Re-locations: Augmenting Personal and Shared Workspaces to Support Remote Collaboration in Incongruent Spaces

DANIEL IMMANUEL FINK, University of Konstanz, Germany
JOHANNES ZAGERMANN, University of Konstanz, Germany
HARALD REITERER, University of Konstanz, Germany
HANS-CHRISTIAN JETTER, University of Lübeck, Germany

556

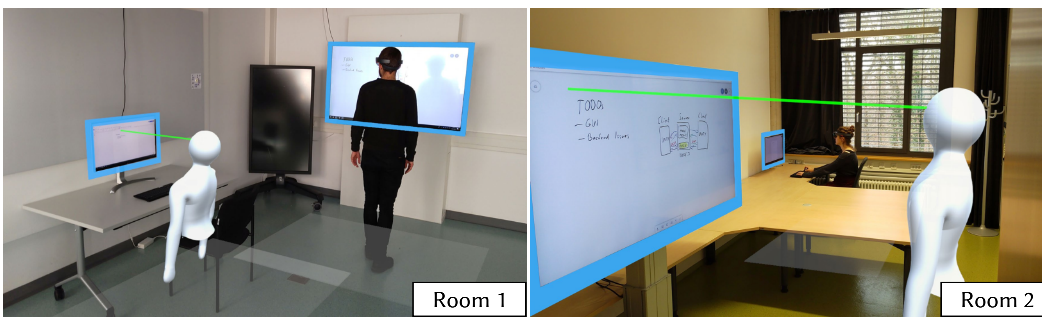

Fig. 1. *Re-locations* are user-defined workspaces that facilitate remote collaboration with augmented reality (AR) in incongruent spaces. The user defines the position and purpose of workspaces (blue borders) in a setup process. Afterward, every *Re-location* uses an independent local coordinate system in front of (or around) the workspace to maintain the common spatial frame of reference at this workspace. This way, remote users can be represented (e.g., visually with avatars and auditory with spatial audio) in the remote space at the position of the corresponding workspace with the same purpose.

Augmented reality (AR) can create the illusion of being virtually co-located during remote collaboration, e.g., by visualizing remote co-workers as avatars. However, spatial awareness of each other's activities is limited as physical spaces, including the position of physical devices, are often incongruent. Therefore, alignment methods are needed to support activities on physical devices. In this paper, we present the concept of *Re-locations*, a method for enabling remote collaboration with augmented reality in incongruent spaces. The idea of the concept is to enrich remote collaboration activities on multiple physical devices with attributes of co-located collaboration such as spatial awareness and spatial referencing by locally relocating remote user representations to user-defined workspaces. We evaluated the *Re-locations* concept in an explorative user study with dyads using an authentic, collaborative task. Our findings indicate that *Re-locations* introduce attributes of co-located collaboration like spatial awareness and social presence. Based on our findings, we provide implications for future research and design of remote collaboration systems using AR.

Authors' addresses: Daniel Immanuel Fink, University of Konstanz, Konstanz, Germany, daniel.immanuel.fink@uni-konstanz.de; Johannes Zagermann, University of Konstanz, Konstanz, Germany, johannes.zagermann@uni-konstanz.de; Harald Reiterer, University of Konstanz, Konstanz, Germany, harald.reiterer@uni-konstanz.de; Hans-Christian Jetter, University of Lübeck, Lübeck, Germany, jetter@imis.uni-luebeck.de.

2573-0142/2022/12-ART556
https://doi.org/10.1145/3567709

CCS Concepts: • **Human-centered computing** → **Mixed / augmented reality**; **Computer supported cooperative work**.

Additional Key Words and Phrases: Augmented reality, remote collaboration, synchronous, symmetric, shared workspace, incongruent spaces, dissimilar spaces.

ACM Reference Format:
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, ISS, Article 556 (December 2022), 30 pages. https://doi.org/10.1145/3567709

1 INTRODUCTION

Recently, there has been an explosive growth in the usage of synchronous remote collaboration platforms: Video conferencing tools enable real-time audio/video communication with colleagues, shared online editors provide access to common documents, and digital whiteboards can help to structure and plan projects. Many of these consumer applications can be sufficient for activities where simulated face-to-face communication is satisfactory. However, these remote work applications often lack a variety of key qualities of synchronous co-located collaboration, such as workspace awareness, social presence, and non-verbal communication. Previous research has investigated different ways to enrich such remote work experiences with attributes of co-located collaboration: For example, creating a common frame of reference by placing video feeds of remote collaborators across-the-table [43], by manipulating the position, scale, and transparency of shared content [18], and also by including shared virtual landmarks [40] have shown to enable spatial referencing via deictic gestures or speech. Especially augmented reality – considered as a promising candidate for next-generation collaborative systems [5] – holds the potential to enrich these remote experiences with the named qualities, e.g., by visualizing remote peers as highly-realistic avatars [44]. Here, recent research has investigated various ways to represent remote users in 3D space (e.g., [9, 35, 53]) in user study settings facilitated by almost architecturally identical remote work spaces (e.g., to ensure the internal validity [40]).

However, in reality, (remote) workspaces can differ fundamentally and incongruent physical and spatial work environments are quite usual. This can lead to different negative side effects such as reduced spatial and workspace awareness or a missing frame of spatial referencing, which can make verbal cues, deictic gestures, or the interpretation of viewing frustums or gaze indicators ambiguous and error prone – reducing the experience of social presence and efficiency of collaborative work as a whole. Additionally, such incongruences prevent the experience of proxemics of social interaction [43]. Recent research has investigated ways to align and merge incongruent remote workspaces for virtual environments (e.g., [10, 56]). For augmented reality, latest research has identified ways to position remote peers as virtual avatars in the local space [62], however, using a non-linear mapping of movements proved to be problematic.

In this paper, we present the concept of ***Re-locations*** – a solution approach to address the problems of incongruent physical and spatial workspaces for synchronous remote collaboration with augmented reality. *Re-locations* reduce the complexity of incongruent physical and spatial workspaces by allowing users to perceive collaborative activities and physical movements only in relation to user-defined virtually shared landmarks (called *Re-locations*). In contrast to previous work that included landmarks as passive objects [40], *Re-locations* can also be interactive: Typical examples include large touchscreens, interactive whiteboards, or desktop computers – locations where local users are „physically“ and remote users are „virtually“. We intentionally decided to include these devices into our concept to create hybrid user interfaces [13] that compensate for the limitations of interacting with augmented reality content [21, 28] while allowing users to stick to common work practices. However, the *Re-locations* concept itself is medium-agnostic, which means that any medium (e.g., 2D content on a display or 3D visualizations in augmented reality) holds the potential to be included as a *Re-location*. Each of these *Re-locations* serves as the origin of a coordinate system to determine the relative positions of the local users that are then used to visualize their virtual counterpart on the remote site. With this, *Re-locations* serve as shared frames of spatial reference for verbal, gestural, or auditory cues. This allows to shift the spatial cognition from individual, incongruent ego-centric reference systems towards shared allocentric reference systems. Users can freely define the local position of all virtually shared landmarks (e.g., placing a virtually shared touchscreen at a position that is locally accessible), while the relative positions of local and remote users towards the landmark are preserved to reproduce proxemics and awareness despite of virtual presence. The goal of these virtually shared landmarks is to preserve natural spatial and social relationships between local and remote collaborators.

The concept of *Re-locations* is the foundation of an abundance of research questions regarding their benefits, drawbacks, or possible disruptions of collaborative activities when applied to real-world situations. Therefore, we decided to implement an initial prototypical setting that allows participants to experience the concept in a basic way. In a user study, we motivated participants to interact with multiple *Re-locations*. Here, we studied emergent user behaviors and classified these collaborative activities using well-established coupling styles [42]. Quantitative data using standardized questionnaires and data loggings as well as subjective feedback by participants helped us to further understand our observations. The purpose of the study was to explore the effects of *Re-locations* on collaborative activities (e.g., classified as coupling styles) and the sensation of presence; rather than proving that *Re-locations* can solve all problems related to incongruent workspaces. Our study was not intended to be a controlled experiment to perform a quantitative comparison of collaboration with and without *Re-locations*. Since our concept contains elements with little or no prior research, we aimed to assess the feasibility of our concept, uncover initial problems, and reveal implications for future research and design.

In the remainder of this paper, we discuss related work as a foundation of the *Re-locations* concept, describe the concept itself, the prototypical implementation for our user study, as well as the user study itself. Finally, we discuss our findings based on our research questions.

In summary, our paper contributes: (1) the concept of *Re-locations*, a solution approach to address the problems of incongruent physical and spatial workspaces for synchronous remote collaboration with augmented reality; (2) an initial user study with a prototypical implementation of the concept; and (3) the identification of opportunities and challenges for future work as an outcome of the discussion of the results of the user study.

2 RELATED WORK

Remote collaboration using mixed reality (MR; including both, augmented reality (AR) and virtual reality (VR) [52]) has the goal "to enable remote people to feel that they are virtually co-located." [32]. In recent research, several examples emerged that demonstrate how MR can be useful for remote collaboration (e.g., [40, 47, 53]). Ens et al. [11] summarized the current state of collaborative MR research: Here, they categorize papers according to dimensions derived from the literature (the CSCW matrix dimensions of space and time [26], symmetry [4], artificiality [3]), and the additional dimensions of focus and scenario. Interestingly, the majority of research on remote collaboration facilitated by MR focuses on asymmetric collaboration, such as remote expert scenarios (e.g. [12, 27]). For symmetric collaboration, different methods emerged that support, e.g., shared workspaces or telepresence scenarios (e.g. [20, 44, 47, 55]). For these scenarios, various representation techniques of remote users in 3D space were developed: While some techniques are rather abstract and low-cost (e.g., virtual viewing frustums [40] with additional gaze ray visualization [46]), others are realistic yet hardware-heavy (e.g., Holoportation from Microsoft Research [44]). A popular representation technique that combines attributes from both, abstract *and* realistic techniques are 3D avatars (e.g., [35, 53, 60]). However, the representation of remote users as avatars in 3D space plays an important yet functional role in our research. While recent research is pushing the attributes of virtual avatars more and more into humanoid resemblances, we see the avatar – in our case – rather as a tool to create the sensation of presence.

In the following, we will describe related work that influenced and inspired our own work: From (1) Remote Shared Workspaces in Mixed Reality, to (2) Remote Collaboration using Interactive Surfaces, and (3) Enabling Tools for Remote Collaboration in Incongruent Spaces.

2.1 Remote Shared Workspaces in Mixed Reality

Remote shared workspaces facilitated by MR technologies can create the illusion that remote users and virtual objects become *virtually* co-located. Besides VR-only applications like Mozilla Hubs [14] or MeetinVR [1] there are also applications like Spatial [54] or Microsoft Mesh [36] that support remote collaboration in augmented reality. AR solutions integrate the physical environment as well as the available physical devices in the space to a limited extent: For example, in Spatial [54], users can define a wall in the room that allows to attach virtual elements (like sticky notes) or share screen content from desktop devices (like laptops) – becoming a part of the AR space. However, provided awareness cues like the position or viewing direction of remote users (represented as 3D avatars) are limited to interactions with virtual elements of the AR space. This means that the avatars do not visualize if a remote peer uses or looks at physical objects or devices (e.g., screens) in their own local environment, which might influence awareness on each other's activities.

Several related work focused on integrating the physical environment in the MR space and allow remote users to see and interact with the collaborator's local surroundings using MR technology. Here, previous research focused on a) using multiple head-mounted cameras to create 360-degree representations in real-time [41], allowing remote users to interact with the local site [55] and b) using HMDs to scan a local environment that can be navigated within an HMD from the remote site [47]. This way, collaborators have to decide on the physical environment of one user that can then be experienced in a VR simulation (remote) and in AR (local) [47]. This asymmetric collaboration allows the local collaborator to interact with their physical environment including digital devices (e.g., laptop, interactive surfaces), but remote collaborators can only interact with virtual content, excluding the potential of using digital devices as additional possibility for remote collaboration (e.g., using Google Docs).

2.2 Remote Collaboration using Interactive Surfaces

Previous research investigated different ways to align the representation of remote users into interactive surfaces like tabletops and wall-displays. They integrated overlays of remote users' arms on an interactive surface [57] in combination with a video stream [45]. Visualizing the remote person as part of the local workspace can create the impression of looking through a screen into the remote room [20, 64] or simulating face-to-face work [38]. However, these methods restrict the workspace to a single device or location in a room and the provided awareness cues are limited to the 2D space of the surfaces. This raises the question how physical screens can be integrated into the augmented reality environment so that awareness cues are provided for both, interactions in the 2D surface space as well as in the augmented 3D space. However, a combination of the described 2D user representations (as part of the interactive surface applications) with 3D user representations (as part of the augmented reality applications) (e.g., avatars) might result in multiple (disconnected) remote user representations per remote user, which might reduce the awareness instead of enhancing it. An 3D user representation could also provide awareness cues for interactions with physical displays, but the differences in room layouts of remote users, including the physical screen positions, prevent the establishment of a common spatial frame of reference when movements of remote users are represented 1:1 in the AR environment. Therefore alignment methods are needed that adapt the remote user representation to the local physical environment.

