Multimodal Technologies and Interaction

MDPI

Article

Mixed-Presence Collaboration with Wall-Sized Displays: Empirical Findings on the Benefits of Awareness Cues

Valérie Maquil *, Adrien Coppens, Lou Schwartz and Dimitra Anastasiou

Luxembourg Institute of Science and Technology, 5, Avenue des Hauts-Fourneaux, L-4362 Esch-sur-Alzette, Luxembourg; adrien.coppens@list.lu (A.C.); lou.schwartz@list.lu (L.S.); dimitra.anastasiou@list.lu (D.A.)
* Correspondence: valerie.maquil@list.lu

Abstract: Collaborative decision-making increasingly involves wall-sized displays (WSDs), allowing teams to view, analyse and discuss large amounts of data. To enhance workspace awareness for mixed-presence meetings, previous work proposes digital cues to share gestures, gaze, or entire postures. While several isolated cues were proposed and demonstrated useful in different workspaces, it is unknown whether results from previous studies can be transferred to a mixed-presence WSD context and to what extend such cues can be used in a combined way. In this paper, we report on the results from a user study with 24 participants (six groups of four participants), testing a mixed-presence collaboration scenario on two different setups of connected WSDs: audio-video link only vs. full setup with seven complementary cues. Our results show that the version with cues enhances workspace awareness, user experience, team orientation and coordination, and leads teams to take more correct decisions.

Keywords: wall-sized display; workspace awareness; user study; awareness cues; collaboration; large interactive displays; large high-resolution displays

check for updates

Citation: Maquil, V.; Coppens, A.; Schwartz, L.; Anastasiou, D. Mixed-Presence Collaboration with Wall-Sized Displays: Empirical Findings on the Benefits of Awareness Cues. *Multimodal Technol. Interact.* **2024**, *8*, 109. https://doi.org/10.3390/mti8120109

Academic Editor: Lars Erik Holmquist

Received: 15 October 2024
Revised: 25 November 2024
Accepted: 1 December 2024
Published: 5 December 2024

1. Introduction

For many of today's contexts and tasks, professionals need to come together and join their forces to find solutions for complex challenges. Such collaborative work can be supported by different shareable devices—for instance, interactive tabletops, tangible user interfaces, or multi-surface environments [1]. Interactive wall-sized displays (WSDs) present a vast amount of pixels over an extensive area, being "at least the size of a human body" [2]. They are beneficial, for instance, in information visualisation and visual data analysis, as more data can be presented in more detail or in different scales and views [3]. In addition, they support collocated collaborative work as demonstrated in application fields, such as medical coordination [4], product-service system design [5], or architectural design [6]. Indeed, the large space accommodates multiple users, who can access and view content at the same time, as well as follow each others' verbal, and non-verbal interactions, such as body postures, hand gestures, gaze, and facial expressions. The "up-to-the-moment understanding of another person's interaction with a shared workspace" [7] is referred to as workspace awareness (WA) and allows collaborators to seamlessly align and integrate activities with those of other group members as well as to assist each other [1].

However, when seeking to limit travel costs or to reduce the carbon footprint, face-to-face collaboration is not always possible, and alternatives are needed to allow teams to work together from a distance. In addition, since the COVID-19 pandemic, several organisations have favored remote working, which had a positive impact on their carbon footprints [8,9]. It is therefore necessary to support these new practices with suitable tools.

In this paper, we focus on mixed-presence collaboration across WSDs, i.e., meeting situations where part of the team is in a collocated situation, and part of the team is attending remotely. While local team members are able to gather WA information easily and naturally, remote collaborators usually have to rely on an audio-video feed and synchronized views to understand what is happening at the distant site. This leads to difficulties in communication and requires additional efforts for staying engaged [10]. The lack of awareness is particularly relevant in the context of decision-making at WSDs. Indeed, when discussing data in front of a large screen, collaborators often use body movements and hand gestures [11,12], which are not accessible remotely with conventional audio-video support.

In the attempt to better mediate WA information and facilitate communication, previous works suggest adding visual cues onto the common workspace or the live video stream. In this paper, we call such cues *workspace awareness (WA) cues* and define them as digital indicators conveying workspace awareness information. Examples of such WA cues are visual pointers, sketches, or visualisations of the users' hand gestures, e.g., [13,14]. Currently proposed cues were mostly designed for smaller or different workspaces (e.g., tabletops [15]), targeting dyads (with one person on each site) [16], or using mixed-reality technology where users wear head-mounted displays [17]. As the interaction with WSDs affords a high amount of non-verbal communication, it is unresearched whether results from such previous studies can be transferred to a mixed-presence WSD context, and to what extend such cues can be used in a combined way, without being overwhelming for the users.

In this work, we seek to advance our understanding on the benefits of WA cues for mixed-presence collaboration on WSDs. We report on a user study (within-subject design) with 24 participants, split into six groups of four persons, who tested a mixed-presence WSD system integrating seven complementary WA cues, in comparison to a system featuring only an audio-video link and synchronized screen views (see Figure 1). Each group had two persons on each site and solved tasks, once with the audio-video link only, and once with the additional WA cues. To understand the impact of the WA cues, for each condition, we collected data on awareness, user experience, perceived workload, preference, and collaboration. We describe our qualitative and quantitative results and present our findings.

This paper makes the following contributions:

- It reports on the design of a complementary set of WA cues to support mixed-presence collaboration for WSDs sharing awareness information on attention, actions, and environment.
- It presents first empirical results about the benefits of WA cues in this context.
- It describes considerations for the interface design of WSD-based mixed-presence collaboration systems.

Figure 1. Both locations of the mixed-presence collaboration scenario on wall-sized displays: the curved 360° Immersive Arena (**left**) and the flat VisWall (**right**).

In the next sections, we first review related works (Section 2), then describe our mixed-presence collaboration system (Section 3), integrating seven complementary workspace awareness cues: list of attendees, miniature view, touch feedback, mid-air pointing feedback, (dis)agreement icon, hand raised icon and gaze icon. After that, we present our user study design with this system (Section 4), as well as report and discuss the results (Sections 5 and 6). Finally, we end with a conclusion section (Section 7).

2. Related Work

In this section, we provide an overview of the related literature to clarify the concepts we rely on, such as workspace awareness. We then discuss existing solutions that have been proposed for enhancing collaboration in remote and mixed-presence scenarios in general, and then focus on existing work to convey awareness information on larger surfaces and WSDs specifically.

2.1. Awareness in Collaborative Work

In mixed-presence settings, effective collaboration requires the sharing of essential information to gather awareness. Awareness in that context may be defined as "an understanding of the activities of others, which provides a context for your own activity" [18], and constitutes a critical concern for collaborative systems. From that perspective, awareness is the accumulation of shared information, the knowledge of group and individual activities and their coordination.

In their work, Endsley presented a model of situation awareness [19] that includes three levels: the perception of elements in the surrounding environment in the current situation, the understanding of their meaning, and the anticipation of the upcoming status of these elements. In a collaborative context, situation awareness for a team is then presented as the degree to which each team member is aware of the aspects of the situation that are relevant to their own individual work. Building on these concepts, Dourish and Bellotti [18] further differentiate awareness information based on whether it is generated explicitly (actively, following a collaborator's deliberate action) or passively (automatically collected by the system).

These two theories have served as a basis for the synthesis work of Gutwin and Greenberg on workspace awareness (WA), which they define as "the up-to-the-moment understanding of another person's interaction with a shared workspace" [7]. Such awareness deals with the presence and location of collaborators, with the actions they are undertaking and the objects they are acting on, but also with where their attention is focused. To understand what kind of information is part of WA and think about how it can be included in a groupware system, Gutwin and Greenberg propose a list of questions covering the "who, what, where, when, and how" of ongoing activities [7]. We note that Gutwin and Greenberg also list WA elements and questions related to past actions and the persistence of information over time. However, as we currently concentrate on synchronous collaboration, we focus on present activities and therefore on the corresponding elements.

Another seminal work on collaborative awareness is that of Carroll et al. [20], who proposed a framework describing four facets of activity awareness to explain how effective a team may be. The first facet is common ground, based on grounding [21] (mutual knowledge and beliefs) through communication. Then comes the communities of practice, that come with shared goals and practices. A third facet is the social capital, which is the accumulation of mutually satisfying interactions in the past, creating a climate that fosters collaboration through cohesion. Finally, the fourth facet relates to human development, the acquired skills and decision-making capabilities of both individual members and teams. The authors further emphasise on the necessity to support such activity awareness with minimal impact on attention and effort as consciously gathering such information can be particularly detrimental to the task at hand [20]. It is therefore necessary to carefully work on the design of ways to support awareness, including in the context of mixed-presence collaboration.

2.2. Awareness Cues for Mixed-Presence Groupware

In order to transmit non-verbal awareness information to remote collaborators, mixed-presence groupware designers typically rely on visual cues. Such cues are instantiated as features or interactive elements that provide answers to the questions of Gutwin and Greenberg [7] and give feedback on the presence, identity or behaviour of users. As they essentially serve to convey WA information, we call them WA cues.

Remote collaboration through standard desktop applications typically involves the sharing of mouse cursor positions [22] as well as video and audio communication. When mid-air gestures are involved, researchers propose to use pointing cues, which relay the positions or objects remote collaborators are referring to with their hands (e.g., [23]). Different aspects related to pointing have been extensively studied. For instance, it was found that the combination of hand gestures with a sketch cue helps completing complex tasks faster, but also led to a higher workload [24], and that the hand-based selection or confirmation [25] can be performed using different gestures.

Another common type of remote awareness information relates to attention. Transmitting head or eye gaze cues has indeed been studied, with visual indicators taking various shapes such as circles [26] or squares [27]. Previous work showed that gaze-based cues can be faster and more precise than pointing cues [27], and that gaze is used complementary to hand gestures [27]. In some situations, such as bimanual operations, they can also replace hand gestures [26]. It should also be noted that a previous study [13] found that seeing the working area, tools, and collaborators' hands was more important than seeing their face or preserving eye contact, in a context where a remote expert is assisting a local worker on a physical task.

A significant portion of the research on gaze and field of view transmission has involved mixed reality and head-mounted displays. Jing et al. [28] experimented with a 60° panoramic-video based mixed reality (MR) system which shares four types of gaze behaviour visualisations consisting of two levels of gaze styles (uni- or bi-directional) and two levels of behaviours (with or without) via three types of interfaces (colour, shape, or both) between a local user and a remote collaborator. They found that distinctive visual feedback in a remote MR environment represents gaze behaviours more effectively in sharing attention and enhancing mutual communication, and that encouraging frequent joint gaze interactions, and gaze behaviour visualisations have improved active mutual reaction to the shared interest. This is consistent with research on more standard displays that has shown the influence of gaze cues on target acquisition and fixation times [29].

2.3. Awareness on Large Interactive Displays

Research on large interactive displays such as WSDs and tabletops indicates that these displays have a positive impact on WA during collocated collaboration [30], because they allow personal parallel work while enabling collaborators to give and receive help or suggestions from others in real time, but also because they improve the understanding of the interactions of others participants with the shared workspace [31]. This translates into changes in collaborative behaviour, that can be observed through indicators such as the ones proposed by Hornecker et al. [32]. These consist of positive WA indicators including interference like reaching for the same object and verbal monitoring ("what did you do there"), but also negative WA indicators, such as reaction without explicit request and parallel work on the same activity without verbal coordination. According to a study on the orientation of large interactive surfaces [33], a good level of awareness can be maintained on both horizontal and vertical surfaces, although verbal comments were more often needed in the vertical screen condition.

Some work has been conducted on studying behaviour related to awareness around large interactive surfaces specifically, including pointing gestures on tabletops in low-awareness situations [34], as well as attention based on gaze using eye tracking devices when interacting with a large vertical display [35]. Recently, Gong et al. [36] suggested to rely on a transparent display so that collocated collaborators may see each other face-to-face through the display. This allows them to interact from both sides of the same physical device, while naturally gathering awareness information. The study they conducted has shown that using such displays may lead to better task performance and has effects on collaboration aspects, compared to traditional WSDs.

