The Effect of Collaboration Styles and View Independence on Video-mediated Remote Collaboration

Seungwon Kim [a], Mark Billinghurst [a], Gun Lee [a]

[a] The School of Information Technology and Mathematical Sciences, The University of South Australia, Australia

Postal address:

[a] School of Information Technology and Mathematical Sciences, University of South Australia, Mawson Lakes Campus, Mawson Lakes, SA 5095

Email address: Seungwon.Kim@unisa.edu.au, Mark.Billinghurst@unisa.edu.au, Gun.Lee@unisa.edu.au

Corresponding author: Gun Lee

Abstract

This paper investigates how different collaboration styles and view independence affect remote collaboration. Our remote collaboration system shares a live video of a local user's real-world task space with a remote user. The remote user can have an independent view or a dependent view of a shared real-world object manipulation task and can draw virtual annotations onto the real-world objects as a visual communication cue. With the system, we investigated two different collaboration styles; (1) remote expert collaboration where a remote user has the solution and gives instructions to a local partner and (2) mutual collaboration where neither user has a solution but both remote and local users share ideas and discuss ways to solve the real-world task. In the user study, the remote expert collaboration showed a number of benefits over the mutual collaboration. With the remote expert collaboration, participants had better communication from the remote user to the local user, more aligned focus between participants, and the remote participants' feeling of enjoyment and togetherness. However, the benefits were not always apparent at the local participants' end, especially with measures of enjoyment and togetherness. The independent view also had several benefits over the dependent view, such as allowing remote participants to freely navigate around the workspace while having a wider fully zoomed-out view. The benefits of the independent view were more prominent in the mutual collaboration than in the remote expert collaboration, especially in enabling the remote participants to see the workspace.

Keywords: Annotation, Augmented Reality, Awareness, Independent View, Remote Collaboration

1 Introduction

Videoconferencing is widely used by people to share a view of themselves and their environment with others (O'Hara et al. 2006; Brubaker et al. 2012). While sharing user's environment, researchers have studied adding virtual cues to the video link for remote collaboration, especially for communicating spatial information, such as pointing at, positioning, or orientating an object (Fussell et al. 2004; Kirk et al. 2007; Alem et al. 2011; Poelman et al. 2012).

Many of these previous studies have explored a remote expert collaboration style (a.k.a. asymmetric collaboration style) where a remote user had a solution for a shared task while the local user did not (e.g. Gauglitz et al. 2014; Fakourfar et al. 2016; Alem et al. 2011). In such tasks, the main concern is how to effectively transmit instructions from the remote user to the local user. Other researchers have focused on a mutual collaboration style (a.k.a. symmetric collaboration style) where both users do not have a solution but discuss together how to find a solution (Kim et al. 2014; Kim et al. 2015). Interestingly, studies exploring these two collaboration styles have found different results in comparing two visual communication cues, drawing annotations and pointer cues. Annotation cues showed better performance and were preferred over pointer cues in a remote expert collaboration (Fussell et al. 2004), but pointer cues were preferred and increased the feeling of being connected in a mutual collaboration (Kim et al. 2014). However, there have been few studies directly comparing these two styles of remote collaboration.

Providing an independent view to a remote user is another important research topic. If a remote user has an independent view, he or she can control his or her view regardless of where the local user is looking. However, if a remote user has a dependent view, he or she has the view controlled by the local user. Fussell et al. (2003) compared independent and dependent views in a remote expert collaboration style and found that the independent view was more effective for collaboration. However, there have been few studies exploring view independence in the mutual collaboration style. We expect that the effect of view independence might differ under different collaboration styles just as the effect of different visual communication cues did.

In this paper, we investigate how different collaboration styles and view independence affect remote collaboration where verbal and drawing annotation cues are shared over video conferencing. The paper makes the following contributions:

- We report on a user study comparing the two remote collaboration styles in video conferencing involving a real-world manipulation task, which is to the best of our knowledge the first study of its kind.
- The study is also the first that investigates the effect of independent view in the mutual collaboration, compared to prior work that involved only remote expert collaborations.
- Based on the study results, we suggest guidelines on how to improve conferencing interfaces for supporting better mutual collaboration

In the rest of the paper, we review prior work, describe the design and implementation details of our prototype system, explain details of the user study design, then followed by results and discuss on the implications. We conclude with summarising the paper and suggesting future research directions.

2 Related Work

In this section we describe the theoretical background and related previous works in remote collaboration. First, we review user awareness in remote collaboration, then previous research in video mediated collaboration with visual communication cues, and independent views.

2.1 Awareness in Remote Collaboration

Collaboration is the process of joint and interdependent activities to achieve a common goal (Hauber 2008). During the collaboration, collaborators make a joint effort to align and integrate their activities in a 'seamless' manner without interrupting each other (Schmidt 2002). In this process it is key to be aware of what is going on in the cooperative work setting and in understanding the collaborators' activities. This is generally referred to as 'awareness' (Dourish and Bellotti 1992), and many researchers use additional adjectives to specify their focus (Schmidt 2002), including 'general awareness' (Gaver 1991), 'collaboration awareness' (Lauwers and Lantz 1990), 'peripheral awareness' (Gaver 1992; Benford et al. 2001), 'background awareness' (Bly et al. 1993), 'reciprocal awareness' (Fish et al 1990) and 'workspace awareness' (Gutwin and Greenberg 1996).

In this paper, since the target collaboration includes sharing a real world workspace, we follow Gutwin and Greenberg (1996) who defined 'workspace awareness' as up-to-the-moment understanding of others' activities in a shared workspace. The workspace is the place where the collaborators' activities are ongoing (Snowdon and Munro 2001). Later, Gutwin and Greenberg (2002) categorized awareness information into three types: *who* (e.g. presence of others, identity and authorship), *what* (e.g. users' activities, intentions, and communication tools), and *where* (e.g. location of activities, gaze of users, view of users, and the reach of users). This was further studied by Antunes and Ferreira (2011) and Ferreira et al. (2011) with an additional emphasis on *how* collaborators interact in the workspace with the given interface.

Selecting a media technology for remote collaboration is significant as different media provide different communication channels and different levels of awareness (Olson and Olson 2000). The media selection should be primarily dependent on the type of task (Cockburn and Greenberg 1993), and it has been presumed that video conferencing should offer a 'rich' medium with audio and video channels (Schmidt 2002). Since this paper focuses on tasks that involve sharing the real world workspace with objects to manipulate, a video conference system is appropriate as it can visually support workspace awareness and show object manipulations in real time (Whittaker 1995).

However, in a video conference system, it is difficult to support the same level of workspace awareness as in face to face collaboration (Gutwin and Greenberg 1999; Schmidt 2002; Greenberg and Gutwin 2016), and several researchers studied better awareness with the video conference system.

Assigning a role (e.g., instructor or worker) is one way of increasing the workspace awareness in remote collaboration (Dourish and Bellotti 1992). An assigned role defines or limits a set of activities of a collaborator, hence the workspace awareness increases as the uncertainty of the collaborator's activities decreases. For example, the collaborators in a remote expert collaboration have explicit roles; the remote instructor mostly sends instruction which the local worker mostly receives. In this one directional communication, the remote instructor can expect that the local worker is in the mode of receiving instructions, so does not need to worry that the local worker conducts any other activities but following instruction. On the other hand, in mutual collaboration, no roles are assigned to remote and local users, so they both send and receive messages. In this bidirectional communication, the remote user cannot easily know that the local user is ready to receive a message or about to start sending a message.

Another example of increasing awareness is to provide a WYSIWIS (What You See Is What I See) user interface (Stefik et al. 1987) in which all collaborators have an identical view. With the interface, collaborators do not need to worry about what the other person might be working on out of the view, and increases the awareness in remote collaboration. For example a remote expert has a same view with a local worker, so the local worker does not worry about the remote expert conducting any activities outside of the view (Kim et al. 2014). However, one limitation with this system is that the remote user can only see where the local user is looking.

Another approach for increasing awareness is to improve the user interface for sharing information. For example, Ishii et al. (1992, 1994) developed ‘Clearboard’ which supported gaze awareness while users shared their workspace, so collaborators could better understand where their partner was looking. Similarly, Reichherzer et al. (2014) developed a system for sharing a panorama image, and including a context compass to show what the other user was looking at and to increase the shared engagement. Others added visual communication cues on the shared workspace view for better awareness. Kuzuoka et al. (1994), Sakata et al. (2003) and Kim et al. (2013) added pointers to share pointing information. Alem et al. (2011) and Kirk et al. (2007) added hand gesture cues on the video to share spatial information. Tang (1991) and Ishii (1992, 1994) developed video conferencing prototypes supporting drawing annotations as a visual cue. Rekimoto and Nagao (1995) developed a portable system that allowed the remote user to draw annotations on the shared environment. Most of these studies found that the shared visual information is useful for understanding messages (or instruction), which agrees with Gergle's study (2013) describing the benefits of sharing visual information for better awareness.

While previous examples focused on explicit representation of information with a pointer, annotations, or hand gestures, some others studied notification of the remote user activity. Gutwin et al. (2011) explored the effectiveness of synthesized dynamic audio notifications, finding that the audio presentations improved collaborator awareness when the user was not looking at the screen. Cidota et al. (2016) introduced an interface with audio or visual notifications indicating what action the remote user had taken, finding that visual notifications were preferred over audio or no notifications, regardless of the task difficulty level.

These studies clearly show that ‘awareness’ is an important factor in collaboration, and assigning roles to collaborators, providing view independency, and creating a good user interface can influence workspace awareness. In the next two sections, we review more interfaces related to our work with an additional visual communication cue, annotation, and supporting an independent view in video mediated remote collaboration.

2.2 Drawing Annotations in Remote Collaboration

Annotation is one of the mostly studied visual communication cues for presenting spatial information. For example, Fussell et al. (2004) and Kim et al. (2013) developed systems that support drawing annotations on a shared video view, and compared using annotations to a pointer cues. They found that annotations were more effective than using pointer cues for communicating spatial information.

However, their systems anchored annotations on the users' screen so they would lose their reference point if the camera viewpoint changed. To overcome this, Billinghurst et al. (2002) demonstrated a teleconferencing system with stabilized annotations created by using Augmented Reality (AR) marker based tracking. Their system calculated the relative position of annotations to the markers, so the annotations were stably displayed at the real world where they were drawn. Gauglitz et al. (2012, 2014) developed annotation stablization on a handheld display and showed that users preferred stablized annotations than unstablized annotations. Similarly, Kim et al. (2014) developed a system that stabilizes annotations in the real world using Simultaneous Localization and Mapping (SLAM). They compared using annotations to pointers in a mutual collaboration and found that a pointer interface was more effective and preferred than an annotation interface; remote participants felt that they more quickly and easily participated in collaboration with a pointer interface. Likewise, Lukosch et al. (2015) developed a system using SLAM to stabilize virtual annotations in the real world, and attach them to real objects. Finally, Adcock and Gunn (2015) developed a system using light projection for drawing annotations and compared stabilized and non-stabilized annotations. The participants in their study preferred the stabilized annotations than non-stabilized annotations.

With the stabilized annotations, Gauglitz et al. (2012, 2014) and Kim et al. (2014) used a manual freeze function to prevent annotations being anchored at a wrong place when a local user unexpectedly changes the viewpoint while the remote user is drawing. The manual freeze function allows a remote user to manually freeze the live video and draw on the still video frame, and then manually return back to the live video again.

Later, researchers (Kim et al. 2015; Fakourfar et al. 2016) introduced an auto freeze function. With the auto freeze function, the freezing and unfreezing interface was integrated with the drawing interaction so that the live video was automatically frozen when a remote user started drawing and was unfrozen when they stopped drawing. They compared auto and manual freeze conditions, and found that the auto freeze condition helped drawing annotations more quickly and was easier to use than the manual freeze condition. In this study, we adopt a stabilized annotation interface with the auto freeze method of Kim et al. (2015).

2.3 Independent View in Remote Collaboration

Providing an independent view for the remote user is another topic that has been actively studied. Kuzuoka et al. (1995) developed a system where a remote user could physically control the camera and have an independent view. Sakata et al. (2003) and Kurata et al. (2004) developed a remote controllable camera mounted on the local user’s shoulder. Gurevich et al. (2015) introduced a remote collaboration system which provides a remote controllable view to a remote user with a projection based drawing annotation tool. Fussell et al. (2003) compared two remote user interfaces having an independent view or dependent view (a.k.a. WYSIWIS interface) in a remote expert collaboration. In their study, the independent view was implemented using a scene camera showing the whole work environment. The dependent view was implemented using a camera mounted on the local user’s head which showed almost an identical view with what the local user saw. Their results showed that the independent view shared a wider area and provided better performance than the dependent view (the WYSIWIS interface). Finally, Tait and Billinghurst (2015) introduced a system that provided an independent view to a remote user while using a three-dimensional reconstruction of the local user’s environment. Their results also showed that increased view independency led to faster task completion time.

These results showed that better awareness with a WYSIWIS interface did not necessarily lead to better performance. This has been studied by Gutwin and Greenberg (1998) who reported the tradeoffs between the effectiveness of individual activities and the awareness of shared activities. The dependent WYSIWIS view restricts the remote user’s view to the same view as the local user, hence collaborators may have a better understanding of the shared activities they are looking at. However, collaboration does not only include understanding shared activities but also individual activities such as independently watching other area. The dependent view does not allow the remote users to freely change their view, and as a result, limits individual activities and reduces the chance of seeing other areas of the workspace. Individual activities require versatile and flexible interfaces for the task performance, but shared activities require a high level of awareness by restricting and aligning the user activities with less flexible interfaces.