2.3 Enabling Tools for Remote Collaboration in Incongruent Spaces

Local physical environments of remote co-workers might differ, which can lead to two problems: (1) Virtual objects might be out of reach or outside of the local walkable space (e.g., in a wall) and (2) local physical objects or devices (like screens) are not available or are not at the same place remotely. The latter problem mainly concerns AR applications. Therefore, methods emerged that try to mitigate those problems.

Sra et al. [56] developed techniques to align rectangular VR environments with different walkable spaces by scaling spaces (and movements) or spatially shifting them. Here, a 1:1 mapping of movements showed higher co-presence scores than scaled movements. In line with this, Congdon et al. [10] developed a technique to merge different walkable VR spaces independent of their layout using pre-defined corresponding reference locations (cf. [25]). Here, the technique shows the VR environment differently, depending on the user's local space, and dynamically maps local movements into the remote user's environment. This solution achieved high co-presence scores, and only a few participants were aware of mapping artifacts.

Yoon et al. [62] developed a technique to position avatars (representing remote users) in different physical environments (i.e., apartments) depending on the remote user's current position in their local environment and further attributes: (1) "Interaction" (avatar must be placed where the current interaction is possible), (2) "Pose" (avatar must be placed where the current pose is possible, e.g., sitting on a chair), and (3) "Space" (space is divided into functional spaces and avatar must be placed in the current functional spaces, e.g., dining area). For their placement algorithm they performed a user survey asking for the preferred avatar position in different situations and environments, identified a set of features, and trained a neural network for generating the plausible position of the avatar. However, for their method, they had to manually define all features of the environments, resulting in a high configuration effort. In a recent work, Yoon et al. [61] adapted their technique to enable remote collaboration in front of a display including avatar placement and gesture retargeting. The configuration effort is still high because the position of seats, screens, and other pointing targets must be configured manually.

The ideas of representing remote users depending on pre-defined areas [10, 25] and semantic attributes [61, 62] of the physical environment, including a 1:1 mapping of movements [56], inspired us to create a concept facilitating remote collaboration with multiple workspaces using AR. The next section presents our resulting concept called *Re-locations*.

3 CONCEPT: RE-LOCATIONS

Using augmented reality to visualize a remote person as part of the physical space can create the illusion that the remote person is present (i.e., virtually co-located). This way, augmented reality can potentially facilitate remote collaboration with spatial activities and, in contrast to virtual reality (VR) that excludes physical surroundings, still allows for interaction with peers and physical objects in the individual local environment (i.e., paper notes, computers, smartphones, tablets) of all collaborators. However, when using augmented reality to visualize remote persons as part of the 3D space (e.g., as avatars), the spatial awareness of each other's activities is limited as physical spaces, including the position of objects and devices, are often incongruent. To explain the problem of incongruent physical spaces, we use the example of the two co-workers Paul and Anna, who want to collaborate remotely in their individual offices using two physical devices: a large screen with an online shared whiteboard and a desktop computer for individual activities. Figure 2 shows Paul's and Anna's offices and highlights the shared whiteboard screen (red border) and the desktop screen (yellow border).

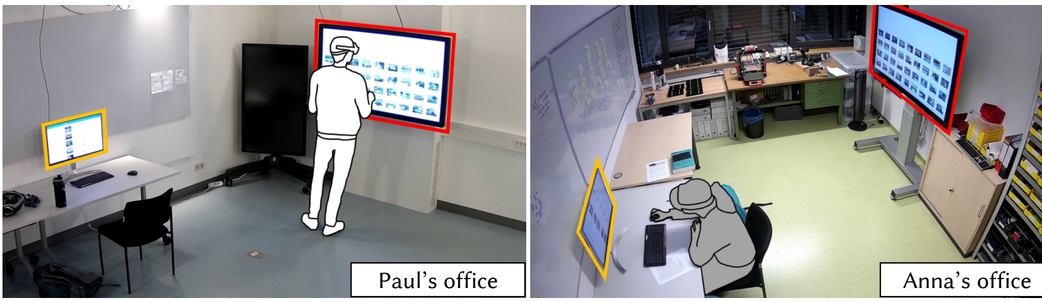

Fig. 2. Two co-workers Paul (white) and Anna (gray) collaborate remotely in their offices using two physical devices: a large screen with an online shared whiteboard (red border) and a desktop computer for individual research (yellow border).

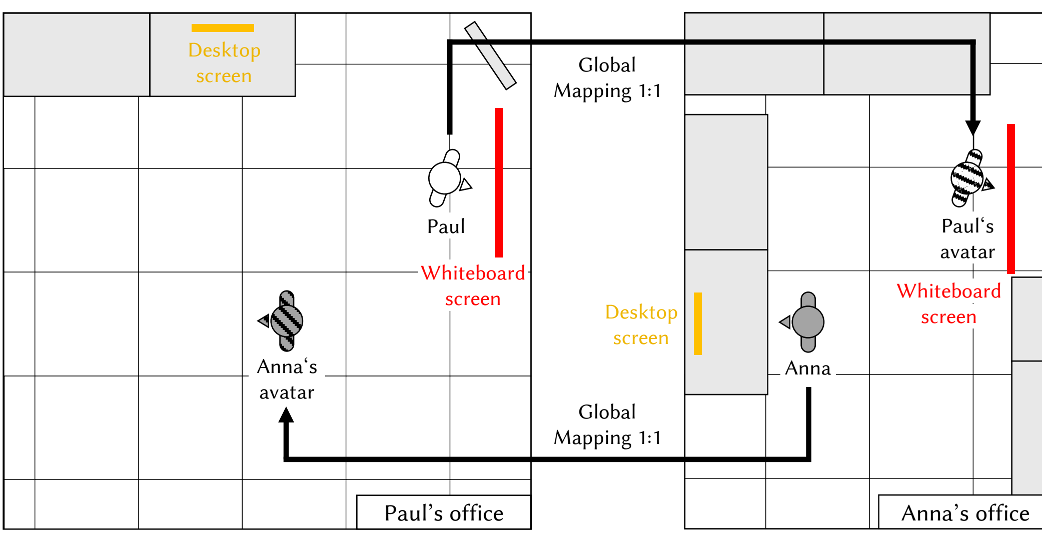

Fig. 3. The use of a global coordinate system for the visualization of remote users (e.g., as avatars) in incongruent physical spaces can lead to the loss of common spatial frames of reference. While Paul's avatar in Anna's office (right) is also in front of the whiteboard screen, Anna's avatar in Paul's office (left) is not in front of the desktop screen.

When using a global 1:1 mapping to visualize Paul (as an avatar) in Anna's office and Anna (as an avatar) in Paul's office, a common spatial frame of reference cannot be established for both areas that are relevant for collaboration (desktop screen and whiteboard screen) at the same time. Figure 3 shows a top-down view of the offices and how a global 1:1 mapping of the movements would affect the position of avatar visualizations representing the remote persons. While Paul's avatar in Anna's office is also in front of the whiteboard screen (like Paul in his office), Anna's avatar in Paul's office (left) is not in front of the desktop screen (like Anna in her office).

This problem of incongruence cannot be solved by translating, rotating, or linearly scaling the global coordinate system in one or both spaces, since the relevant shared landmarks (i.e., the desktop screen and the whiteboard screen) are aligned differently to each other in the different spaces. Therefore, incongruent spaces can make it difficult to establish common spatial frames of reference that are essential for nonverbal communication like deictic gestures or viewing directions.

Based on the findings of previous work, we developed the concept of *Re-locations*, a method for facilitating remote collaboration with augmented reality in incongruent spaces. The general idea of the following concept is to enrich remote collaboration activities on (multiple) physical devices with attributes of co-located collaboration such as spatial awareness and spatial referencing: Bridging the gap between *re*mote and co-*located* collaboration by locally *relocating* remote user representations to user-defined workspaces to establish a common spatial frame of reference.

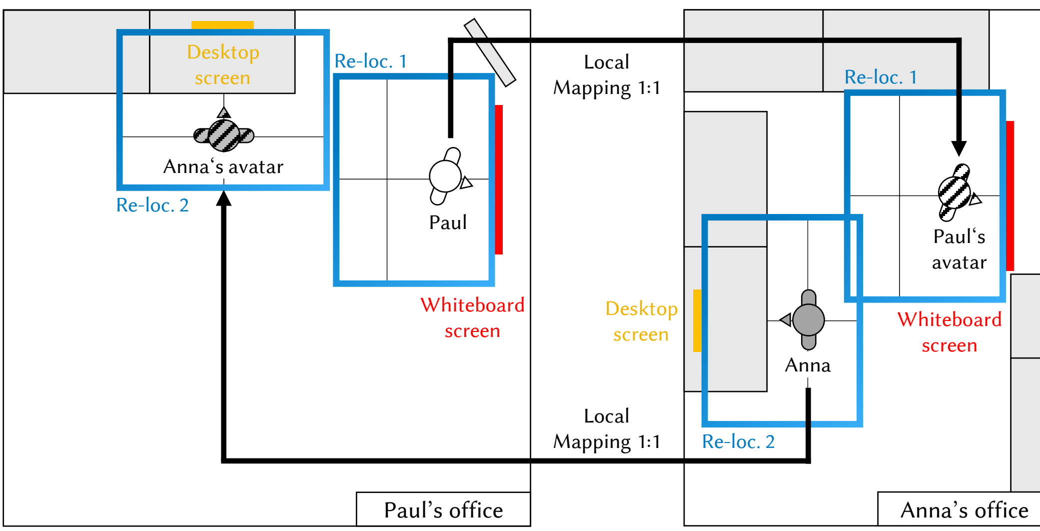

Fig. 4. The *Re-locations* concept uses local coordinate systems in front of (or around) user-defined workspaces to maintain the common spatial frame of reference at all collaboration relevant areas in space. This way, Paul’s avatar is in Anna’s office (right) like Paul (in his office) in front of the whiteboard screen and Anna’s avatar is in Paul’s office (left) like Anna (in her office) in front of the desktop screen.

Re-locations are user-defined workspaces that can be used to facilitate remote collaboration (please also refer to our supplemental video for animated explanations of the concept). Each *Re-location* has its specific purpose (e.g., a type of personal workspace such as a desktop computer or a type of shared workspace such as a shared whiteboard) that is defined in a setup process. In our example scenario, Paul and Anna use two physical devices for collaboration, so they define two *Re-locations*, one for the desktop screen and one for the whiteboard screen. Figure 4 shows the top-down view of their offices including the defined *Re-locations*. Instead of using a global coordinate system for the whole space (like in Figure 3), the *Re-locations* concept uses for every *Re-location* an independent local coordinate system in front of (or around) the user-defined workspace to maintain the common spatial frame of reference at this workspace. This way, avatars can be visualized in the remote space inside the corresponding *Re-locations* (with the same purpose).

With this method, Anna’s avatar (in Paul’s office) is like Anna (in her own office) located in front of the desktop screen (see Figure 4 left), and Paul’s avatar (in Anna’s office) is like Paul (in his own office) located in front of the whiteboard screen (see Figure 4 right). The avatars are located inside the corresponding remote *Re-location* at the same position as the users in their own local *Re-location* – imitating their orientation and movements. The prerequisite for this is the definition of a matching *Re-location* with the same purpose at the remote site.

Remote users are only represented as avatars inside *Re-locations*. When a user leaves a *Re-location*, the avatar fades out at the remote site. When a user enters a *Re-location*, the avatar fades in again. For example, when Anna leaves the desktop screen *Re-location* and enters the whiteboard screen *Re-location* in her local office, her avatar will fade out at the desktop and appear at the whiteboard in Paul's office. While this reduces visual clutter, it also allows for a 1:1 mapping of movements (cf. Sra et al. [56]). With this, collaborators can transition between tightly-coupled collaboration to loosely-coupled parallel work while working in different places with different devices in their own room, although the remote site might have an entirely different room layout.

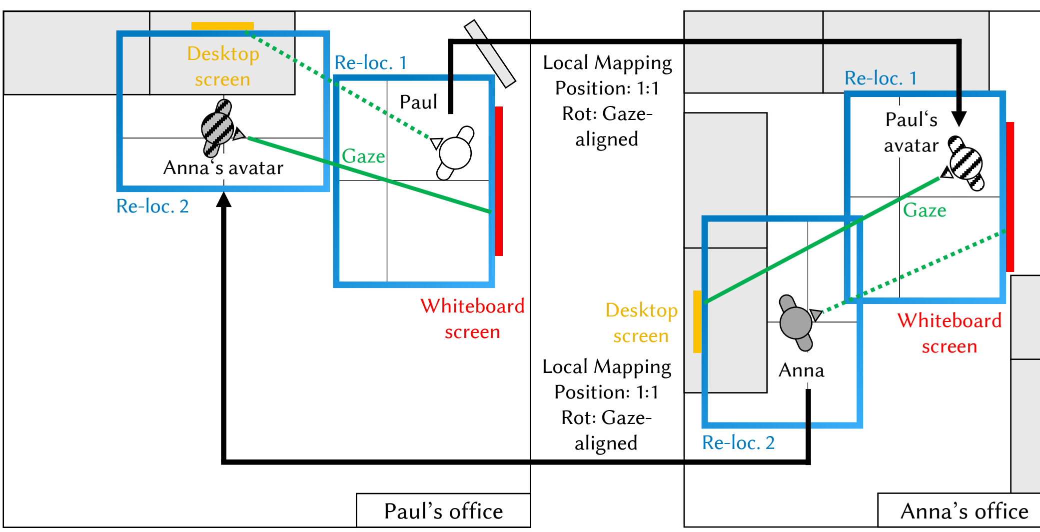

Fig. 5. The *Re-locations* concept allows to provide gaze visualizations across *Re-locations*. When Anna is standing inside the desktop screen *Re-location* while looking at the whiteboard screen (right) the 1:1 mapping of the rotation will be interrupted, and her avatar rotates towards the whiteboard in Paul's office (left). The same behavior applies to Paul (left) and his avatar (right) who is standing inside the whiteboard screen *Re-location* while looking at the desktop screen.