To convey awareness information to remote collaborators, some systems have opted to include body parts, such as hands [37] or arms [15], to provide both pointing and gestural information. In a study involving collocated participants on a tabletop setup [38], these types of body embodiments have been shown to be preferred by participants compared to more abstract visualisations. However, objective measures did not show any difference in terms of awareness. Some systems have even gone further by opting to include larger parts or even full body representations of remote collaborators [39,40]. Yet, as the representation becomes larger, it starts occluding more and more of the view and may thereby hinder the remote collaborator's ability to view the rest of the content in the shared workspace.

Interestingly, in the aforementioned study focusing on tabletop settings [38], Pinelle et al. found that larger embodiments have not been considered as more distracting as compared to smaller embodiments.

Overall, except for a few works mentioned in the present section, the existing literature on mixed-presence groupware has mostly focused on either traditional desktop systems and more recently on augmented reality and virtual reality-based collaboration, including questions related to the combination of WA cues [41] in that context. The scarce literature regarding mixed-presence groupware in the particular context of WSDs is of an exploratory nature and has provided first insights on the role and design of a video feed to connect users from two locations [42,43]. In this paper, we build on these works and study how such a video feed can be complemented with additional visual indicators on awareness information.

3. Our Mixed-Presence Collaboration System

To study the benefits of WA cues in mixed-presence collaboration with WSDs, we developed a decision-making scenario involving the analysis of several types of data (See https://www.youtube.com/playlist?list=PL0kgBrqpGd3HF_ndMY-EySvQbqVQRGWIL, accessed on 15 October 2024, for a video illustrating the full system). We instantiated the scenario on two physically-distributed WSDs and added conventional support for mixed-presence collaboration: the views on both sites are synchronized and connected through an audio-video feed. As part of a user-centered design process, we then designed seven complementary WA cues (list of attendees, miniature view, touch feedback, mid-air pointing feedback, (dis)agreement icon, hand raised icon and gaze icon) to provide additional information on what the remote team members are talking about, where their attention is focused, and what they are doing and intending to do.

In the following subsections, we will give more information about the decision-making scenario, the conventional remote collaboration support, as well as the aforementioned WA cues.

3.1. Decision-Making Scenario

Our scenario involves the use of a control tower that supports experts and decision-makers in managing medical supply chains during distress times (e.g., a pandemic) [12]. It addresses a problem experienced during the first wave of COVID-19, where hospitals had difficulties obtaining the needed personal protective equipment (e.g., surgical masks, goggles, gowns) on time. The hospitals had to take care of all the tasks related to ordering new stock by themselves as the usual intermediary suppliers were no longer delivering. The scenario therefore describes a team of decision-makers from a hospital that has to ensure that the stock of protective equipment meets the hospital's needs. The team needs to estimate the upcoming new COVID-19 infections and hospital occupancy (task 1), select one type of equipment where the strongest stock shortage is to be expected (task 2), select an offer to replenish the stock (task 3), and select a delivery method (task 4). For each one of the steps, different types of data are provided (see Figure 2) and juxtaposed onto the WSD. Participants need to interpret and analyse the data, and identify the best solution given the existing constraints.

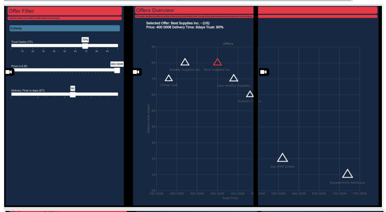
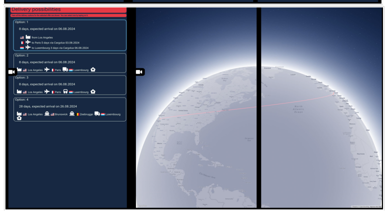
Tasks		Data	Roles *
	**Task 1: estimate future COVID numbers** • read the news • estimate hospital and ICU occupancy, • and adapt the sliders according to their estimation.	Text, sliders, graphs	Head of ICU
	**Task 2: select the protective equipment to restock** • compare the numerical data in the table, • discard items • and choose the one with the lowest number.	Table with icons, text, and numerical data	CEO
	**Task3: Select one offer in the overview** • filter offers • and then select the cheapest offer that fulfils all criteria (price, provider's trust factor, delivery time)	Sliders, scatterplot	Head of ICU, Head of finance, Head of logistics, CEO
	**Task 4: select delivery option** • select the delivery option that fulfils the criteria (time, not passing through some areas, avoid some towns as stage)	Icons, text, and numerical data, map	Head of finance, head of logistics

Figure 2. The four tasks used in the decision-making scenario. * Predominant roles (the four roles are always involved).

This scenario is based on a national project to react to COVID-19 and has involved several hospital professionals to define its first version (Demonstration video of the result of the ACTING NoW project with an overview of the initial scenario; here, we focused on the collaborative decision-making scenario on the WSD: https://www.youtube.com/watch?v=qPlrT63K5qU, accessed on 15 October 2024). We chose this scenario for its understandability by non-experts, the variety of data visualisations that are used and the necessity to discuss to find the best solution. As we recruited participants with different backgrounds from our research campus, we adapted the control tower to be accessible for them. Different roles have been defined related to a real situation: the CEO of the hospital, the head of logistics, the head of ICU and the head of finance. They all have different information and constraints, displayed on a card, that they can share with the other participants at the right moment.

3.2. Basic Support for Remote Collaboration

As basic support for remote collaboration, the system includes an audio-video link such as users expect in a remote collaboration context. It was implemented by positioning a single camera in the middle of each WSD, filming the participants from the front. The raw video feed from that camera is shown to the remote users through a window that is also placed in the middle of the WSD (yellow dotted rectangle in Figure 3). As for the sound, a microphone placed in the middle of each WSD is used to send the recorded audio, that is then played through a speaker on the distant site.

In addition to the audio-video link, the system integrates synchronized views, meaning that modifications (e.g., selecting a row or moving a slider) on one site are visible in real time on the second site.

Figure 3. The Immersive Arena with the two first tasks of the scenario displayed and the visualisation of the video feeds, respectively, for the remote (yellow dotted line) and local (green solid line, additionally highlighted by an arrow) teams.

3.3. Workspace Awareness Cues

To design the WA cues, we followed a process consisting of five steps (see Figure 4). We first started with a test session (1) where participants role-played the decision-making scenario using only the basic remote collaboration support (synchronized views and audio-video feed). Seven participants were divided into two groups that had to collaborate from two separate (and distant) WSDs. The same participants were then invited to take part in a brainstorming session (2) where they identified and discussed their encountered problems related to awareness as well as their ideas for solutions.

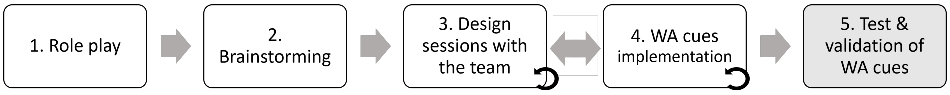


Figure 4. The design process of the WA cues. The rounded arrow shows an iterative step.

We analysed and structured the emerged problems and ideas using Gutwin and Greenberg's WA framework [7], leading to an adjusted framework on WA components (see Table 1 and [44]) that covers the range of WA information needed for our context. By making use of this framework, we then iteratively designed (3) and implemented (4) seven complementary WA cues (see Figure 5), using walkthroughs in our research team to test intermediate prototypes and adjust the design. Decisions were taken based on our own experiences during the tests, but also by considering established guidelines from cognitive ergonomics (e.g., [45]). As part of the design process, we adapted the shapes and colours used for visualising the WA cues, as well as their location on the screens. While results from the first four steps have already been published elsewhere [44,46], in this paper we focus on the final design and the evaluation of the WA cues (5).

Table 1. Components of awareness information, including the related WA elements, refined questions and corresponding WA cues. The * marking denotes differences with the original table of Gutwin and Greenberg [7].

WA Component	WA Elements	Questions	WA Cues
Environment The configuration of the WSD environment(s).	Presence Identity View * Reach	Is anyone in the workspace? Who is participating? Who is that? What is available to be viewed? What is available to be manipulated?	List of attendees Miniature view
Action Actions that are done as part of the collaborative work. Here: manipulating, pointing, speaking (dis)agreeing, willing to speak.	Action Authorship Artefact Location	What are they doing? Who is doing that? What object are they working on? From which location are they working?	Touch feedback Mid-air pointing feedback (Dis)agreement icon Hand raised icon
Attention The state of others' attention while an action is being done.	Gaze	Where are they looking?	Gaze icon

In the following subsections, we give more details on the design of the seven WA cues, structured by WA component (Environment, Action and Attention).

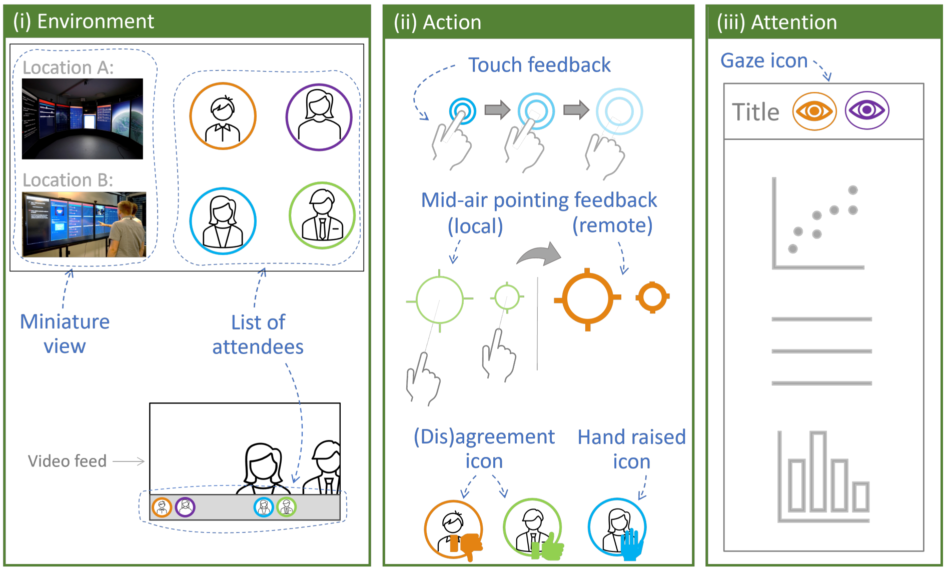


Figure 5. The integrated WA cues, grouped according to our adjusted framework. *Environment* cues (miniature views and list of attendees) are displayed on a dedicated screen and underneath the video feed. To clarify the *Action* cues, the manipulation feedback animation rapidly fades away while growing in size, the mid-air pointing feedback grows as the pointing user stands further back, and the gesture icons are shown next to the profile pictures, both near the miniatures and the video feed. As for the *Attention* cue, gaze icons are placed in the title bar on top of the display area.

3.3.1. Environment WA Cues

The *list of attendees* shows a picture of each of the attendees currently in the workspace, along with a dedicated colour. It provides information on the *Presence* ("Is anyone in the workspace?") and *Identity* ("Who is in the workspace?") of participants. The picture is taken automatically by the system and updated regularly to represent at each moment who is currently present. We chose not to display names or roles, to avoid having to configure the system in advance so that we could support ad hoc meetings. In our intermediate tests, the list of attendees was found to be relevant at the beginning of the meeting as well as sporadically. That is why we decided to put it in the peripheral field of vision, on a separate screen. We also made the choice to identify authors of *Actions* with colours. To make the colour attributions visible at a glance, we additionally displayed the list of attendees through coloured badges placed under the video which was frequently consulted in our intermediary test sessions.

