all the cards into 3 groups of 4 animals based on common features	4† + 3	2† + 3	0

† Handwritten physical cards.
⋆ Tasks were counter-balanced between sessions.
‡ Printed-cards are presented first.
§ Hand-written cards are presented first.

Training: Participants arranged 8 colored virtual cards while familiarizing with hand tracking interactions using the Hololens. Once comfortable with the interactions, they grouped the virtual cards by color on the table.

Hybrid: Half of the physical cards were distributed across two rooms. For task 2, participants turned over 3 cards in each room along with 2 virtual cards, arranging the 8 words alphabetically. For task 3, participants turned over 4 more cards in each room along with 3 virtual cards, arranging the 17 words by habitat.

Control: Control of all physical cards was provided for a single room, while the other room was provided volumetric cards. The room with control was swapped and counterbalanced across sessions. Participants were asked to arrange the 22 animal words (8 physical and 4 virtual) by either geographical region or by number of limbs.

Negotiation: Aimed to stimulate negotiation between participants through co-creating content using physical cards; some were printed and others were created during the session. The printed cards were a new set of animal names divided equally between the two rooms. For task 6, each room was asked to turn over 1 printed card, then as a group, write the names of two animal words that would create a group of 4 with the printed cards, placing them in a group on the table. This process was repeated two more times. For task 7, participants were asked to remove the 6 printed cards. Using their 6 handwritten words, they were asked to turn over the remaining six cards and create 3 new categories of 4 words. An example illustration of the task is presented in Figure 6.

For all tasks, participants were asked to work together as a team to achieve the requested goal. If they did not know an answer, they were asked to work with their available resources to achieve a result. It was made clear to participants that they were not being measured on completion time, but rather on the quality of the outcome. However after four minutes, the experimenter would contact them on the intercom system to alert them of a minute remaining before they would have to submit their solution. This was to avoid instances were indecision would inhibit further progress.

4.4 Data

For the duration of the study, two researchers observed and recorded participants' behavior, communication, and body language through written notes. The audio and video of the tasks being completed in each room was recorded for later analysis by multiple ceiling mounted cameras. The volumetric capture sessions from both rooms were also recorded.

5 FINDINGS

Our data included a total of 15 hours of video recordings and audio recordings from interviews and training studies, as well as 20 pages of researcher observation notes across all sessions. Observations were conducted by a combination of the first three authors. We followed a deductive thematic analysis approach outlined by Braun and Clarke [5]. To support the data familiarization process, the lead author transcribed all audio recordings, then re-watched all session recordings, creating time-synchronized footage from each room using video-editing software. The lead author then undertook coding across all data. The research team met during this process to discuss preliminary outcomes and the identified codes, to reach agreement on the following set of themes. We observed behaviors from our participants as they undertook tasks that explored their perception, interaction, and communication as a team in a hybrid workspace. The following sections present our findings on: proxemic behaviors, collaborating in telepresence, interacting in remote-physical experiences, and strategies for effective communication in hybrid environments.

5.1 Respecting each other's space

Proxemic behaviors were on the most part respected by the participants throughout the study sessions. With this work being the first to consider volumetric avatars in a partially-distributed collaboration, our observations demonstrate similar outcomes as shown in prior work involving human representations [12, 69]. Participants continuously re-positioned themselves around the workspace and each other, commonly in a 2-1 configuration, with two collaborators on one side of the table standing a social distance apart next to one another regardless of their locality. Co-located and remote participants alike respected each other's personal space as they continued to maintain effective collaborative partnerships, *"we can sense the presence of B_{2-P6}. So we don't want to move to their side, instead we could work together each on different sides, so we can communicate with each other"* (B_{1-P4}).

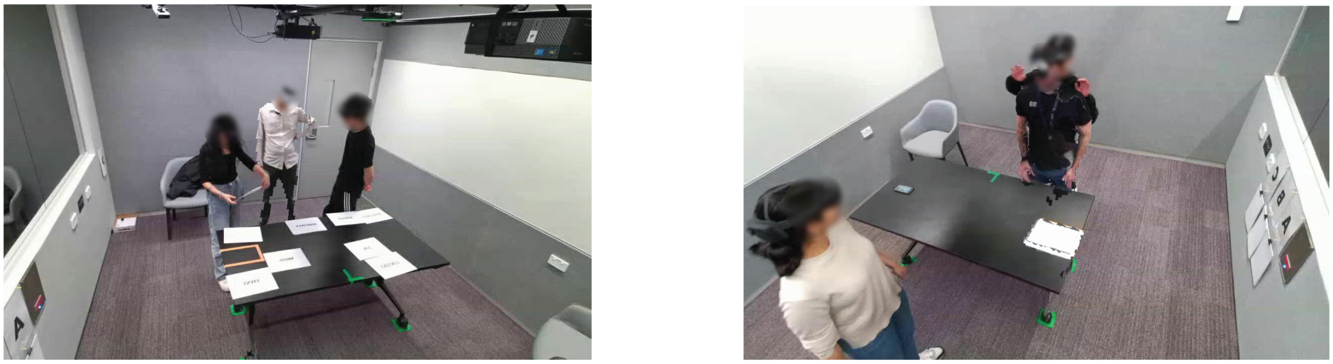
On some occasions however, this social proxemic distance was violated by remote participants as they re-positioned their view of the workspace. In one instance, C_{1-P8} was unaware of their collaborator's positioning as they were in discussion with their co-located partner. As they verbally directed their question to C_{2-P7}'s whereabouts, they scanned the surrounding area and realized that they were standing in an intimate distance to their remote partner. In the immediate moments after this realization: (i) C_{1-P8} leaned backwards in a mix of shock and surprise, (ii) C_{1-P8} took a step backwards but also opened their body stance to include C_{2-P7} into their conversation with C_{1-P9}, and (iii) C_{2-P7} took two steps to their right, stepping into an open space between the two co-located collaborators in Room 1. For the next few minutes, the three collaborators discussed several words positioned in this social triangle on the same side of the table (Figure 3a). C_{2-P7} made a gesture of a bird flapping its wings, keeping their elbows tucked in to avoid hitting their collaborators, while C_{1-P8} partially leaned over the table as they stepped forward to place cards into the workspace, ensuring they at least remained within a social proximity to C_{2-P7}. During interviews, when participants were asked about their perception of their collaborators inside the workspace, responses mimicked the behavior exhibited in this example. F_{1-P18} explained:

> *"We all respected each other's physical space, like even the virtual, and the real ones even don't have to do it. We could literally be standing like this, (makes a gesture demonstrating occupying the same space) and we can still work. But even then, we had this inherent sense of: Okay. That's their space. I'm going to move immediately."* (F_{1-P18})

The observed behaviors are generally in congruent with the findings by Wilcox et al. [69] and Fink et al. [12] where participants would make a conscious effort to respect their collaborators space. In the prior example, once *Group C* were consciously aware of each other's proximity, all members expressed body language which was representative of co-located tight coupling [49], even though they were distributed across different rooms. Unlike the workstation approach devised in *Re-locations* [12], participants were unable to spatially associate a collaborator's positioning based on their interaction with a particular display. As such, visual awareness played a key role in supporting the proxemic phenomenon; when participants were focused solely on the task space, they would sometimes lose spatial understanding to their colleagues positioning. While this information remained unavailable, the collaboration could continue uninterrupted. This was observed in several sessions, as unbeknownst to multiple participants, they would be standing in an intimate proximity to one another. However, if awareness was restored, they would immediately shift their positioning to reclaim social distances (Figure 3b), with adjectives used to explain the feeling as *"a bit of a shock"* (F_{1-P17}), *"weird"* (E_{2-P13}), and *"very intrusive"* (B_{2-P6}).