Therefore, it is the best to have an interface which balances the effectiveness of individual activities with the awareness of shared activities. The independent view (i.e. the scene camera view) in Fussell's study (Fussell et al. 2003) is an example of the compromise between them. Their independent view had a wider view which allowed a remote user to see the whole workspace. The remote user could freely watch any part of the workspace while still being able to see what the local user was doing. This wide independent view that still includes the other user's view is called a "relaxed WYSIWIS" (Stefik et al. 1987; Beaudouin-Lafon and Karsenty 1992; Roseman and Greenberg 1996).

Overall, collaboration style, view independency, and user interface (including visual communication cues) are key factors influencing remote collaboration. Many researchers have studied the remote expert collaboration style, but there are few studies comparing remote expert and mutual collaborations. In this paper, we compare these two styles of collaboration while a remote user has independent or dependent views.

3 System Prototype

To investigate how remote collaboration is affected by collaboration style and view independence, we built a prototype video-conferencing system based on prior work (Kim et al. 2014; Kim et al. 2015). Using the prototype, a local user can share a live video of their task space with a remote user, and the remote user can draw annotations on the shared video with the auto freeze function. For the user study the two users were in the same room separated by a partition, so were able to talk with each other.

The style of collaboration mostly depends on the task and situation rather than a specific system feature, so we did not develop any additional feature for comparing the collaboration styles. However view independence is not achievable without the system supporting it. We added a feature to our prototype to provide either dependent or independent views to the remote user. In the following subsections we explain more details of the prototype, first for the setup with the dependent view and then with the independent view.

3.1 Supporting the Dependent View (WYSIWIS interface)

For the dependent view setup we adopted the interface from previous studies that supported stabilized annotations with an auto-freeze function (Kim et al. 2014; Kim et al. 2015). The local user wears a Vuzix Wrap 1200DX-VR head mounted display (HMD) with a Logitech C920 webcam attached to it (see Figure 1a). The live video from the webcam is displayed on the

HMD and also shared with a remote user with the resolution of 640 by 480 pixels at 24 frames per second (fps). The webcam and HMD are connected to a PC (NVIDIA GeForce GTX 670 GPU, 16 GB RAM, and 3.4 GHz Dual Core CPU) on which a robust monocular SLAM system is visually tracking the scene from the live video feed (Tan et al. 2013) and shows virtual annotations stabilized on the real-world scene. SLAM provides robust and reliable tracking even when parts of the scene are occluded by the users' moving hands. The algorithm is highly reliable (95% success ratio with 1,000 iterations) and runs at 25 fps (Tan et al. 2013).

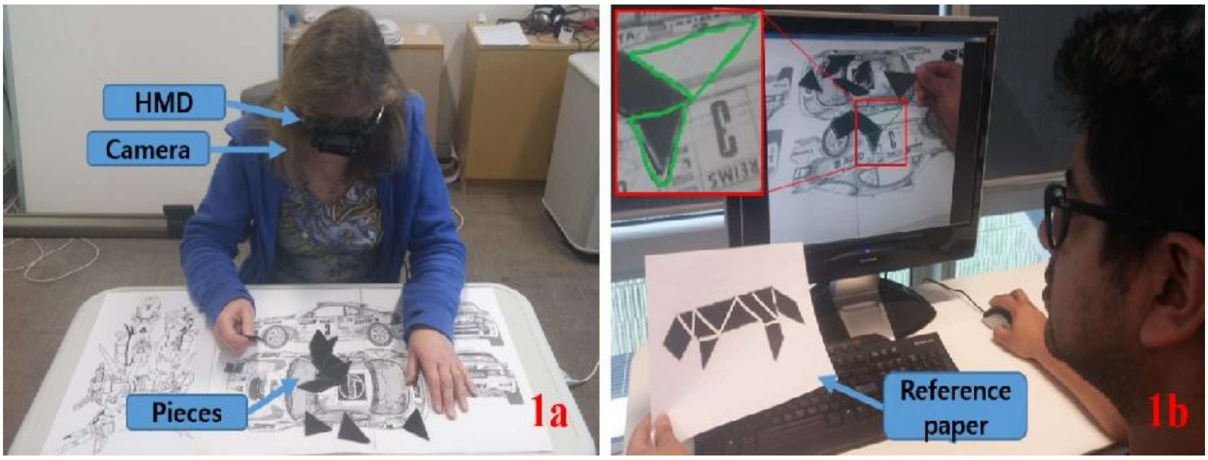

Figure 1: System setup for the remote expert collaboration with a dependent view. A remote user draws annotations on the shared view with a mouse while holding a reference instruction picture (1b), and a local user wears an HMD with a camera and shares the view and annotations from a remote partner (1a).

The remote user watches the same live video view of the local user, but displayed on a desktop monitor (see Figure 1b). The remote user can draw annotations on the shared view with mouse input by pressing the left mouse button and dragging, and the annotations are immediately shown to both users. To stabilize virtual annotations in the real-world, the system calculates the position of the mouse cursor projected into the real world using a ray casting method and draws lines between the sequential points on trajectory of the projected mouse cursor. To calculate the projected mouse cursor points, the system creates an invisible ray in the direction of the mouse cursor, and the three feature points from SLAM tracking that are closest to the ray are selected to create a plane. The collision point between the ray and the plane is derived to form an annotation in the real-world space.

3.2 Supporting the Independent View (navigable scene camera view)

To provide an independent view to a remote user, we used the same scene camera set up as Fussell's study (2003), placing a camera for the remote user on a tripod (see Figure 2a). The scene camera was positioned beside the local user to show the workspace as if the remote user was sitting beside the local user. Like the dependent view set up, the local user wears a video see-through HMD with an attached camera and can watch the workspace. The remote user can draw annotations on the independent view with the same method using in the dependent view. The virtual annotations were simultaneously shown in both users' views by sharing the SLAM tracking data.

Figure 2: System setup for the mutual collaboration with the independent view. A local user wears an HMD to see the workspace and remote partner's annotations while assembling a model. A reference paper is placed at top right corner of the workspace, and another camera for remote partner's independent view is placed beside the local user (2a). A remote user watches the workspace with the independent view while drawing annotation with the mouse interaction (2b).

While our system does not allow control of the physical position of the scene camera, we use a software based method to support a navigable independent view. Our system captures a view of the whole workspace and crops it according to digital zoom and pan applied by remote user's input. The scene camera is placed at a position where it can capture the whole workspace (a desk, 93cm by 50cm) as a basic live video feed, which is about seventy centimeters away from the center of the workspace.

To decide the panning and zooming range, we measured how users watched the workspace while wearing an HMD and performing an experimental task, solving a customized Tangram[1] puzzle (see section 4.1). We recruited 12 participants and measured the distance between the HMD camera and a 3D point in the workspace where the center of the view was pointing at.

[1] https://en.wikipedia.org/wiki/Tangram

To calculate the distance, the system first calculated the position of the camera and the 3D point in the workspace using SLAM and the method described earlier which is used for drawing annotations. The measured distance in the SLAM coordinate system was then converted into metric measures. The distance was measured every half a second and users spent 90 seconds on average to complete the task.

Figure 3 shows the results of how far away the participants watched the workspace from. Through this test, we collected 2117 samples and found that the average distance between the camera and the workspace was 38.69 centimeters, with a range of 18 to 60 centimeters.

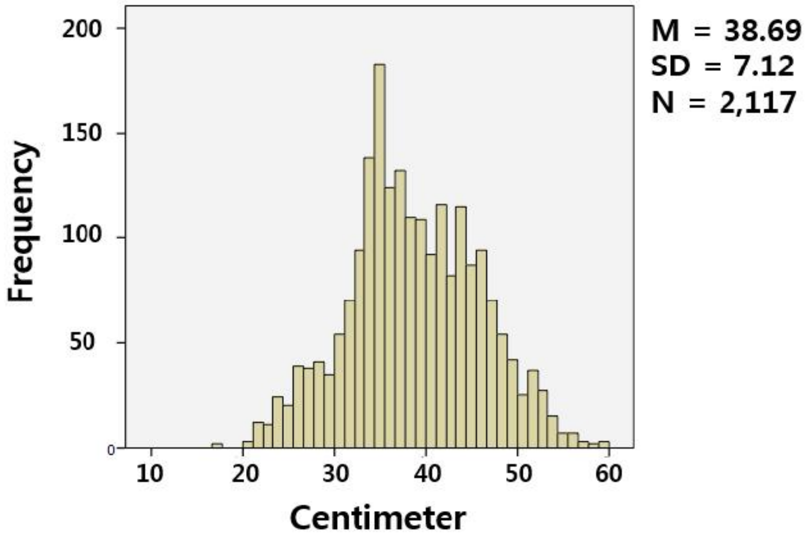

Figure 3: The distance between the camera and the workspace (*M* = Mean, *SD* = Standard Deviation, *N* = number of samples)

Based on this result, we decided the range of the digital zoom levels from the base video feed which showed the whole workspace from 70 centimeters away. The fully zoomed out view was set to be equivalent to watching the workspace from sixty centimeters away, while the fully zoomed in view was set to be equivalent to watching it from twenty centimeters away. There were 60 zoom levels that users can choose from. The remote users used mouse and keyboard interactions for controlling digital zooming and panning (see Table 1). Pressing the ‘+’ or ‘-’ key on the keyboard increased or decreased the zoom by one level, and holding the keys down continuously increases or decreases the zoom level. Users could also use the mouse scroll wheel to control the zoom level. The default zoom level was set to a view from 38 centimeters away, which was the average distance measured.

Table 1 : Interaction mapping for navigation in an independent view

	Panning				**Zoom**	
	Left	Down	Right	Up	In	Out
Keyboard	‘a’	‘s’	‘d’	‘w’	‘+’	‘-’
Mouse	Dragging. (The panning follows the mouse cursor movement while the mouse right button is kept pressed.)				Scroll down	Scroll up

For panning, the system calculates the position of the remote user’s independent view in the base video feed. The system calculates the four corners of the independent view according to the current panning and zoom level, and then displays the area inside the four corners. The ‘a’, ‘s’, ‘d’, and ‘w’ keys on the keyboard are mapped for panning the view left, down, right, and up, respectively (see Table 1). Users can press two keys simultaneously to pan diagonally. Alternatively, users can drag (move while pressing the right button) the mouse for panning, and the direction of panning is exactly the opposite of where the cursor is dragged to. The initial position of the independent view is the center of the base video feed.

Additionally, the system draws a brown tetragonal border on both the remote and local views, representing the area of the partner’s view (see Figure 4). The four 2D corner points in the partner’s screen are projected into the workspace to calculate the corresponding 3D corner points of the tetragon using the same method used for drawing annotations.

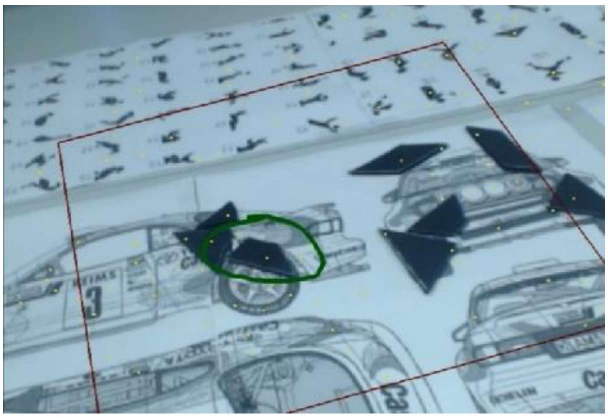

Figure 4: A brown tetragon in the local user’s view representing the area of the remote user’s view.

4 User Study Design

To investigate how different collaboration styles and view independency affect remote collaboration, we designed and conducted a user study with the prototype system. In this section, we describe the design of the user study and how we conducted it.

4.1 Experimental Task and Environment

In preparing the experimental task, we followed a prior work (Kim et al. 2014) and used Tangram puzzles. Tangrams have a set of flat puzzle pieces in various shapes, which the player needs to assemble the pieces together (without overlaying on top of each other) to form a target shape. Tangrams were chosen as the experimental task for three reasons. First, assembling a Tangram includes object manipulation so it encourages remote users to draw annotations to share spatial information. Second, local participants can mainly focus on communication or collaboration as Tangram pieces are light and small, so it is easy to hold and manipulate them. Third, solving a Tangram can be done on a desk sized workspace which does not require large viewpoint changes, possibly causing dizziness in the remote collaborator (Tsubaki et al. 2005;). On the other hand, the task space is big enough to investigate the difference between dependent and independent views as it still requires a moderate amount of view change and camera movement.

We modified the Tangram puzzle to reduce the effect of participants' previous experience. The modified Tangram included ten puzzle pieces that had different sizes or shapes compared to those in a conventional Tangram. We prepared five Tangram puzzles (one for the face-to-face training session and four for the experiment trial sessions) and we adjusted and balanced the level of difficulty by modifying the reference papers. To prevent participants from getting stuck in mutual collaboration tasks, the level of difficulty was controlled by providing hints on a reference paper, marking six border lines between the pieces. We conducted a pilot test with five people who did not participate in the main study, and each Tangram puzzle was solved in about two minutes with the provided hints on the reference papers.