Integrating additional awareness cues like gaze visualizations further support spatial awareness and spatial referencing (cf. Piumsomboon et al. [46]). The *Re-locations* concept also allows to provide additional awareness cues not only within but also across *Re-locations*. Figure 5 shows this concept, where Anna is standing inside the desktop screen *Re-location* while looking at the whiteboard screen. In this situation, the 1:1 mapping of the rotation will be interrupted, and her avatar rotates towards the whiteboard in Paul's office. The gaze ray of her avatar points to the same spot on the whiteboard she is looking at. The same behavior applies to Paul and his avatar who is standing inside the whiteboard screen *Re-location* while looking at the desktop screen (see Figure 5).

This is possible as the positions and dimensions of the *Re-locations'* workspaces are defined in the setup process. First it is checked if the gaze intersects a workspace area (screen area). Then, the 3D intersection point between gaze and workspace area is calculated. Afterward, the 3D intersection point is converted into the 2D coordinate system of the workspace (e.g., the screen) and finally normalized (bottom left of the workspace = (0,0), top right of the workspace = (1,1)), which allows to capture the gaze position independent of screen sizes and resolutions. At the remote side, the normalized 2D intersection point is converted (at the local position of the corresponding workspace) into the 3D coordinate system. Finally, the avatar interrupts the local 1:1 mapping of the rotation (inside the *Re-location*) and rotates the body and head towards the calculated 3D position.

However, depending on the screen orientations in the rooms, situations can occur where the gaze ray might point to the back of a screen because the screen cannot be viewed from the position of the avatar. In these situations, the system should recognize that the gaze visualization would result in a mismatch with the real world and disable the gaze ray or replace it with an alternative gaze visualization to avoid ambiguities.

As multiple *Re-locations* might be overlapping (see Figure 5 right), the gaze direction is used to decide inside which *Re-location* the avatar at the remote site will be shown. When a user is not looking at a *Re-location* workspace, the smallest angle to a *Re-location* workspace decides.

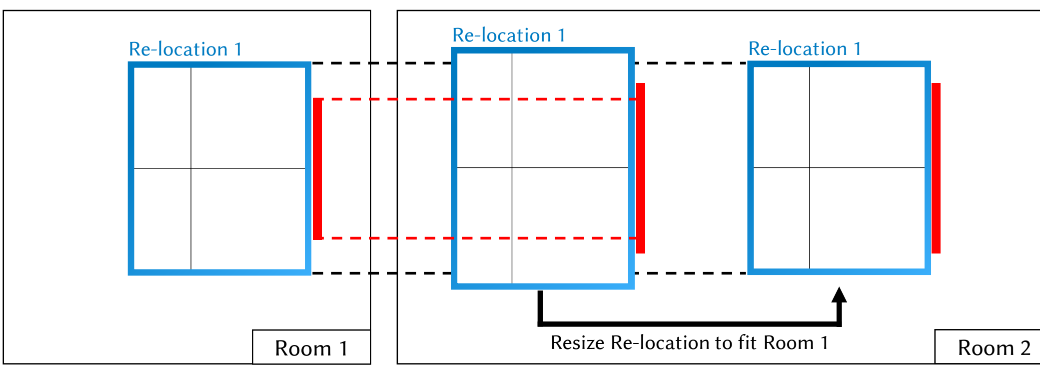

Fig. 6. When the size of *Re-locations* with the same purpose differ between rooms (left and middle), to maintain the 1:1 mapping, the bigger *Re-location* (middle) is resized to fit the smaller *Re-location* (right).

Additionally, matching *Re-locations* with the same purpose can differ in their size. Similarly to previous work [56], we decided to shrink matching *Re-locations* to the size of the smallest *Re-location* (cf. Figure 6). This allows for a 1:1 mapping of avatar position, rotation, and movement. Alternatively, scaling *Re-locations* would result in scaling of movements, which might be irritating. However, the gaze-ray will be scaled to support different display sizes. For example, when a local user with a large screen looks at the top right corner of the screen, then their avatar at the remote site also looks at the top right corner of the smaller screen. In this case, the 1:1 mapping is adapted, similarly to the gaze visualization across multiple *Re-locations* (see the normalization of the workspace's 2D coordinate system described above).

Finally, the *Re-locations* concept is not limited to the the visual representation of the remote user but also supports the use of spatial audio to playback the voice of the remote user at the position of the avatar, which potentially further supports the spatial awareness of remote co-workers. Furthermore, it is necessary to mention that the *Re-location* concept is not bound to a specific application or activity. The concept works *as is* as a facilitating technology for remote collaboration in augmented reality, e.g., with standard collaboration tools (e.g., whiteboard applications, Google Docs, Overleaf, etc.), custom-developed software, or even as a stand-alone technology, e.g., to increase the sensation of presence in shared activities using physical devices.

3.1 Prototype

For the exploration of the *Re-locations* concept, we developed an initial prototype using the AR HMD Microsoft HoloLens (first generation). The *Re-locations* concept is not only limited to screens. For example, it would be possible to define a table *Re-location* and display 3D objects on it. However, this first implementation only supports *Re-locations* in front of physical screens to investigate the basic concept. The prototype was implemented with Unity3D [59], the Mixed Reality Toolkit [37], Vuforia [48], and Node.js [15].

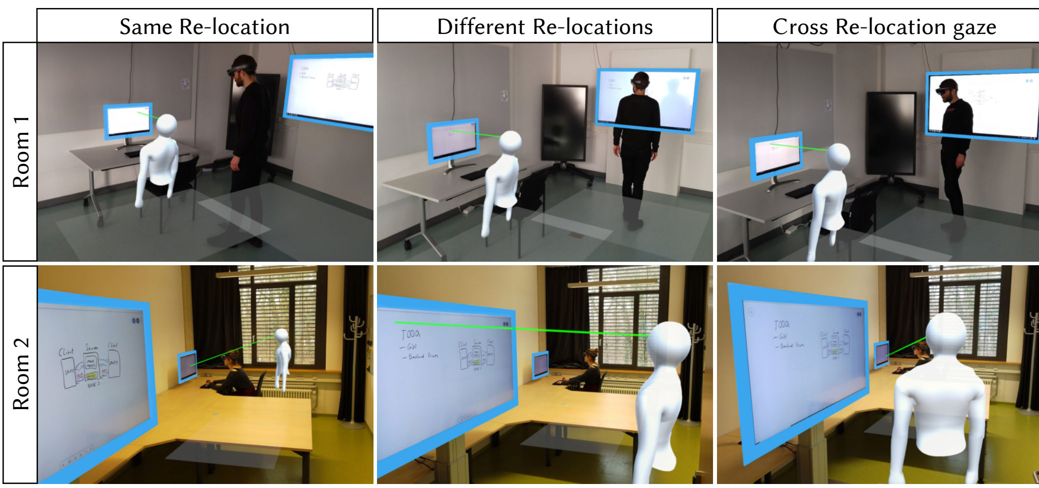

Fig. 7. The initial *Re-locations* prototype from a third person view. On the left, both users are inside the same *Re-location* (in front of a desktop screen). In the middle, the users are inside different *Re-locations*. On the right, the user in room 1 is inside the whiteboard *Re-location* and looks at the desktop screen. Therefore, the avatar in room 2 also looks at the desktop screen. All images were captured with a third HoloLens from the first person view. Due to a problem with the HoloLens capturing function, augmented content is displayed offset in HoloLens capturings. We, therefore, post-processed and adjusted the augmented content of the images (e.g., blue frames) for presentation clarity afterwards.

For the implementation of the *Re-locations* concept, a setup process is required, in which the position and size of the screens in the room are calibrated. Additionally, the prototype should provide the ability to save the performed calibration to avoid a setup process before each use. For these reasons, after the start of the prototype, the user can either join an already calibrated room or calibrate a new room. For the room calibration, the user needs two printed Vuforia markers [48]: a room marker and a calibration marker. The room marker is unique for every room, while the calibration marker is always the same. First, the room marker is placed at a fixed position in the room. The room marker identifies the room and acts as an anchor, and therefore, its position must not change after calibration. At the beginning of the calibration, the user needs to look at the floor and the room marker. Afterward, the user can calibrate a screen in the room by selecting a *Re-location* type (e.g., desktop, whiteboard) and placing the calibration marker on the bottom left and top right corner of the screen. The last part of the procedure can be repeated several times to calibrate several screens. When the room calibration is finished, it will be saved on a Node.js server, bound to the used room marker. Therefore, it only needs to be performed once per room (or if the room layout or devices are changed).

For joining an already calibrated room, the user only needs to look at the room marker, and the room layout and configuration will be loaded from the server. Figure 7 shows the prototype

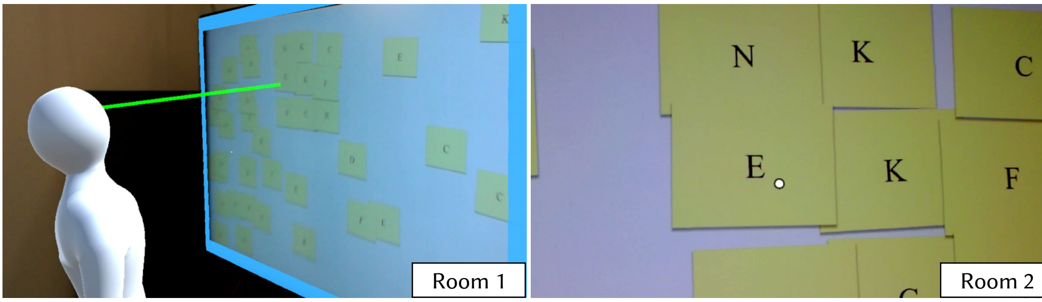

Fig. 8. The initial *Re-locations* prototype from the first person view showing a trial task data set on the screen. On the left, the user in room 1 stands in front of the whiteboard screen and looks at the avatar of the user in room 2. On the right, the user in room 2 stands in front of the whiteboard screen and looks at it.

from a third-person view and Figure 8 from the first person view. After users joined a room, they can hear each other and when inside a *Re-location* (in front of a calibrated screen) also see each other as a white three-dimensional upper-body avatar. For this, the position, rotation, and voice data are sent to the Node.js server and distributed to the clients (i.e., HMD). The representation as an upper-body avatar seems sufficient [35] and solves the problem that the HoloLens can not track the user's feet. The avatar is abstract and has no facial properties (no eyes, nose, mouth), but the front of the face is recognizable by a head gaze ray (cf. Figure 7). We intentionally decided to include abstract avatars, as they play an important yet rather functional role in our research. However, previous work has shown that abstract avatars provide high social presence scores with conversation patterns similar to co-located collaboration [53].

When the remote user is inside a *Re-location*, the avatar hovers at the same position in the remote *Re-location*. The head of the avatar imitates the head movement of the remote user. The body starts to rotate with the head when the angle between the body and head (front direction) is higher than 40 degrees. We used empirical data sampling to find a threshold for the maximum angle between body and head – 40 degrees proved to be a good fit. An exception is when the user looks from one *Re-location* at an avatar or screen in another *Re-location* (see Figure 7 right). In this case, the avatar rotates towards the viewing point of the remote user in the local room (cross *Re-location* gaze). Additionally, the head gaze is visualized by a green ray when the remote user is looking at an avatar or screen (see Figure 8). Blue borders around the screens indicate that they are calibrated. The size of a *Re-location* is visible by a white semi-transparent area on the floor. Furthermore, the prototype supports spatial audio. This allows users to hear the remote user's voice from the direction of the corresponding avatar in the local room. Existing voice chat solutions only rarely support spatial audio or transmit the voice data over an external server. Therefore, the voice chat was self-implemented to provide a high quality (raw audio data is transmitted), stability, and spatial audio. The *Re-locations* prototype is available as open-source project on GitHub[1].

4 STUDY

We conducted an exploratory user study to investigate the *Re-locations* concept with an initial prototype (with basic functionality) to answer our research questions. The overall goal of our study was to understand how *Re-locations* can facilitate remote collaboration in incongruent spaces. We address this overall research goal by investigating the following main and sub research questions, both qualitatively and quantitatively:

[1] https://github.com/hcigroupkonstanz/Re-locations

RQ1: How does the use of *Re-locations* influence collaborative behavior?