Table 3: Animal words printed on the cards. Cards are either Physical (P) or Virtual (V).

Hybrid & Control Tasks				Negotiate Tasks	
Word	*State*	*Word*	*State*	*Word*	*State*
Werewolf	P	Penguin	P	Octopus	P
Possum	P	Snake	P	Cheetah	P
Frog	V	Zebra	P	Orangutan	P
Shark	P	Bee	P	Dolphin	P
Cricket	P	Spider	P	Frog	P
Octopus	V	Platypus	P	Mosquito	P
Kangaroo	P	Centaur	P	Eagle	P
Dragon	P	Drongo	V	Unicorn	P
Draco	V	Bilby	P	Koala	P
Pangolin	P	Centipede	P	Emu	P
Tadpole	P	Stegosaurus	P	Reindeer	P
				Wombat	P



(a) Three participants organized themselves into a social space while negotiating the placement of physical and volumetric objects.

(b) Two participants occupy the same *physical* location, one local and one remote, creating an intimate proxemic relationship.

Figure 3: Interactions with partially distributed teams. (a) Two co-located users in Room 1 *(black shirts)*. One remotely present user from Room 2 *(white shirt)*. (b) Two co-located users in Room 1 *(white shirt, black shirt - hands up)*. One remotely present user from Room 2 *(black shirt - hands down)*. — Images are the projection of the volumetric frame captured from Room 2 into a color frame at the same moment in Room 1.

5.2 Collaborating in the same physical environment

The hybrid telepresence workspace allowed for the effective collaboration of remotely located users undertaking a physical task as if they were co-located with their peers. Throughout the study, participants exhibited behaviors that would be common in co-located scenarios, even though there was an awareness to the remote locality of their colleagues, *"it's hard to say whether it's a shared virtual space or a shared real space. Because I think, in my mind, I know that we're in separate rooms"* (E_{2-P13}). This stance was shared across many sessions, *"I felt like we sort of felt like we're in the same room together even though I knew that it wasn't the case"* (A_{2-P3}), *"I feel it is pretty real to me. Especially when we are moving cards together"* (B_{1-P5}), and *"I treated everyone naturally and equally. I don't feel there is a difference, even though B_{2-P6} is in another room"* (B_{1-P4}). Considering collaboration based on coupling styles [27, 49, 66], the observed collaboration shifted between two tight couplings; *DISC* [49]: where participants were in active discussion about the task but were not actively interacting with objects, and *VE* [27, 49, 66]: One person is actively interacting inside the workspace, while other participants view and engage in conversation, and one loose coupling; *V* [66]: a participant views the task without being sufficiently engaged to provide help or suggestions. For the loose coupling *V*, this was observed when a single participant took over during the *control* task, such as in Groups A, B, and D. During these scenarios, the two additional participants only occasionally interjected, preferring to stand back and observe as their dominant collaborator took full control of completing the task independently. For the remaining tasks however, groups maintained tight coupling throughout, staying either in active discussion or view engaged. Interestingly, while our hybrid workspace is a remote collaboration tool, the original co-located tabletop coupling styles expressed by Tang et al. [66] are more closely aligned to the behaviors we observed. Based on the adapted categorization from Neumayr et al. [49], *VE* is a coupling style restricted to co-located users in a partially distributed team, however in our workspace, this tight coupling was a predominant coupling observed across all groups between remote locations. This commonly involved clarification of words, proposed positioning of objects, and decision making on creating content. For clarification and proposing positioning, participants used spatial deixis, combining commands of *'this'*, *'that'*, and *'there'* with body gestures such as pointing, shaking of objects, or spatial references to known locations, such as a previously discussed word. An example of this form of interaction is presented in Figure 4. When questions were raised, they were addressed to both co-located and remote colleagues simultaneously, *"I would walk towards the word and then I would put my hand over it and say what does it say?"* (C_{1-P8}). Participants agreed that there was an innate ability in understanding current context, *"I think immediately once we all got into like this sort of zone. Then it just became like a second nature thing"* (F_{1-P18}), and *"I do like, the idea of being in the same space and looking at the same like table working on the same, like reference point on the same task, it feels very natural to me"* (E_{2-P13}). The surprise to the ease of communication was summarized as:

> *"It didn't feel like a chess game or having to wait for the other person to do the step. It didn't feel that way. I think it was more real time. Like I could immediately say this or ask a question and get a response immediately as compared to waiting for the other person."* (C_{1-P8})

For the final tasks, participant's created new content by writing words onto blank sheets of card with a black marker (Figure 5). This action introduced a new form of interaction into the hybrid workspace. Participants revealed that the experience brought about unique emotions, heightening the sense of immersion and being together in the same space. This connection was observed as remote participants would verbalize words as they were being written, help correct spelling mistakes, laugh at the misfortune of miscalculating

(a) Room 1 **(b) Room 2**

Figure 4: Two users directed attention to an object during an open discussion with one co-located user. The object being referenced from Room 1 in the left frame is physical, while from Room 2 in the right frame, it is volumetric.

the space required to write a word, or commending penmanship. Reflecting on this experience, participant's appreciated the ease of use, *"(Creating content) was easier than typing in VR"* (B_{2-P6}), the increased sense of immersion, *"I think it was much more immersive than other types (of digital technologies)"* (B_{1-P5}), and also the associated emotional connections:

> *"(When) two of us wrote two different animal names together, and she was writing on the other side. I was thinking, she's just next to me writing. I found it really interesting the way I felt attached (to her). I felt something. I felt it was really interesting to have that feeling. Okay, she's writing with me even though for a second I forgot that shes in the other room. The speed was the same and it felt real."* (D_{1-P11})
>
> *(In response to D_{1-P11} when reflecting on there being a single table) "When you were writing down the words, it felt like we were at the exact same table."* (D_{1-P10})