For remote expert collaboration, we also used the same Tangram but prepared different reference papers that included the assembly solutions by marking all the required border lines between pieces. Since the collaboration in the remote expert style is providing spatial information for piece assembly from a remote user to a local user and each puzzle includes the same number of pieces, the assembling task had a very similar level of difficulty between the puzzles. To ensure that the remote user knows the solution in the remote expert collaboration, we let the remote user keep the reference paper with all border lines and let him or her solve the puzzle alone before performing the experimental task. The order of experiment conditions was counter balanced using a balanced Latin square design.

In the experiment, each participant sat in front of a desk, and the local participant wore an HMD while the remote participant used a desktop computer. The desk for the local participant was prepared as the workspace with puzzle pieces which were captured and shared to the remote participant through a live video. The local and remote participants were separated with a partition so were not able to see each other directly, yet they were still able to talk to each other.

4.2 Experimental Conditions

The study used a factorial design with two independent variables: *view independency* (V) and *collaboration style* (C). The *view independency* (V) factor had two levels, dependent (V_D) or independent (V_I) views, and the *collaboration style* (C) was either remote expert collaboration (C_R) or mutual collaboration (C_M) (see Table 2). Combining the two factors with two levels, there were four conditions:

(1) C_RV_D: Remote expert collaboration with Dependent view

(2) C_RV_I: Remote expert collaboration with Independent view

(3) C_MV_D: Mutual collaboration with Dependent view

(4) C_MV_I: Mutual collaboration with Independent view

Table 2 : Two independent variables (view types and collaboration style) forming four experiment conditions highlighted in gray background.

	Remote Expert Collaboration (C_R)	**Mutual Collaboration (C_M)**
Independent View (V_I)	Remote expert collaboration with Independent view (C_RV_I)	Mutual collaboration with Independent view (C_MV_I)
Dependent View (V_D)	Remote expert collaboration with Dependent view (C_RV_D)	Mutual collaboration with Dependent view (C_MV_D)

Local participants always saw a view captured from the camera attached to the HMD they were wearing, while remote participants saw the same view as their local partner in the condition with V_D, or had an independent view from the scene camera placed on a tripod in front of the workspace in the condition with V_I. In the condition with C_R, remote participants were provided with a reference paper that had all the border lines between the puzzles pieces marked as a solution. In conditions with C_M, the reference paper including only partial solutions (six border lines marked for balancing the level of difficulty) was placed at the top right corner of the shared workspace (i.e. on the local participant's desk), so both remote and local participants could take information from it.

4.3 Experimental Procedure

Twenty-four participants were recruited in pairs (i.e. 12 pairs), 20 males and 4 females with ages ranging from 20 to 35 years old (*M* = 27.6; *SD* = 4.4). They all had prior experience of using videoconferencing regularly (more than once a month), and each pair knew each other as either friends or colleagues (eleven pairs) or family (one pair). We recruited participants in pairs who already knew each other, and they chose their partner prior to the experiment.

The experiment started with an introduction explaining the user study, and the participants answered a questionnaire about demographic information. Then the experiment continued with a face-to-face training session where participants were asked to solve a sample Tangram puzzle. This ensured that the participants understood the task well. After the training session, the researcher explained the four conditions and how the prototype system worked.

After explaining the system, we randomly assigned the local or remote user role to each participant which was kept the same throughout the experiment, then they tried four puzzle solving sessions using the prototype system under different conditions. Each session consisted of two minutes of training to get used to the given interfaces and five minutes of experimental task. After each session, participants answered a questionnaire asking their experience with the given condition. After all four experimental sessions, users had final interviews, asking about their experience with four experimental conditions. To reduce the impact of mistakenly dictation and hearing the other user's response if they verbally answered the interview questions, we asked them to write their answers on the back of the questionnaire. The experiment took about 90 minutes for each pair of participants.

4.4 Experiment Measurements

To assess collaboration, Gutwin and Greenberg (1999) suggested three types of experimental measurements: product, process, and satisfaction. Product measures are mostly for assessing collaboration outcomes in terms of quality (e.g. accuracy) or efficiency (e.g. task completion time). Process measures mostly assess patterns of collaboration and user communication, and are often obtained by system log data, observation, and video/audio analysis. Satisfaction measures focus on assessing the participants' subjective opinions on the quality of their collaboration, collected through interviews and questionnaires.

In this study, we collected data from all three types of measurements, including task performance (task completion time), questionnaire, interview, system log, video/audio screen recordings, and observation. We note that the task performance was only measured in the conditions with the remote expert collaboration, because it is not relevant in the mutual collaboration case as there are too many uncontrolled variables (e.g. participant's ability to solve the puzzle, or even taking chance to find a solution) that could affect the task performance.

The questionnaire (see table 3) was based on prior work by Kim (2014) and Gutwin (1999), and included questions asking about three types of awareness information: who (presence of others – Question 7 asking user feeling of togetherness with the collaborating partner), what (users' activities – Question 2 asking the effectiveness in sending messages and Question3 asking the effectiveness in receiving messages), and where (location of activities – Question 5 asking whether seeing work space properly and Question 6 asking the level of having a same focus with a partner). Moreover, the questionnaire included questions about overall collaboration (Question 1 asking the level of enjoyment, Question 4 asking the level of mental stress in communication with partner, and Question 8 asking whether collaborating well or not) The questions were answered by rating on a Likert-scale ranging from 0 (Strongly disagree) to 10 (Strongly agree). After answering the questions, each participant answered one open question asking what he or she liked or disliked about the condition.

Table 3 : Questions in the questionnaire after each session (Q5, highlighted in grey, was only for the remote user).

Q1	I enjoyed assembling a model.
Q2	I was able to express my idea properly.
Q3	I easily understood what my partner tried to do.
Q4	The communication with my partner was mentally stressful.
Q5	I was able to see the work environment properly.
Q6	We focused on the same piece or area.
Q7	I felt we were together. (staying in the same environment)
Q8	I felt we collaborated well.

After the participant completed all four experimental sessions, they ranked the two view types based on their preference under each collaboration style. We did not ask participants' their preference between the two collaboration styles as this may depend on the task or the situation rather than being a matter of choice. At the end of the experiment, we asked what difference they found between the conditions in a final interview.

In addition to subjective feedback, we also collected objective measures through analysing system logs which recorded the remote participants’ activities with the mouse and keyboard. The data included the amount of mouse movement for drawing annotations and the view navigation information in the conditions using independent view. We also observed the pattern of collaboration from the video and audio recordings.

5 Experiment Results

In this section, we describe the results in three aspects: how users communicated, how they watched the workspace, and the overall collaboration experience. We separately explain the effect of each independent variable, collaboration styles and view types, and separately analysed the remote and local participant’s data because their experience would be different in the experiment as they were using different user interfaces.

Overall, both remote and local participants felt that there was better communication from the remote end to local end in the C_R conditions than in the C_M conditions as they had the same focus on a piece or an area in the C_R conditions than in the C_M conditions. With these benefits of the C_R conditions, remote participants enjoyed the collaboration and had stronger feeling of togetherness, but local participants had conflicting opinions in terms of enjoyment and togetherness. Moreover, providing an independent view to remote participants helped the remote participant have better communication and better see the workspace so they had more enjoyment and an increased feeling of togetherness. However the results from the local participants were not the same, even though they still preferred the conditions with the independent view.

5.1 Ratings and Interview Results

We first analysed internal consistency of all rating items with the scale reversed for Q4. The rating questions answered by the remote participants had a very high level of consistency, with Cronbach’s α value of 0.902. The Cronbach’s α with one deleted item ranged from 0.873 to 0.910. The rating questions answered by local participants also had a high level of consistency with the Cronbach’s α value of 0.839, and the range of it with one deleted item was from 0.798

to 0.852. To analyse the Likert-scale rating results, we used a repeated measures ANOVA[2] (α = .05) with Aligned Rank Transform (ART)[3] as proposed by Wobbrock et al. (2011). The ARTool (http://depts.washington.edu/madlab/proj/art/) software was used to apply ART to the raw data, then the aligned data was analysed using ANOVA with the IBM SPSS Statistics[4].

To visualize data, we present box plots with the interquartile range[5] (at Figure 5, 6, 7, 8, 10, and 11). The box shows the interquartile range divided into two quartiles at the median (50^{th} percentile). The upper half of the box (from median to the third quartile) is green, and the lower half (from the first quartile to median) is brown or orange. Tukey's fences were used to identify the outlier by setting the fence range as below:

Lower Fence = the first quartile – (1.5 × interquartile range)

Upper Fence = the third quartile + (1.5 × interquartile range)

Despite some of the ratings being identified as statistical outliers based on their Tukey fence score, we included them in the analysis to capture a holistic view. Since participants gave rating scales for a series of questions without any significant environmental change, and there was no participant giving ratings identified as outliers consecutively in the series of questions, we assume that the outliers were not originated from erroneous nor inattentive responses. In addition, given that the measures were subjective ratings in ordinal scale, the method we used for analysing the results tends to be robust to these outliers by transforming the ordinal measure to meet the requirements of parametric analysis.

---

[2] Analysis of variance (ANOVA) is a statistical method to determine whether the means of groups are different using F-statistics. F-statistics is based on the ratio of mean squares, estimate of population variance that accounts for the degrees of freedom (DF) used to calculate that estimate. By calculating the ratio of variation between sample means and variation within the samples by F-statistics, ANOVA checks whether means are different or not.

[3] ART was introduced to analyse nonparametric data from a multiple-factors experiment without complex process. It is a preprocessing step to transform and align nonparametric data before analysing it using repeated measures ANOVA

[4] www.ibm.com/analytics/au/en/technology/spss/

[5] The interquartile range is middle 50 percent among the provided series of data. It usually means the range between lower quartiles (25^{th} percentiles) and upper quartiles (75^{th} percentiles).

5.1.1 Communication

Figure 5 shows the results of the Likert scale rating questions that relate to the participants' communication (Q2, Q3, and Q4 in Table 3).

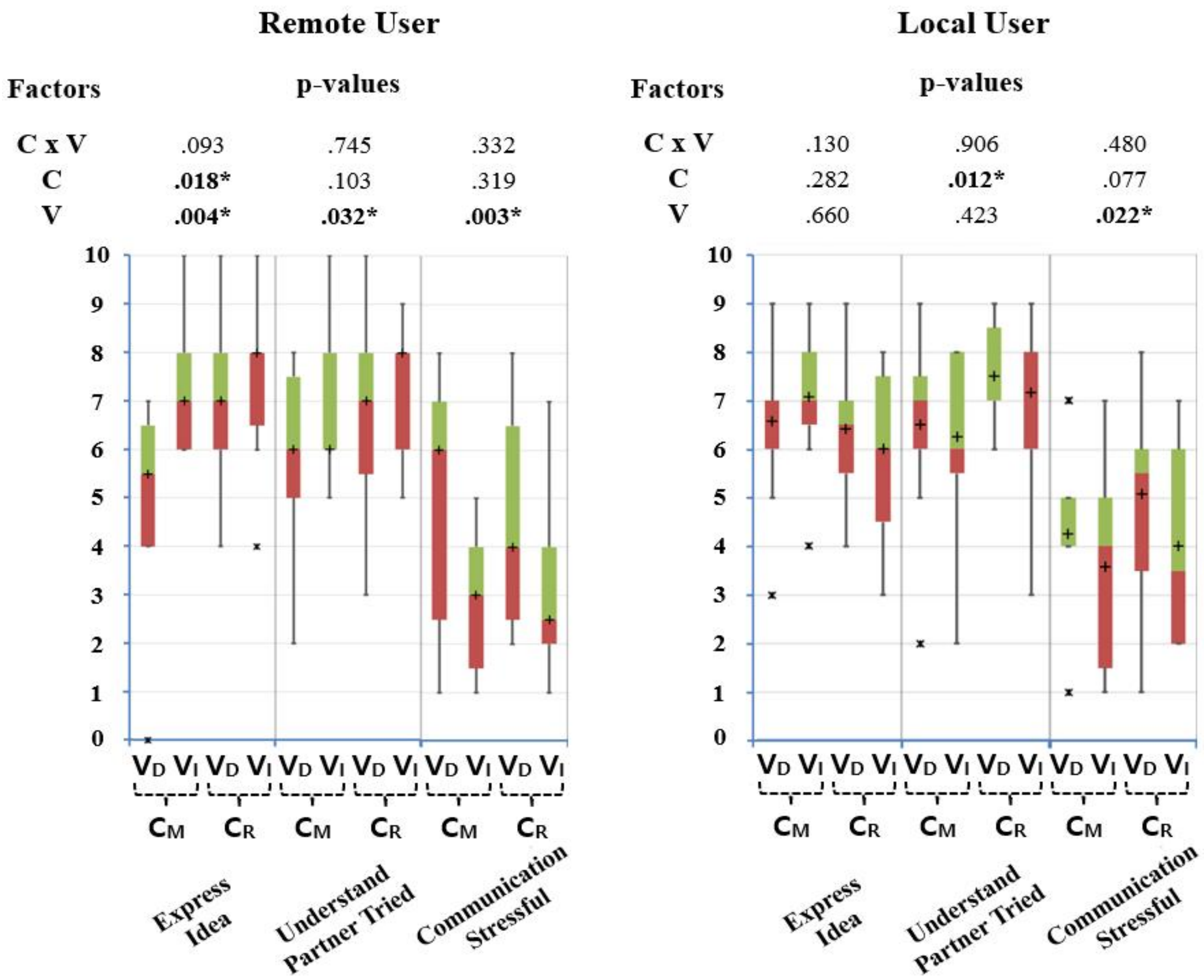

Figure 5: Results of Likert scale ratings from the remote (left) and local (right) participants for Q2: Express idea, Q3: Understand partner, and Q4: Communication was stressful (0: strongly disagree ~ 10: strongly agree; *: statistically significant difference, +: mean, x: outlier).