The second WA cue, the *miniature view*, shows a static view of each of the sites, to provide information on how the sites look like, what participants can see on the screens, and how they can interact with the displayed content. It therefore concerns WA elements of *View* ("What is available to be viewed?") and *Reach* ("What is available to be manipulated?"). We opted to place it next to the list of attendees as we similarly found this information to be most useful at the beginning of the collaborative session and occasionally. We decided not to use additional back or top video cameras to capture this information, to avoid occluding further the screen view and opted for a simple and single image that is shown next to the list of attendees.

3.3.2. Action WA Cues

Previous works showed that there were five different types of actions that participants are interested to be kept aware of in this type of context: (a) touch pointing actions, (b) mid-air pointing actions, (c) interface manipulations, (d) (dis)agreement actions, and

(e) when participants are willing to speak [44]. This resulted in the design of four different WA cues that are pictured in the (ii) Action block of Figure 5. Each of these WA cues provides information on *Action* ("What are they doing?") and on *Authorship* ("Who is doing that?"), that is complemented by additional information on the *Artefact* ("What object are they working on?") and the *Location* ("From which location are they working?") when appropriate.

The first type of action WA cues, *touch feedback*, covers (a) pointing gestures that touch the screens and (c) interface manipulations. They are shown through an animated double circle depicted in the colour code associated to the participant who made the action. In case of a longer touch gesture (e.g., a slider movement) the whole path is highlighted through smaller circles that are appearing regularly throughout the movement. The colour is used to convey the information on the *Authorship*. The double circle was chosen because it attracts more attention than a simple circle. Our rationale was that, as found during our intermediate tests, when participants touch the screen, they either perform a manipulation that other participants might benefit from being notified about, or intentionally try to attract the attention of all participants (particularly the remote ones). Both cases call for an immediate and visible indicator, hence our design decision. We additionally left the center of the circles empty to keep the target data visible (e.g., a word or a number, which would end up circled instead of obfuscated if the circle was filled).

Mid-air pointing feedback is provided when users point without touching the screen and is shown as a target symbol on the artefact that is pointed at. We have chosen a different shape to the one used to represent touch gestures, to make it clear that they are not the same *Action*. The center of the shape is also empty to allow the target data to be read. To provide an indication of location, we used a thicker stroke for remote cursors and adapt the size of the symbol based on the distance from the user's hand to the target artefact (the symbol decreases in size as the hand approaches the target). This indicates that the pointing is more accurate.

Agreeing and disagreeing icons appear when participants make thumbs up or thumps down gestures, and are placed next to the profile picture on the attendees list as well as underneath the video feed. We chose to include these gestures as they are widely used in many parts of the world, including East Asia and the USA (even if thumbs up is offensive in Iran, Afghanistan, and Middle East), and because of their omnipresence in social networks as well as in conventional videoconferencing tools [47].

In a similar way, a *hand raised icon* is shown next to the profile picture if a participant physically raises their hand. We opted to add this gesture as well since it is customary to raise one's hand to request the floor [48] both in real life and through conventional videoconferencing tools.

The action WA cues enable to answer the following questions: "What are they doing?" by displaying a cue which is different regarding the nature of the action, "Who is doing what?" by adding a colour code for the authorship related to the colour around badges. Furthermore, for touch and pointing "What object are they working on?" is answered as the both sites have synchronized views and the WA action cues are displayed where they are performed, as well as "From which location are they working?" as the remote and local cursors are visualised in a different way, and become smaller when the user is close to the screen.

3.3.3. Attention WA Cues

Finally, regarding attention WA cues, we use a *gaze icon*, coloured according to the participant, on top of the display area (e.g., screen or window) where their head is directed at. It provides information on *Gaze* ("where are they looking?") based on the participants' general head direction as an indicator of where their attention is focused. We decided to not use accurate eye tracking, to avoid the use of specific devices that the participants need to wear, which could feel obtrusive, especially in a group situation. Subsequent mentions of "gaze" therefore refer to "head gaze", unless otherwise specified. We further decided to limit the precision of the gaze cue to "target screen", instead of trying to indicate the precise (computed) target of attention. This was because some of the participants from the brainstorming were concerned about privacy and the possibility that precise gaze information might be overwhelming or distracting [44]. We considered using a coloured border to surround a target screen but that approach raises questions regarding the combination of different borders when multiple persons are looking at the same screen: should "mutually exclusive" dashed lines be used? Should the colours be composed somehow? Because that might be too complex for users to interpret and since we generally considered that option to likely be too distracting, we chose to rely on the gaze icons on top of the target screen.

3.4. Technical Implementation

We instantiated the aforementioned WA cues through the combination of two systems: a multi-device orchestration framework called DeBORAh [49] and a human behaviour tracking pipeline [46].

DeBORAh is a web-based software layer that handles the front-end part on the wall-sized displays, driving the scenario, its synchronisation between different sites, and serving the views with the related data visualisations. It does so through a flexible approach based on containers and webpages so that the multimedia content may be displayed differently on separate locations and setups, and handles multi-display multi-device systems. The WA cues are also rendered within DeBORAh based on the output of the tracking pipeline, on top of the scenario's content.

In order to feed the tracking pipeline, we rely on commodity depth cameras (Azure Kinect), which naturally provide pointing and (head) gazing information through the associated body tracking capabilities. We put that information in relation to the display arrangement following a prior scene calibration step [44], in order to turn Kinect-space deictic lines into screen-space points to be used by DeBORAh. We further attribute the resulting WA cues to specific persons thanks to an identification procedure based on continuous face recognition, which leads to differentiated WA cue visuals with colour attributions. We also augment the tracking capabilities of the system with a hand gesture recognition model, to provide the hand raised and agreement/disagreement gestures mentioned in Section 3.3.2.

4. User Study

We conducted an exploratory user study to learn how groups of participants would collaborate in mixed-presence with our system, and how the WA cues would impact their experience, their collaboration, and their performance. The study was conducted in Q4 2023, with a total of 24 participants, divided into six groups of four. The members of a group were equally distributed across the two sites. The participants had to solve a series of tasks as part of a decision-making scenario, involving the interpretation of data visualisations and information shown on the displays.

4.1. Research Questions

The research questions of our user study are the following:

- Do the WA cues actually enhance workspace awareness as compared to an audio-video feed only? (RQ1)
- Which WA cues do users prefer and perceive as most useful? (RQ2)
- Do the WA cues provide users with a good UX? (RQ3)
- Do the WA cues improve the collaborative decision making as compared to an audio-video feed only? (RQ4)

With RQ1 we wanted to verify whether the proposed WA cues actually fulfil the purpose they were designed for, i.e., to provide users with a better workspace awareness as compared to an audio-video link. The rationale of RQ2 is to further investigate differences in how the individual WA cues are perceived, in order to obtain first insights into which types of WA cues are the most relevant. With RQ3, we wanted to find out how the additional information that users receive would impact their user experience and, in particular, if it would be overwhelming or distracting. Finally, with RQ4, we were interested in the impact that the WA cues would have on collaboration and decision-making. The rationale behind this research question is that a high degree of awareness is considered crucial for collaborative work, hence, we assume that the collaboration experience and quality would improve with the WA cues.

To be able to analyse the impact of the WA cues system onto participants' perceptions and collaboration, we designed a comparative user study based on two different conditions: (C1) solving tasks using the full system with an audio-video link and the seven WA cues, and (C0) solving tasks with an audio-video link only, without the WA cues. The experimental design was constructed as within-subject design with each group performing two tasks in condition 0, and two tasks in condition 1. To minimize order effects, the conditions were counterbalanced with every second group starting with another condition, see Figure 6.

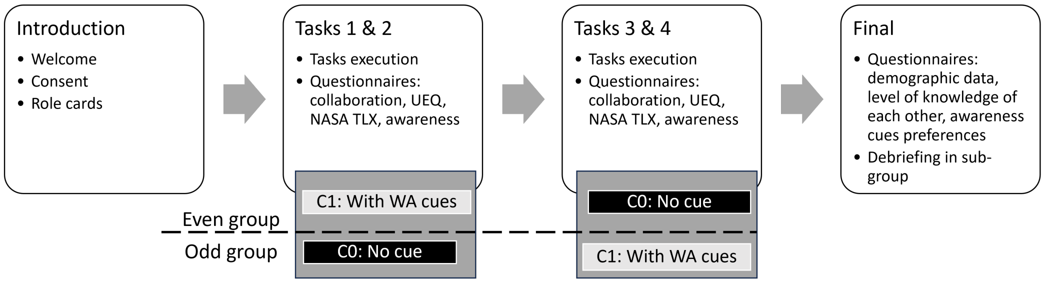
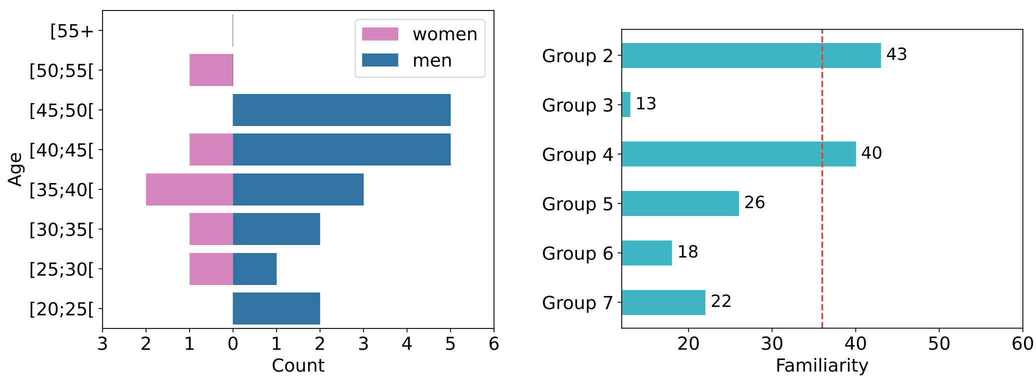


Figure 6. Overview of the protocol followed during the user study.

4.2. Participants

In total, 24 participants agreed to take part in the user study (6 females and 18 males) with an average age of 38 (min = 23 and max = 52), see Figure 7a. They were researchers, engineers or administrative staff, recruited from our research institute through convenience sampling.

As the experiments were conducted in English, we asked them to self-evaluate their English level. The majority of them rated themselves as advanced English speakers (16 of 24), some were intermediate English speakers (5 of 24) and three of them have English as their mother tongue (3 of 24). In addition, the majority of the participants use a large interactive display one or two times per year (14 of 24), the rest never used this kind of equipment (10 of 24).



(**a**) Age pyramid.

(**b**) Familiarity score for each group.

Figure 7. Study participants demographics and intra-group familiarities.

According to the timeline of enrolment and their availabilities, participants were assigned to groups of four with a specific timeslot for a session. In addition, each group was further subdivided into a subgroup of two, and assigned to one particular physically-distant WSD setup (either one from the two on Figure 1).

We asked them how much they know each other within each group using a Likert scale of 5, where 1 means "not at all familiar" and 5 means "very familiar". The minimum resulting familiarity score for a given group is 12 (the four group members evaluate their familiarity with the three other team members at 1, that means: $4 * (3 * 1) = 12$), the maximum possible is 60 ($4 * (3 * 5) = 60$) with the middle being at 36. Groups 1 and 3 were above average familiar (43 and 40, respectively), group 4 and 6 were below average familiar (26 and 22, respectively) while groups 2 and 5 were not familiar, see Figure 7b.

4.3. Apparatus

The study was conducted in two different connected lab environments, physically located in two different buildings (see Figure 1). In each of the labs, we used a WSD to display the data visualisations and information as explained in Section 3.1. For this study, even though the displayed content itself was synchronised across the two WSDs, it varied in terms of layout and placement of elements on the screens, due to a difference in the setup and technical specifications of the used WSDs. To be more specific, one lab used a circular WSD that we call the *Immersive Arena* and the other one used a more traditional flat WSD called the *Viswall*. The Immersive Arena is a 360° circular arrangement of 12 screens of 4 K resolution each, extending to a diameter of 3.64 m and reaching a height of 2 m. The Viswall is itself formed by a 12×2 arrangement of Full HD screens, that spans across 7 m and also reaches 2 m high. To create as similar conditions as possible in both settings (similar size and layout of the content) we only use part of the available display space on both WSDs. More specifically, we used nine screens in the circular WSD and scaled the content on the flat WSD so that the size is the same as on the other site.