5.3 Interaction across the seam

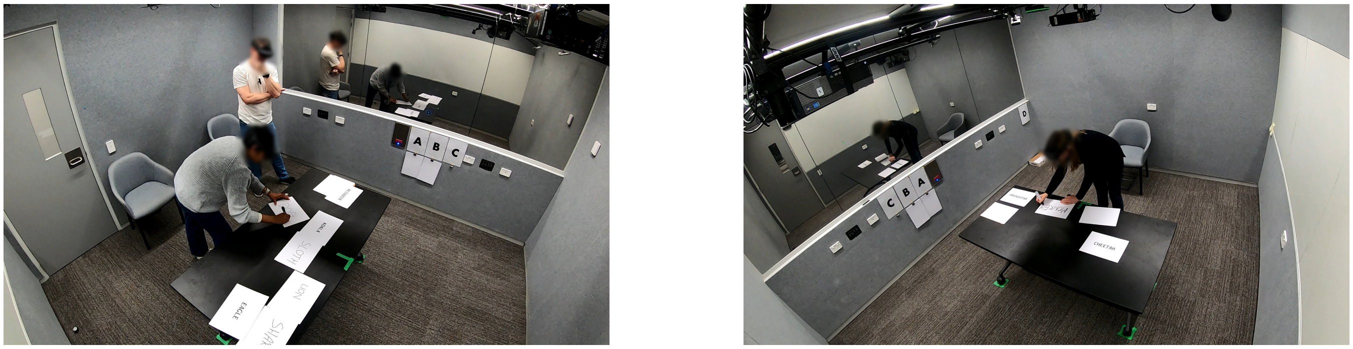
Interaction in our study design involved participants working with three different forms of information; (1) physical cards, (2) volumetric cards, and (3) virtual cards. Physical cards were movable by users inside the same room, while volumetric cards were their representation to remote peers. Virtual cards could be moved by all users. Participants were asked about the number of interaction objects in the environment, giving a clear consensus to virtual cards being unique, while the view on physical and volumetric cards were less clear, *"there were two sets where it's a virtual set and a physical set we can perceive differently"* (B_{1-P4}). The virtual cards, however, were seen as an important facilitator for feeling connected, *"I think that was more collaborative with E_{2-P13} because we could all touch them and move them around. Whereas the physical ones, obviously, we were kind of bound by the physical space"* (E_{1-P14}).

There was a high level of immersion for participants, who commonly reported losing the ability to discern the differences between physical and volumetric cards; every session included moments where participants mistakenly reached out to manipulate volumetric cards. Upon reflection, participants explained *"It was strange wanting to move the cards and then realizing that they weren't your cards"* (E_{1-P14}). As D_{1-P10} expressed their frustration to the group that they had realized they had turned over two physical cards instead of one, D_{1-P12} referred to having a *"disjointed feeling when realizing it (the volumetric card) wasn't there."* Expressing their mutual understanding to this effect, D_{1-P10} further elaborated, *"it felt like it was realistic and you were going to turn it over yourself."* After experiencing both the hybrid tasks and control tasks, towards the end of the study, one participant reflected on the blurred line between real and virtual:

> *"There was a point where in the very last part, I forgot that I could move my cards. I was like, I thought it was one of those scenarios where only they can move their cards. So, I realized that it was more immersive than I thought, because I'm getting things mixed up now. Because they were kind of like, move, move it and I was like, why aren't they moving it?* (B_{2-P6})

At the end of the session, participants were asked of their preference on interaction items inside the space. Every participant expressed their preference towards physical cards, allowing them to maintain interactive control. Surprisingly, the majority second preference was volumetric. Observations showed that a majority of groups would leave interaction with the virtual cards until the very end of the task, with participants revealing three primary reasons: (1) they forgot about their existence as they were *"out of sight"* (D_{1-P12}), (2) virtual cards were difficult to use to assist in coordination alongside the other two form factors, and (3) interaction with virtual cards were difficult. This second point is an example of group territories as described by [68], where virtual elements were used to attempt to assist coordination inside the workspace. This coordination strategy was short lived however, as multiple groups expressed their difficulties in combining the three different form factors of objects as well as their poor use as category markers. newline

> *"Honestly, it wasn't easy to move them and rotate them, and if you use them as markers when you put real ones on top, you just can't read anything. So like it will be*



(a) Room 1 (b) Room 2

Figure 5: Participants simultaneously created physical content and volumetric content through physical inscription, writing the words (a) 'Dog' in Room 1, and (b) 'Monkey' in Room 2. Both images are captured at the same moment in time, with their actions communicated within the Volumetric Hybrid Workspace.

> *nice to have them as a way to say okay, this is the area for this (topic), but then when you try to place the other cards, you can't read (anymore). They could be made more transparent but even then it just, felt like they're not really there."* (F_{2-P16})

5.4 Strategies to support coordination

Strategies to assist in coordination emerged throughout the study sessions, highlighting the importance on ensuring that everyone felt included. Unlike prior work which highlighted local site territories being used for co-located participants alongside the remote group territory [48], the nature of the workspace required inclusiveness due to the separation of interaction across remote sites, *"you had to trust the other person (to do their interaction)"* (D_{1-P12}), and *"it comes down to communication skills"* (D_{2-P10}). F_{1-P18} explained their need to continually include their remote partner in their decision making process, *"I found myself asking a lot about his opinion because he was not physically with me. So, every time we would make a change, I would make a change. I would ask for F_{2-P16}'s opinion to say is that okay with you."* The entire workspace become a group territory, with all members of the partially distributed team utilizing the space as if they were co-located while maintaining a cohesive group audio territory [48]. The most common approach was to lay out all the cards flat in the work area, ensuring that every card was visible to all collaborators. As sorting was undertaken, more specific sites were proposed by team members, and cards were shifted into smaller piles. Alongside the verbal channel, one strategy for coordination was through the holding of a card facing towards the team. Conversation was then directed towards this particular word, and its positioning inside the interaction volume, *"I think like once we also kind of worked out like we held (the cards), the cameras picked it up a bit better also made it a bit easier to work out"* (E_{1-P14}). This behavior was noted in 4/6 sessions, although only one group also employed this strategy across all control tasks.