5.1.1.1 Effect of Collaboration Styles

Our results showed that assigning a role to participants may have an effect on communication. From the rating results, the C_R condition assigning remote expert and local worker roles to remote and local participants showed significantly better communication from remote participants to local participants compared to the C_M condition where no role was assigned. Remote participants felt that they expressed their ideas significantly better in the C_R conditions than in the C_M conditions ($F(1,11)=7.781$, $p=.018$), and local participants better rated their understanding of remote participants' expression in C_R conditions than in C_M conditions

($F(1,11)= 9.061$, $p=.012$). Correspondingly, several participants commented that the communication in the C_R conditions was fluent and active because it was simply providing and following instructions. For example, R5 (remote participant in group 5) felt that he could express his instructions well because his partner followed well, and R1 and R2 felt that they confidently gave instructions in the C_R conditions. L1, L5, and L11 (local participant in group 1, 5, and 11) commented that their partners were clear, and the communication was easy in the C_R conditions.

For the communication from local participants to remote participants, we did not find a significant difference in rating scales between the C_R and C_M conditions (remote participants understand local partner: $F(1,11)=.223$, $p=.103$, local participants express idea: $F(1,11)=1.282$, $p=.282$). Moreover, they did not feel significant difference in the aspect of communication stressful between the C_R and C_M conditions (remote participants: $F(1,11)=1.091$, $p=.319$, local participants: $F(1,11)=3.801$, $p=.077$).

5.1.1.2 Effect of View Type

Providing an independent view to remote participants showed a significant main effect on the remote participant's communication. The remote participants rated the V_I conditions significantly higher than the V_D conditions in terms of expressing ideas and understanding what their partners tried ($F(1,11)=13.626$, $p=.004$ and $F(1,11)=6.011$, $p=.032$ respectively). In the interview, remote participants mentioned that the independent view provided a stable and large (wide angle) view while the dependent view was moving and showing a smaller area. The local participants' ratings were not significantly affected by the remote partner's view type in expressing ideas ($F(1,11)=0.205$, $p=.660$) and understanding their partners ($F(1,11)=1.377$, $p=.423$).

Interestingly, both local and remote participants felt that the communication was significantly more stressful in the V_D conditions than in the V_I conditions (remote participants: $F(1,11)=14.175$, $p=.003$; local participants: $F(1,11)=7.126$, $p=.022$). In the interview, remote participants (R5, R8, and R11) mentioned that unexpected changes of viewpoint in the V_D conditions was annoying, and asking local partners to move their viewpoint was uncomfortable. Local participants (L3, L7, and L12) reported that remote partners asking where to look at was annoying.

5.1.1.3 Summary

To summarise, communication from remote participants to local participants was perceived to be more efficient in the C_R conditions than in the C_M conditions. Moreover, providing an independent view to remote participants helped them to have better communication while both the remote and the local participants suffered from moving and showing a small area with the dependent view. Furthermore, asking and being asked to change the viewpoint was annoying to both remote and local participants. No significant interaction between collaboration styles and view types were found (remote participant - express idea: $F(1,11)=3.382$, $p=.093$), understand partner: $F(1,11)=.111$, $p=.745$), communication stressful: $F(1,11)=1.028$, $p=.332$); local participant - express idea - $F(1,11)=2.672$, $p=.130$), understand partner: $F(1,11)=.015$, $p=.906$), communication stressful: $F(1,11)=.535$, $p=.480$)).

5.1.2 Watching the Workspace

Figure 6 shows the results of the Likert scale rating questions (Q5 and Q6 in Table 3) that relate to how participants watched the workspace and their partners' actions.

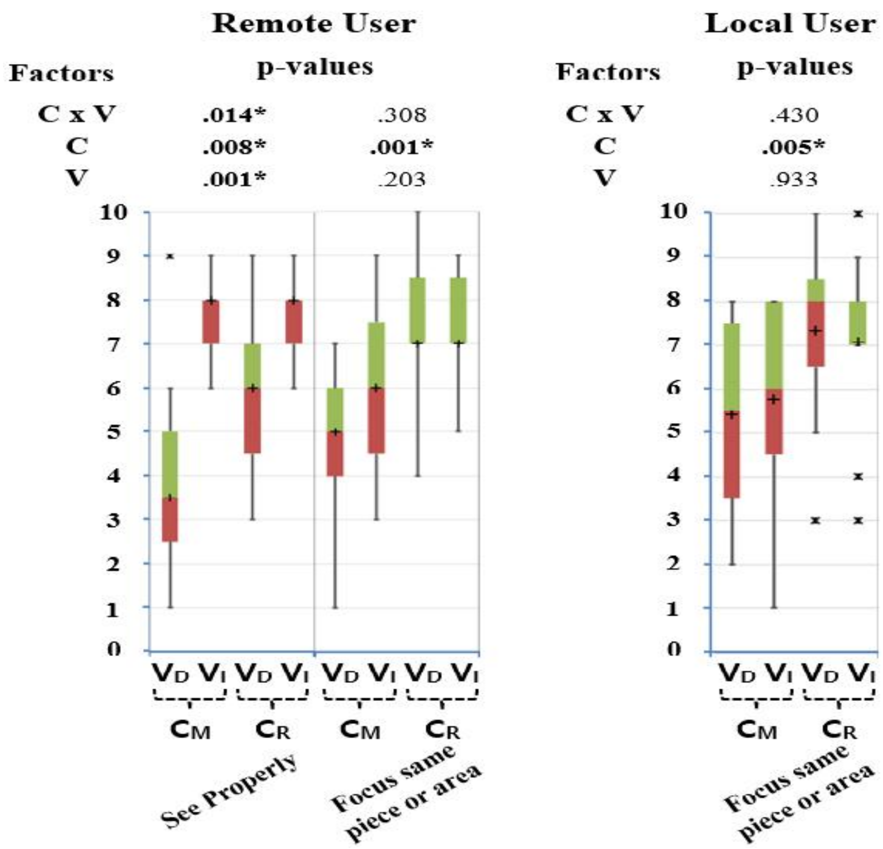

Figure 6: Results of Likert scale ratings from the remote (left) and local (right) participants for Q5: See properly, and Q6: Focus on the same piece or area (0: strongly disagree ~ 10: strongly agree; *: statistically significant difference, +: mean, x: outlier).

5.1.2.1 Effect of Collaboration Style

The results showed that assigning a role to participants may have an effect on watching the workspace on both the remote and local ends. Remote participants felt that they saw the workspace significantly better ($F(1,11)=10.226$, $p=.008$) in the C_R conditions than in the C_M conditions. Additionally, both local and remote participants rated the C_R conditions higher than the C_M conditions in terms of focusing on the same puzzle piece or area with their partners (remote participants: $F(1,11)=21.764$, $p=.001$, local participants: ($F(1,11)=12.645$, $p=.005$). From the interview, participants (R1, R4, R10, R12, L2 and L3) commented that in the C_R conditions remote participants mostly instructed what to do and the local participants only followed the instruction, so they had same viewpoint and focus. However, in the C_M conditions, they sometimes have a different focus, for example L3 reported that they (R3 and L3) manipulated pieces more independently (compared to the trials in the C_R conditions).

5.1.2.2 Effect of View Type

Providing an independent view to the remote participants had a significant main effect on the remote participants' seeing the workspace. The remote participants felt that they saw the workspace significantly better in V_I conditions than in V_D ($F(1,11)=80.084$, $p<.001$). In the interviews, we found relevant comments from the participants. Two thirds of the remote participants (R1, R2, R3, R4, R6, R7, R9, and R12) commented that they were able to see the workspace better with the fully zoomed out view in the V_I conditions as they could see more, and keep watching an interesting piece or area regardless of the local participant's view change. Additionally, this advantage of the V_I conditions was more prominent in the C_M condition compared to C_R ($F(1,11)=8.574$, $p=.014$).

However, both local and remote participants were not significantly affected by the view types in terms of focusing on the same piece or area with their partner (remote participant: $F(1,11)=1.829$, $p=.203$; local participants: $F(1,11)=.007$, $p=.933$).

5.1.2.3 Summary

In summary, both local and remote participants felt more focusing on the same piece or area in the C_R conditions than in the C_M conditions. Interestingly, both remote and local participants did not feel a significant difference between the view types in focusing on the same piece or area. In terms of seeing the workspace properly, remote participants rated the C_R conditions higher than the C_M conditions, and the V_I conditions higher than the V_D conditions.

Additionally, the advantage of the V_I view compared to the V_D view for seeing the workspace properly was more prominent in the C_M conditions than in the C_R conditions.

5.1.3 Overall Collaboration

Figure 7 shows the results of the Likert scale rating questions (Q1, Q7, and Q8 in Table 3) that relate to how participants enjoyed assembling the model, felt they were together, and collaborated with each other.

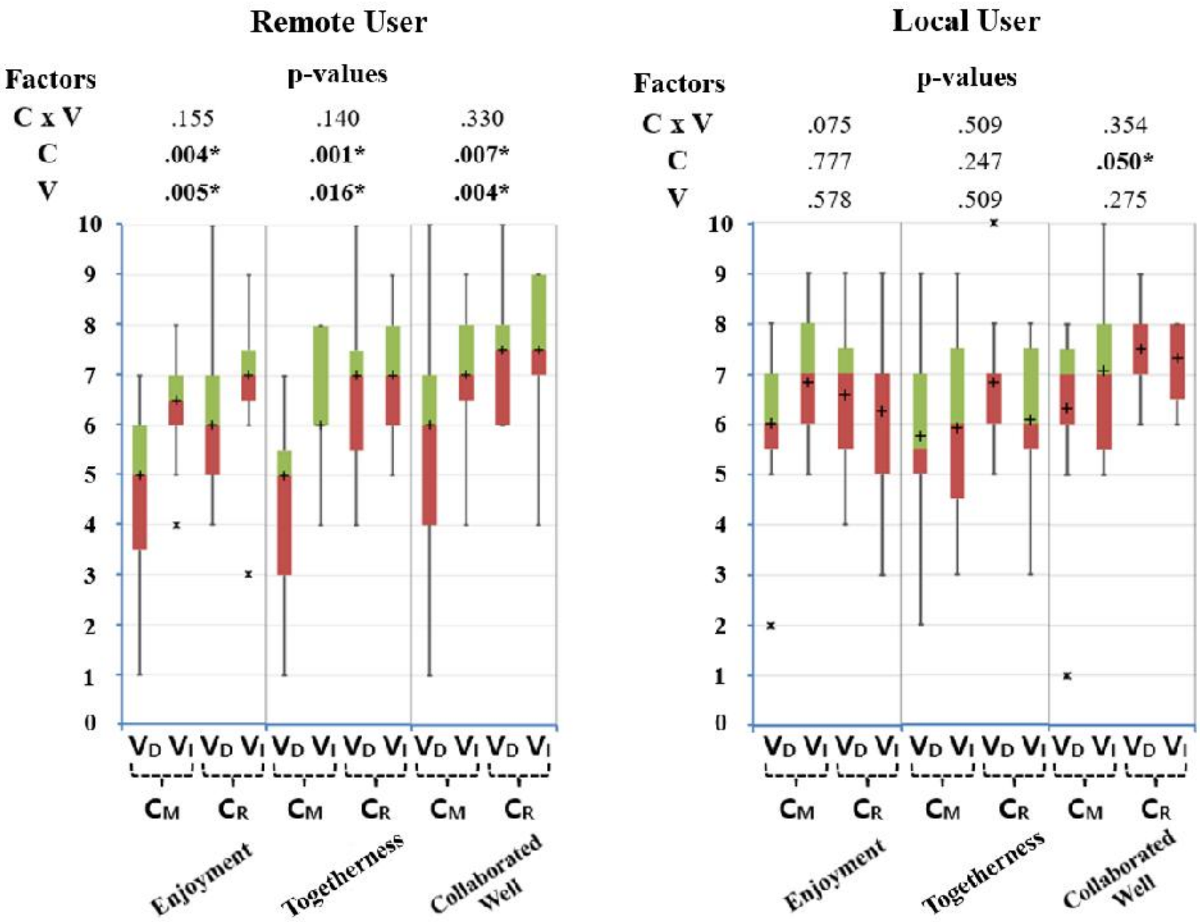

Figure 7: Results of Likert scale ratings from the remote (left) and local (right) participants for Q1: Enjoyment, Q7: Togetherness, and Q8: Collaborated well (0: strongly disagree ~ 10: strongly agree; *: statistically significant difference, +: mean, x: outlier).