4.4. Procedure and Tasks

As described in Section 3.1, participants played the role of experts and decision-makers in managing medical supply chains during the COVID-19 pandemic in a hospital. The scenario was divided into four tasks: (1) forecasting COVID-19 cases and hospitalizations, (2) identifying equipment shortages and selecting the most problematic, (3) selecting the optimal restocking offer, and (4) choosing a delivery method.

The 24 participants were divided into six groups of four. Each session started with an introduction phase with a short explanation of the research project and its objectives. We welcomed the participants, explained them the purpose of the experiment, how their data will be used and collected their consent. Then, we introduced them to the scenario (see Section 3.1), the environment (see Section 4.3) and explained the role cards. We asked them to keep the information on their cards private, and only share it when it is relevant. We also showed them the audio-video feed and explained that their environment are synchronised. Each participant was randomly assigned one of four roles (CEO, Head of ICU, Head of Finance, Head of Logistics) within the scenario and received the corresponding role card on which we indicated extra information that was not shown on the displays, see Section 3.1. For instance, the head of ICU had the information that with the projected peak, at least 10 ICU patients can be expected. This simulated background information and expertise, and was put in place to push participants to collaborate both locally and with the distant site, based on principles of positive interdependency [50]. Since not all roles were involved in all tasks, the role cards led to a predominance of some roles for each task (see Figure 2).

Then, we activated one of the two conditions, according to the planning. In case of the C1 condition (with WA cues) the facilitator then demonstrated each of the seven WA cues. Independently of the condition, the facilitator then asked participants to execute the first two tasks and left the space to avoid further disruptions. After having completed the tasks, participants filled in a questionnaire about their perception on the collaboration, the user experience (UEQ [51]), the workload (NASA-TLX [52]), and their awareness (see Section 4.5.1 for more details). After that, they had to complete the last two tasks under the other condition (and if C1 condition, an introduction to the WA cues was performed) and filled in the same questionnaires. At the end we asked them their demographic data and their preferences in regard of the WA cues. In addition, a debriefing was conducted to discuss which WA cues they have used. The full protocol and checklist is available on a data repository (see Data Availability Statement).

We also designed the questions and data in a way that the chosen equipment in tasks 1 and 2 did not affect the subsequent input for tasks 3 and 4, which could therefore be solved correctly in any case.

Prior to the main study described here, we conducted a pilot study with one group to check if the procedure and the prototype were working well.

4.5. Measures

We used both qualitative and quantitative measures to collect data for our study (see Table 2). We video recorded all the sessions using three or four cameras on each site, covering multiple angles. The recorded videos added up a total amount of 375 min of recordings. In addition, we logged participants activities and installed a room microphone to record the audio on each of both sites. We used OpenAI's Whisper large-v3 model [53] to perform the automatic transcription of speech data to obtain a transcript for each site. As part of the automatic transcription, short moments of another language (that occurred despite the instructions) were translated into English. After that, we manually separated the transcript from one site into turns, assigned the appropriate speaker, and complemented it with missing turns from the second site.

As part of the separation process, we also corrected obvious and impactful transcription errors, typically resulting from hallucinations from the transcription model (e.g., repeated words or even sentences for portions of the audio without any speech). Due to time constraints, the few remaining minor errors (e.g., "head of *silence*" instead of "head of *finance*") were not corrected as it seemed fair to assume such minor mistakes in the automated transcription were spread somewhat evenly, leading to a negligible impact on the analysis. Furthermore we analysed and manually annotated the video recordings using the ELAN software [54] (Version 6.2). The individual team members were pseudo-anonymized through an ID. In ELAN, each ID was then assigned a "tier" where their corresponding pointing and touch gestures were manually annotated by a human coder.

Table 2. Indicators used to evaluate participants experience and behaviours.

Measure	Data Source	Indicators
Awareness	Questionnaire	WA results
	Video	Communication efficiency (spoken words per second)
	Video	Interaction conflicts
Preferences	Questionnaire	WA cue preferences results
	Debriefing	WA cue usage feedback
User experience	Questionnaire	UEQ-S results
	Questionnaire	NASA TLX results
Collaboration	Questionnaire	Teamwork assessment results
	Video	Speaking time distribution
	Video	Interaction distribution
	Video	Pointing gesture distribution
	Logs and video	Time to solve
	Logs and video	Correctness of solution
	Logs	Number of interface manipulations

4.5.1. Awareness Measurement

To understand participants' perception of their *awareness*, we created a similar awareness questionnaire as in [55] and defined 23 statements covering the WA components and elements from our framework as indicated in Table 1, e.g., "I knew how the screens of the remote team looked like." (Environment—View) or "I knew what remote team members were pointing at." (Action/Pointing—Artefact). Participants rated these items on a 5-point Likert scale ranging from 1 (never) to 5 (always).

Furthermore, we considered objective measures on awareness. Since previous work had observed that in situations of high awareness, users coordinate actions with little verbal communication [1], we collected these from communication aspects. *Communication efficiency* refers to the level of communication effort participants need to apply in order to solve the task and is often measured through the number of words or utterances per group or the number of physical deixis [56]. Since in our study, there are some differences in the tasks (type of visualisations displayed and distribution of private information), we used the average number of spoken words per turn as well as the number of spoken words per second.

In addition, we extracted the number of interaction conflicts, i.e., moments where several group members attempt to access or modify a shared resource at the same time. Such interaction conflicts are considered to measure coordination breakdowns and indicate a lack of awareness [56]. We extracted this information from the video annotations, by looking for overlapping touch annotations related to interface manipulations. As video recordings (and therefore the resulting annotations) from different sites were not necessarily started at the exact same time, we had to apply an offset to each annotation before evaluating whether there are overlaps.

To obtain more specific insights on each of the seven WA cues individually, we included in the post-test a question where participants could indicate a rating ranging from 1 (very useless) to 5 (very useful). This quantitative data was complemented by an open-ended question asking about a justification of the most useful and useless WA cue, as well as a debriefing interview where we asked participants questions about which condition they preferred and which WA cues they were using.

4.5.2. User Experience Measurement

To measure the *user experience*, we used the short version of the user experience questionnaire (UEQ-S) [51], complemented by the NASA-TLX questionnaire [52] to obtain additional insights on the workload that the WA cues generate for participants. We only used the first part of the NASA-TLX and evaluated all six subscales, without asking participants to compare subscales pairwise. This version is referred to as "raw TLX" [52] in the literature.

4.5.3. Collaboration Measurement

To obtain insights on participants' collaboration, we used objective metrics (equity of participation and communication efficiency) and combined it with subjective metrics (questionnaires). *Verbal participation* is often measured based on either the speaking time, or the number of turns or spoken words (e.g., [57–59]). In this work, we follow the same approach and counted the number of spoken words for each person based on the transcripts. Similar to verbal participation, we calculated *physical participation* and *gestural participation* through cumulative counting of the number of interface manipulations and mid-air pointing gestures as in [57–59], respectively. Interface manipulations could be, for instance, short taps on buttons, but also longer drag movements on a slider. We then transformed all three levels in relative proportions of the total number of spoken words, interface manipulations, and mid-air pointing gestures, and used the standard deviation as a measure of equity [60].

As a subjective measure on collaboration, we collected participants perceptions, and created a own survey based on the teamwork assessment questionnaire [61] with seven questions, covering the three dimensions of team orientation (i.e., team members' ability to agree on goals, tasks, and concepts involving the mission), communication behaviour (i.e., team members' ability to provide important information to others), and coordination behaviour (i.e., team members' ability to coordinate their behaviour). Responses were made on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

4.5.4. Decision-Making Quality Measurement

Finally, to measure the efficiency of decision-making, we extracted the *time to solve* and *number of clicks*, starting from after the introduction of the facilitator and ending when the groups hits the "next" button. We also took into account the *correctness of the solution* by giving a score of 1.0 in case all parameters were correctly set, and a proportionally lower score for each subtask that was not correctly solved (e.g., 0.75 in case 3/4 subtasks were correctly solved). To make sure that the tasks in the two conditions can be evaluated independently with regard to the correctness, we defined the parameters and constraints in a way that the correct solution for tasks 3 and 4 remains the same, independently from the selection in task 1 and 2.

4.5.5. Data Analysis

To analyse the differences between experimental conditions, we used several statistical tests. Likert data from the questionnaire were analysed using Wilcoxon signed-rank tests and for the objective measures (logging data and transcript analysis), we used paired samples t-tests.

For some of the measurements (user experience, perceived collaboration and workload), we also compared our results with the questionnaire data from a previous study [12], with six groups of four participants, involving the same scenario, group size, and roles, but in a collocated setting. To identify differences, we used Mann–Whitney U tests on this between-subject experiment data set. None of the participants of our study had participated in the mentioned prior study.

5. Results

The study has provided us with multiple types of data on awareness, preference, user experience, and collaboration, presented in the following subsections.

5.1. Awareness Aspects Analysis

The results of the awareness questionnaire are shown in Figure 8. On average, participants rated the awareness in all categories higher in the C1 condition as compared to the C0 condition. The Wilcoxon signed-rank tests showed that the differences of three categories are significant, more specifically the *Environment* category ($Z = 3.2874$, $p < 0.001$), the *Pointing Actions* category ($Z = 2.5553$, $p < 0.01$), and the *Attention* category ($Z = 2.4124$, $p < 0.05$). The differences for *Speech Actions* ($Z = 1.6561$, $p = 0.1003$), *Manipulation Actions* ($Z = 1.834$, $p = 0.0674$), *(Dis)agreeing actions* ($Z = 0.5897$, $p = 0.574$), and *Willing to speak Actions* ($Z = 0.8448$, $p = 0.4053$) were not significant.

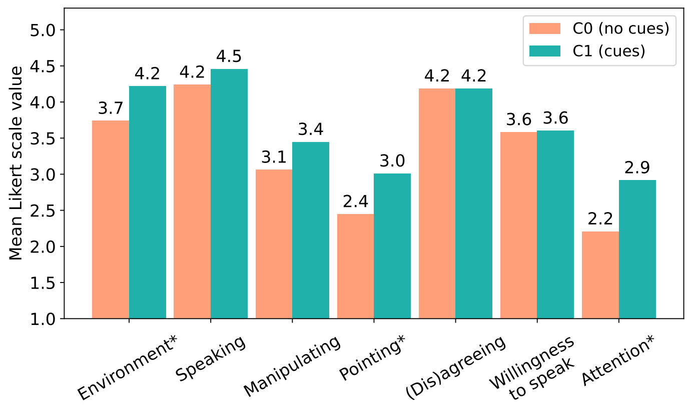


Figure 8. The level of awareness for each WA component in the condition C0 vs. C1. * indicates statistical significance between C0 and C1.

Concerning communication efficiency, we found that the groups had in C1 turns with an average length of 9.73 (SD = 1.69) and spoke 1.89 (SD = 0.22) words per second, whereas in C0 their turns were on average 11.71 words long (SD = 2.01) and they spoke 1.78 (SD = 0.36) words per second. Neither the difference in the length of turns (t(5) = 2.2923, p = 0.0705), nor in the words per second (t(5) = 0.94602, p = 0.3876) was significant.

Looking at the number of conflicts, we identified between 0 (G7, C0) and 5 (G4, C1) conflicts for each group and condition. Despite our expectations, there were more conflicts occurring in the C1 condition as compared to the C0 condition. On average, there were 2.33 conflicts in C1 in comparison to 2.0 on average in C0, but a T-test showed that the difference is not significant (t(5) = −0.5, p = 0.6383).