With the inclusion of volumetric avatars in our distributed collaboration, body movement inside the space became a tool for supporting awareness alongside the group audio territory, either changing the local environment to match the remote, "*there was one (volumetric) card here that I was trying to move, but instead of moving that one out, I'd switched the others around to get a group back together*" (D_{1-P12}), or following a collaborator's movement and reacting, *"usually when I saw them moving and say like we're going to move these ones here. I just moved mine, I didn't have to wait for them to tell me, I can see them moving and I know where they are"* (F_{2-P16}). Additionally, with every user wearing an optical see-through head-mounted display to experience the hybrid workspace, their eyes were occluded from both co-located and remote colleagues. When asked about the impact the device had on their underlying ability to communicate, there was a consensus that the device did not have much of an impact, with the groups relying on verbal and non-verbal communication channels, *"yeah, we were wearing headsets (but) it wasn't like a very striking thing to me that we didn't see the eyes (D_{1-P11})."*

> *"To be honest, I don't think I thought it was any different than just working in the same room… we all rely a lot on voice communication."* (F_{2-P16})
>
> *"I often look at facial cues like if I say something and I'm like do you agree with me I like look at you and see how you react, but I never did that with F_{1-P18} throughout the entire study."* (F_{1-P17})
>
> *"If it was physical, you probably mostly rely on body language because you don't need to look (at one another). That's why you didn't look at each other."* (F_{2-P16})

To assist in maintaining a coherent frame of reference across the group, participants provided verbal directions in their partner's viewpoint, *"I always like speak from your side. Like saying on your left, on your right"* (C_{1-P9}). Alternative approaches to creating a consistent frame of reference among the group included dividing the work area in half, with each side representing a single site, cardinal directions, with an agreement on *North*, or incremental directions, with one side of the table representing *Low*, and the other representing *High*. Incremental approaches still created difficulties when a case was uncovered which exceeded one of the pre-existing boundary cases, *"we're going to put the highest limbs there but then there was something with even higher limbs so we had to move everything. And it was a mix of real cards here, real cards (there), and virtual cards and that was difficult to cooperate to like, move this there,"* (C_{1-P8}).

6 DISCUSSION

6.1 Partially Distributed Telepresence

Rendering quality, especially of the users' body, has been the focus of improvement for volumetric telepresence due to its importance in maintaining immersion, presence, and social interaction quality [39, 59, 72]. The volumetric telepresence system used in this paper has significantly better rendering quality in real time (see Figure 6), compared to other prior work in volumetric telepresence [25]. As participants were wearing a headset with reflective covering that affected the capture sensors, there were gaps in the volumetric data of the face and head, which was similarly reported in prior work [25]. During the tasks, participants reported that the gaps did not affect their interaction and communication with remote users, nor did the existence of the Hololens impact interactions with co-located users, as it was explained that the face was less important to them than body cues, verbal interactions, and gestures. Prior work suggested that this might be an issue [13, 51]; however, we did not observe similar outcome in this study.

6.2 Co-located and Remote Collaboration

Participants did not alter their collaborative behaviors towards co-located and remote users, with observations noting that participants had a strong perception of working together in a shared space (Section 5.2). Contrary to prior work on partially distributed teams using traditional workstation devices [49], our volumetric hybrid workspace exhibited strong tight couplings between remote locations that were highlighted as being only possible in co-locateds due to it being impossible to view the same physical artifact. One advantage of volumetric telepresence is the capability to capture photo-realistic, real-time, body representations of users, compared to the more common approach of cartoon-like 3D virtual avatars. This results in co-located users being present *physically* inside the room, while remote users were displayed as recognizable, real-time reconstructed *volumetric* bodies. The underlying experience created by the volumetric workspace resulted in the perception that all users co-existed within the singular physical workspace. This phenomenon can be quantified through the question, '*How many tables were in the exercise?*', in which all respondents answered 'only one'. One participant further commented that *"I feel that was a trick question" (*E_{1-P13}*)*. Another doubted their answer F_{2-P18} and outlined that there were physically two tables, one in each room. These responses highlight *a strong perception of shared ownership of the workspace*, enabled by the design of a hybrid workspace using the volumetric telepresence platform (see Section 3.2).

The study was designed to incorporate two identical rooms that were equally captured and rendered using the volumetric telepresence system, while also involving tasks such that activities were equally divided across both rooms. From the perspective of each user, the volumetric system brings remote users into their local physical space, with asymmetric telepresence systems often using the term 'visitors' to describe remote users [64]. However, when asked about the sense of space ownership, four out of six groups reported that the solo participant in Room 2 consistently felt like a 'guest', even for tasks where all the physical cards were in the solo room. Participant C_{1-P8} described that *"he (referring to* C_{2-P7}*) felt like a guest"*. From the same room, participant C_{1-P9} stated that *"I felt like he (referring to* C_{2-P7}*) is like an NPC (non-playable character)!"* The solo participant in this group C_{2-P7} commented when he walked back to Room 1 (for the interview) that he did not realize C_{1-P8} and C_{1-P9} were together physically in the same room, and that he felt like he was *"joining a virtual conference"*. The solo participation from the last group F_{2-P16} said that he *"always felt like I was joining the team"*. This study demonstrated the sense of space ownership was not symmetrical in a symmetric telepresence system. While prior work in group-to-group telepresence [4, 52, 53] has not reported such imbalances, primary room dominance is a common artifact of traditional video conferencing meeting setups where remote participants feel less included [60, 61]. Our study also highlighted a primary room dominance in favor of the co-located pair of participants (2 versus 1); however, the negative connotations attached to the room dominance were not conveyed, such as feeling left out or incapable of being involved.

6.3 Object Ownership

The final two tasks created a new sense of object ownership, one in which participants created the objects they interacted with by writing on blank cards (Figure 6). Participant F_{1-P18} commented that the task *"required a different type of coordination"* and described that she was fascinated by being observed by the remote participant as she was writing. Another participant B_{1-P5} mentioned that writing was easier than typing in VR, suggesting that she expected to be able to create content in a virtual format as well, a feature that was not yet created for the study. Prior work [58] has shown that relatedness and psychological ownership of virtual objects are strengthened if placed in a physical environment. The study in this paper shows that a volumetric telepresence system can provide a platform to create content and increase ownership of both physical and virtual objects. Auda et al. [1] explored a mechanism for sharing ownership of objects during remote collaborations, showing that sharing ownership of virtual objects decreased social interactions. We observed the opposite effect with created and shared content with volumetric representations encouraging participants to engage in tightly coupled interactions.

6.4 Signal, Negotiation and Fidelity of Object Control

Across the three formats of objects used in the tasks, the level of control that a user can exert on the object differs. A local user has full control over physical objects within the same environment, while only has 'view-access' to volumetric objects in remote spaces. Virtual objects, on the other hand, can be fully controlled by both local and remote users through provided gestures. During the study, we observed multiple processes surrounding object control.