5.1.3.1 Effect of Collaboration Style

Remote participants felt more enjoyment ($F(1,11)=12.787$, $p=.004$), more togetherness with local partners ($F(1,11)=18.817$, $P=.001$), and better collaboration ($F(1,11)=10.786$, $p=.007$) in the C_R conditions than in the C_M conditions. However, local participant did not feel significant effect between the collaboration styles in terms of ‘enjoyment’ ($F(1,11)=.084$, $p=.777$) and ‘togetherness’ ($F(1,11)=1.495$, $p=.247$). They only rated ‘collaborating well’ significantly higher in the C_R conditions than in the C_M conditions.

In the interview, some remote participants (R2, R3, R10, and R11) mentioned that they enjoyed the C_R conditions more than the C_M conditions because they could actively participate in the collaboration. Local participants had split opinions on which collaboration style they enjoyed more. Some of them (L1 and L3) reported they enjoyed C_R more because it was simple and easy. On the contrary, some others (L2, L4, and L12) reported that they enjoyed C_M because they could contribute ideas rather than only following the remote participant's instruction. We also found comments about 'togetherness'. Remote participants (R1, R2, R3, and R7) mentioned that their activities were aligned well with their partner's activities in the C_R conditions because of the clear roles as an instructor and a worker which led to a higher feeling of togetherness. Local participants had split opinions. Some of them (L1 and L7) felt togetherness because they worked as remote participants' hands following their instructions, but some others (L3 and L10) reported that they were merely following instructions like a mindless robot in the C_R conditions while being more equal as they also could suggest ideas in the C_M conditions.

5.1.3.2 Effect of View Type

Between the two view types, remote participants felt a significant difference in enjoyment, togetherness, and collaborating well. They enjoyed more ($F(1,11)=12.343$, $p=.005$), felt more togetherness ($F(1,11)=8.015$, $P=.016$) and felt better collaboration ($F(1,11)=12.733$, $p=.004$) in the V_I conditions than the V_D conditions. However, local participants did not feel significant effect between the view types in the aspects (enjoyment: $F(1,11)=.328$, $p=.578$; felt togetherness: $F(1,11)=1.495$, $p=.247$; felt collaborating well: $F(1,11)=1.320$, $p=.275$).

In the interview, half of the remote participants (R1, R2, R4, R8, R10, R11) mentioned that they enjoyed the V_I conditions because it was like having their own space and freedom from local partners. Moreover, R1 and R12 reported that they enjoyed navigating the workspace. Interestingly, some remote participants (R4 and R12) mentioned they felt more together with the brown tetragon in the V_I conditions that showed where the local partner was watching. In contrast, none of the remote participants commented about togetherness in the V_D conditions.

5.1.3.3 Summary

To summarise, remote participants felt more enjoyment, more together with local partners, and better collaboration in the C_R conditions than in the C_M conditions, and in the V_I conditions than in the V_D conditions. Local participants had conflicting opinions on the collaboration style in terms of enjoyment and togetherness, and did not feel significantly different between the V_I conditions and the V_D conditions in terms of enjoyment and togetherness. There was no significant interaction between the collaboration style and view type (remote participant - enjoyment: $F(1,11)=2.021$, $p=.155$), togetherness: $F(1,11)=2.525$, $p=.140$), well collaboration: $F(1,11)=1.040$, $p=.330$); local participant - enjoyment: $F(1,11)=3.865$, $p=.075$), togetherness: $F(1,11)=1.033$, $p=.331$), well collaboration: $F(1,11)=.935$, $p=.354$)).

5.2 Observation and Log data

In this section, we describe the results of log data which recorded the remote participants' activities and observations from the video recordings. Unfortunately, we accidently forgot to record videos of group 1's C_MV_D condition and group 6's C_MV_I condition, so they are not included in the results. In the video analysis, the communication between remote and local participants was transcribed and coded, and we found that any spatial information in the communication was for piece selection, position, or/and orientation.

5.2.1 Communication

We firstly describe the different communication patterns between the C_R and C_M conditions. In the C_R conditions remote and local participants played roles of an expert and a worker, and the communication was mostly one way from the remote participants to the local participants. Remote participants provided all required spatial information to assemble a target model and the local participants rarely suggested any ideas. In the C_M conditions, the communication was bidirectional as both remote and local participants contributed their ideas to assemble the target model. From the video analyses, we found that on average local participants suggested 68.45 percent of the ideas (SD = 11.43%) in the C_M conditions.

This difference between the C_R and C_M conditions might influence the use of graphical annotation between the remote and local participants. The system log data revealed that the remote participants drew annotations significantly more in the C_R conditions than in the C_M conditions ($F(1,11)=32.728$, $p<.001$), and had significantly more mouse drawing strokes in the C_R conditions than in the C_M conditions ($F(1,11)=8.731$, $p=.013$) (see Figure 8). However, between two collaboration styles, there was no significant difference in the drawing amount per stroke (calculated as dividing all mouse drawing amount by the number of drawing strokes, $F(1,11)$ =.613, p==.450).

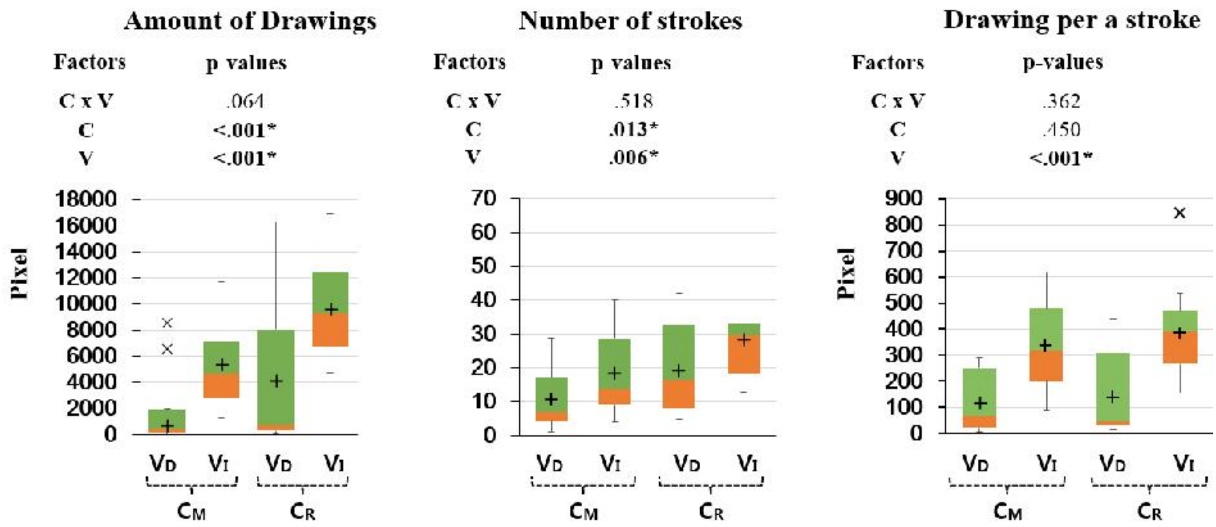

Figure 8. The amount of drawings, number of strokes, and the amount of drawing per a stroke (*: statistically significant difference, +: mean, x: outlier)

In the rest of this section, we describe the participants' communication under each collaboration style. We report on the remote expert collaboration and the mutual collaboration separately as participants showed different communication patterns according to the collaboration style.

5.2.1.1 Communication in the Remote Expert Collaboration (in C_R conditions)

First, we analysed recorded videos of the C_R conditions to check how the participants exchanged spatial information. As mentioned, the remote participants provided all required information for every piece manipulation such as piece selection, position, and orientation. Annotation and verbal cues were the main two communication channels and there were two communication patterns depending on the use of these cues and the changing viewpoint (see Figure 9). In the first pattern, the remote participant used both annotation and verbal cues when they indicated one specific piece by drawing a circle or a tick mark with a verbal demonstrative pronoun such as 'this' and 'it'. Then they changed their viewpoint to where the piece should be placed to describe the position and/or orientation by drawing annotations. In the second pattern, they only verbally described a shape and/or a size of a piece (such as 'a small triangle' or 'a square') for identification without any annotation while watching where that piece should go. Then at the same time, they described the position and orientation by drawing annotations without changing their viewpoint.

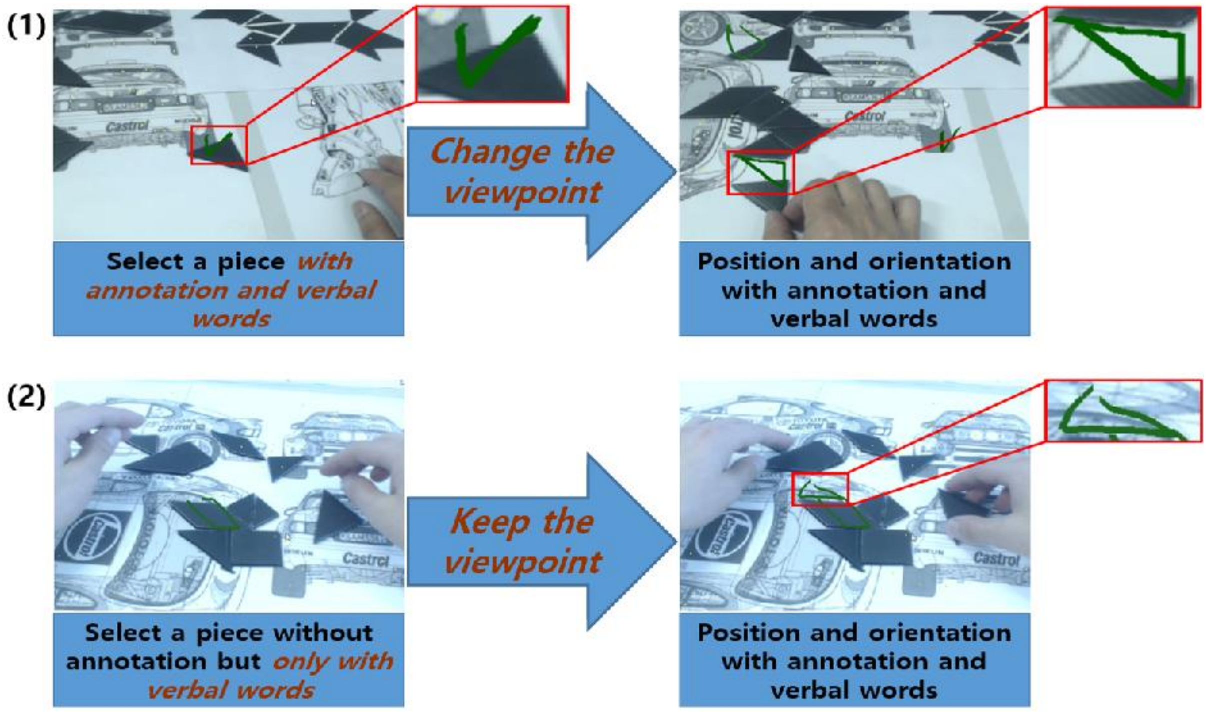

Figure 9: Two communication patterns used by the remote participants in remote expert collaboration conditions (C_R)

These two patterns were found in both the dependent view (first pattern: M = 59.81%, SD = 33.69%; second pattern: M = 40.19%, SD = 33.69%) and the independent view (first pattern: M = 61.89%, SD = 36.70%; second pattern: M = 38.11%, SD = 36.70%) conditions, and there was no significant difference between the dependent and independent views in the paired t-test (α = .05, $t(11)=0.384$, p=.708).

Regardless of the patterns, the position and orientation information was mostly described by drawing annotations while the verbal description mostly used deictic phrases to describe the user's action such as 'put it here' or 'leave it like this'. Describing the position and orientation by annotations was predominant in both V_D (M=97.5%, SD= 4.52%) and V_I (M=96.74%, SD=4.81%) conditions and there was no significant difference between the two views found by the paired t-test ($t(11)= -0.399$, p=.697). The remote participants rarely described the position and orientation using only verbal communication.

The local participants' feedback in the C_R conditions was mostly showing their activities through the shared live video, and sometimes verbally asking for confirmation from the remote partners.

5.2.1.2 Communication in Mutual Collaboration (in C_M conditions)

Unlike the C_R conditions, both participants contributed their ideas on how to solve the shared task in the C_M conditions. The local participants provided 62.68 (*SD*=11.83%) and 75.5 percent (*SD*=7.06%) of the piece selection in V_I and V_D conditions, respectively. Note, we excluded the result from L2 as an outlier as he mentioned that he just followed his remote partner in C_M condition, in the same manner as he did previously in two C_R conditions. The percentage of local participants leading the selection showed a significant difference ($t(9)= -2.439$, $p=.037$) between the V_I and V_D conditions as local participants selected a piece more often when the remote partners had a dependent view (in the V_D conditions) than when they had an independent view (in the V_I conditions).

After initial positioning, participants sometimes had a discussion, and the verbal communication in the discussion was more descriptive compared to those held in the C_R conditions, as they explained why they thought the move was right or wrong. However, from observation and video analysis we found that three remote participants (R1, R3, and R11) were sometimes not included in the C_M collaboration. This happened when local participants quickly tried several actions (e.g. manipulating puzzle pieces) in a sequence without any verbal description, and their remote partners merely watched without giving any feedback. When local participants verbally described their actions, remote participants actively provided feedbacks and ideas.