- **RQ1.1:** Which styles of collaborative coupling do participants apply when using *Re-locations*? To what extent are they using coupling styles that have been reported exclusively for co-located collaboration?
- **RQ1.2:** How does the use of *Re-locations* influence participants in their individual use of space, the provided personal and shared workspaces, and the role of furniture?
- **RQ1.3:** Do users expose natural social behavior that indicates an illusion of co-located collaboration such as (deictic) gestures?

RQ2: How does the use of *Re-locations* influence the subjective perception of presence?

- **RQ2.1:** What subjective levels of social presence, spatial presence, and social richness do users report?
- **RQ2.2:** Which factors do users report as supporting or hindering the feeling of presence?

It is important to notice that our study was not intended to be a controlled experiment to perform a quantitative comparison of collaboration with and without *Re-locations*. Instead, our study was aimed at identifying and analyzing the variety of working styles and coupling styles with *Re-locations* and to what extent they create the illusion of co-located collaboration. Our goal was to identify challenges and potentials for further improvements of the solution approach given the novelty of *Re-locations* and the sparsity of research in this research area.

4.1 Study Environments and Apparatus

Two labs of our research facility served as distributed work environments for our participants (see Figures 2 and 9). Both rooms were facilitated with a personal and a shared workspace. The personal workspaces were represented each by an office desk and desktop computer with a 27-inch display (4k resolution), mouse, and keyboard – a standard setup for a single-user workplace. Their content was un-synced yet alike (i.e., a list of items). Participants were able to sit down at the personal workspace. The shared workspaces were represented each by a vertically mounted 65-inch multi-touch display (4k resolution) – a size comparable to small whiteboards (cf. [23]). Their content was synced and alike (e.g., items and their positions). To ensure for better comparability and a

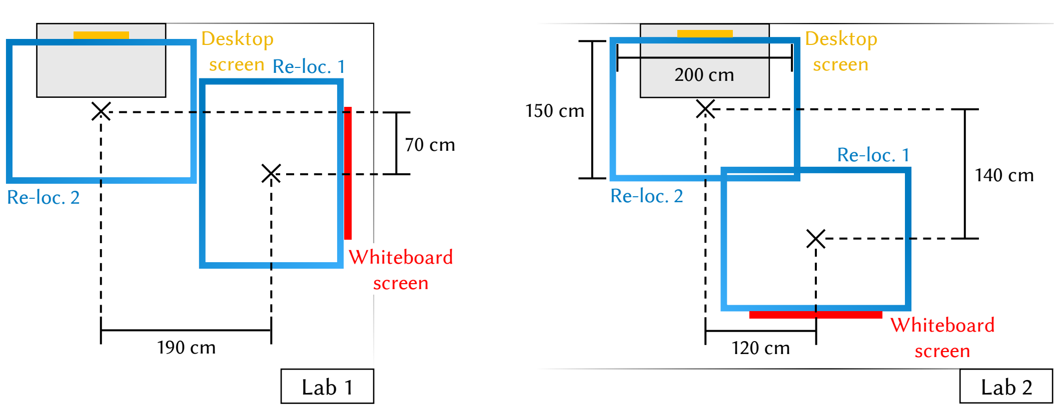

Fig. 9. A top-down view of the relative positions of the workspaces and their corresponding *Re-locations*. The workspaces in lab 1 were arranged in an L-shape. The workspaces in lab 2 were placed back-to-back and had an offset. All *Re-locations* had the same size of 200 x 150 cm.

comfortable viewing height, we positioned the personal and shared displays in each lab at the same height (from the floor to the center): ~110 cm for the personal workspace screen and ~150 cm for the shared workspace screen. However, the individual positions of the workspaces in the rooms and their floor plans differed across the two study environments. We chose two common room layouts for this: An L-shaped layout and a back-to-back layout. This allowed us to study the main purpose of the *Re-locations* concept: The facilitation of remote collaboration in AR with incongruent room layouts. Figure 9 shows the relative positions of workspaces and their corresponding *Re-locations*. We implemented the study task as web apps that run in Google Chrome on Windows 10.

One examiner per lab was needed to run the user study. In addition to the two task-related workspaces in each room, each examiner sat at a separate office desk with all documents, a laptop to control the procedure, and a tablet (Apple iPad Pro 9.7 inch) for digital questionnaires that were built as web apps. In each room, there was a Microsoft HoloLens (HMD) running our prototype and an individual room marker positioned at a wall. Participants had to wear the HMD during task completion. We mounted wide-angle cameras at the ceiling of each room to record participants' movement and behavior during the user study. Additionally, a microphone was used to record the concluding interview with both participants. All devices were connected to the same local network to guarantee instantaneous synchronization and communication of devices.

4.2 Task

The choice for the study task was a trade-off decision between complexity, abstractness, and real-world relation. Common sensemaking tasks (e.g., the "Stegosaurus" task and data set of the 2006 VAST challenge [17]) were used by previous research to study co-located and remote collaboration (e.g., [42, 63]). However, these tasks are rather complex, and their task completion times are long (up to 90 minutes and more), which might result in discomfort and fatigue when wearing current off-the-shelf AR HMDs [8]. While rather abstract tasks (e.g., [30]) can be adapted to fit the study setting (e.g., to control for study duration) and were used by previous research to study collaboration types [2, 29], they lack intellectual communication and negotiation between participants. A task that includes discussions and negotiations of participants is the hotel search task by Jetter et al. [24]. Here, participants have to decide on an accommodation for their mutual vacation. However, each participant has a different set of personal criteria (e.g., "the hotel has at least 4 stars," or "it may cost a maximum of 150€"). The task is designed that a simple combination of all criteria does not result in a solution (i.e., no hotel fits all criteria). Thus, participants have to negotiate, discuss, and ponder their personal criteria. Therefore, this task leads to a realistic yet controllable scenario of natural collaboration, which is adaptable by, e.g., the number of personal criteria or available accommodations. We see this task as a valid representative of authentic yet casual collaborative activities, where peers equally compare different options, negotiate, and agree on a final solution. Here, participants do not require specialized prior knowledge. Awareness on both, each other's activity and location (supported by our *Re-locations* prototype) is beneficial yet not mandatory to solve the task.

As the hotel search task by Jetter et al. [24] was originally designed for co-located collaboration with a single tabletop display, we took it as an inspiration and adapted it as follows: For the task, two participants worked together in two separate rooms, using our *Re-locations* prototype. Here, the rooms differed in their floor plans and location of devices, which allowed us to investigate our research questions. However, each room was provided with the same personal and shared workspace (cf. section 4.1). At the personal workspace, participants had a custom website with a list of all hotels and their attributes (e.g., name, image, or price per night) in the look and feel of state-of-the-art hotel booking websites (cf. Figure 10 left). We disabled the browser's search functionality as our pre-tests have shown that this can heavily reduce task completion times, which

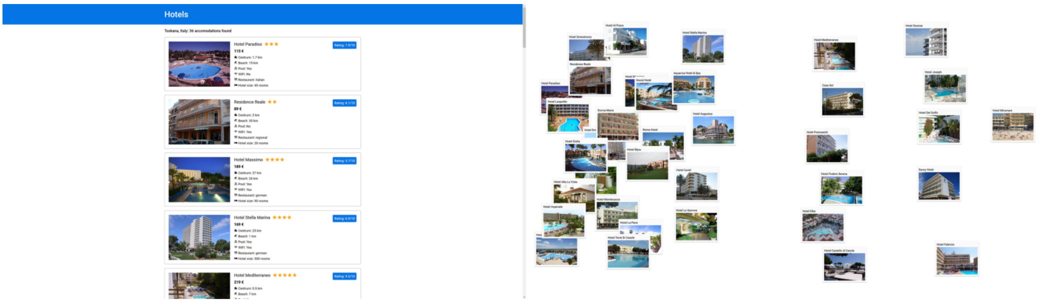

Fig. 10. The web apps displayed on the personal and shared workspace. At the personal workspace (left), participants had a website with a list of all hotels and their attributes. At the shared workspace (right), all hotels were represented as sticky notes mimicking a whiteboard. Before the task, the sticky notes were arranged in a grid. The figure shows the result of a dyad after finishing the task.

in turn might prevent any collaborative flow to establish. At the shared workspace, participants worked with a different view on the same data: Here, all hotels were represented as sticky notes – visualized by their image and name – mimicking a traditional whiteboard with attached sticky notes or photos (cf. Figure 10 right). The positions of all sticky notes were synchronized between the two rooms to foster this impression of a co-located whiteboard session. This separation of two workspaces – personal and shared – in both rooms allowed participants to smoothly transition between closely coupled collaboration and loosely coupled parallel work [22, 42].

To solve the study task, participants had to agree on a hotel from a subset of 36 fictive yet realistic hotels from the hotel set by Jetter et al. [24]. The subset allowed us to control the overall duration of the user study. Participants were given a printed sheet with their personal criteria (4 for each participant). As no hotel fulfilled all criteria, they had to negotiate, discuss, and agree on a hotel. Here, participants were asked to make as few concessions as possible. They were allowed to move freely in the provided rooms and solve the task as they liked. Furthermore, they were informed that they could use the shared workspace to organize, cluster, or prioritize hotels (e.g., by sorting them from left to right, cf. Figure 10 right) and the personal workspace, e.g., for individual search activities. Participants notified the examiners as soon as they agreed on a hotel.

4.3 Procedure

Participants were welcomed in the corridor by two examiners and separately brought to one of the two labs. This prevented participants from spotting any differences in the two study environments. To achieve a consistent study procedure in the two separate rooms, each examiner followed a predefined procedure protocol, and the two examiners communicated current states via an instant messenger on their laptops. First, a welcome sheet with general information on the purpose and procedure of the user study was given to participants. After signing the consent form, participants filled out a demographic questionnaire on a provided tablet. Then, the two examiners individually introduced participants to the applications running on the personal and shared workspace, using a test dataset. Then, the *Re-locations* prototype was started, and participants individually put on the HMD. Participants were tasked to look at the room marker to log in. Then, they were guided to the shared workspace *Re-location* and asked whether they can hear and see their colleague (i.e., the avatar). After agreeing on feeling comfortable in using the HMD and the application, the actual hotel dataset was loaded and both participants were handed their task description with individual criteria on a printed sheet of paper. The data logging and video recording were centrally started for both rooms by one examiner. After finishing the task, participants notified the examiners and presented their decision. Participants took off the HMD, and the data logging and video recording were stopped centrally for both study environments. Participants were then asked to fill out a questionnaire regarding the perceived presence (Temple Presence Inventory (TPI) [31]) on a tablet. After that, one participant was brought to the other study environment for a concluding joint interview with both participants that was audio recorded. Each study session lasted about one hour in total, and participants were compensated for their time. Pilot studies were used to test the applications and communication of the two examiners to guarantee for a smooth procedure. We followed all ethical and sanitary guidelines provided by our university and informed participants on all aspects of the study (e.g., purpose and procedure). However, we decided not to make them aware of any differences in the layout of their rooms, or any additional feature of the *Re-locations* prototype (e.g., visibility of the avatar, gaze ray, or spatial audio) to minimize bias.

4.4 Participants

20 participants (10 dyads) were recruited for the user study. We excluded two dyads for further analysis: One dyad violated the task regulations by actively working against each other and another dyad had to be excluded as the audio recording failed for their session. The remaining 16 participants (10 female, 6 male – forming 8 dyads) were aged between 20 and 40 years ($M = 25.25$, $SD = 4.65$). 13 participants were undergraduate students – the other participants were employees of the university. Participants had a mixed background ranging from psychology to physics, educational science, or law. Three dyads were female-only, 1 dyad male-only, and 4 dyads were female-male. All dyads knew their partner before the user study. Six participants were visually impaired (corrected by glasses ($n = 3$), corrected by contact lenses ($n = 3$)). Only two participants had made experiences with augmented reality applications before. In contrast, eleven participants used virtual reality applications before. Fifteen participants had made experiences with remote collaboration applications (e.g., Skype, Google Docs) before. The participants were asked to rate on a five-point Likert scale (1 (very rare) - 5 (very often)) how often they use these remote collaboration applications ($M = 2.93$, $SD = 1.22$).

4.5 Data Collection and Analysis

We used multiple data collection methods to answer our research questions: Audio and video recordings, semi-structured interview, questionnaires, and data logging. We explain the use of the different methods in the following subsections.

4.5.1 Audio and Video Recordings. All study sessions were audio recorded with the built-in microphones of the individual HMDs and video recorded with wide-angle cameras in birds-eye view in both study environments (cf. Figure 2). For the analysis, the two audio and two video recordings of each study session were synchronized and merged, resulting in a single video file per study session. This combination of observable interactions, gestures, spatial behavior, and verbal communication allowed us to study and analyze meaningful collaborative activities in depth.