The ***signal of object control***, i.e., *'who could control what'*, was not clear within the volumetric telepresence system, especially when there were visual similarities among objects used in multiple

(a) (b)

Figure 6: An example of the creation of information during the study. (a) Content can be physically created in the hybrid workspace. (b) The created content can be visualized remotely in real-time.

locations. We designed the task which involved mixed groupings of physical, volumetric, and virtual cards to understand how participants distinguished the three object formats. For the paper cards used in the study, the rendering quality of volumetric cards was not perfect with clear jagged edges compared to physical cards, see Figure 6. We posited that this visual quality could provide clear signal to indicate the format of the object and the possible control afforded to the user. However, with all six groups, we observed that participants often reached out to move volumetric cards, only to realize that the volumetric cards belonged to participant(s) in the other room. This finding was in direct contrast with prior research in shared virtual objects in a co-located immersive environment where Lee et al. [38] found that participants did not interact with virtual objects that did not belong to them. The key difference was that the objects used in this experiment were of different formats, 'physical' vs 'volumetric', compared to a more homogeneous nature of virtual objects used by Lee et al. [38].

At the interview, participants explained that because they could easily pick up and move the physical cards into groups, they expected the same affordance with volumetric cards. A framework of affordances presented by Steffen et al. [65] highlighted that *'amplifying reality'* is one of the key affordances of augmented reality systems, in which augmented content provide extra realistic context and capability to the physical environment. In this study, we have observed the opposite effect where *the affordances of physical objects amplified the perceived affordances of volumetric objects*, albeit being false affordances. We did not observe the same confusion between physical cards and virtual cards in the study; thus the same statement could not be said between physical and virtual objects. Shin [62] proposed a model of affordance actualization that in mixed reality systems, the more the users show empathy and embodiment with mixed reality environments, the more the users feel encouraged to interact with mixed reality objects, i.e., increased affordance actualization. In this study, we observed a similar effect where participants were inclined to transfer the affordances from physical objects to volumetric: *"I kept trying to pinch and pick it (volumetric card) up and bring it closer, even if it was, you know, on his part of the world in (the other) room"*.

Once the signal of object control was understood, we observed the process of ***negotiation of object control*** between two rooms to complete the tasks. A typical approach is to request the participants who were in control of the cards to move the cards on the table to create a common pile. Section 5.4 outlined some of the control negotiation techniques, including separation of cards layout on the table, body movements, verbal, and pointing gestures. For example, participant D_{1-P12} positioned himself in the corner of the table, where he wanted participant D_{2-P10} to move the cards in Room 2 to the same corner of the table. There were instances when prompting for action was not required as one room moved their common cards into one pile, and the other room followed without being asked. Furthermore, we noted at least three instances where participants resorted to moving their own cards instead of waiting for remote participants to move theirs. This strategy was seen in both rooms with solo or duo participants. Participant B_{2-P6} commented in the interview that she was frustrated that she did not have the control to move volumetric cards, even causing her to break the sense of being there in the hybrid workspace. These negotiation strategies navigated the territoriality created by different object formats [48].

One aspect of object control that sets virtual objects apart from physical and volumetric objects was the ***fidelity of object control***. The fidelity of object control applied to the three types of objects such that full control for physical object, gesture control for virtual object via hand tracking, and no control for volumetric object (or rather indirect control via remote users). When asked about their preference of interacting with the three types of objects, there were two main answers that preferred volumetric objects over virtual and vice versa, with physical objects being the most preferred. Section 5.3 showed that participants preferred asking remote users to actualize cards movement (indirect control) rather than moving the virtual cards through gestures. For those who preferred virtual cards over volumetric, the reason was given that readability was far superior for virtual compared to volumetric cards. Overall, we suggest that *the balance of fidelity of object control versus fidelity of visual renderings should be carefully considered in volumetric telepresence systems.*

6.5 Limitations and Future Work

The mixed reality telepresence system aimed to support a productive collaborative working relationship between partially distributed teams. While the system draws upon prior work of mixed reality collaborative systems, its bi-directional design required that we identified specific tasks that are appropriate for mixed reality representations and meaningful for partially distributed teams. This resulted in our approach demanding an evaluation be done with a relatively small number of participants with relevant experiences in distributed meetings and the analysis be done qualitatively. Future work will broaden the possible activities for bidirectional mixed reality collaborations between partially distributed teams. The activities are likely to include dynamic tasks that require participants to shift between tight and loose couplings while engaging. This would allow quantitative analysis of volumetric hybrid workspace with a focus on coupling styles [27, 49, 66].

The format of the tasks may also be broadened. The tasks in this study used cards based on physical characteristics, resulting in virtual cards being a simulacrum of the physical cards. New activities would be designed to better understand the interplay between physical, volumetric, and virtual interaction. Future work is needed to exploit the virtual characteristics of the content. Approaches could involve using virtual content as a form of private territory, or allow virtual content to be created, allowing for it to be used for more effective coordination practices.

Finally, the bidirectional telepresence approach only considered a single congruent environment. This allowed interactions to be captured around the single point of interest. Future work could explore several variations to the Volumetric Hybrid Workspace telepresence model drawing on prior work including increasing spatial complexity through adding an additional area of interest, incorporating users in virtual reality [25], combining virtual avatars [19] alongside volumetric reconstructions, or considering the application of collaboration in incongruent spaces [12].

7 CONCLUSION

In this paper we presented a *Volumetric Hybrid Workspace*, a real-time, multi-site volumetric telepresence system which supports symmetrical, grounded communication experiences within a shared physical workspace. The system aligns multiple remote physical locations around a shared physical artifact, enabling co-located and remote collaborators alike to interact both physically and virtually with three levels of objects: physical, virtual, and volumetric. We conducted a qualitative study which explored the use of volumetric hybrid workspaces involving six groups of three users located across two sites incorporating the three forms of interaction objects. Participants completed sorting tasks using physical and virtual cards.

We found that participants respected social proximity conventions within the shared hybrid workspace. Participants reported a strong sense of shared ownership of the hybrid workspace, employing similar communication and collaboration techniques across co-located and remote users. This work demonstrated tight coupling styles which were previously considered for co-located only in a remote setting. The classification of three object formats of physical, volumetric, and virtual introduces nuances in the level of control afforded to the users. Volumetric objects create a false sense of control and affordance due to visual similarities with physical objects. Negotiation of object control among co-located and remote collaborators requires multiple strategies. Users' preferences across three types of objects differ with regards to the fidelity of control and the fidelity of visual renderings. This study adds important insights to the resulting experience of symmetrical, multi-user, partially distributed, volumetric telepresence experiences.