5.2. Preference and Utility Analysis

To analyse the results of the WA cues preferences questionnaire, we counted the number of positive, negative and neutral ratings. As seen in Figure 9, we found that the most positive ratings were given to the touch feedback (17/24), followed by the list of attendees (16/24) and the mid-air pointing feedback (15/24). The most negative ratings were given to the agree/disagree icons (12/24) and to the hand raised icon (11/24). Overall, opinions varied considerably with each WA cue having ratings on the entire scale (1 to 5).

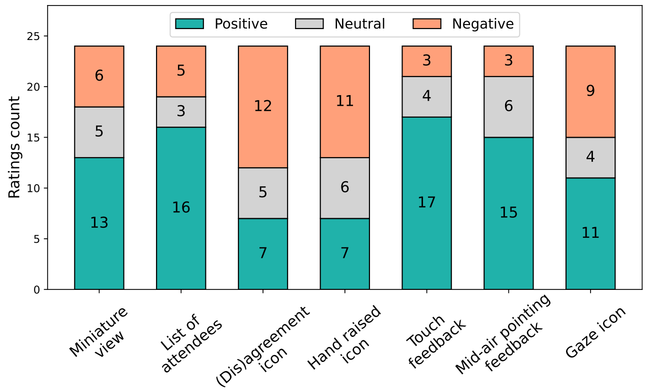


Figure 9. Individual cue ratings.

In the open-ended questions and the debriefing interview, participants provided some reasons for these ratings. They argued that they did not need the WA cues about the environment as they knew how the room was set and because it was a small group (G7-ICU and entire G3). Furthermore, participants explained that the gestures for showing agreement and taking the floor were unnecessary, because it was easier and quicker to express this through verbal communication with the audio-video link (G3-FIN, G3-CEO, G3-ICU, G5-CEO, G6-ICU). For example, G3-ICU noted: "Given the number of participants, well, we still manage to collaborate orally". However, most of them thought that these WA cues would be useful in a meeting with more participants (all groups but G2), especially if they do not know each other well (G5-ICU, G6-FIN). Regarding touch and pointing WA cues, the feedback was most positive, mentioning how it helped them to understand what team members were talking about and what is important to look at. For instance, G3-ICU found that "When someone was pointing the screen somewhere, somehow it gave, some feedback about if somebody was trying to point out attention about something". The few participants, that found touch feedback unnecessary, explained that they only used it to change parameters, and that this change was anyway visible through the synchronized views (G3-LOG). G4-FIN also indicated that they would prefer having a camera filming from the back to see the pointing gestures. A point that was generally raised about the touch and pointing WA cues was the difficulty to distinguish and remember the associations of the different colours and then the authorship of each action.

Finally, we found that the opinions about the gaze icons were rather divided. Many did not see the information or had problems following it (G3-LOG, G3-ICU, G4-FIN, G4-CEO, G4-ICU, G6-CEO) and explained, for instance: "I didn't check it much because it was on top of the screen and it changed rapidly so that's maybe even a distraction." (G6-CEO).

Nevertheless, others pointed out that it can be very powerful to know where people are focusing their attention (G2-ICU), and that it can help to understand the context people are talking about (G6-ICU).

Overall, some participants were very positive about the usefulness of WACs, finding that they create a sense of being in the same workspace and that they promote understanding of what remote participants are doing. For example, G3-ICU argues "I have the impression compared to, for example, Teams meetings here, you have more the impression to be in some, I wouldn't say casual, but you have somehow the impression to be a bit in the same room".

5.3. User Experience Aspects Analysis

The results of the UEQ questionnaire are shown in Figure 10. Overall, the user experience was rated higher in the C1 condition (M = 1.77; SD = 0.59) than in the C0 condition (M = 1.33; SD = 1.19). This is also the case for the subscales *pragmatic quality*, and *hedonic quality*. In comparison to the available benchmark data, the overall UX can be considered as "good" in the C0 condition, and as "excellent" in the C1 condition. A Wilcoxon signed-rank test showed that the difference in hedonic quality (Z = 2.3151, $p < 0.05$) as well as the overall user experience is significant (Z = 2.3931, $p < 0.05$), whereas the difference in pragmatic quality is not significant (Z = 1.828, $p = 0.0683$). A closer look at the mean values of the individual items revealed that all pairs were rated better in the C1 condition as compared to the C0 condition, except for one pair (confusing-clear), which received lower ratings in the C1 condition (see Figure 11).

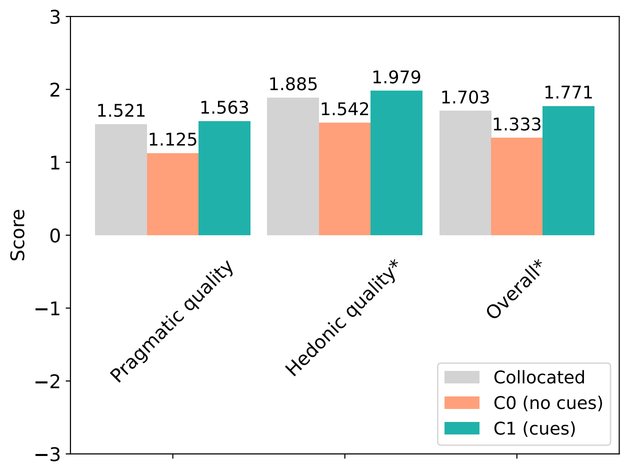


Figure 10. Grouped mean scores for the UEQ questionnaire in condition C0 vs. C1. Statistical significance between conditions C0 and C1 is indicated by *. For comparison, the grouped results from a prior collocated study are also displayed.

Concerning the results of the NASA-TLX questionnaire, participants on average rated the overall workload in the C1 condition 7.63 (SD = 5.11) and in the C0 condition 7.94 (SD = 5.15). The difference between both conditions is not significant (Z = −0.9293, $p = 0.363$). Also, the average scores on the individual subscales per condition do not vary much (see Figure 12).

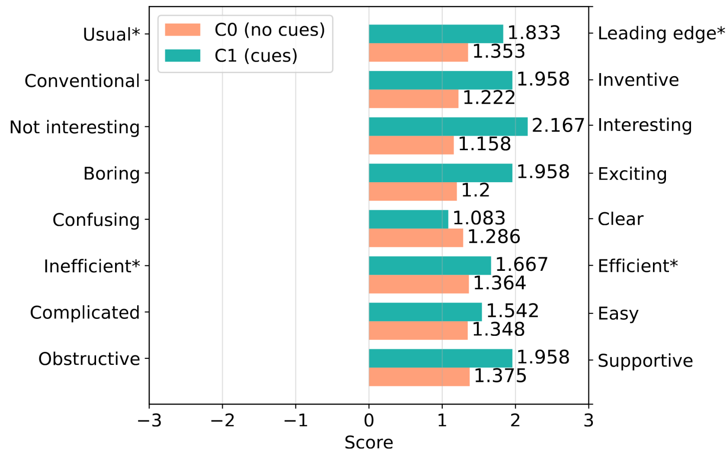


Figure 11. Mean scores for individual UEQ scales in condition C0 vs. C1. Statistical significance between conditions C0 and C1 is indicated by *.

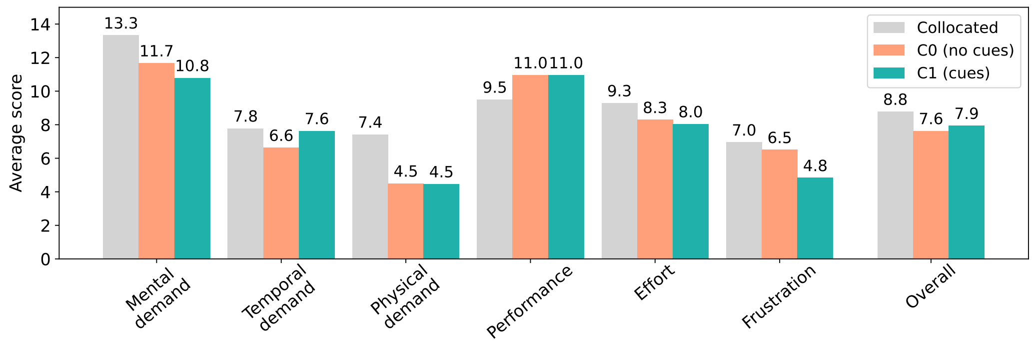


Figure 12. Average scores from NASA-TLX, comparing condition C0 vs. C1.

Compared to the results of a prior study (with the same scenario but in collocated work mode and without WA cues [12]) and based on descriptive statistics, we can see that the UEQ results for the C0 condition were generally rated lower as in the collocated setting. In contrast, in the C1 condition, they were rated even higher as in the collocated condition. Regarding the perceived workload, surprisingly, it turned out to be, on average, slightly higher in the collocated condition as in both mixed-presence conditions. The Mann–Whitney U tests showed that neither the difference in the user experience (Collocated vs. C0: Z = 219.5, $p = 0.1603$; Collocated vs. C1: Z = 283.5, $p = 0.9341$) nor the difference in the workload (Collocated vs. C0: Z = 232.5, $p = 0.2566$; Collocated vs. C1: Z = 214, $p = 0.1294$) are significant. So based on the sample data, we cannot say that there is a difference between the collocated and mixed-presence conditions.

5.4. Collaboration and Decision-Making Aspects Analysis

The results of the collaboration questionnaire revealed that participants in general had a very good perception of the collaboration experience with almost exclusively positive ratings. When comparing the two conditions, we found that participants rated the collaboration experience better in the condition C1 as in the condition C0 (see Figure 13). The mean of the summed collaboration ratings was 4.31 (SD: 0.27) in the WA cues (C1) condition, as compared to 3.96 (SD: 0.35) in the control (C0) condition. A Wilcoxon signed-rank test showed that this difference was significant (Z = -2.6357, $p < 0.01$). A statistical analysis of the individual sub-scales additionally revealed that the significant differences are observable for the categories *Team Orientation* (t(23) = 2.98, $p = 0.007$) and *Coordination Behaviour* (t(23) = 2.41, $p = 0.024$). Figure 13 shows participants average scores on their perceptions in both conditions.

In comparison to a prior collocated study, the collaboration in C0 was rated lower in all three dimensions as in the collocated, and the collaboration in C1 was rated slightly higher as in the collocated setting (see Figure 13). According to a Mann–Whitney U test, these differences are not significant (Collocated vs. C0: Z = 331, $p = 0.3785$; Collocated vs. C1: Z = 218.5, $p = 0.1532$), therefore, we fail to show evidence that there is a difference between these two conditions.

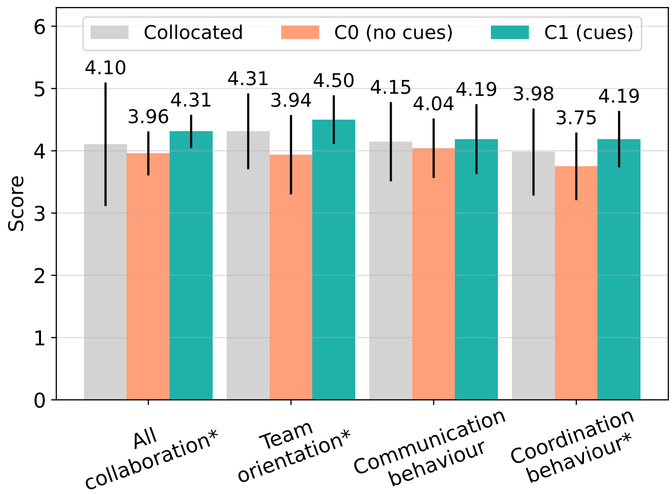


Figure 13. Comparison of answers given to the collaboration questionnaire, between condition C0 vs. C1. * indicates statistical significance between C0 and C1. For comparison, the results from a prior collocated study using the same questionnaire are also displayed.

Regarding the equity of participation, we calculated the distribution of speaking time, mid-air pointing gestures, and interface manipulations with regards to each group member and site, and used the standard deviation to calculate the equity. The results show that the average equity values are between 11.8 (manipulation equity of sites) and 19.64 (verbal equity of speakers).