The synchronized video files (more than 3 hours final video material) were manually coded to create detailed transcripts describing the collaborative behavior of each dyad at each point in time. To code participants' behavior, we followed a complementary approach of describing (1) collaborative and (2) individual activities. Therefore, we coded all (1) collaborative activities following the coupling styles for hybrid collaboration provided by Neumayr et al [42], as they include both, co-located and remote behaviors – in contrast to previous descriptions of coupling styles focusing on co-located collaboration (e.g., [7, 22, 58]). However, we included a differentiation (also known as modifier) to two of the coupling styles (SIDC and SIDD) as this allowed us to better

Table 1. Mapping of Neumayer et al.'s hybrid coupling styles [42] and the collaborative behavior of the participants. Colors indicate tightly (red) and loosely (blue) coupled activities. C & R indicate coupling styles that were documented by Neumayr et al. for co-located (C) or remote collaboration (R).

DISC	**Discussion.** Participants discuss how to start or proceed with solving the task (strategy) or which hotel they should choose.	C & R
VE	**View Engaged.** One participant is working at the whiteboard and the other participant observes him/her (e.g., sits at the desktop screen and looks at the whiteboard) and comments on the observed activities.	C
SV	**Sharing of the same view.** Both participants are standing or working at the whiteboard (and looking at it).	C
SIDC	**Sharing of the same information but using different physical displays for coordinated exploration.** Participants are working or looking at different screens or both participants are working or looking at the desktop screen (personal workspace). Thereby, they are engaged in some conversation. We distinguish between participants looking at different screens (SIDC-DW) or both looking at their individual desktop screen (SIDC-DD).	C & R
SIDD	**Sharing of the same information but using different physical displays.** Corresponds to SIDC but participants are not engaged in active conversation. Also includes the modifier to distinguish between SIDD-DW and SIDD-DD.	C & R
SSP	**Work is shared to solve the same specific problem.** Participants are actively dividing the task (searching for a hotel matching the most criteria). For example, they are searching for different hotels or reading different hotel information.	C & R
SGP	**Work on the same general problem but from different starting points.** Participants are working on the task, but have not (yet) figured out that they have different criteria when searching for hotels (and are not engaged in active conversation).	C & R
DP	**Work on different problems.** Participants are working on different problems. As there is only a single problem given by the task, this happens only if the task is not performed correctly.	C & R
D	**Disengaged.** One participant is distracted by the system or study environment and therefore disengaged from the task.	C & R

distinguish between the provided workspaces (desktop and whiteboard). Our mappings of dyads' collaborative behavior and Neumayr et al.'s hybrid collaboration coupling styles are shown in Table 1. The coded collaborative behavior was matched with our coding for (2) individual activities that described (a) spatial locations, (b) viewing directions, and (c) deictic gestures (e.g., pointing; although not covered by our prototypical implementation) of individual participants – reflecting on our research questions. Additional minor behaviors that are not task-related and therefore not covered by the named codes (e.g., on-boarding activities) were noted as N (no coupling style).

As suggested by Saldaña [49], the entire coding was performed by a single researcher to ensure consistency across all study sessions. Additionally, a team of three researchers regularly met to discuss ambiguities, review the coding scheme, and coding results.

4.5.2 Semi-structured Interview. We conducted a semi-structured interview at the end of each session to gain deeper qualitative insights into participants' collaborative behavior and their perception of presence. The interviewer clarified that all questions refer to the visualizations (e.g., *Re-locations* and the avatar) and communication with their team partner and are not related to the hotel applications on the screens. Participants were asked about strengths, weaknesses, and comparability to well-known remote conferencing tools (e.g., Skype or Zoom). Then, they were asked about different aspects of the *Re-locations* prototype: a) Awareness of team partner's activities, b) different room layouts, c) avatar visualizations (including gaze visualizations), and d) voice chat.

For all topics, participants were asked to provide their assessments, positive and negative aspects, and suggestions for improvement.

For the analysis, we followed the thematic analysis method of Braun and Clarke [6]: We transcribed all interviews, coded interesting statements, which were then used to generate themes. The resulting themes were (1) *Re-locations*, (2) presence, (3) avatar visualization, (4) voice chat, and (5) technology. We assigned all codes to these themes and counted the frequency of the statements in the codes. All interviews were coded by one person to ensure consistency.

4.5.3 Questionnaires. In the beginning, participants filled out a questionnaire about demographics, prior experiences with AR or VR, and remote video conferencing tools (e.g., Skype or Zoom). After finishing the study task and putting off the HMDs, participants filled out the Temple Presence Inventory (TPI) [31] to investigate, e.g., spatial or social presence.

4.5.4 Data Logging. A data logging functionality was implemented into the *Re-locations* prototype that recorded the position, rotation, and viewing direction (head gaze) of participants inside *Re-locations* during the study. The obtained data were used to evaluate the spatial user behavior (regarding collaborative activities) and the avatar viewing time (regarding the presence). Data was only logged while participants were inside *Re-locations*.

5 FINDINGS

In this section, we report our findings in relation to our research questions. Although the work results of the task played a minor role for our evaluation, all dyads were able to solve the task successfully by deciding on the hotel that met most requirements. In the following, references on a dyad x of participants is termed as Dx (e.g., $D1$ for dyad 1). Individual participants of a dyad x are termed as Dxa and Dxb (e.g., $D1a$ and $D1b$ for the participants in dyad 1).

5.1 RQ1: Collaborative Behavior

We present our findings regarding participants' collaborative behavior from multiple complementary perspectives: (1) quantitative findings regarding the coded coupling styles, (2) qualitative descriptions of selected example situations (i.e., based on coupling styles), (3) quantitative findings based on the coded individual behaviors, (4) the results of the thematic analysis of subjective feedback as part of the interview, and (5) logged activities.

5.1.1 Coupling Styles. Overall, the collaboration was predominantly tightly coupled (T) with 91.1 % of the total duration (7.3 % loosely coupled (L), 1.6 % no coupling style (N)). Table 2 provides an overview of the absolute total duration and relative total duration of individual coupling styles. The most frequent coupling style was SIDC (61.8 %). During the coupling style SIDC, participants were looking at different screens for slightly longer (SIDC-DW: 35.1 % of the total duration) than they were both looking at their individual desktop screen (SIDC-DD: 26.7 % of the total duration). The second most frequent coupling style was DISC with a percentage of 17.6 % followed by SGP with 6.6 %. Furthermore, the coupling styles VE (4.9 % of the total duration) and SV (4.1 % of the total duration) were also observed. Interestingly, VE and SV were – as of yet – only documented for co-located collaboration (cf. Neumayr et al. [42]). The participants divided the task (SSP) 2.7 % of the total duration. One participant (of a dyad) was disengaged 0.7 % of the total duration. The coupling styles SIDD and DP never occurred.

In the following, we qualitatively describe example situations for the most frequent coupling styles (SIDC-DW and SIDC-DD) and coupling styles that were as of yet only documented for co-located collaboration [42] (VE and SV) to further break down the activities.

Table 2. The total duration ($\sum(s)$) and the percentage of total duration ($\sum(\%)$) at which the individual coupling styles occured. Additionally, the percentage of all tightly coupled (T) and loosely coupled (L) coupling styles are combined (T/L). Minor behaviors that are not covered by the coupling styles (e.g., on-boarding activities) are noted as no coupling style (N).

	DISC	VE	SV	SIDC			SIDD			SSP	SGP	DP	D	N
				$\sum$	DD	DW	$\sum$	DD	DW					
$\sum(s)$	2000	552	465	7003	3029	3974	0	0	0	308	747	0	81	179
$M(s)$	250	69	58.13	875.38	378.63	496.75	0	0	0	38.5	93.38	0	10.13	22.38
$SD(s)$	132.95	82.18	69.47	240.88	272.87	318.31	0	0	0	67.82	31.33	0	7.04	33.29
$\sum(\%)$	17.6 %	4.9 %	4.1 %	61.8 %	26.7 %	35.1 %	0 %	0 %	0 %	2.7 %	6.6 %	0 %	0.7 %	1.6 %
T/L	91.1 %										7.3 %			1.6 %

Figure 11 shows an example of a situation which often occurred during the coupling style SIDC-DD (see also section 5.1.2). Both participants go through the list of hotels on the desktop screen while talking to each other. The participant in lab 1 (*D6a*) sits on the chair while the participant in lab 2 (*D6b*) does not sit down but stands next to the chair to interact with the desktop computer – a typical behavior when a seat is occupied by another person. The participant in lab 1 was already sitting on the chair when the participant in lab 2 moved to the desktop computer.

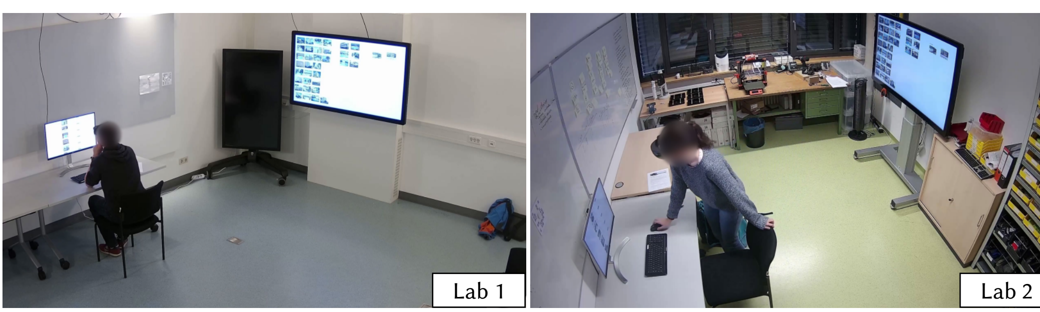

Fig. 11. Example situation in which the coupling style SIDC occurred while both participants looked at the desktop screen (SIDC-DD). Both participants go through the list of hotels on the desktop screen while talking to each other. The participant in lab 1 (*D6a*) sits on the chair while the participant in lab 2 (*D6b*) does not sit down but stands next to the chair to interact with the desktop computer.

Figure 12 shows an example situation which resulted in the most frequent coupling style SIDC-DW. Here, the dyad *D7* worked closely together by splitting the responsibilities: The participant in lab 1 (*D7a*) sorts the hotel sticky notes on the whiteboard screen based on the information communicated by the participant in lab 2 (*D7b*) who reads hotel information on the desktop screen – a strategy several dyads have followed.

A behavior which occurred several times during the coupling style VE was that participants were working on different screens, while one of them observed the other (e.g., *D2b* sitting at the desktop screen and *D2a* interacting with the whiteboard screen) and commented or guided on their activities – indicating spatial awareness complemented by spatial referencing activities. In Figure 13, the participant in lab 1 (*D2a*) sorts the hotel sticky notes and searches for a hotel on the right side of the whiteboard. The participant in lab 2 (*D2b*) watches this and comments with a verbal spatial cue: "*No, [hotel name] is on the left side [of the screen].*".

Fig. 12. Example situation in which the coupling style SIDC occured while both participants looked at different screens (SIDC-DW). The participant in lab 1 (*D7a*) sorts the hotel sticky notes on the whiteboard screen based on the information communicated by the participant in lab 2 (*D7b*) who reads hotel information on the desktop screen.

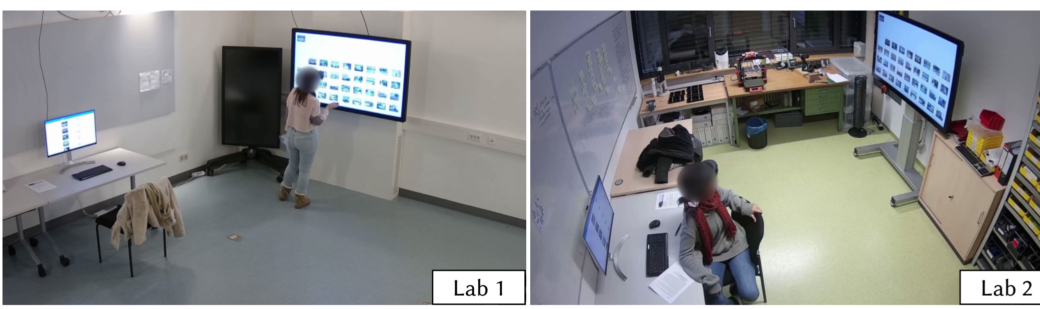

Fig. 13. Example situation in which the coupling style VE occured. The participant in lab 1 (*D2a*) sorts the hotel sticky notes and searches for a hotel on the right side of the whiteboard. The participant in lab 2 (*D2b*) watches this and comments: "*No, [hotel name] is on the left side [of the screen].*"

Fig. 14. Example situation in which the coupling style VE occured. The participant in lab 1 (*D3a*) sits at the desktop screen and uses a deictic gesture to point to the whiteboard screen and show the participant in lab 2 (*D3b*) where the hotel sticky note should be placed.

Similarly, in Figure 14, the participant in lab 1 (*D3a*) sits at the desktop screen and uses a deictic gesture to point to the whiteboard screen and tell the participant in lab 2 (*D3b*) where the hotel sticky note should be placed – "[The hotel] has only 3 stars which means it goes [there]." (*D3a*).