REFERENCES

[1] Jonas Auda, Leon Busse, Ken Pfeuffer, Uwe Gruenefeld, Radiah Rivu, Florian Alt, and Stefan Schneegass. 2021. I'm in Control! Transferring Object Ownership Between Remote Users with Haptic Props in Virtual Reality. In *Proceedings of the 2021 ACM Symposium on Spatial User Interaction* (Virtual Event, USA) *(SUI '21)*. ACM, New York, NY, USA, Article 10, 10 pages. https://doi.org/10.1145/3485279.3485287

[2] Steven Baker, Jenny Waycott, Romina Carrasco, Ryan M Kelly, Anthony John Jones, Jack Lilley, Briony Dow, Frances Batchelor, Thuong Hoang, and Frank Vetere. 2021. Avatar-Mediated Communication in Social VR: An In-Depth Exploration of Older Adult Interaction in an Emerging Communication Platform. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411764.3445752

[3] Till Ballendat, Nicolai Marquardt, and Saul Greenberg. 2010. Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment. *ACM International Conference on Interactive Tabletops and Surfaces*, 121–130. https://doi.org/10.1145/1936652.1936676

[4] Stephan Beck, André Kunert, Alexander Kulik, and Bernd Froehlich. 2013. Immersive Group-to-Group Telepresence. *IEEE Transactions on Visualization and Computer Graphics* 19, 4 (2013), 616–625. https://doi.org/10.1109/TVCG.2013.33

[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3 (2006), 77–101. Issue 2.

[6] S Cho, S w. Kim, J Lee, J Ahn, and J Han. 2020. Effects of volumetric capture avatars on social presence in immersive virtual environments. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 26–34. https://doi.org/10.1109/VR46266.2020.00020

[7] Rafael Kuffner dos Anjos, Maurício Sousa, Daniel Mendes, Daniel Medeiros, Mark Billinghurst, Craig Anslow, and Joaquim Jorge. 2019. Adventures in hologram space: exploring the design space of eye-to-eye volumetric telepresence. *25th ACM Symposium on Virtual Reality Software and Technology*, 1–5. https://doi.org/10.1145/3359996.3364244

[8] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. 2019. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies* 131 (11 2019), 81–98. https://doi.org/10.1016/j.ijhcs.2019.05.011

[9] Fazliaty Edora Fadzli, Ajune Wanis Ismail, Mohamad Yahya Fekri Aladin, and Nur Zuraifah Syazrah Othman. 2020. A Review of Mixed Reality Telepresence. *IOP Conference Series: Materials Science and Engineering* 864 (5 2020), 012081. Issue 1. https://doi.org/10.1088/1757-899X/864/1/012081

[10] A J Fairchild, S P Campion, A S García, R Wolff, T Fernando, and D J Roberts. 2017. A Mixed Reality Telepresence System for Collaborative Space Operation. *IEEE Transactions on Circuits and Systems for Video Technology* 27 (2017), 814–827. Issue 4. https://doi.org/10.1109/TCSVT.2016.2580425

[11] Andreas Rene Fender and Christian Holz. 2022. Causality-preserving Asynchronous Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. ACM, New York, NY, USA, Article 634, 15 pages. https://doi.org/10.1145/3491102.3501836

[12] Daniel Immanuel Fink, Johannes Zagermann, Harald Reiterer, and Hans-Christian Jetter. 2022. Re-Locations: Augmenting Personal and Shared Workspaces to Support Remote Collaboration in Incongruent Spaces. *Proc. ACM Hum.-Comput. Interact.* 6 (11 2022). Issue ISS. https://doi.org/10.1145/3567709

[13] Christian Frueh, Avneesh Sud, and Vivek Kwatra. 2017. Headset Removal for Virtual and Mixed Reality. In *ACM SIGGRAPH 2017 Talks* (Los Angeles, California) *(SIGGRAPH '17)*. ACM, New York, NY, USA, Article 80, 2 pages. https://doi.org/10.1145/3084363.3085083

[14] Susan R. Fussell, Robert E. Kraut, and Jane Siegel. 2000. Coordination of communication: effects of shared visual context on collaborative work. *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, 21–30. https://doi.org/10.1145/358916.358947

[15] Danilo Gasques, Janet G Johnson, Tommy Sharkey, Yuanyuan Feng, Ru Wang, Zhuoqun Robin Xu, Enrique Zavala, Yifei Zhang, Wanze Xie, Xinming Zhang, et al. 2021. ARTEMIS: A collaborative mixed-reality system for immersive surgical telementoring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14. https://doi.org/10.1145/3411764.3445576

[16] Steffen Gauglitz, Cha Lee, Matthew Turk, and Tobias Höllerer. 2012. Integrating the physical environment into mobile remote collaboration. *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 241–250. https://doi.org/10.1145/2371574.2371610

[17] William Gibson. 1984. *Neuromancer.* Ace Books. 271 pages.

[18] Jens Emil Grønbæk, Banu Saatçi, Carla F. Griggio, and Clemens Nylandsted Klokmose. 2021. MirrorBlender: Supporting Hybrid Meetings with a Malleable Video-Conferencing System. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA, Article 451, 13 pages. https://doi.org/10.1145/3411764.3445698

[19] Jens Emil Sloth Grønbæk, Ken Pfeuffer, Eduardo Velloso, Morten Astrup, Melanie Isabel Sønderkær Pedersen, Martin Kjær, Germán Leiva, and Hans Gellersen. 2023. Partially Blended Realities: Aligning Dissimilar Spaces for Distributed Mixed Reality Meetings. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3544548.3581515

[20] K Gupta, G A Lee, and M Billinghurst. 2016. Do You See What I See? The Effect of Gaze Tracking on Task Space Remote Collaboration. *IEEE Transactions on Visualization and Computer Graphics* 22 (2016), 2413–2422. Issue 11. https://doi.org/10.1109/TVCG.2016.2593778

[21] Carl Gutwin and Saul Greenberg. 1996. Workspace awareness for groupware. In *Conference Companion on Human Factors in Computing Systems* (Vancouver, British Columbia, Canada) *(CHI '96)*. ACM, New York, NY, USA, 208–209. https://doi.org/10.1145/257089.257284

[22] Carl Gutwin and Saul Greenberg. 2002. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)* 11 (2002), 411–446. https://doi.org/10.1023/A:1021271517844

[23] Robert A. Heinlein. 1999. Waldo (1942). , 125–212 pages.