When comparing the two conditions using descriptive statistics, we found that the mean values of all three types of participation equity between *group members* are higher in C1 as compared to C0. In contrast, the mean values the equities between *sites* are lower in C1 as compared to C0 (see Figure 14). These differences, however, are not statistically significant. A closer look at the data per group shows that there are large differences between the groups, so the number of data points might not be sufficient to detect significant differences.

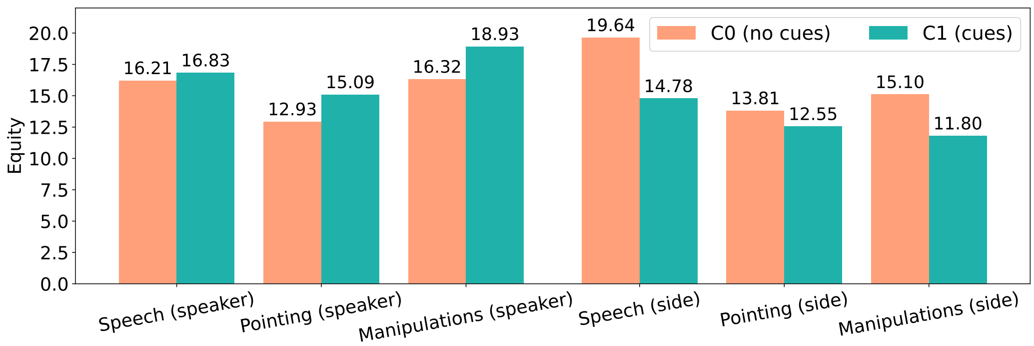


Figure 14. Equity of participation in both conditions.

We also looked at the performance of the groups, and in particular the time and interface manipulations they needed to solve the tasks, as well as the correctness of the chosen solution (see Figures 15–17). The results show that, on average, groups needed 640.17 s (SD = 179.80) to solve the tasks in the C1 condition in comparison to 787.50 s (SD = 227.83) in the C0 condition. Furthermore, they used 29.17 (SD = 13.98) interface manipulations in C1, and 33.33 (SD = 8.42) in C0 on average. However, these differences are not significant with the results from paired samples t-tests showing t(5) = 1.0857, $p = 0.3272$ for duration and t(5) = 0.7725, $p = 0.4747$ for manipulations. Concerning the correctness of the solution, we found that proposed solution were only incorrect in the C0 condition whereas all tasks were correctly solved in the C1 condition. The paired samples t-test showed that the difference between the two conditions is significant (t(5) = −2.74, $p < 0.01$). Overall, given the small number of groups (6), these results need to be interpreted carefully.

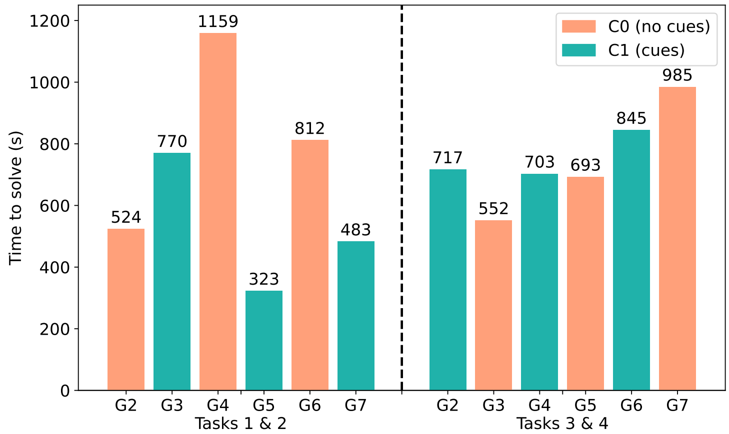


Figure 15. Completion times per group and set of tasks.

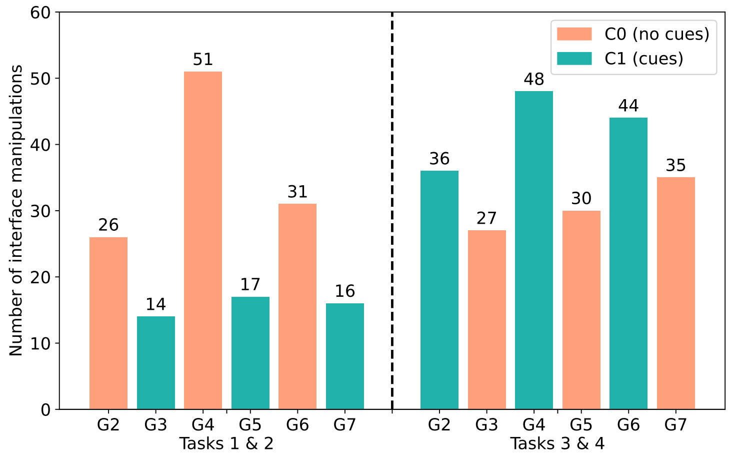


Figure 16. Interface manipulations per group and set of tasks.

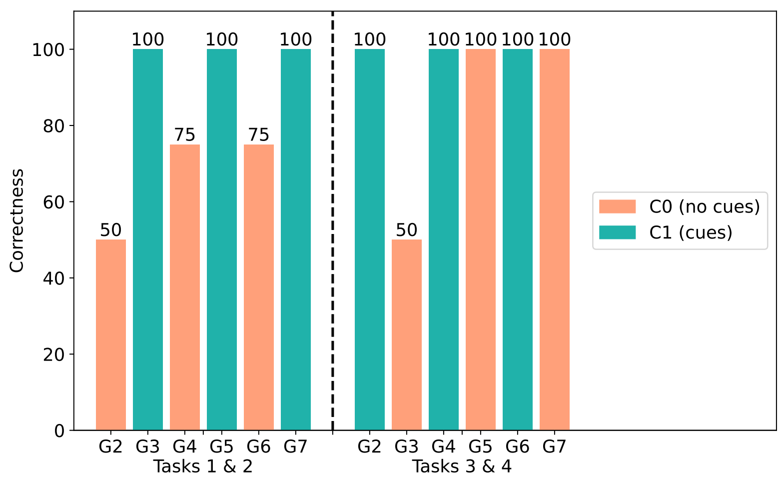


Figure 17. Correctness of answers given by the participants.

6. Discussion

6.1. How WA Cues Enhance Awareness in Mixed-Presence Collaboration

From our results, we can deduce that the WA cues enabled a better awareness of the *environment*, the *pointing actions* and the *attention*. In contrast, there was no significant difference regarding awareness on *speech*, *manipulation*, *(dis)agreeing* and *hand raising* actions. These results most likely indicate that some of the implemented WA cues have indeed been effective in enhancing WA whereas others not. Taking into account that we designed the WA cues to provide a specific type of WA information, it would suggest that both *environment* cues, the *mid-air pointing feedback*, and the *gaze icons* can be considered as effective. Concerning the *touch feedback*, the results suggest that it was effective in providing awareness on *pointing actions*, but not on *interface manipulation*. A possible explanation for this is that the feedback provided by the interface elements (e.g., moving sliders, or highlighted cells) was sufficient for the users to understand the manipulations and that the touch feedback did not provide much additional information.

Participants' feedback via the open questions and the debriefing interview confirm these prior reflections by indicating that the *touch feedback* and the *mid-air pointing feedback* were on average found to be most useful, followed by the environment cues (list of attendees and miniature views) and the *gaze icons*. The other icons (agreement, disagreement and hand raised) were in line with the results of the awareness questionnaire, found as least useful.

However, participants' opinions vary largely with each cue being both rated as "very useful" and "very useless" by at least one participant. One possible explanation for this might be that the WA cues are more versatile than anticipated and that, in the end, users used different WA cues to obtain the same awareness information. For example, to refer to information on the screen, users could either use mid-air pointing from a distance, or touch the screen and make use of the touch feedback. Some users' might just prefer one way over the other to do such an action. In addition, the specific needs for non-verbal communication might be different for the users, as some people feel more comfortable in using verbal approaches for indicating their thoughts and intentions, while others prefer to use gestures or other system features as a complement [62].

Regarding the objective measures of communication efficiency and interaction conflicts, however, we could not find any significant difference between both conditions. Based on descriptive statistics, the communication efficiency was slightly better in the C1 condition, which was manifested through less spoken words over time and shorter turns. However, the difference is not significant. Regarding interaction conflicts, we found that the number of conflicts was even higher in the C1 condition as compared to the C0 condition. The lack of statistical significance, on one hand, might be due to the small sample size, which was not sufficient taking into account the differences between the groups. On the other hand, it might also indicate that there is no difference in conflicts and that the type of awareness that is provided by the WA cues is not helpful for avoiding conflicts. In particular, the WA cues were designed based on a list of WA elements that relate to the *present*, however, to avoid conflicts, knowledge about the future (i.e., what users are going to do) would be needed. As pointed out by Gutwin and Greenberg [7], the design for maintaining awareness on future actions is particularly challenging.

To sum up, we can conclude about RQ1 that the WA cues do significantly enhance workspace awareness as compared to an audio-video feed only, in particular with regard to the *environment*, the *pointing actions* and the *attention*. In addition, to answer RQ2, our observations lead us to affirm that users prefer and perceive as most useful the *touch and mid-air pointing feedback*, followed by *environment* and *gaze icons*.

6.2. Design Considerations for Supporting WA

One of the design tensions of the WA cues is related to the vast quantity of awareness information that can be captured in a WSD environment and how it can be displayed on the screens without being overwhelming. In this comparative study, users were faced with a full setup of seven cues that were shown to them simultaneously, hence, providing them with a vast amount of information. Nevertheless, this did not create any particular difficulties for the users. According to the UEQ questionnaire, the UX was even better with the cues as compared to setting with only basic remote collaboration support. Additionally, the workload was not significantly different in both situations and lower as compared to the collocated setting. This shows that the users did not feel disturbed by the various WA cues, and instead appreciated the experience even more. A similar result was already observed in a previous study on awareness support on tabletop interfaces where users did not feel distracted by large embodiments of hand gestures [38].

Furthermore, participants' ratings of the cues have shown that they see merits for each of them, depending on their preferences and their prior familiarity of the sites and the other group members. This, in turn, speaks in favor for providing a full setup of the seven WA cues, in order to suit the preferences of each individual or group or to let groups customize the interface and decide which WA cues they would like to display. However, in case of other constraints (e.g., technical limitations or a lack of space on the screens) designers may need to prioritize some cues. Our results show that these should be the *mid-air pointing* and *touch* feedback to indicate pointing information, followed by cues to inform about the *environment* (*list of attendees* and *miniature view*) and the *gaze* icon.

However, the results of the present study also show that the design of the WA cues is not yet optimal and that there are further issues that need to be addressed. In the UEQ questionnaire, users rated the system with WA cues (C1) as more "confusing" as the one in the control condition (C0). In the feedback they mentioned that they did not see some of the WA cues, that they had difficulties in memorizing and allocating the colours, and that they would prefer to have some information in other places. Therefore, more work is needed to refine the WA cues and to ensure a better clarity of the WA information.

In conclusion and as an answer to RQ3, the WA cues provide users with a good UX, however, the design is not yet optimal.

6.3. The Role of WA Cues for Facilitating Collaboration

Ultimately, the goal of our system is to facilitate collaboration in mixed-presence settings. Therefore, in this study, we also collected data on collaboration aspects. The results indicate that, from the perception of the users, collaboration was enhanced when the WA cues were available. More specifically, we found team orientation and coordination to be improved. In contrast, there was no significant difference in the communication behaviour. Overall, the subjective collaboration with the WA cues (C1) obtained similar ratings as in the collocated situation, suggesting that our system is effective in generating a similar collaboration experience as when working collocated.

Regarding the objective measures on collaboration, we found a significant difference with regard to the correctness of solution. All teams did solve all the tasks correctly when using the system with cues, in contrast to the control condition where only 75% of the subtasks were correctly solved. However, we found no significant differences in equity of participation, nor time to solve or number of clicks, which might be because of the small number of groups (6), but also because of the difference in the number of predominant roles involved in each task.