5.2.1.3 Summary

In summary, the remote participants provided all required information to local partners for the assembly task in the C_R conditions. There were two communication patterns for instructing piece selection in the C_R conditions, and the remote participants mostly drew annotations describing the position and orientation of a piece. In the C_M conditions, local participants led more than half of the piece selection task, especially in the V_D conditions compared to the V_I conditions. Remote participants drew more and made more mouse drawing strokes under the C_R conditions than in the C_M conditions.

5.2.2 View Coordination

In this section, we describe the benefit of the independent view and how the participants used and managed the views.

5.2.2.1 The benefits of Independent View

In the V_D conditions, the local participants usually decided the shared viewpoint, so sometimes there was a conflict between the local and remote participants on what they wanted to see. The remote participants sometimes asked the local partners to move the viewpoint to where they wanted it. With the V_I conditions, remote participants freely navigated around the workspace and did not bother local partners by asking them to look at a certain part.

In the V_I conditions the remote participants could freely draw annotations in any part of the workspace regardless of the local user's view. It might influence the use of annotations (see Figure 8 in the section 5.2.1). Remote participants drew annotations significantly more in the V_I conditions than in the V_D conditions ($F(1,11)=24.135$, $p<.001$), and had significantly more mouse strokes (for drawing) in the V_I conditions than in the V_D conditions ($F(1,11)=11.271$, $p=.006$). They also drew significantly more, in pixels, per stroke with the V_I condition than with the V_D condition ($F(1,11)=25.381$, $p<.001$).

With the V_I conditions, all remote participants, except one, fully zoomed out to see the larger area at the early stage of a session, and mostly stayed in that view for the rest of the time. The one exception was R4 who actively zoomed in or out for the first half of the sessions (in both C_M and C_R) but zoomed out (close to full zoom out) for the second half of the sessions. An interesting observation with the fully zoomed out view was that remote participants sometimes noticed pieces outside of the local partners' view, and helped them to find the piece.

However, the independent view in the V_I conditions sometimes caused a misalignment of focus between the local and remote participants. They sometimes focused on different areas or pieces and could not understand each other. For example, Local participants (L4, L5, and L8) sometimes did not know where their remote partners drew annotations because the annotations were drawn outside of their views.

5.2.2.2 Amount of Navigation in Independent View

Since the V_D conditions did not include a navigation function, we describe the navigation interaction only with the V_I conditions. To calculate the amount of navigation in the V_I conditions, we measured how much the view moved in pixels on the remote participant's screen. Since remote participants mostly had a fully zoomed out view, there was little navigation (see Figure 10), but we still found some interesting results. Participants navigated significantly more in the C_M conditions than in the C_R conditions ($F(1,11)=24.799$, $p<.001$). They used the mouse for navigation significantly more than the keyboard ($F(1,11)=62.748$, $p<.001$) as they always placed their dominant hand on the mouse. Moreover, the difference between mouse and keyboard navigation in the mutual collaboration was significantly greater than that in the remote expert collaboration ($F(1,11)= 27.107$, $p<.001$).

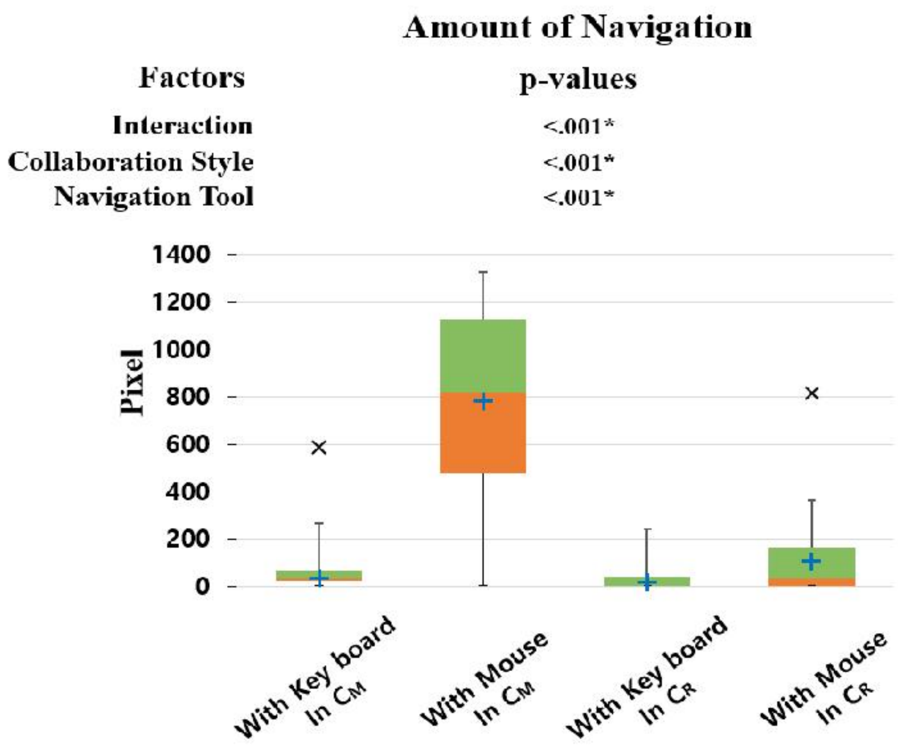

Figure 10: The average amount of navigation by a keyboard and a mouse (C_M: mutual collaboration, C_R: remote expert collaboration)

5.2.2.3 Summary

To summarise, remote participants were able to freely navigate the workspace in the V_I conditions while keeping the fully zoomed-out view. The V_I resolved the issue of V_D where local participants controlled the shared viewpoint so that the remote participants sometimes could not see what they wanted to see. This advantage of the independent view influenced the annotation use as remote participants drew more. Moreover, remote participants navigated significantly more in the C_M than in the C_R conditions. However, with the V_I conditions, local participants sometimes had a problem of not seeing the annotations when remote participants drew outside of their view.

5.3User Performance and Preference

The task performance (task completion time) in the C_M conditions is highly dependent on various factors (such as participants previous experience and ability), so we did not measure the performance in the C_M conditions, but participants mostly completed the task within five minutes. We did compare the task completion time of the C_R conditions between the two different view types: dependent view and independent view (see Figure 11) using a paired *t*-test (α = .05), and found that participants completed the task significantly faster with an independent view than with a dependent view ($t(11)=3.041$, $p=.011$).

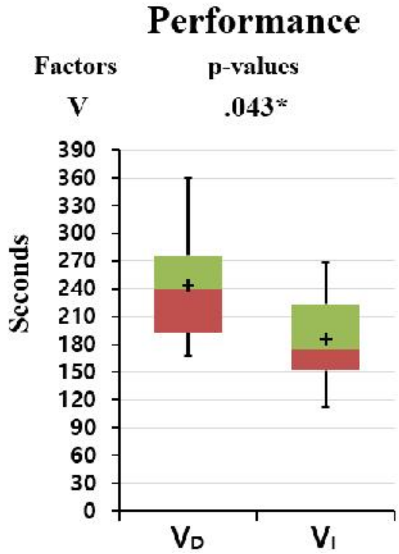

Figure 11: User performance with the remote expert collaboration (V: view type, V_I: independent view, V_D: dependent view)

We asked participants their preference between the dependent and independent views for each collaboration style (see Table 4). In the C_R conditions, eleven remote participants (92%) preferred the independent view while only one preferred the dependent view (8%). Half of the local participants (50%) preferred the independent view while the other half (50%) preferred the dependent view. For the C_M conditions, ten remote participants (83%) preferred the independent view while two (17%) preferred the dependent view. Eleven local participants (92%) preferred the independent view while one (8%) preferred the dependent view.

Table 4 : Participants' preference for the remote expert collaboration and mutual collaboration

	Remote Expert Collaboration		**Mutual Collaboration**	
	Remote Participants	**Local Participants**	**Remote Participants**	**Local Participants**
Independent View	11	6	10	11
Dependent View	1	6	2	1

A Wilcoxon Signed-Rank Test (α = .05) was used to analyse their preferences. For the remote expert collaboration, the remote participants significantly preferred the independent view over the dependent view (Z=-2.887, p=.004), while no significant difference was found in the preference of local participants (Z=0, p=1.0). For the mutual collaboration, the independent view was significantly more preferred than the dependent view by both remote (Z=-2.309, p=.021) and local participants (Z=-2.887, p=.004).

6 Discussion

We investigated two collaboration styles and two view types in the remote collaboration while verbal and annotation cues were available. Here we discuss the results of the user study.

Overall, the effect from collaboration styles and view independency was more obvious on the remote end than the local end. For collaboration styles, the data from remote participants was consistent while local participants had conflicting answers and were less consistent. For view independency, as the interfaces of independent and dependent views were on remote end, the effect was more significant on remote participants than on local participants.

6.1 Effect of Collaboration Style

The conditions with remote expert collaboration (C_R) showed several benefits over the conditions with mutual collaboration (C_M). First, participants felt that communication from the remote participant to local participant was more effective in the C_R than in the C_M conditions. Second, remote participants more properly saw the workspace and were more focused on the same piece or area with local partners in the C_R than in the C_M conditions. Third, remote participants felt more enjoyment, more togetherness with local partners, and better collaboration in the C_R than in the C_M conditions. However, local participants had conflicting opinions in terms of enjoyment and togetherness.

6.1.1 The Benefits of Remote Expert Collaboration

The benefits of the C_R conditions may come from defining or limiting user activities by assigning roles. In the C_R conditions, both remote and local participants mostly focused on a single one-way communication activity as remote participants sent information while local partners received the information. However, in the C_M conditions, both participants had mutual communication involving three activities: sending and receiving activities, and finding a task solution. As a result, participants simply focused on one type of activity in the C_R conditions, but they conducted three activities and switched among them in the C_M conditions. For this switching process, both remote and local participants should have known when and how they switch to have a collaboration. This difference between C_R and C_M conditions may have affected our results.

This argument agrees with Dourish and Bellotti (1992) who suggested that assigning a clear role improves the user awareness. Since there were fewer user activities in the C_R conditions than the C_M conditions, participants could be more easily aware of their partner's activity. For example, since remote participants mostly sent information in the C_R conditions, the local participants were ready to receive the information. However, in the C_M conditions, a collaboration partner was engaged in more activities and so had a lower level of awareness about what activity the collaborator will do, and they may not have been ready to properly react to the other user's activity.

6.1.2 Effect of Collaboration Style in Seeing the Workspace

One of the interesting results is that 'seeing the workspace properly' depends not only on view independency but also on the collaboration style. To discuss it, we firstly discussed the benefit of independent view for 'seeing the workspace properly'. Similar to our results of the V_I conditions, Gurevich et al. (2015) and Tait (2015) both reported that providing an independent view to a remote user helps them to freely see where he/she wanted to. This is the main benefit of the independent view, and if a collaboration style supports freely looking around, it would be possible to have the same benefit, seeing the workspace properly.

In the C_R conditions, remote participants would mostly decide where they looked because only they knew the solution. To allow the remote user to see what he/she wants to see, there are two main considerations. First, reduce the number of times that the local user decides where the remote user sees. Second, let the remote user decide what he/she sees. In the C_R conditions, the local participants would not even know what they should show, so should follow the instructions of remote participants in choosing a viewpoint. However, in the C_M conditions, both local and remote participants could suggest ideas to solve the solution, and local participants could take more control of where to look compared to the C_R conditions. In short, there is more possibility of remote participants choosing what and where to see in the C_R conditions than in the C_M conditions, and so 'seeing the workspace properly' depends not only on view independency but also on the collaboration style.

6.1.3 Local Participants' Enjoyment and Togetherness in Remote Expert Collaboration

Even though local participants felt that they collaborated better in the C_R conditions than in the C_M conditions, they had conflicting opinions on which condition they felt more enjoyment in and togetherness in. Some local participants felt more enjoyment and togetherness in the C_R condition which provided better receiving messages, but others felt less because they just had to follow the instructions without contributing their own ideas. This difference might have come from the different preference of the local participants between the performance of collaboration (such as understanding messages well) and contributing ideas.

6.1.4 Insufficiency of Mutual Collaboration

In the C_M conditions, some local participants quickly tried several piece manipulations in a row without giving any verbal description, and their remote partners merely watched without giving any feedback. This may have come from poor management between individual and shared activities. The local participants' action of piece manipulation was a shared activity as it was directly shown in the shared video and transferred spatial information to the remote partners. However, this shared activity was performed as an individual activity as the manipulation happened quickly (without providing enough time for the remote partners to understand them) and mostly alone without verbal communication (which could be used for providing better understanding). To solve this, we argue that the interface should encourage users to be more descriptive and collaborative in the shared activities in order to enable users to feel more together.

The distance between the participants and the task space also affected the collaboration. The local participants physically stayed in the task space so they could directly manipulate task objects. However, remote participants had to understand the task space through a medium (the shared video), and always had to delegate the manipulation of task objects to the local partners. This might not significantly affect remote expert collaboration because the local participants could not perform the task alone without the remote participants' instruction. However, in mutual collaboration, no one had the solution, and the local partners were able to directly try out their own ideas. As a result, in the C_M conditions, local participants led 68.5 percent of the piece selection (62.7 and 75.5 percent in V_D and V_I conditions, respectively) rather than having a balanced collaboration. This suggests that to have better balanced collaboration between local and remote collaborators in mutual collaboration, the interface should encourage the remote collaborator to be more active and the local collaborator to be better aware of the presence of a remote partner.