[24] Jaylin Herskovitz, Yi Fei Cheng, Anhong Guo, Alanson P. Sample, and Michael Nebeling. 2022. XSpace: An Augmented Reality Toolkit for Enabling Spatially-Aware Distributed Collaboration. *Proceedings of the ACM on Human-Computer Interaction* 6 (11 2022), 277–302. Issue ISS. https://doi.org/10.1145/3567721

[25] Andrew Irlitti, Mesut Latifoglu, Qiushi Zhou, Martin N Reinoso, Thuong Hoang, Eduardo Velloso, and Frank Vetere. 2023. Volumetric Mixed Reality Telepresence for Real-time Cross Modality Collaboration. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3544548.3581277

[26] Andrew Irlitti, Thammathip Piumsomboon, Daniel Jackson, and Bruce H Thomas. 2019. Conveying spatial awareness cues in xR collaborations. *IEEE Transactions on Visualization and Computer Graphics* 25 (11 2019), 3178–3189. Issue 11. https://doi.org/10.1109/TVCG.2019.2932173

[27] P Isenberg, D Fisher, M R Morris, K Inkpen, and M Czerwinski. 2010. An exploratory study of co-located collaborative visual analytics around a tabletop display. *2010 IEEE Symposium on Visual Analytics Science and Technology*, 179–186. https://doi.org/10.1109/VAST.2010.5652880

[28] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon. 2011. KinectFusion. *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 559–568. https://doi.org/10.1145/2047196.2047270

[29] Allison Jing, Kieran William May, Mahnoor Naeem, Gun Lee, and Mark Billinghurst. 2021. EyemR-Vis: Using Bi-Directional Gaze Behavioural Cues to Improve Mixed Reality Remote Collaboration. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411763.3451844

[30] T Kanade, P Rander, and P J Narayanan. 1997. Virtualized reality: constructing virtual worlds from real scenes. *IEEE MultiMedia* 4 (1997), 34–47. Issue 1. https://doi.org/10.1109/93.580394

[31] Seungwon Kim, Gun Lee, Mark Billinghurst, and Weidong Huang. 2020. The combination of visual communication cues in mixed reality remote collaboration. *Journal on Multimodal User Interfaces* 14 (12 2020), 321–335. Issue 4. https://doi.org/10.1007/s12193-020-00335-x

[32] K. Kiyokawa, M. Billinghurst, S.E. Hayes, A. Gupta, Y. Sannohe, and H. Kato. 2002. Communication behaviors of co-located users in collaborative AR interfaces. *Proceedings. International Symposium on Mixed and Augmented Reality*, 139–148. https://doi.org/10.1109/ISMAR.2002.1115083

[33] M Kowalski, J Naruniec, and M Daniluk. 2015. Livescan3D: A Fast and Inexpensive 3D Data Acquisition System for Multiple Kinect v2 Sensors. *2015 International Conference on 3D Vision*, 318–325. https://doi.org/10.1109/3DV.2015.43

[34] Balasaravanan Thoravi Kumaravel, Fraser Anderson, George Fitzmaurice, Bjoern Hartmann, and Tovi Grossman. 2019. Loki: Facilitating Remote Instruction of Physical Tasks Using Bi-Directional Mixed-Reality Telepresence. *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, 161–174. https://doi.org/10.1145/3332165.3347872

[35] Gregorij Kurillo, Ruzena Bajcsy, Klara Nahrsted, and Oliver Kreylos. 2008. Immersive 3D Environment for Remote Collaboration and Training of Physical Activities. *2008 IEEE Virtual Reality Conference*, 269–270. https://doi.org/10.1109/VR.2008.4480795

[36] Marc Erich Latoschik, Daniel Roth, Dominik Gall, Jascha Achenbach, Thomas Waltemate, and Mario Botsch. 2017. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*. ACM, New York, NY, USA, 1–10. https://doi.org/10.1145/3139131.3139156

[37] Jason Lawrence, Danb Goldman, Supreeth Achar, Gregory Major Blascovich, Joseph G. Desloge, Tommy Fortes, Eric M. Gomez, Sascha Häberling, Hugues Hoppe, Andy Huibers, Claude Knaus, Brian Kuschak, Ricardo Martin-Brualla, Harris Nover, Andrew Ian Russell, Steven M. Seitz, and Kevin Tong. 2021. Project starline. *ACM Transactions on Graphics* 40 (12 2021), 1–16. Issue 6. https://doi.org/10.1145/3478513.3480490

[38] Benjamin Lee, Xiaoyun Hu, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. 2021. Shared Surfaces and Spaces: Collaborative Data Visualisation in a Co-located Immersive Environment. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 1171–1181. https://doi.org/10.1109/TVCG.2020.3030450

[39] Kyungjin Lee, Juheon Yi, and Youngki Lee. 2023. FarfetchFusion: Towards Fully Mobile Live 3D Telepresence Platform. *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*, 1–15. https://doi.org/10.1145/3570361.3592525

[40] Sang-Yup Lee, Ig-Jae Kim, Sang C Ahn, Heedong Ko, Myo-Taeg Lim, and Hyoung-Gon Kim. 2004. Real time 3D avatar for interactive mixed reality. In *Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry - VRCAI '04*. ACM Press, New York, New York, USA, 75. https://doi.org/10.1145/1044588.1044602

[41] David Lindlbauer and Andy D. Wilson. 2018. Remixed Reality: Manipulating Space and Time in Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173703

[42] Andrew Maimone and Henry Fuchs. 2011. Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras. *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, 137–146. https://doi.org/10.1109/ISMAR.2011.6092379

[43] Andrew Maimone, Xubo Yang, Nate Dierk, Andrei State, Mingsong Dou, and Henry Fuchs. 2013. General-purpose telepresence with head-worn optical see-through displays and projector-based lighting. *2013 IEEE Virtual Reality (VR)*, 23–26. https://doi.org/10.1109/VR.2013.6549352

[44] Microsoft Mesh. 2023. https://www.microsoft.com/en-us/mesh

[45] Nicholas Michael, Maria Drakou, and Andreas Lanitis. 2017. Model-based generation of personalized full-body 3D avatars from uncalibrated multi-view photographs. *Multimedia Tools and Applications* 76, 12 (jun 2017), 14169–14195. https://doi.org/10.1007/s11042-016-3808-1

[46] Paul Milgram and Fumio Kishino. 1994. Taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems* E77-D (1994), 1321–1329. Issue 12.

[47] Jens Müller, Roman Rädle, and Harald Reiterer. 2016. Virtual Objects as Spatial Cues in Collaborative Mixed Reality Environments. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 1245–1249. https://doi.org/10.1145/2858036.2858043

[48] Thomas Neumayr, Mirjam Augstein, and Bettina Kubicek. 2022. Territoriality in Hybrid Collaboration. *Proc. ACM Hum.-Comput. Interact.* 6 (11 2022). Issue CSCW2. https://doi.org/10.1145/3555224

[49] Thomas Neumayr, Hans-Christian Jetter, Mirjam Augstein, Judith Friedl, and Thomas Luger. 2018. Domino: A Descriptive Framework for Hybrid Collaboration and Coupling Styles in Partially Distributed Teams. *Proc. ACM Hum.-Comput. Interact.* 2 (11 2018). Issue CSCW. https://doi.org/10.1145/3274397

[50] Ohan Oda and Steven Feiner. 2012. 3D referencing techniques for physical objects in shared augmented reality. *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 207–215. https://doi.org/10.1109/ISMAR.2012.6402558