Nevertheless, by comparing the mean values of the conditions and groups, the average equity of participation improved in all three dimensions (verbal, gestures and interactions) with regard to the distribution between sites, but not with regard to the distribution between speakers. Although no significant difference was found, this observation suggests that the WA cues might allow for improving intergroup collaboration (i.e., between the two sites, but less for supporting intragroup collaboration (i.e., within one site)). To be able to better understand if and how enhanced awareness can necessarily lead to improved collaboration, taking into account the inter- and intra-group configurations, more studies need to be conducted with a larger sample size and less varying conditions.

Fan et al. [63] suggest that participation can be considered as "*most equity*" when the equity value is less than 5, "*some equity*" when the value is in between 5 and 12.5, and "*unequal*" when the value is higher than 12.5. Building on this scale and in comparison to previous studies in tabletop settings [63,64], the equity of participation in any of our conditions can be considered as mostly unequal, with only a few groups having accomplished tasks with some equity. This might possibly be explained through the difference in the task and the device. Indeed, both mentioned studies were using tabletops and a problem solving task requiring frequent system interactions to achieve a solution. In contrast, our task is based on data analysis requiring mainly information exchange and the construction of a common understanding. System interactions are less frequent and performed to explore the data or validate the final choices. For such tasks and settings, a more appropriate scale might therefore be needed.

To conclude, about RQ4, our observations conduct us to say that WA cues improve the perceived collaboration, in particular team orientation and coordination behaviour. Furthermore, the effectiveness of decision making was improved with WA cues as the chosen solution was always correct with WA cues, which is not the case with an audio-video feed only.

6.4. Limitations

Because of the complex setup of the study, there were some limitations which might have impacted the result of this study. The number of groups and participants is small and there might be differences because of the familiarity with each other, the personality of the group members and their English level. Furthermore, the tasks were slightly different with the second part having more constraints and more predominant roles involved. So, in the second part, participants might have felt more induced to contribute and talk. Also, the participants needed some time to become familiar with the system and the group, with some of them being more reserved at the beginning, and contributing more in the second part of the study. Therefore, the learning effect between the conditions was high and the results were varying considerably between the groups and tasks, making the comparison in the small sample size difficult.

While the use of a Kinect-based system and the reliance on (facial and gesture) recognition models enabled unobtrusive behaviour tracking capabilities, it should be noted that these technical choices also led to tracking inaccuracies. In particular, the (dis)agreement icons were more often subject to such problems, as they results from a combination of the Kinect-based body tracking system and an image-based hand gesture recognition model, which has to make decisions from sometimes subtle or (partially) occluded gestures. This, together with potential technical delays regarding the transmission of the tracking data, could have impacted the results.

Finally, some of the participants were already familiar with the WSDs used as part of the study, while others were experiencing such systems for the first time. Similarly, some of the participants knew other group members they were teamed up with, while others had to collaborate with persons they had not met before. This difference in familiarity with both the system and the users might have impacted the participants' perception of the overall system and the value of (some of) the WA cues, especially those related to the *Environment*.

7. Conclusions

In this paper, we have reported on the results from a user study with 24 participants, testing a mixed-presence collaboration scenario on two connected WSDs in two different conditions. Our aim was to explore how WA cues are understood and experienced by users and to find out if complementing conventional remote collaboration features with WA cues makes a difference to their user experience, awareness and collaborative decision-making.

With our results, we have shown that our complementary set of WA cues, in addition to an audio-video feed, is beneficial for mixed-presence collaboration with WSDs. The main results are:

- The WA cues enhance awareness of *pointing actions*, *environment* and *gaze*, as well as improve the user experience, team orientation and coordination, and lead to more correct decisions.
- The most useful WA cues are *touch feedback* and *mid-air pointing feedback*, followed by the two environment cues (*list of attendees* and *miniature views*) and *gaze icons*.
- Users have individual preferences and needs for making use of the WA cues.

Our results provide valuable information for designing mixed-presence collaboration systems for WSDs, highlighting the need to provide multiple WA cues in addition to an audio-video link. Those WA cues should primarily cover WA information on *environment*, *touch and mid-air pointing actions*, and *attention*. In addition, our results suggest that a system with WA cues might be most beneficial for enhancing intragroup collaboration and the efficiency of communication and decision-making, subject to being confirmed through larger studies, with other systems and different contexts.

Future research can build upon our results and conduct further design-oriented and empirical studies in light of optimizing the design of such WA cues, ultimately enhancing their clarity and visibility while being useful for collaborative work. By making use of our proposed WA components and elements, individual WA cues can be positioned and compared in order to understand their individual impact on awareness, workload, and collaboration. The adjusted framework itself will also provide opportunities to assess the current coverage of the landscape of WA cues and identify under-explored areas. The procedures and measurements from our user study offer a solid starting point for future studies on mixed-presence collaboration with wall-displays. Such studies could address the limitations of our study by, for instance, using a mixed-methods research approach or by investigating how the results of this paper apply to other types of tasks (e.g., collaborative design) and in the presence of larger groups or asymmetric setups.

Overall, with this work, we contribute to a next generation of mixed-presence decision-making WSDs, where users can collaborate smoothly, and enjoy a similar experience and quality as in collocated collaboration.

Author Contributions: Conceptualization, V.M.; methodology, V.M., A.C., L.S. and D.A.; software, A.C.; validation, V.M., A.C., L.S. and D.A.; formal analysis, V.M., A.C. and L.S.; investigation, V.M., A.C., L.S. and D.A.; resources, V.M., A.C., L.S. and D.A.; data curation, V.M., A.C., L.S. and D.A.; writing—original draft preparation, V.M., A.C., L.S. and D.A.; writing—review and editing, V.M., A.C., L.S. and D.A.; visualization, A.C. and LS; supervision, V.M.; project administration, V.M.; funding acquisition, V.M., A.C., L.S. and D.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Luxembourg National Research Fund (FNR) under the FNR CORE ReSurf project (Grant nr C21/IS/15883550).

Institutional Review Board Statement: The study was conducted in accordance with the fundamental ethical principles of the European Code of Conduct for Research Integrity and the Code of Ethics of the Luxembourg Institute of Science and Technology. The planned procedures were reviewed by LIST's Ethical Committee and by the Data Protection Officer to ensure GDPR compliance. The approval ID is 2021I00209, approval date is 23 March 2021.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data used are publicly available under the following links, accessed on 15 October 2024. Screenshots of the scenario used for the user study: https://zenodo.org/doi/10.5281/zenodo.12664295 Protocol of the user study and role cards used: https://zenodo.org/doi/10.5281/zenodo.12663836. Collaboration questionnaire and awareness questionnaire: https://zenodo.org/doi/10.5281/zenodo.12664007. Transcripts (+attributions): https://doi.org/10.5281/zenodo.13920252. Questionnaires answers and data processing: https://zenodo.org/doi/10.5281/zenodo.12664084.

Acknowledgments: The authors would like to thank all participants of the user study.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

UX	User experience
WA	Workspace awareness
WSD	Wall-sized display

References

1. Yuill, N.; Rogers, Y. Mechanisms for collaboration: A design and evaluation framework for multi-user interfaces. *ACM Trans.-Comput.-Hum. Interact. (TOCHI)* **2012**, *19*, 1–25. [CrossRef]
2. Belkacem, I.; Tominski, C.; Médoc, N.; Knudsen, S.; Dachselt, R.; Ghoniem, M. Interactive Visualization on Large High-Resolution Displays: A Survey. *Comput. Graph. Forum* **2022**, *43*, e15001.
3. Rajabiyazdi, F.; Walny, J.; Mah, C.; Brosz, J.; Carpendale, S. Understanding researchers' use of a large, high-resolution display across disciplines. In Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces, Funchal, Portugal, 15–18 November 2015; pp. 107–116. [CrossRef]
4. Simonsen, J.; Karasti, H.; Hertzum, M. Infrastructuring and participatory design: Exploring infrastructural inversion as analytic, empirical and generative. *Comput. Support. Coop. Work (CSCW)* **2020**, *29*, 115–151. [CrossRef]
5. Wall, J.; Bertoni, M. The Decision Arena: A model-centric interactive workspace for Product-Service System design. In Proceedings of the DS 101: Proceedings of NordDesign 2020, Lyngby, Denmark, 12–14 August 2020; pp. 1–10.
6. Kubicki, S.; Guerriero, A.; Schwartz, L.; Daher, E.; Idris, B. Assessment of synchronous interactive devices for BIM project coordination: Prospective ergonomics approach. *Autom. Constr.* **2019**, *101*, 160–178. [CrossRef]
7. Gutwin, C.; Greenberg, S. A descriptive framework of workspace awareness for real-time groupware. *Comput. Support. Coop. Work (CSCW)* **2002**, *11*, 411–446. [CrossRef]
8. Maipas, S.; Panayiotides, I.G.; Kavantzas, N. Remote-Working Carbon-Saving footprint: Could COVID-19 pandemic establish a new working model with positive environmental health implications? *Environ. Health Insights* **2021**, *15*, 11786302211013546. [CrossRef] [PubMed]
9. Orzeł, B.; Wolniak, R. Digitization in the design and construction industry—Remote work in the context of sustainability: A study from Poland. *Sustainability* **2022**, *14*, 1332. [CrossRef]
10. Fullwood, C.; Doherty-Sneddon, G. Effect of gazing at the camera during a video link on recall. *Appl. Ergon.* **2006**, *37*, 167–175. [CrossRef] [PubMed]
11. Langner, R.; Kister, U.; Dachselt, R. Multiple Coordinated Views at Large Displays for Multiple Users: Empirical Findings on User Behavior, Movements, and Distances. *IEEE Trans. Vis. Comput. Graph.* **2019**, *25*, 608–618. [CrossRef] [PubMed]
12. Maquil, V.; Anastasiou, D.; Afkari, H.; Coppens, A.; Hermen, J.; Schwartz, L. Establishing Awareness through Pointing Gestures during Collaborative Decision-Making in a Wall-Display Environment. In Proceedings of the Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023; pp. 1–7. [CrossRef]
13. Fussell, S.R.; Setlock, L.D.; Parker, E.M. Where do helpers look? Gaze targets during collaborative physical tasks. In Proceedings of the CHI'03 Extended Abstracts on Human Factors in Computing Systems, Ft. Lauderdale, FL, USA, 5–10 April 2003; pp. 768–769. [CrossRef]
14. Kunz, A.; Nescher, T.; Küchler, M. Collaboard: A novel interactive electronic whiteboard for remote collaboration with people on content. In Proceedings of the 2010 International Conference on Cyberworlds, Singapore, 20–22 October 2010; pp. 430–437. [CrossRef]
15. Yamashita, N.; Kaji, K.; Kuzuoka, H.; Hirata, K. Improving visibility of remote gestures in distributed tabletop collaboration. In Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, Hangzhou, China, 19–23 March 2011; pp. 95–104. [CrossRef]
16. Edelmann, J.; Mock, P.; Schilling, A.; Gerjets, P. Preserving non-verbal features of face-to-face communication for remote collaboration. In Proceedings of the Cooperative Design, Visualization, and Engineering: 10th International Conference, CDVE 2013, Alcudia, Mallorca, Spain, 22–25 September 2013; Proceedings 10; Springer: Berlin/Heidelberg, Germany, 2013; pp. 27–34. [CrossRef]
17. Bai, H.; Sasikumar, P.; Yang, J.; Billinghurst, M. A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–13. [CrossRef]
18. Dourish, P.; Bellotti, V. Awareness and coordination in shared workspaces. In Proceedings of the 1992 ACM Conference on Computer-Supported Cooperative Work, Toronto, ON, Canada, 1–4 November 1992; pp. 107–114. [CrossRef]
19. Endsley, M.R. Toward a theory of situation awareness in dynamic systems. *Hum. Factors* **1995**, *37*, 32–64. [CrossRef]