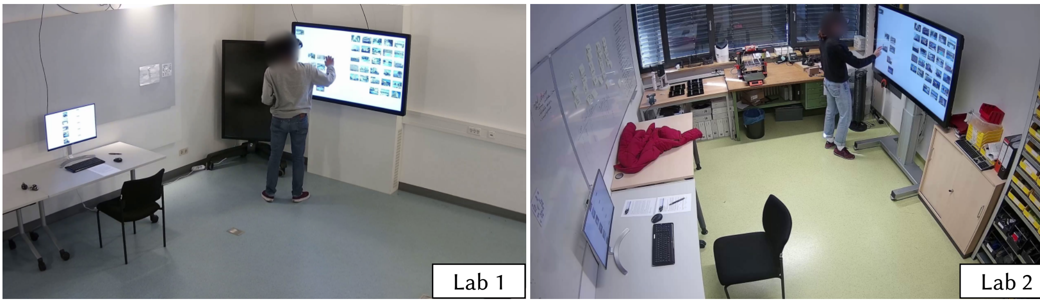

Fig. 15. Example situation in which the coupling style SV occured. Both participants (of dyad *D*4) are located at the whiteboard screen and collaboratively sort hotel sticky notes while sharing the same view.

Figure 15 shows an example situation during the coupling style SV. Both participants (of dyad *D*4) are located at the whiteboard screen and collaboratively sort hotel sticky notes while sharing the same view. While working at the same screen does not necessarily reap the whole potential of the *Re-locations* concept (e.g., spatial referencing for incongruent spaces), it supports collaborative activities that were as of yet only documented for co-located activites [42] by providing spatial awareness cues (i.e., visual and auditory) about the collaborator's location and activities.

5.1.2 Individual behavior. Participants switched the workspace (i.e., moved to the other workspace), in the mean, 17 times ($SD = 21.15$). They switched the viewing direction (between looking in the direction of the desktop or whiteboard screen), in the mean, 54.19 times ($SD = 39.6$). Thereby, they looked, in the mean, 16.88 times ($SD = 21.24$) from the desktop workspace at the whiteboard screen and 2.88 times ($SD = 4.43$) from the whiteboard workspace at the desktop screen.

At the desktop workspace, we coded the behavior of the participants, regarding the available chair (possibility to sit), in detail. In the mean, participants sat on the free chair (remote participant is not sitting) 1.44 times ($SD = 1.31$) and on the occupied chair (remote participant is sitting) 1.94 times ($SD = 1.57$). They stood next to or behind the chair when it was free, in the mean, 0.31 times ($SD = 0.6$) and when it was occupied 5.44 times ($SD = 8.63$). Thereby, in the mean, they interacted with the desktop screen while standing 3.19 times ($SD = 5.26$) when the chair was occupied, but never when the chair was free ($M = 0$, $SD = 0$).

Although the prototypical implementation of *Re-locations* does not support the transmission of gestures to the remote site (i.e., the avatars' arms were not moving), we have coded their occurrence based on the video material. In the mean, participants made 3.06 gestures ($SD = 5.62$) (excluding interaction gestures). They performed, in the mean, 2.44 pointing gestures ($SD = 4.62$) at a screen of which 1.81 pointing gestures ($SD = 3.78$) were directed to the screen of the same workspace the participant was located and 0.63 pointing gestures ($SD = 2$) were directed to the screen of the other workspace.

5.1.3 Semi-structured Interview. In the following, results of the thematic analysis on collaborative behavior with the themes *Re-locations*, *Voice chat*, *Avatar visualization*, and *Technology* are presented.

Re-locations. Participants were asked if they noticed that the room layout in the other room was different. Fifteen participants did not notice the different room layout. *D4a* said to their team partner: "*Well, it looked to me like your room was shaped exactly like mine.*" However, one participant noticed the different room layout. Participant *D2a* stated: "*At my place, she was so fast at the PC. So poof, and yes, it could not have been when the room was exactly like mine.*" Participants were also asked if they noticed that the other person was only visible in front of the screens (i.e., inside a *Re-location*). Ten participants did not notice that the avatar was not visible all the time. Six participants noticed that the avatar was sometimes not visible. *D2b* said: "*...when she said, "I'm gonna run over there now," she was gone for a second, and then she came back.*"

Voice chat. Eight participants mentioned the spatial audio feature during the interview by themselves when being asked for the quality of the voice communication (i.e., they were not informed about the spatial audio feature before). For example, *D4a* said: "*Once you sat at the PC and came closer to me, the sound also changed, so it feels like you're really behind me now.*" Five participants explicitly said they liked the spatial audio functionality. *D1a* stated: "*It felt more in the room because the sound came from both sides somehow.*" Participant *D2a* did not like the spatial audio and said: "*I didn't quite understand why [it should be useful].*" All 16 participants liked the sound quality of the voice chat. Two participants mentioned that they liked the low latency. However, six participants noticed a reverb effect. *D3b* said: "*There was a slight echo for me.*"

Avatar visualization. Four participants rated the avatar visualization as helpful. Participant *D3b* answered: "*Yes, definitely. I think it is cool because it makes things a little bit more personal.*" All other 12 of the 16 participants rated the avatar visualization as not helpful. Eight participants mentioned the task as a reason. For example, *D1b* said: "*Maybe it has advantages in other tasks where it is important to see where the other person stands or something.*" Participants rated the avatar as too abstract ($n = 15$), wished for a more realistic representation ($n = 7$), or did not like the missing lower body ($n = 2$). In line with our observations, participants ($n = 10$) stated that they were missing gestures. *D1b* stated: "*If it['s body parts] really moved the way the person moved, I think it would have been helpful.*" Ten participants rated the gaze ray as helpful, six as not helpful. Two participants described the gaze ray as (sometimes) disturbing. For example, *D3b* said: "*...annoying when you have a neon light shining through you all the time*", while being at the same position as the avatar.

Technology. Five participants mentioned during the interview that the HMD was uncomfortable. For example, *D1b* said: "*Uncomfortable to wear.*" Two participants even got slight headaches. Participant *D1a* said: "*...with this task you could have phoned. Then I wouldn't have had a headache.*" Five participants mentioned as negative point the small field of view of the HMD. Four participants liked the freedom of movement of the device, or that the hands were free.

5.1.4 Data Logging. The data logging was used to calculate the overall task completion time, the time participants were inside *Re-locations*, and the time participants were looking at screens (according to the head gaze ray). The task completion time ranged between 744 (12:24 min) and 1940 seconds (32:20 min) ($M = 1421.75$, $SD = 400.13$). They were, in the mean, 1387.31 seconds inside *Re-locations* ($SD = 360.94$, 98 % of task completion time) and 34.44 seconds outside ($SD = 136.15$, 2 % of task completion time). In the mean, the participants of a dyad were 666.25 seconds inside the same *Re-location* ($SD = 349.09$, 47 % of task completion time), 686.88 seconds inside different *Re-locations* ($SD = 384.28$, 48 % of task completion time), and 68.63 seconds were only one participant inside a *Re-location* ($SD = 192.49$, 5 % of task completion time). It never happened that both participants of a dyad were simultaneously outside of *Re-locations*.

We used participants' individual head gaze to calculate screen times (how long they looked at screens). In the mean, participants looked 995.88 seconds at a screen ($SD = 309.54$, 70 % of task completion time). The screen time can be differentiated by the time they looked at the screen of the *Re-location* they were inside ($M = 956.81$, $SD = 285.82$, 96 % of screen time), or at the screen of the other *Re-location* ($M = 39.06$, $SD = 86.23$, 4 % of screen time).

5.2 RQ2: Presence

Presence was assessed using questions of the Temple Presence Inventory (TPI) [31], a semi-structured interview, and data logging. In the following, the results of the different methods are presented.

5.2.1 Temple Presence Inventory (TPI). The questions of the TPI [31] measure different dimensions of presence. The three item sets spatial presence, social presence (actor within medium (parasocial interaction)), and social richness of the inventory were used. Every question is answered on a seven-point scale, ranging from −3 (negative option) to 3 (positive option).

Spatial presence. The answers to all questions are ambiguous. Except for two questions, the results of the mean values are neutral or slightly positive. The question “How much did it seem as if the objects and people you saw/heard had come to the place you were?” reached a quite neutral mean value of 0.69 (SD = 1.08). However, compared to the other results of the spatial presence questions, the question reached the highest mean value. Also, the question “How much did it seem as if you could reach out and touch the objects or people you saw/heard?” achieved a neutral mean value of −0.06 (SD = 1.57). The question “How often when an object seemed to be headed toward you did you want to move to get out of its way?” reached a mean value of −2.06 (SD = 1.39). The question “To what extent did you experience a sense of being there inside the environment you saw/heard?” reached a neutral mean value of 0.06 (SD = 1.48). Also, the question “To what extent did it seem that sounds came from specific different locations?” achieved a neutral mean value of 0.06 (SD = 1.43). The question “How often did you want to or try to touch something you saw/heard?” reached a mean value of −1.63 (SD = 1.75). A wide answer range and a mean value of −0.13 (SD = 1.93) achieved the question, “Did the experience seem more like looking at the events/people on a movie screen or more like looking at the events/people through a window?”

Social presence (actor within medium (parasocial interaction)). Except for two questions, the mean values are rather positive. The question “How often did you have the sensation that people you saw/heard could also see/hear you?” reached a mean value of 2.44 (SD = 0.73). Also, the question “To what extent did you feel you could interact with the person or people you saw/heard?” reached a high mean value of 2.38 (SD = 0.72). A slightly negative mean value of −0.63 (SD = 1.63) achieved the question, “How much did it seem as if you and the people you saw/heard both left the places where you were and went to a new place?” However, the question “How much did it seem as if you and the people you saw/heard were together in the same place?” reached a mean value of 0.94 (SD = 1.88). A mean value of 1.75 (SD = 1.57) achieved the question, “How often did it feel as if someone you saw/heard in the environment was talking directly to you?” A negative mean value of −1.5 (SD = 1.63) reached the question “How often did you want to or did you make eye-contact with someone you saw/heard?” The question “Seeing and hearing a person through a medium constitutes an interaction with him or her. How much control over the interaction with the person or people you saw/heard did you feel you had?” achieved a mean value of 1.94 (SD = 0.93).

Social richness. The item set rates the media experience with pairs of bipolar adjectives. The media experience was rated as more “personal” than “impersonal” (M = 2, SD = 0.97) and more “sociable” than “unsociable” (M = 2, SD = 0.89), followed by more “responsive” than “unresponsive” (M = 1.94, SD = 0.85), more “lively” than “dead” (M = 1.75, SD = 0.86), more “immediate” than “remote” (M = 1.19, SD = 1.6), more “sensitive” than “insensitive” (M = 0.94, SD = 1.29), and more “emotional” than “unemotional” (M = 0.81, SD = 1.33).

5.2.2 Semi-structured Interview. In the following, the *presence* theme (cf. section 4.5.2) is presented. Thirteen participants said that they always were aware of where the partner was. Three participants did not know it the entire time. One participant (*D4a*) explicitly mentioned that spatial audio helped to be aware of the partner's position. Twelve participants had the impression that they were in the same room. For example, participant *D5b* said: "*I feel like you are just sitting there next to me somehow. Like we are in the same room.*" Furthermore, for example, participant *D1a* said: "*Well, it certainly felt more like the person was here than on the phone.*"

5.2.3 Data Logging. Regarding the presence, the data logging was used to calculate how long participants looked at the avatar. For the calculation, the head gaze was used. In the mean, participants looked 165.38 seconds at the avatar of the remote participant (SD = 90.79, 12 % of task completion time). The time the participants, in the mean, looked at each other (i.e., simulated face-to-face communication) was 3 seconds (SD = 6.95, 2 % of the avatar viewing time).

6 DISCUSSION AND IMPLICATIONS

We discuss our findings based on our research questions on collaborative behavior and presence. Furthermore, we provide implications for future research and design of remote collaboration systems using augmented reality.

6.1 RQ1: Collaborative behavior

The overall goal of our user study was to understand how *Re-locations* can facilitate remote collaboration in incongruent spaces. Therefore, as part of the concluding interview, we asked participants if they noticed that their spaces were incongruent (i.e., differed in their spatial layout). Fifteen participants did not notice that their room layouts differed, which is notable as the positions, distances, and orientations of the two *Re-locations* (desktop screen and whiteboard screen) actually differed. This shows that the general idea of enabling remote collaboration with augmented reality in incongruent spaces including multiple physical devices – using our study apparatus – worked. The only exception was participant *D2a*, who remarked that the two study environments have to be different due to the time it took their team partner to move from one *Re-location* to another.

The transition between two *Re-locations* was named by half of the participants: When leaving a *Re-location*, avatars became invisible. This was a design decision to e.g., reduce visual clutter and unnatural avatar movements, which might evoke an uncanny valley effect [39]. However, assuming that the distances between *Re-locations* are larger, this might reduce spatial awareness when collaborators are outside of *Re-locations*. Future research should therefore investigate ways to represent collaborators between or outside of *Re-locations*: (1) Visually, by varying the level of abstractness of the user representation [51] (e.g., comparable to off-screen visualizations) or (2) Auditory, by providing auditory feedback, e.g., when entering or leaving a *Re-location*.