6.2 Effect of View Type

The independent view (V_I conditions) showed several benefits over the dependent view (V_D conditions). The independent view resolved the issue of the dependent view where the local participants dominated the navigation of the shared viewpoint and caused an unstable moving view. The remote participants freely navigated their view around the workspace, and had a wider fully zoomed-out view. This helped remote participants to feel that they could better communicate, have more enjoyment and be more together with local partners, and have better collaboration in the V_I conditions than in the V_D conditions. Interestingly, the effect of the independent view was more prominent in the C_M conditions than in the C_R conditions. For remote participants, the independent view helped 'seeing the workspace properly' more prominently in the C_M conditions than in the C_R conditions. Local participants preferred the independent view over the dependent view in the C_M conditions but not in the C_R conditions.

6.2.1 Advantage of Independent View

Our independent view implementation supported a 'relaxed WYSIWIS' interface as the remote participants fully zoomed out their views. According to Stefik et al. (1987) and Roseman and Greenberg (1996), this may help remote participants perform individual activities (such as watching other areas rather than merely watching the same area with partners) while keeping a certain level of awareness of their partner's activities as it still showed the area where the other participant was watching. However, interestingly, our V_I conditions did not only keep the level of awareness on par with V_D, but made further improvements as the remote participants felt better understanding of local partners (awareness) with the V_I conditions than the V_D conditions. This may be because of the explicit representation (the brown tetragon) of the region local partners were looking at. In addition, we also mention that stability of the scene camera. The dependent view was from the head mounted camera on local participants' head, and the dependent view might be affected by the head movement and less stable compared to the independent view from the steady scene camera. It may be a factor influencing awareness because the remote participants would better understand the workspace with the stable view.

With our implementation of the independent view, remote participants could have supernatural abilities. The local participants were required to stay within a comfortable distance to use their hands for manipulating task objects, and their pose was constrained by the table workspace and chair where they sat on. However, the remote participants used a digital medium to collaborate, which could remove physical constraints and provide abilities such as watching the workspace from further away but still being able to draw on it.

6.2.2 Different impact of Independent View in Difference Collaboration Styles

Tait and Billinghurst (2015) and Gurevich et al. (2015) compared independent and dependent views under remote expert collaboration, and their results agreed with ours. The main difference between their studies and ours is that we investigated view independence in both remote expert collaboration and mutual collaboration while they only investigated remote expert collaboration.

The effect of the independent view was more prominent in the C_M conditions than in the C_R conditions. For remote participants, V_I helped ‘seeing the workspace properly’ more prominently in the C_M conditions than in the C_R conditions. In the local participants’ preference, V_I was superior to V_D in the C_M conditions but it was not in the C_R conditions. These results agree with Carey’s (1989) that suggests a user only receiving instructions would allow the instructor to control the user. In the C_R conditions, the local participants manipulated pieces according to the instructions provided by the remote participants, so they might have less actively changed their viewpoint until the remote participants asked them to do so. On the contrary, in the C_M conditions, the local participants tried out their own ideas and actively changed the viewpoint with their own will. In other words, the local participants followed the remote participant and tried to align their focus with the remote partner in the C_R condition, so the effect of the dependent view might not have been that serious compared to that in the C_M condition where the local participants did not have to follow the remote participants’ opinion.

7 Limitations and Future Works

In this section, we describe the limitations of the experiment and direction for future works. First, we discuss on limitations of our study in recruiting participants and scope of this study in the remote collaboration. Next, we discuss possible future works.

7.1 Limitation in Recruiting Participants

In our study, each pair of participants were friends or colleagues, which would have affected their communication and attitude towards each other. This may have had an influence on the results compared to the case where they did not know each other in advance. However, we note that in general people video conference more with acquaintances rather than with an unknown person. Another limitation in terms of participants is that most of the participants were male (20) and only four of them were female, and their age group was mostly in twenties and thirties. Recruiting unbalanced gender and age may affect the results and it requires further investigation with wider age groups and with more balanced gender.

7.2 Limited Scope of Remote Collaboration Study

This study explores two special styles of remote collaboration, however, remote collaboration can vary according to the amount of information each user has, such as when a remote user has some but not all the information necessary to complete the task. Moreover, collaborators could have different types of information to solve the task, such as a remote user having color information while a local user has size information. Given that our study only had two extreme cases, it might be difficult to generalize our results, but they can still help better understand the impact of remote collaboration styles. We note that most previous studies (Fussell et al. 2004; Kirk et al. 2007; Gurevich et al. 2015; Tait and Billinghurst 2015) only focused on the remote expert collaboration while our study covered a wider range by including mutual collaboration. In the future we could explore different collaboration styles.

The communication cues in our study were drawn annotations and speech, which could potentially be factors affecting the results and limiting the area of the study. Using different communication cues could lead to different results and lessons learned from them. For example, showing hand gestures may be an easier way to communicate and increase the feeling of togetherness with the other collaborators rather than drawing an annotation.

Moreover, the collaboration task could have been more diverse. In this study, the experimental task included physical objects that are small, easy to hold in a small task space area. If the task objects were larger, heavier, and requiring a larger task space, then the results could be different. Furthermore, if the task is manipulating virtual artifacts rather than physical objects, the system may require an interface for manipulating the virtual artifacts corresponding with physical hand manipulation by local participants in this study, and the effect of the interface should be considered. There are a wide range of other tasks possible, such as an intellectual task (e.g. writing an article or guiding a presenter to have better presentation).

7.3 Future Works

Considering the limitations discussed in the previous section, there are many possible directions for future work, including exploring different styles of collaboration, such as users having different amounts of information, users using different visual communication cues (pointers or hand gesture), and tasks including different kinds of physical or virtual objects. This research could also be extended by allowing the remote user to control a shared view, switch between both independent and dependent views or even viewing both at the same time.

Mutual collaboration may be more complicated compared to remote expert collaboration as it includes more types of user activities changing over time. With this complexity, one user can ignore the other user's shared activity (e.g. drawing an annotation) while conducting an individual activity (e.g. looking for a piece for next manipulation at different area). To solve this issue, the system (or an interface) could provide better awareness of other user's activities, so a user can easily know the other user's current or next activity. For example, the interface could show a visual notification when a local user picks up a piece or when a remote user draws an annotation, so that collaborating partners can be aware of when the partner starts certain activities and decide whether to pause his or her own activities to see the partners' activities. The interface could also change the background color when the users are not focusing on the same piece or area (by using eye gaze tracking), so that they could avoid performing shared activities when they are not looking at the same piece or area. Another method for better awareness is to design an interface that encourages users to have more communication that can help to understand the other user. For example, displaying the other user in a small window just next to the gaze point may increase the feeling of togetherness and encourage verbal description about their activities.

The awareness we focused on in this study was understanding the other user's activities while sharing spatial information. However, user awareness should not be limited to external shared information but should also include internal users' factors such as feelings or emotions. In the future we could study user interfaces that share users' emotion, so that collaborators can properly react to their partners' emotional cues. This would include technology for capturing emotion, designing and representing cues for sharing emotion, and research on the effect of the shared emotion.

Broadly, the future work could also include multi-party collaboration with more than two collaborators, and asynchronous collaboration.

8 Conclusion

In this study, we explored how different remote collaboration styles (remote expert collaboration and mutual collaboration) and view types (dependent and independent view) affect remote collaboration in a video-mediated shared physical object manipulation task with a drawing annotation visual communication cue. For remote expert collaboration, users have assigned roles as a remote expert provides instructions to a local worker. In mutual collaboration, users do not have an assigned role and they need to find a solution and discuss ideas together. With a dependent view, the remote user has the same view as the local user and the view is controlled by the local user. With an independent view, the remote user can control his or her own view regardless of where the local user is looking at.

We found that remote expert collaboration provided some benefits compared to mutual collaboration, including better communication from a remote participant to a local participant, more focus on the same object or area between participants, and remote participants' increased feeling of enjoyment and togetherness. We argued that these benefits might come from defining or limiting user activities by assigning roles as participants simply focused on one type of activity (one-way communication from a remote participant to a local participant), while they conducted three activities (sending and receiving activities, and finding a task solution) and switched among them in the mutual collaboration.

The independent view also had several benefits over the dependent view, including resolving the issue of the local participants dominating the navigation of the shared viewpoint in the dependent view, and allowing the remote participants to freely navigate around the workspace while having a wider fully zoomed-out view. The remote participants felt better communication with the local partners, while seeing the workspace more properly, having more enjoyment and feeling of togetherness with the independent view. Interestingly, the benefit of the independent view was more prominent in mutual collaboration than in remote expert collaboration, especially in terms of remote participants seeing the workspace properly and the local participants' preference.

This study has a limitation in terms of exploring the two extreme cases of remote collaboration styles, but the findings could still be useful for understanding attributes of remote collaboration and to identify the future research directions.

For future work, we suggest several potential research topics. First, extending this study with different collaboration conditions, such as different collaboration styles where each user has a different amount or types of information, different communication cues, or different tasks. Second, exploring user interfaces for improving awareness, such as using gaze cues and activity notifications. Third, studying user interfaces that encourage more frequent communication by showing the collaboration partner next to the user's view, so that users may have better awareness. Finally, exploring interfaces that support sharing of user's emotions.

Reference

Adcock, Matt; and Chris Gunn (2015). Using Projected Light for Mobile Remote Guidance. *Computer Supported Cooperative Work (CSCW)*, vol.24, no.6, December 2015 pp. 591–611.

Alem, Leila; Franco Tecchia; and Weidong Huang (2011). Remote Tele-assistance System for Maintenance Operators in Mines. In Naj Aziz (ed): *Coal2011. Proceedings of Underground Coal Operators' Conference, Wollongong, NSW, Australia, 10-11 Febrary 2011*. Wollongong: University of Wollongong, pp. 171-177.

Antunes, Pedro; and Antonio Ferreira (2011). Developing collaboration awareness support from a cognitive perspective. In Jr. Ralph H. Sprague (ed): *HICSS'44. Proceedings of the 44th Hawaii International Conference on System Sciences, Kauai, HI, USA, 4-7 January 2011*. Washington DC: IEEE computer society, pp. 1–10.

Beaudouin-Lafon, Michel; and Alain Karsenty (1992). Transparency and Awareness in a Real-time Groupware System. In Jock Mackinlay and Mark Green (eds): *UIST '92. Proceedings of the 5th annual ACM symposium on User Interface Software and Technology, Monteray, California, 15-18 November 1992*. New York: ACM Press, pp. 171–180.

Benford, Steve; Chris Greenhalgh; Tom Rodden; and James Pycock (2001). Collaborative virtual environments. *Communications of the ACM*, vol.44, no.7, July 2001, pp. 79–85.

Billinghurst, Mark; Adrian Cheok; Simon Prince; and Hirokazu Kato (2002). Real world teleconferencing. *IEEE Computer Graphics and Applications*, vol.22, no.6, December 2002, pp.11–13.

Bly, Sara A.; Steve R. Harrison; and Susan Irwin (1993). Media spaces: bringing people together in a video, audio, and computing environment. *Communications of the ACM*, vol.36, no. 1, January 1993. pp. 28–46.

Brubaker, Jed R.; Gina Venolia; and John C. Tang (2012). Focusing on Shared Experiences : Moving Beyond the Camera in Video Communication. In *DIS'12. Proceedings of the Designing Interactive Systems Conference, Newcastle Upon Tyne, United Kingdom, 11 - 15 June 2012*. New York: ACM, pp. 96–105

Cidota, Marina; Stephan Lukosch; Dragos Datcu; and Heide Lukosch, (2016). Comparing the Effect of Audio and Visual Notifications on Workspace Awareness Using Head-Mounted Displays for Remote Collaboration in Augmented Reality, *Augmented Human Research (AH Research)*, vol.1, no.1, October 2016, pp. 1-15

Cockburn, Andrew; and Saul Greenberg (1993). Making Contact: Getting the Group Communicating with Groupware. In Simon Kaplan (ed):*COCS '93. Proceedings of the Conference on Organizational Computing Systems, Milpitas, California, USA, 01 - 04 November 1993*. New York: ACM, pp. 31–41.

Dourish, Pau; and Victoria Bellotti (1992) Awareness and Coordination in Shared Workspaces. In Marilyn Mantel and Ron Baecker (eds):*CSCW '92. Proceedings of the 1992 ACM conference on Computer-Supported Cooperative Work, Toronto, Ontario, Canada, 01-04 November 1992*. New York: ACM, pp. 107–114.

Fakourfar, Omid; Kevin Ta; Richard Tang; Scott Bateman; and Anthony Tang (2016) Stabilized Annotations for Mobile Remote Assistance. In Jofish Kaye, Allison Druin, Cliff Lampe, Dan Morris, and Juan Pablo Hourcade (eds):*CHI'16. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems ACM Press, San Jose, CA, USA, 07 - 12 May 2016*. New York: ACM pp 1548–1560

Ferreira, Antonio; Pedro Antunes; and Valeria Herskovic (2011) Improving Group Attention: An Experiment with Synchronous Brainstorming, *Group Decision and Negotiation*, vol. 20, no. 5, February 2011, pp.643–666

Fish, Robert S.; Robert E. Kraut; and Barbara L. Chalfonte (1990) The VideoWindow system in informal communication. In Frank Halasz (ed):*CSCW '90. Proceedings of the 1990 ACM conference on Computer-Supported Cooperative Work, Los Angeles, California, USA, 07 - 10 October 1990*. New York : ACM, pp.1–11

Fussell Susan R; Leslie D. Setlock; Jie Yang; Jiazhi Ou; and Elizabeth Mauer (2004) Gestures Over Video Streams to Support Remote Collaboration on Physical Tasks. *Human-Computer Interaction*, vol.19, no.3, September 2004. pp. 273–309.