20. Carroll, J.M.; Rosson, M.B.; Convertino, G.; Ganoe, C.H. Awareness and teamwork in computer-supported collaborations. *Interact. Comput.* **2006**, *18*, 21–46. [CrossRef]
21. Clark, H.H.; Brennan, S.E. Grounding in communication. In *Perspectives on Socially Shared Cognition*; American Psychological Association: Washington, DC, USA, 1991; pp. 127–149. [CrossRef]
22. Osawa, N. Aggregate pointers to support large group collaboration using telepointers. In Proceedings of the CHI'06 Extended Abstracts on Human Factors in Computing Systems, Montréal, QC, Canada, 22–27 April 2006; pp. 1169–1174. [CrossRef]
23. Fussell, S.R.; Setlock, L.D.; Yang, J.; Ou, J.; Mauer, E.; Kramer, A.D. Gestures over video streams to support remote collaboration on physical tasks. *Hum.-Comput. Interact.* **2004**, *19*, 273–309. [CrossRef]
24. Huang, W.; Kim, S.; Billinghurst, M.; Alem, L. Sharing hand gesture and sketch cues in remote collaboration. *J. Vis. Commun. Image Represent.* **2019**, *58*, 428–438. [CrossRef]
25. Schwaller, M.; Lalanne, D. Pointing in the air: Measuring the effect of hand selection strategies on performance and effort. In Proceedings of the Human Factors in Computing and Informatics: First International Conference, SouthCHI 2013, Maribor, Slovenia, 1–3 July 2013; Proceedings; Springer: Berlin/Heidelberg, Germany, 2013; pp. 732–747. [CrossRef]
26. Piumsomboon, T.; Dey, A.; Ens, B.; Lee, G.; Billinghurst, M. The effects of sharing awareness cues in collaborative mixed reality. *Front. Robot. AI* **2019**, *6*, 5. [CrossRef]
27. Higuch, K.; Yonetani, R.; Sato, Y. Can eye help you? Effects of visualizing eye fixations on remote collaboration scenarios for physical tasks. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 2016; pp. 5180–5190. [CrossRef]
28. Jing, A.; May, K.; Matthews, B.; Lee, G.; Billinghurst, M. The impact of sharing gaze behaviours in collaborative mixed reality. *Proc. ACM-Hum.-Comput. Interact.* **2022**, *6*, 1–27. [CrossRef]
29. Albiz, J.; Viberg, O.; Matviienko, A. Guiding visual attention on 2d screens: Effects of gaze cues from avatars and humans. In Proceedings of the 2023 ACM Symposium on Spatial User Interaction, Sydney, NSW, Australia, 13–15 October 2023; pp. 1–9.
30. Tong, L.; Serna, A.; George, S.; Tabard, A. Supporting Decision-making Activities in Multi-Surface Learning Environments. In Proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017), Setúbal, Portugal, 21–23 April 2017; Volume 1, pp. 70–81. [CrossRef]
31. Mateescu, M.; Pimmer, C.; Zahn, C.; Klinkhammer, D.; Reiterer, H. Collaboration on large interactive displays: A systematic review. *Hum.-Comput. Interact.* **2021**, *36*, 243–277. [CrossRef]
32. Hornecker, E.; Marshall, P.; Dalton, N.S.; Rogers, Y. Collaboration and interference: Awareness with mice or touch input. In Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work, San Diego, CA, USA, 8–12 November 2008; pp. 167–176. [CrossRef]
33. Tong, L.; Tabard, A.; George, S.; Serna, A. Horizontal vs. Vertical: How the orientation of a large interactive surface impacts collaboration in multi-surface environments. In Proceedings of the Human-Computer Interaction—INTERACT 2017, Mumbai, India, 25–29 September 2017; Springer: Berlin/Heidelberg, Germany, 2017; pp. 202–222. [CrossRef]
34. Anastasiou, D.; Afkari, H.; Maquil, V. "You move THIS!": Annotation of Pointing Gestures on Tabletop Interfaces in Low Awareness Situations. In Proceedings of the LREC2020 Workshop "People in Language, Vision and the Mind" (ONION2020), Marseille, France, 11–16 May 2020; pp. 22–27.
35. Wisiecka, K.; Konishi, Y.; Krejtz, K.; Zolfaghari, M.; Kopainsky, B.; Krejtz, I.; Koike, H.; Fjeld, M. Supporting Complex Decision-Making. Evidence from an Eye Tracking Study on In-Person and Remote Collaboration. *ACM Trans.-Comput.-Hum. Interact.* **2023**, *30*, 1–27. [CrossRef]
36. Gong, J.; Sun, J.; Chu, M.; Wang, X.; Luo, M.; Lu, Y.; Zhang, L.; Wu, Y.; Wang, Q.; Liu, C. Side-by-Side vs. Face-to-Face: Evaluating Colocated Collaboration via a Transparent Wall-sized Display. *Proc. ACM-Hum.-Comput. Interact.* **2023**, *7*, 1–29. [CrossRef]
37. Stafford, A.; Piekarski, W.; Thomas, B.H. Implementation of god-like interaction techniques for supporting collaboration between outdoor AR and indoor tabletop users. In Proceedings of the 2006 IEEE/ACM International Symposium on Mixed and Augmented Reality, Santa Barbara, CA, USA, 22–25 October 2006; pp. 165–172. [CrossRef]
38. Pinelle, D.; Nacenta, M.; Gutwin, C.; Stach, T. The effects of co-present embodiments on awareness and collaboration in tabletop groupware. In Proceedings of the Graphics Interface 2008, Windsor, ON, Canada, 28–30 May 2008; pp. 1–8. [CrossRef]
39. Kuechler, M.; Kunz, A.M. Collaboard: A remote collaboration groupware device featuring an embodiment-enriched shared workspace. In Proceedings of the 2010 ACM International Conference on Supporting Group Work, Sanibel Island, FL, USA, 6–10 November 2010; pp. 211–214. [CrossRef]
40. Apperley, M.; McLeod, L.; Masoodian, M.; Paine, L.; Phillips, M.; Rogers, B.; Thomson, K. Use of Video Shadow for Small Group Interaction Awareness on a Large Interactive Display Surface. In *Proceedings of the AUIC '03*; Biddle, R., Thomas, B., Eds.; Australian Computer Society, Inc.: New Nork, NY, USA, 2003; Volume 18, pp. 81–90. [CrossRef]
41. Kim, S.; Lee, G.; Huang, W.; Kim, H.; Woo, W.; Billinghurst, M. Evaluating the combination of visual communication cues for HMD-based mixed reality remote collaboration. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; pp. 1–13. [CrossRef]
42. Avellino, I.; Fleury, C.; Beaudouin-Lafon, M. Accuracy of deictic gestures to support telepresence on wall-sized displays. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Republic of Korea, 18–23 April 2015; pp. 2393–2396.

43. Avellino, I.; Fleury, C.; Mackay, W.E.; Beaudouin-Lafon, M. Camray: Camera arrays support remote collaboration on wall-sized displays. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 6718–6729. [CrossRef]
44. Coppens, A.; Schwartz, L.; Maquil, V. Workspace Awareness Needs in Mixed-Presence Collaboration on Wall-Sized Displays. In Proceedings of the Cooperative Design, Visualization, and Engineering, Valencia, Spain, 15–18 September 2024; Luo, Y., Ed.; Springer Nature: Cham, Switzerland, 2024. [CrossRef]
45. Scapin, D.L.; Bastien, J.C. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behav. Inf. Technol.* **1997**, *16*, 220–231. [CrossRef]
46. Coppens, A.; Hermen, J.; Schwartz, L.; Moll, C.; Maquil, V. Supporting Mixed-Presence Awareness across Wall-Sized Displays Using a Tracking Pipeline based on Depth Cameras. *Proc. ACM Hum.-Comput. Interact.* **2024**, *8*, 1–32. [CrossRef]
47. Matsumoto, D.; Hwang, H.C. Cultural similarities and differences in emblematic gestures. *J. Nonverbal Behav.* **2013**, *37*, 1–27. [CrossRef]
48. Millikan, R.G. Language conventions made simple. *J. Philos.* **1998**, *95*, 161–180. [CrossRef]
49. Vandenabeele, L.; Afkari, H.; Hermen, J.; Deladiennée, L.; Moll, C.; Maquil, V. DeBORAh: A Web-Based Cross-Device Orchestration Layer. In Proceedings of the 2022 International Conference on Advanced Visual Interfaces, Rome, Italy, 6–10 June 2022; pp. 1–3. [CrossRef]
50. Johnson, D.W.; Johnson, R.T. Making cooperative learning work. *Theory Pract.* **1999**, *38*, 67–73. [CrossRef]
51. Schrepp, M.; Hinderks, A.; Thomaschewski, J. Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *Int. J. Interact. Multimed. Artif. Intell.* **2017**, *4*, 103–108. [CrossRef]
52. Hart, S.G. NASA-task load index (NASA-TLX); 20 years later. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2006**, *50*, 904–908. [CrossRef]
53. Radford, A.; Kim, J.W.; Xu, T.; Brockman, G.; McLeavey, C.; Sutskever, I. Robust Speech Recognition via Large-Scale Weak Supervision. *arXiv* **2022**, arXiv:2212.04356. [CrossRef]
54. Wittenburg, P.; Brugman, H.; Russel, A.; Klassmann, A.; Sloetjes, H. ELAN: A professional framework for multimodality research. In Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006), Genoa, Italy, 22–28 May 2006; pp. 1556–1559.
55. Schröder, J.H.; Schacht, D.; Peper, N.; Hamurculu, A.M.; Jetter, H.C. Collaborating across realities: Analytical lenses for understanding dyadic collaboration in transitional interfaces. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023; pp. 1–16.
56. Wallace, J.R.; Scott, S.D.; Stutz, T.; Enns, T.; Inkpen, K. Investigating teamwork and taskwork in single-and multi-display groupware systems. *Pers. Ubiquitous Comput.* **2009**, *13*, 569–581. [CrossRef]
57. Hornecker, E.; Marshall, P.; Rogers, Y. From entry to access: How shareability comes about. In Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces, Helsinki, Finland, 22–25 August 2007; pp. 328–342.
58. Scott, S.D.; Carpendale, M.S.T.; Inkpen, K. Territoriality in collaborative tabletop workspaces. In Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, Chicago, IL, USA, 6–10 November 2004; pp. 294–303.
59. Rogers, Y.; Lim, Y.k.; Hazlewood, W.R.; Marshall, P. Equal opportunities: Do shareable interfaces promote more group participation than single user displays? *Hum.-Comput. Interact.* **2009**, *24*, 79–116. [CrossRef]
60. Morris, M.R.; Cassanego, A.; Paepcke, A.; Winograd, T.; Piper, A.M.; Huang, A. Mediating group dynamics through tabletop interface design. *IEEE Comput. Graph. Appl.* **2006**, *26*, 65–73. [CrossRef]
61. MacMillan, J.; Paley, M.J.; Entin, E.B.; Entin, E.E. Questionnaires for distributed assessment of team mutual awareness. In *Handbook of Human Factors and Ergonomics Methods*; CRC Press: Boca Raton, FL, USA, 2004; pp. 510–520.
62. Hall, J.A.; Horgan, T.G.; Murphy, N.A. Nonverbal communication. *Annu. Rev. Psychol.* **2019**, *70*, 271–294. [CrossRef]
63. Fan, M.; Antle, A.N.; Neustaedter, C.; Wise, A.F. Exploring how a co-dependent tangible tool design supports collaboration in a tabletop activity. In Proceedings of the 2014 ACM International Conference on Supporting Group Work, Sanibel Island, FL USA, 9–12 November 2004; pp. 81–90. [CrossRef]
64. Maquil, V.; Afkari, H.; Arend, B.; Heuser, S.; Sunnen, P. Balancing Shareability and Positive Interdependence to Support Collaborative Problem-Solving on Interactive Tabletops. *Adv.-Hum.-Comput. Interact.* **2021**, *2021*, 6632420. [CrossRef]