With the concept of *Re-locations* and our prototypical implementation we tried to reduce the complexity of incongruent spatial workspaces, which might eventually lead to a shared frame of spatial reference. This in turn allows for qualities that are usually attributed to co-located collaboration: spatial awareness and spatial referencing (e.g., via deictic speech or gestures). To better understand these collaborative activities, we carefully coded them based on established coupling styles [42] (RQ1.1) and matched these observations with emergent individual behavior (RQ1.2) regarding spatial awareness and spatial referencing (RQ1.3). The most dominant coupling style in our analysis was SIDC, which was previously documented for co-located and remote collaboration [42]. However, we further distinguish between participants looking at different screens (SIDC-DW) and both participants looking at their individual desktop screen (SIDC-DD). For SIDC-DW, we did not observe behaviors that usually only occur during co-located collaboration. Yet, for SIDC-DD, we observed participants in several situations not sitting down but standing next to the chair at the desktop screen, as the chair was *virtually* occupied by the remote participant. This is interesting, as this imitates natural spatial and social relationships (e.g., proxemics and spatial awareness) between local and remote collaborators that are typical for co-located collaboration.

Possible causes for these behaviors could be the perception of the avatar sitting on the chair or that the remote participant's gaze ray disturbs when both participants are simultaneously sitting on the chair (as described by two participants for situations where both participants were at the same position). Although the second reason is not desirable, in both cases these behaviors would have been observed because the local participant was aware that the remote participant was sitting at the desktop. We interpret these behaviors (during SIDC-DD) as indicators that attributes of co-located collaboration (e.g., spatial awareness) were introduced by *Re-locations*. Our findings, however, are ambiguous for this: While, in the mean, participants similarly frequently sat on a free chair as on an virtually occupied chair, they stood next to the chair much more often when it was occupied than when it was free, and they never interacted with the desktop screen while standing when the chair was free. Future research, therefore, should further investigate this by (1) systematically comparing the influence of different attributes of user representations (e.g. visual or auditory cues) on spatial awareness, (2) controlling session duration as this might influence collaborative flow, and (3) using different study tasks that further increase the demand for spatial awareness and coordination (e.g., creative tasks).

From a methodological perspective, being able to classify remote collaboration facilitated by *Re-locations* with a coding scheme based on hybrid collaboration (Neumayr et al.'s coupling styles [42]) with only minor adaptations (adding a modifier for styles SIDC and SIDD) leads to at least two explanations: Either the definitions of the used coupling styles are general enough that they also cover remote collaboration facilitated by *Re-locations* or the prototypical implementation of the concept introduced attributes of co-located activities. Indeed, we observed coupling styles VE and SV, which as of yet were only documented for co-located collaboration (cf. Neumayr et al. [42]).

A common behavior that occurred several times during VE was that participants were working on different screens, while one of them observed the other and commented or guided on their activities, which further indicated spatial awareness including spatial referencing activities. The content of the whiteboard screen was synced and therefore allowed participants to catch their partner's attention (cf. [18]) e.g., by wiggling with sticky notes – enabling some form of awareness. However, we also observed situations in which participants commented on the behavior of the remote participant without interaction at the whiteboard screen of the remote participant (see example situation in Figure 13). Furthermore, participants used pointing gestures, although the prototypical implementation of *Re-locations* does not support the transmission of gestures to the remote site (i.e., the avatars' arms were not moving) – a behavior that probably does not occur when users usually share their screen remotely (e.g., using Zoom). Additionally, the majority of participants mentioned missing gestures (e.g., as pointing technique) during the interview. Future research, therefore, should support pointing gestures if possible. Building on recent research that studied the interpretability of deictic gestures in VR [34], it would be interesting to study pointing gestures (1) within same *Re-locations*, (2) across different *Re-locations*, but also (3) on other task-related aspects (e.g., user representations or physical devices).

The coupling style SV only occurred while both participants were located at the whiteboard screen. Here, participants benefited from both, the awareness cues provided by the *Re-locations* (e.g., gaze ray and spatial audio) and by the synchronization of the whiteboard screen that allowed to grab attention by e.g., wiggling with sticky notes. Therefore, findings can not be clearly attributed to the *Re-locations* alone. Additionally, working at the same screen does not necessarily reap the whole potential of the *Re-location* concept (e.g., spatial referencing for incongruent spaces), it still supports collaborative activities on physical devices that were as of yet only documented for co-located activities [42] by providing spatial awareness cues (i.e., visual and auditory) about the collaborator's location and activities. While the majority of participants rated the gaze ray as a helpful tool, two participants described the gaze ray as sometimes disturbing in situations where both participants were at the same position (a typical situation for SV). Therefore, future research should investigate ways to create awareness on collaborators' gaze by dynamically transitioning to an alternative gaze visualization (e.g., a simple dot on the screen) when collaborators are to close or "inside" each other (i.e., both participants have the same position and orientation).

6.2 RQ2: Presence

The intended illusion of the *Re-locations* concept was that the participants had the feeling they work in the same space (i.e., virtually co-located). Indeed, the TPI questions regarding the feeling of being in the same place were rated positively, indicating the general potential of the Re-locations concept. However, future work is necessary to investigate whether this finding can be attributed to spatial audio, the avatar representation, or the spatial arrangement of workspaces.

Participants rated the perceived sense of social presence and social richness using the TPI (RQ2.1): Results indicate that they had the sensation that they could see, hear, and interact with each other. Also, results show that they had the feeling they had much control over the interaction. In the interview, twelve participants also mentioned they felt as they were in the same room, indicating some sense of social presence. That the participants did not make eye contact with people they saw or heard, is very likely due to the fact that the avatar misses facial attributes (e.g., eyes).

However, the ratings on the perceived sense of spatial presence have a wide range, indicating mixed experiences (RQ2.1). Most participants answered in the interview that they always knew where the team partner was, which indicates a sense of spatial presence. Also observed situations, in which participants stood next to the chair while it was (virtually) occupied, indicate a sense of spatial presence. Interestingly, the spatial presence question regarding spatial audio was answered rather positively. This result is consistent with the interview statements, where over half of the participants noticed the spatial audio feature. One participant mentioned explicitly that spatial audio was the reason they knew where the team partner was (RQ2.2). Further, many participants felt (more) as they were in the same room but could not precisely articulate why. Therefore, spatial audio was possibly the reason for this feeling. Moreover, the participants had the feeling that the remote person was talking with them directly. Therefore, spatial audio might have heavily impacted different dimensions of presence. We see spatial audio as a promising approach to complement the visual components of AR – e.g. as an auditive off-screen technique to raise awareness to other users' locations in a room. Future research, therefore, could further investigate spatial audio (1) by studying the benefits of head-mounted spatial audio (i.e., dynamic) vs. stationary spatial audio (i.e., fixed speakers at workspaces) on spatial awareness, (2) by studying its influence on hybrid collaboration (i.e., a combination of co-located and remote collaborators), and (3) by systematically analyzing the effects of spatial audio on different aspects of presence (e.g., social or spatial presence).

7 LIMITATIONS AND FUTURE WORK

The overall goal of our user study was to understand how *Re-locations* can facilitate remote collaboration in incongruent spaces. For this, we implemented an initial prototype that allowed us to explore how collaboration can unfold in such environments. Therefore, the user study was intentionally explorative rather than comparative. This limits the generalizability of the findings of our user study. In future work, we want to further investigate the concept (as described in section 6) by (1) systematically comparing the influence of different attributes of the concept on spatial awareness, (2) controlling session duration, and (3) using study tasks that further increase the demand for spatial awareness and coordination. For the initial exploration of the concept, we chose to use two common room layouts for our user study: L-shape and back-to-back. However, in reality, room layouts might differ more heavily in their position, distance, and orientation. Here, larger distances between *Re-locations* could lead to a reduced spatial awareness when collaborators are outside or between *Re-locations*. For these situations, future work should investigate ways to increase spatial awareness by e.g., providing audio feedback when entering or leaving a *Re-location* or visually representing users that are currently not within a *Re-location* (e.g., inspired by common off-screen visualizations).

As part of the user study, we invited mainly students that knew each other before to solve an authentic collaborative task. Collaboration with unfamiliar collaborators might be different – e.g., it might include longer phases of getting to know each other but it might also lead to rather loosely-coupled parallel work, depending on the task at hand. Here, a rather abstract task (e.g., as used by [2, 29]) might be suitable to control session duration – yet missing negotiation between participants might also lead to loosely-coupled parallel work, reducing the need for spatial awareness. A potential alternative task that might lead to longer discussion and negotiation phases, which was studied in-depth by previous research (e.g., [42, 63]) is the Stegosaurus task [17]. We deliberately decided against using this task for our user study due to extended task completion times that might result in discomfort and fatigue when wearing current off-the-shelf AR HMDs. However, with the advent of more comfortable AR HMDs, future work should investigate how a rather high task difficulty and extended task duration might influence collaborative activities facilitated by *Re-locations*.

Additionally, the size of *Re-locations* for our user study was the same for both screens (see Figure 9). While our pre-tests showed that this size allows both participants to comfortably stand or sit within the same *Re-location*, it does not necessarily reflect the influence of the size and semantic meaning of personal and shared workspaces. Here, future research should carefully investigate (1) different or even dynamic sizes for *Re-locations* (based on aspects of proxemics [16, 19, 33]) and (2) how participants can further configure and utilize personal and shared workspaces (based on aspects of territoriality [50]).

8 CONCLUSION

In this paper, we presented the concept of *Re-locations*, a method for enabling remote collaboration with augmented reality in incongruent spaces including multiple physical devices. We conducted an exploratory user study with dyads using an authentic, collaborative task to investigate the influence of *Re-locations* on the collaborative behavior and the subjective perception of presence. The results indicate that *Re-locations* introduce attributes of co-located collaboration like spatial awareness and social presence. Furthermore, the results suggest that with *Re-locations*, participants adopted collaborative activities that could be classified with coupling styles previously only documented for co-located collaboration [42] and they had the feeling to work in the same space, although the room layouts differed. Therefore, future research should further investigate the concept and its individual components by systematically comparing the influence of the different attributes on spatial awareness. Finally, the study showed the potential of spatial audio for remote collaboration in AR. It could be used as an auditive off-screen technique to raise awareness on the remote user's position in the room and should be further investigated in future studies.

ACKNOWLEDGMENTS

This research was funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation) – Project-ID 251654672 – TRR 161 (Project C01). This work is a part of the X-PRO project. The project X-PRO is financed by research subsidies granted by the government of Upper Austria.

REFERENCES

[1] MeetinVR ApS. 2022. *MeetinVR*. MeetinVR ApS. Retrieved February 16, 2022 from https://www.meetinvr.com/

[2] Ignacio Avellino, Cédric Fleury, Wendy E. Mackay, and Michel Beaudouin-Lafon. 2017. CamRay: Camera Arrays Support Remote Collaboration on Wall-Sized Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6718–6729. https://doi.org/10.1145/3025453.3025604

[3] Steve Benford, Chris Greenhalgh, Gail Reynard, Chris Brown, and Boriana Koleva. 1998. Understanding and Constructing Shared Spaces with Mixed-Reality Boundaries. *ACM Trans. Comput.-Hum. Interact.* 5, 3 (Sept. 1998), 185–223. https://doi.org/10.1145/292834.292836

[4] M. Billinghurst, S. Bee, J. Bowskill, and H. Kato. 1999. Asymmetries in Collaborative Wearable Interfaces. In *Digest of Papers. Third International Symposium on Wearable Computers*. IEEE Comput. Soc, San Francisco, CA, USA, 133–140. https://doi.org/10.1109/ISWC.1999.806693

[5] Mark Billinghurst and Hirokazu Kato. 2002. Collaborative Augmented Reality. *Commun. ACM* 45, 7 (July 2002), 64–70. https://doi.org/10.1145/514236.514265

[6] Virginia Braun and Victoria Clarke. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[7] Frederik Brudy, Joshua Kevin Budiman, Steven Houben, and Nicolai Marquardt. 2018. *Investigating the Role of an Overview Device in Multi-Device Collaboration*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173874

[8] Marco Cavallo, Mishal Dolakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Immersive Insights: A Hybrid Analytics System ForCollaborative Exploratory Data Analysis. In *25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 9, 12 pages. https://doi.org/10.1145/3359996.3364242

[9] SungIk Cho, Seung-wook Kim, JongMin Lee, JeongHyeon Ahn, and JungHyun Han. 2020. Effects of Volumetric Capture Avatars on Social Presence in Immersive Virtual Environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Atlanta, GA, USA, 26–34. https://doi.org/10.1109/VR46266.2020.00020

[10] Ben J. Congdon, Tuanfeng Wang, and Anthony Steed. 2018. Merging Environments for Shared Spaces in Mixed Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 11, 8 pages. https://doi.org/10.1145/3281505.3281544

[11] 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 (2019), 81 – 98. https://doi.org/10.1016/j.ijhcs.2019.05.011 50 years of the International Journal of Human-Computer Studies. Reflections on the past, present and future of human-centred technologies.