Fussell Susan R; Leslie D. Setlock; and Robert E. Kraut (2003). Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In Gilbert Cockton and Panu Korhonen (eds): *CHI '03. Proceedings of the SIGCHI conference on Human factors in computing systems, Ft. Lauderdale, Florida, USA, 05 - 10 April 2003*. New York: ACM, pp. 513 - 520

Gauglitz, Steffen; Cha Lee; Matthew Turk; and Tobias Hollerer, (2012). Integrating the physical environment into mobile remote collaboration. In Elizabeth Churchill, Sriram Subramanian, Patrick Baudisch, and Kenton O'Hara (eds):*MobileHCI '12. Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services. San Francisco, California, USA, 21 - 24 September 2012*. New York: ACM. pp. 241-250

Gauglitz, Steffen; Benjamin Nuernberger; Matthew Turk; and Tobias Hollerer (2014). World-stabilized annotations and virtual scene navigation for remote collaboration. In Hrvoje Benko, Mira Dontcheva, and Daniel Wigdor (eds): *UIST '14, Proceedings of the 27th annual ACM symposium on User Interface Software and Technology, Honolulu, Hawaii, USA, 05 - 08 October 2014*. New York:ACM pp.449–459.

Gaver, William W (1991). Sound Support For Collaboration. In: Bannon L., Robinson M., Schmidt K. (eds):*ECSCW'91. Proceedings of the Second European Conference on Computer-Supported Cooperative Work, Amsterdam, Netherlands, 1991*. Dordrecht:Springer Netherlands, pp.293–308.

Gaver, William W (1992). The affordances of media spaces for collaboration. In Marilyn Mantel and Ron Baecker (eds): *CSCW '92. Proceedings of the 1992 ACM conference on Computer-Supported Cooperative Work, Toronto, Ontario, Canada, 01 - 04 November 1992*. New York: ACM pp. 17–24

Gergle, Darren; Robert E. Kraut; and Susan R. Fussell (2013). Using visual information for grounding and awareness in collaborative tasks. *Human–Computer Interaction*, vol.28, no.1, 27 Mar 2012, pp. 1–39.

Greenberg, Saul; and Carl Gutwin (2016). Implications of We-Awareness to the Design of Distributed Groupware Tools. *Computer Supported Cooperative Work (CSCW),* vol. 25, no. 4-5, October 2016, pp. 279–293.

Gurevich, Pavel; Joel Lanir; and Benjamin Cohen (2015). Design and Implementation of TeleAdvisor: a Projection-Based Augmented Reality System for Remote Collaboration, *Computer Supported Cooperative Work (CSCW)*, vol.24, no.6, December 2015, pp. 527-562

Gutwin, Carl; and Saul Greenberg (1996). Workspace awareness for groupware. In Bonnie Nardi, Gerrit C. Veer, and Michael J. Tauber (eds): *CHI '96, Conference Companion on Human Factors in Computing Systems*, Vancouver, British Columbia, Canada, 13 - 18 April 1996. New York: ACM. pp. 208–209.

Gutwin, Carl; and Saul Greenberg (2002). A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, vol.11, no. 3-4, September 2002, pp. 411–446.

Gutwin, Carl; and Saul Greenberg (1999). The Effects of Workspace Awareness Support on the Usability of Real-Time Distributed Groupware. *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol.6, no. 3, Sept. 1999, pp. 243–281

Gutwin, Carl; and Saul Greenberg (1998). Designfor Individuals,Design for Groups: Tradeoffs BeWeen Power and Workspace Awareness. In Steven Poltrock and Jonathan Grudin (eds): *CSCW '98. Proceedings of the 1998 ACM conference on Computer supported cooperative work, Seattle, Washington, USA, 14 - 18 November 1998*. New York: ACM. pp.207–216

Gutwin, Carl; Oliver Schneider; Robert Xiao; and Stephen Brewster (2011). Chalk Sounds : The Effects of Dynamic Synthesized Audio on Workspace Awareness in Distributed Groupware. In Pamela Hinds, John C. Tang, Jian Wang, Jakob Bardram, and Nicolas Duchenenaut (eds): *CSCW'11. Proceedings of the ACM 2011 conference on Computer supported cooperative work, Hangzhou, China, 19-23 March 2011*. New York: ACM, pp. 85–94.

Ishii, Hiroshi; Minoru Kobayashi; and Kazuho Arita (1994) Iterative design of seamless collaboration media. *Communications of the ACM*, vol.37, no.8, Aug. 1994, pp.83–97

Ishii, Hiroshi; Minoru Kobayashi; and Jonathan Grudin (1992) Integration of Interpersonal Space and Shared Workspace - ClearBoard Design and Experiments. *ACM Transactions on Information Systems (TOIS)*. Volume 11 Issue 4, Oct. 1993, pp. 349-375

Hauber, Joerg (2008) *Understanding Remote Collaboration in Video Collaborative Virtual Environments*. Christchurch New Zealand: University of Canterbury Press. Doctoral dissertation from University of Canterbury

Kim, Seungwon; Gun A. Lee; and Nobuchika Sakata (2013) Comparing pointing and drawing for remote collaboration. In Mark Billinghurst and Bruce H. Thomas (eds):*ISMAR'13. Proceedings of 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Adelaide, SA, Australia, 1-4 October 2013*. Washington DC: IEEE computer society pp. 1-6

Kim, Seungwon; Gun A. Lee; Sangtae Ha; Nobuchika Sakata; and Mark Billinghurst, (2015) Automatically Freezing Live Video for Annotation during Remote Collaboration. In Bo Begole, Jinwoo Kim, Kori Inkpen, and Woontack Woo (eds): *CHI EA '15. Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, Seoul Korea. 18 - 23 April 2015*. New York:ACM, pp 1669–1674

Kim, Seungwon; Gun A. Lee; Nobuchika Sakata; and Mark Billinghurst (2014). Improving co-presence with augmented visual communication cues for sharing experience through video conference. In Gudrun Klinker and Nassir Navab (eds): *ISMAR'14. Proceedings of 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). Munich, Germany, 10-12 September 2014*. Washington DC: IEEE computer society, pp 83–92

Kirk, David; Tom Rodden; and Danaë Stanton Fraser (2007). Turn It This Way: Grounding Collaborative Action with Remote Gestures. In Mary Beth Rosson and David Gilmore (eds): *CHI'07. Proceedings of the CHI conference on Human Factors in Computing Systems, San Jose, California, USA, 28 April - 03 May, 2007*. New York : ACM. pp. 1039–1048

Kurata, Takeshi; Nobuchika Sakata; Masakatsu Kourogi; Hideaki Kuzauoka; and Mark Billinghurst (2004). Remote collaboration using a shoulder-worn active camera/laser. In :*ISWC'04. Proceeding of the Eighth International Symposium on Wearable Computer. Arlington, Virginia, 31 October - 03 November 2004*. Washington DC: IEEE computer society, pp.62–69.

Kuzuoka, Hideaki; Gen Ishimoda; Yushi Nishimura; Ryutaro Suzuki; and Kimio Kondo (1995) Can the GestureCam be a surrogate? In H. Marmolin, Y. Sundblad, and K. Schmidt (eds): *ECSCW'95. Proceedings of the Fourth European Conference on Computer-Supported Cooperative Work, Stockholm, Sweden, 10 - 14 September 1995*. Norwell, MA, USA: Kluwer Academic Publishers, pp. 181–196.

Kuzuoka, Hideaki; Toshio Kosuge; and Masatomo Tanaka (1994) GestureCam: A video communication system for sympathetic remote collaboration. In John B. Smith, F. Don Smith, and Thomas W. Malone (eds): *CSCW '94. Proceedings of the 1994 ACM conference on Computer supported cooperative work, Chapel Hill, North Carolina, USA, 22 - 26 October 1994*. New York: ACM, pp.35–43.

Lauwers, J. Chris; and Keith A. Lantz (1990) Collaboration awareness in support of collaboration transparency: requirements for the next generation of shared window systems. In Jane C. Chew and John Whiteside (eds): *CHI '90. Proceedings of the SIGCHI conference on Human factors in computing systems, Seattle, Washington, USA, 01 - 05 April 1990*. New York: ACM, pp.303–311

Lukosch, Stephan; Heide Lukosch; Dragos Datcu; and Marina Cidota (2015) Providing Information on the Spot: Using Augmented Reality for Situational Awareness in the Security Domain. *Computer Supported Cooperative Work (CSCW)* vol.24, no. 6, December 2015. pp.613–664

O'Hara, Kenton; Alison Black; and Matthew Lipson (2006) Everyday practices with mobile video telephony. In Rebecca Grinter, Thomas Rodden, Paul Aoki, Ed cutrell, Robin Jeffries, and Gary Olson (eds): *CHI '06. Proceedings of the SIGCHI conference on Human Factors in computing systems, Montréal, Québec, Canada, 22 - 27 April 2006*.

New York : ACM. pp.871–880

Olson, Gary M.; and Judith S. Olson (2000). Distance matters. *Human-Computer Interaction*, vol.15, no.2, September 2000, pp.139–178

Poelman, Ronald; Oytun Akman; Stephan Lukosch; and Pieter Jonker, (2012) As if being there: mediated reality for crime scene investigation, In Steven Poltrock and Carla Simon (eds): *CSCW '12. Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, Seattle, Washington, USA, 11 - 15 February 2012*. New York: ACM pp.1267–1276

Reichherzer, Carolin; Alaeddin Nassani; and Mark Billinghurst (2014) [Poster] Social panoramas using wearable computers. In Gudrun Klinker and Nassir Navab (eds): *ISMAR'14. Proceedings of 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany, 10-12 September 2014*. Washington DC: IEEE computer society, pp. 303–304

Rekimoto, Jun; and Katashi Nagao (1995) The World through the Computer: Computer Augmented Interaction with Real World Environments. In George Robertson (ed):*UIST'95. Proceedings of the 8th annual ACM symposium on User interface and software technology, Pittsburgh, Pennsylvania, USA, 15 - 17 November 1995*. New York: ACM, pp. 29–36

Roseman, Mark; and Saul Greenberg (1996) Building real-time groupware with GroupKit, a groupware toolkit. *ACM Transactions on Computer-Human Interaction (TOCHI)* vol.3, Issue 1, March 1996. pp.66–106

Sakata, Nobuchika; Takeshi Kurata; Takekazu Kato; Masakatsu Kourogi; and Hideaki Kuzuoka, (2003). WACL: supporting telecommunications using wearable active camera with laser pointer. In *ISWC'03. Proceedings Seventh IEEE International Symposium Wearable Computer 2003, White Plains, NY, USA, 21-23 October 2003*. Washington DC: IEEE computer society, pp.53–56.

Schmidt, Kjeld (2002) The Problem with `Awareness': Introductory Remarks on `Awareness in CSCW'. *Computer Supported Cooperative Work (CSCW)* vol.11, no. 3, 2002, pp.285-298

Snowdon, Dave; Elizabeth F. Churchill; and Alan J. Munro (2000) *Collaborative Virtual Environments: Digital Spaces and Places for CSCW*. London:Springer

Stefik, M; Bobrow DG, Foster G (1987) WYSIWIS Revised : Early Experiences with Multiuser Interfaces. *ACM Transactions on Information Systems (TOIS)*, Vol.5, No.2, April 1987, pp.147-167

Tait, Matthew; and Mark Billinghurs (2015) The Effect of View Independence in a Collaborative AR System. *Computer Supported Cooperative Work (CSCW)*, vol.24, no.6, Springer Netherlands, pp.563-589

Tan, Wei; Haomin Liu; Zilong Dong; Guofeng Zhang; and Hujun Bao, (2013) Robust monocular SLAM in dynamic environments, In Mark Billinghurst and Bruce H. Thomas (eds):*ISMAR'13. Proceedings of 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Adelaide, SA, Australia, 1-4 October 2013*. Washington DC: IEEE computer society, pp. 209-218

Tang, John C.; and Scott L. Minneman (1991) Videodraw: a video interface for collaborative drawing. *ACM Transactions on Information Systems (TOIS)*, vol.9, no.2, April 1991, pp.170–184

Tsubaki, Ikuko; Toshiya Morita; Takahiro Saito; and Kiyoharu Aizawa (2005) An adaptive video stabilization method for reducing visually induced motion sickness. In *ICIP 2005. Proceedings of IEEE International Conference on Image Processing. Genova, Italy, 14 September 2005*. Washington DC: IEEE computer society, pp. 497-500

Whittaker, Steve (1995) Rethinking video as a technology for interpersonal communications: theory and design implications, *International Journal of Human-Computer Studies (IJHCS)*, vol.42, no.5, May 1995, pp.501–529.

Wobbrock, Jacob O.; Leah Findlater; Darren Gergle; and James J. Higgins (2011) The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In Desney Tan (ed): *CHI '11. Proceedings of CHI conference on human factors in computing systems, Vancouver, BC, Canada, 07 - 12 May 2011*. New York :ACM pp.143–146.