[51] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L Davidson, Sameh Khamis, Mingsong Dou, Vladimir Tankovich, Charles Loop, Qin Cai, Philip A Chou, Sarah Mennicken, Julien Valentin, Vivek Pradeep, Shenlong Wang, Sing Bing Kang, Pushmeet Kohli, Yuliya Lutchyn, Cem Keskin, and Shahram Izadi. 2016. Holoportation: Virtual 3D Teleportation in Real-Time. *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, 741–754. https://doi.org/10.1145/2984511.2984517

[52] Kazuhiro Otsuka. 2016. MMSpace: Kinetically-augmented telepresence for small group-to-group conversations. In *2016 IEEE Virtual Reality (VR)*. 19–28. https://doi.org/10.1109/VR.2016.7504684

[53] Kazuhiro Otsuka. 2018. Behavioral Analysis of Kinetic Telepresence for Small Symmetric Group-to-Group Meetings. *IEEE Transactions on Multimedia* 20, 6 (2018), 1432–1447. https://doi.org/10.1109/TMM.2017.2771396

[54] Oliver Otto, Dave Roberts, and Robin Wolff. 2006. A Review on Effective Closely-Coupled Collaboration Using Immersive CVE's. *Proceedings of the 2006 ACM International Conference on Virtual Reality Continuum and Its Applications*, 145–154. https://doi.org/10.1145/1128923.1128947

[55] Tomislav Pejsa, Julian Kantor, Hrvoje Benko, Eyal Ofek, and Andrew Wilson. 2016. Room2Room: Enabling Life-Size Telepresence in a Projected Augmented

Reality Environment. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1716–1725. https://doi.org/10.1145/2818048.2819965

[56] Thammathip Piumsomboon, Gun A Lee, Jonathon D Hart, Barrett Ens, Robert W Lindeman, Bruce H Thomas, and Mark Billinghurst. 2018. Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3173574.3173620

[57] Thammathip Piumsomboon, Gun A Lee, Andrew Irlitti, Barrett Ens, Bruce H Thomas, and Mark Billinghurst. 2019. On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3290605.3300458

[58] Lev Poretski, Ofer Arazy, Joel Lanir, Shalev Shahar, and Oded Nov. 2019. Virtual Objects in the Physical World: Relatedness and Psychological Ownership in Augmented Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland) *(CHI '19)*. ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300921

[59] Daniel Roth, Jean-Luc Lugrin, Dmitri Galakhov, Arvid Hofmann, Gary Bente, Marc Erich Latoschik, and Arnulph Fuhrmann. 2016. Avatar realism and social interaction quality in virtual reality. In *2016 IEEE Virtual Reality (VR)*. 277–278. https://doi.org/10.1109/VR.2016.7504761

[60] Banu Saatçi, Kaya Akyüz, Sean Rintel, and Clemens Nylandsted Klokmose. 2020. (Re)Configuring Hybrid Meetings: Moving from User-Centered Design to Meeting-Centered Design. *Computer Supported Cooperative Work (CSCW)* 29 (12 2020), 769–794. Issue 6. https://doi.org/10.1007/s10606-020-09385-x

[61] Banu Saatçi, Roman Rädle, Sean Rintel, Kenton O'Hara, and Clemens Nylandsted Klokmose. 2019. *Hybrid Meetings in the Modern Workplace: Stories of Success and Failure*. 45–61. https://doi.org/10.1007/978-3-030-28011-6_4

[62] Donghee Shin. 2022. The actualization of meta affordances: Conceptualizing affordance actualization in the metaverse games. *Computers in Human Behavior* 133 (2022), 107292. https://doi.org/10.1016/j.chb.2022.107292

[63] Misha Sra, Aske Mottelson, and Pattie Maes. 2018. Your Place and Mine: Designing a Shared VR Experience for Remotely Located Users. *Proceedings of the 2018 Designing Interactive Systems Conference*, 85–97. https://doi.org/10.1145/3196709.3196788

[64] A Steed, W Steptoe, W Oyekoya, F Pece, T Weyrich, J Kautz, D Friedman, A Peer, M Solazzi, F Tecchia, M Bergamasco, and M Slater. 2012. Beaming: An Asymmetric Telepresence System. *IEEE Computer Graphics and Applications* 32 (2012), 10–17. Issue 6. https://doi.org/10.1109/MCG.2012.110

[65] Jacob H. Steffen, James E. Gaskin, Thomas O. Meservy, Jeffrey L. Jenkins, and Iopa Wolman. 2019. Framework of Affordances for Virtual Reality and Augmented Reality. *Journal of Management Information Systems* 36, 3 (2019), 683–729. https://doi.org/10.1080/07421222.2019.1628877

[66] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. 2006. Collaborative Coupling over Tabletop Displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1181–1190. https://doi.org/10.1145/1124772.1124950

[67] Franco Tecchia, Leila Alem, and Weidong Huang. 2012. 3D helping hands: a gesture based MR system for remote collaboration. In *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry* (Singapore, Singapore) *(VRCAI '12)*. ACM, New York, NY, USA, 323–328. https://doi.org/10.1145/2407516.2407590

[68] Philip Tuddenham and Peter Robinson. 2009. Territorial Coordination and Workspace Awareness in Remote Tabletop Collaboration. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2139–2148. https://doi.org/10.1145/1518701.1519026

[69] Laurie M Wilcox, Robert S Allison, Samuel Elfassy, and Cynthia Grelik. 2006. Personal Space in Virtual Reality. *ACM Trans. Appl. Percept.* 3 (10 2006), 412–428. Issue 4. https://doi.org/10.1145/1190036.1190041

[70] Julie Williamson, Jie Li, Vinoba Vinayagamoorthy, David A Shamma, and Pablo Cesar. 2021. Proxemics and Social Interactions in an Instrumented Virtual Reality Workshop. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411764.3445729

[71] Andrew D. Wilson. 2017. Fast lossless depth image compression. *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces, ISS 2017* (10 2017), 100–105. https://doi.org/10.1145/3132272.3134144

[72] Yuanjie Wu, Yu Wang, Sungchul Jung, Simon Hoermann, and Robert W. Lindeman. 2021. Using a Fully Expressive Avatar to Collaborate in Virtual Reality: Evaluation of Task Performance, Presence, and Attraction. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.641296

[73] Michael Zollhöfer, Patrick Stotko, Andreas Görlitz, Christian Theobalt, Matthias Nießner, Reinhard Klein, and Andreas Kolb. 2018. State of the Art on 3D Reconstruction with RGB-D Cameras. *Computer Graphics Forum* 37 (5 2018), 625–652. Issue 2. https://doi.org/10.1111/cgf.13386