[12] Omid Fakourfar, Kevin Ta, Richard Tang, Scott Bateman, and Anthony Tang. 2016. Stabilized Annotations for Mobile Remote Assistance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1548–1560. https://doi.org/10.1145/2858036.2858171

[13] Steven Feiner and Ari Shamash. 1991. Hybrid User Interfaces: Breeding Virtually Bigger Interfaces for Physically Smaller Computers. In *Proceedings of the 4th annual ACM symposium on User interface software and technology - UIST '91 (UIST '91)*. ACM Press, Hilton Head, South Carolina, USA, 9–17. https://doi.org/10.1145/120782.120783

[14] Mozilla Foundation. 2022. *Mozilla Hubs*. Mozilla Foundation. Retrieved February 16, 2022 from https://hubs.mozilla.com/

[15] OpenJS Foundation. 2020. *Node.js website*. OpenJS Foundation. Retrieved September 16, 2020 from https://nodejs.org

[16] Saul Greenberg, Nicolai Marquardt, Till Ballendat, Rob Diaz-Marino, and Miaosen Wang. 2011. Proxemic Interactions: The New Ubicomp? *Interactions* 18, 1 (jan 2011), 42–50. https://doi.org/10.1145/1897239.1897250

[17] Georges Grinstein, Theresa O'Connell, Sharon Laskowski, Catherine Plaisant, Jean Scholtz, and Mark Whiting. 2006. VAST 2006 Contest - A Tale of Alderwood. In *Visual Analytics Science And Technology, 2006 IEEE Symposium On*. IEEE, Baltimore, MD, USA, 215–216. https://doi.org/10.1109/VAST.2006.261420

[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)*. Association for Computing Machinery, New York, NY, USA, Article 451, 13 pages. https://doi.org/10.1145/3411764.3445698

[19] E.T. Hall. 1966. *The Hidden Dimension*. Doubleday.

[20] Keita Higuchi, Yinpeng Chen, Philip A. Chou, Zhengyou Zhang, and Zicheng Liu. 2015. ImmerseBoard: Immersive Telepresence Experience Using a Digital Whiteboard. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2383–2392. https://doi.org/10.1145/2702123.2702160

[21] Sebastian Hubenschmid, Johannes Zagermann, Simon Butscher, and Harald Reiterer. 2021. STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, Article 469, 14 pages. https://doi.org/10.1145/3411764.3445298

[22] P. Isenberg, D. Fisher, S. A. Paul, M. R. Morris, K. Inkpen, and M. Czerwinski. 2012. Co-Located Collaborative Visual Analytics around a Tabletop Display. *IEEE Transactions on Visualization and Computer Graphics* 18, 5 (2012), 689–702. https://doi.org/10.1109/TVCG.2011.287

[23] Mads Møller Jensen, Roman Rädle, Clemens N. Klokmose, and Susanne Bodker. 2018. Remediating a Design Tool: Implications of Digitizing Sticky Notes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173798

[24] Hans-Christian Jetter, Jens Gerken, Michael Zöllner, Harald Reiterer, and Natasa Milic-Frayling. 2011. Materializing the Query with Facet-Streams: A Hybrid Surface for Collaborative Search on Tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 3013–3022. https://doi.org/10.1145/1978942.1979390

[25] Dongsik Jo, Ki-Hong Kim, and Gerard Jounghyun Kim. 2015. SpaceTime: adaptive control of the teleported avatar for improved AR tele-conference experience. *Computer Animation and Virtual Worlds* 26, 3-4 (2015), 259–269. https://doi.org/10.1002/cav.1645

[26] Robert Johansen. 1988. *Groupware: Computer Support for Business Teams*. The Free Press, USA.

[27] Seungwon Kim, Mark Billinghurst, and Gun Lee. 2018. The Effect of Collaboration Styles and View Independence on Video-Mediated Remote Collaboration. *Computer Supported Cooperative Work (CSCW)* 27, 3-6 (2018), 569–607. https://doi.org/10.1007/s10606-018-9324-2

[28] Pascal Knierim, Dimitri Hein, Albrecht Schmidt, and Thomas Kosch. 2021. The SmARtphone Controller: Leveraging Smartphones as Input and Output Modality for Improved Interaction within Mobile Augmented Reality Environments. *i-com* 20, 1 (apr 2021), 49–61. https://doi.org/10.1515/icom-2021-0003

[29] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, and Eric Lecolinet. 2016. Shared Interaction on a Wall-Sized Display in a Data Manipulation Task. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2075–2086. https://doi.org/10.1145/2858036.2858039

[30] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, Eric Lecolinet, and Wendy E. Mackay. 2014. Effects of Display Size and Navigation Type on a Classification Task. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 4147–4156. https://doi.org/10.1145/2556288.2557020

[31] Matthew Lombard, Theresa B Ditton, and Lisa Weinstein. 2009. Measuring (Tele)Presence: The Temple Presence Inventory. In *Proceedings of the 12th Annual International Workshop on Presence*. International Society for Presence Research, Los Angeles, California, USA, 1–15.

[32] Stephan Lukosch, Mark Billinghurst, Leila Alem, and Kiyoshi Kiyokawa. 2015. Collaboration in Augmented Reality. *Computer Supported Cooperative Work (CSCW)* 24, 6 (Dec. 2015), 515–525. https://doi.org/10.1007/s10606-015-9239-0

[33] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-Device Interaction via Micro-Mobility and f-Formations. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA) *(UIST '12)*. Association for Computing Machinery, New York, NY, USA, 13–22. https://doi.org/10.1145/2380116.2380121

[34] Sven Mayer, Jens Reinhardt, Robin Schweigert, Brighten Jelke, Valentin Schwind, Katrin Wolf, and Niels Henze. 2020. Improving Humans' Ability to Interpret Deictic Gestures in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376340

[35] Rachel McDonnell, Michéal Larkin, Benjamín Hernández, Isaac Rudomin, and Carol O'Sullivan. 2009. Eye-Catching Crowds: Saliency Based Selective Variation. *ACM Trans. Graph.* 28, 3, Article 55 (July 2009), 10 pages. https://doi.org/10.1145/1531326.1531361

[36] Microsoft. 2022. *Mesh for Microsoft Teams*. Microsoft. Retrieved February 16, 2022 from https://www.microsoft.com/en-us/mesh

[37] Microsoft. 2022. *Mixed Reality Toolkit for Unity*. Microsoft. Retrieved February 22, 2022 from https://github.com/microsoft/MixedRealityToolkit-Unity

[38] Shinya Minatani, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta. 2007. Face-to-Face Tabletop Remote Collaboration in Mixed Reality. In *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE, Nara, Japan, 1–4. https://doi.org/10.1109/ISMAR.2007.4538823

[39] Masahiro Mori, Karl MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine* 19, 2 (2012), 98–100. https://doi.org/10.1109/MRA.2012.2192811

[40] Jens Müller, Roman Rädle, and Harald Reiterer. 2017. Remote Collaboration With Mixed Reality Displays: How Shared Virtual Landmarks Facilitate Spatial Referencing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6481–6486. https://doi.org/10.1145/3025453.3025717

[41] Shohei Nagai, Shunichi Kasahara, and Jun Rekimoto. 2015. LiveSphere: Sharing the Surrounding Visual Environment for Immersive Experience in Remote Collaboration. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction* (Stanford, California, USA) *(TEI '15)*. Association for Computing Machinery, New York, NY, USA, 113–116. https://doi.org/10.1145/2677199.2680549

[42] 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, CSCW, Article 128 (Nov. 2018), 24 pages. https://doi.org/10.1145/3274397

[43] Kenton O'hara, Jesper Kjeldskov, and Jeni Paay. 2011. Blended Interaction Spaces for Distributed Team Collaboration. *ACM Trans. Comput.-Hum. Interact.* 18, 1, Article 3 (May 2011), 28 pages. https://doi.org/10.1145/1959022.1959025

[44] 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. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 741–754. https://doi.org/10.1145/2984511.2984517

[45] A. Pauchet, F. Coldefy, L. Lefebvre, S. Louis Dit Picard, A. Bouguet, L. Perron, J. Guerin, D. Corvaisier, and M. Collobert. 2007. Mutual Awareness in Collocated and Distant Collaborative Tasks Using Shared Interfaces. In *Human-Computer Interaction – INTERACT 2007*, Cécilia Baranauskas, Philippe Palanque, Julio Abascal, and Simone Diniz Junqueira Barbosa (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 59–73. https://doi.org/10.1007/978-3-540-74796-3_8

[46] Thammathip Piumsomboon, Arindam Day, Barrett Ens, Youngho Lee, Gun Lee, and Mark Billinghurst. 2017. Exploring Enhancements for Remote Mixed Reality Collaboration. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications* (Bangkok, Thailand) *(SA '17)*. Association for Computing Machinery, New York, NY, USA, Article 16, 5 pages. https://doi.org/10.1145/3132787.3139200

[47] Thammathip Piumsomboon, Youngho Lee, Gun Lee, and Mark Billinghurst. 2017. CoVAR: A Collaborative Virtual and Augmented Reality System for Remote Collaboration. In *SIGGRAPH Asia 2017 Emerging Technologies* (Bangkok, Thailand) *(SA '17)*. Association for Computing Machinery, New York, NY, USA, Article 3, 2 pages. https://doi.org/10.1145/3132818.3132822

[48] PTC. 2020. *Vuforia website*. Parametric Technology GmbH. Retrieved September 16, 2020 from https://www.ptc.com/products/vuforia

[49] Johnny Saldaña. 2021. *The Coding Manual for Qualitative Researchers*. SAGE Publications.

[50] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori Inkpen. 2004. Territoriality in Collaborative Tabletop Workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) *(CSCW '04)*. Association for Computing Machinery, New York, NY, USA, 294–303. https://doi.org/10.1145/1031607.1031655

[51] Sofia Seinfeld, Tiare Feuchtner, Antonella Maselli, and Jörg Müller. 2021. User Representations in Human-Computer Interaction. *Human-Computer Interaction* 36, 5-6 (2021), 400–438. https://doi.org/10.1080/07370024.2020.1724790

[52] Richard Skarbez, Missie Smith, and Mary C. Whitton. 2021. Revisiting Milgram and Kishino's Reality-Virtuality Continuum. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.647997

[53] Harrison Jesse Smith and Michael Neff. 2018. Communication Behavior in Embodied Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173863

[54] Inc. Spatial Systems. 2022. *Spatial*. Spatial Systems, Inc. Retrieved February 16, 2022 from https://spatial.io/

[55] Maximilian Speicher, Jingchen Cao, Ao Yu, Haihua Zhang, and Michael Nebeling. 2018. 360Anywhere: Mobile Ad-Hoc Collaboration in Any Environment Using 360 Video and Augmented Reality. *Proc. ACM Hum.-Comput. Interact.* 2, EICS, Article 9 (June 2018), 20 pages. https://doi.org/10.1145/3229091

[56] Misha Sra, Aske Mottelson, and Pattie Maes. 2018. Your Place and Mine: Designing a Shared VR Experience for Remotely Located Users. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 85–97. https://doi.org/10.1145/3196709.3196788

[57] Anthony Tang, Carman Neustaedter, and Saul Greenberg. 2007. VideoArms: Embodiments for Mixed Presence Groupware. In *People and Computers XX — Engage*, Nick Bryan-Kinns, Ann Blanford, Paul Curzon, and Laurence Nigay (Eds.). Springer London, London, 85–102. https://doi.org/10.1007/978-1-84628-664-3_8

[58] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. 2006. Collaborative Coupling over Tabletop Displays. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI 2006, April 22–27, 2006, Montréal, Canada)*. ACM Press, New York, 1181–1190. https://doi.org/10.1145/1124772.1124950

[59] Unity Technologies. 2020. *Unity Website*. Unity Technologies. Retrieved September 16, 2020 from https://unity.com

[60] Boram Yoon, Hyung-il Kim, Gun A. Lee, Mark Billinghurst, and Woontack Woo. 2019. The Effect of Avatar Appearance on Social Presence in an Augmented Reality Remote Collaboration. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, Osaka, Japan, 547–556. https://doi.org/10.1109/VR.2019.8797719

[61] Leonard Yoon, Dongseok Yang, Choongho Chung, and Sung-Hee Lee. 2021. A Full Body Avatar-Based Telepresence System for Dissimilar Spaces. *CoRR* abs/2103.04380 (2021). https://doi.org/10.48550/arXiv.2103.04380 arXiv:2103.04380

[62] Leonard Yoon, Dongseok Yang, Jaehyun Kim, ChoongHo Chung, and Sung-Hee Lee. 2022. Placement Retargeting of Virtual Avatars to Dissimilar Indoor Environments. *IEEE Transactions on Visualization and Computer Graphics* 28, 3 (2022), 1619–1633. https://doi.org/10.1109/TVCG.2020.3018458

[63] Johannes Zagermann, Ulrike Pfeil, Roman Rädle, Hans-Christian Jetter, Clemens Klokmose, and Harald Reiterer. 2016. When Tablets Meet Tabletops: The Effect of Tabletop Size on Around-the-Table Collaboration with Personal Tablets. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5470–5481. https://doi.org/10.1145/2858036.2858224

[64] Jakob Zillner, Christoph Rhemann, Shahram Izadi, and Michael Haller. 2014. 3D-Board: A Whole-Body Remote Collaborative Whiteboard. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 471–479. https://doi.org/10.1145/2642918.2647393