 DOI: 10.5278/ojs.quivirr.v3.2022.a0005

Experiences in Collecting 360° Video Data and Collaborating Remotely in Virtual Reality

Anna Vatanen[1]
University of Oulu
0000-0002-8236-657X

Heidi Spets
University of Oulu
0000-0003-0570-9986

Maarit Siromaa
University of Oulu
0000-0002-3015-5682

Mirka Rauniomaa
University of Oulu
0000-0003-0180-4257

Tiina Keisanen
University of Oulu
0000-0002-2425-7412

Abstract

This paper reports on a pilot project called *Remote Research and Collaboration Using VR and 360° Video* (RReCo) that was carried out in late Spring 2021 at the University of Oulu, Finland. The project explored new ways of collecting, viewing and analysing video data for the purposes of engaging in remote, collaborative research on social interaction and activity. Here we share our experiences in collecting different types of video data, especially 360°, and relate those to our user experiences in analysing the data together in virtual reality. Our remote multisite data sessions were organised using software for immersive qualitative analytics, virtual reality and live streaming. In this paper, we also reflect on the similarities and differences between our data sets, especially with view to how awareness of different technical setups may help in making informed choices, and thereby increase the reliability of research on social interaction.

[1] All authors were equally involved in carrying out the pilot project reported in this paper. Authors are listed here in reverse alphabetical order by their surnames.

Keywords: 360° video, Data collection, Immersive qualitative analytics, Reliability, Remote collaboration, Research on social interaction, Virtual reality

1. Introduction

This paper reports on a pilot project called *Remote Research and Collaboration Using VR and 360° Video* (RReCo) that set out to explore new ways of collecting, editing, viewing, transcribing and analysing video data for the purposes of engaging in video-based research on social interaction and activity. The pilot project built on our past and ongoing, joint and individual research projects, in which we have collected data from different contexts and settings. Our research topics vary, but we are all interested in language, multimodality and social interaction, and, methodologically, we draw mainly on ethnomethodological conversation analysis (EMCA; see, for example, Eilittä et al. accepted). We received funding for the RReCo pilot project from the University of Oulu, Finland, which enabled us to use the LeaF infrastructure at the University for data collection and editing.

In the past couple of years, exploring new practices of collaboration has become especially topical due to the COVID-19 pandemic, which has made it challenging for researchers, along with other professionals, to travel and work together in person. Furthermore, despite the existence of multiple online solutions for remote collaboration, such as video-conferencing tools like Zoom and Teams and augmented platforms or spatial audio interfaces like Magic Leap and High Fidelity, there is a growing need for more immersive multi-user virtual environments. In this paper, we wish to contribute to existing knowledge on the use of 3D sites for remote collaboration (see, for example, Sra et al. 2018; Wienrich et al. 2018; Yoon et al. 2019; Boonbrahm et al. 2020) by reflecting on our experiences in collecting both two-dimensional (2D)[2] and 360° video data and in analysing the data in virtual reality (VR).

Coinciding with the increasing demand for remote collaboration solutions, we have, since 2019, worked in collaboration with the BigSoftVideo team at Aalborg University, Denmark. The team has developed two pieces of software: AVA360VR ("Annotate, Visualise, Analyse 360 Video in Virtual Reality") and CAVA360VR ("Collaborate, Annotate, Visualise, Analyse 360 Video in Virtual Reality") (see McIlvenny and Davidsen 2017; McIlvenny 2020a). As beta testers for both software, during the RReCo pilot project, we held regular CAVA360VR-based "Nord360" data sessions[3] with researchers at Aalborg. In addition to functioning as a venue for both initial and more advanced explorations of various data (on data sessions, see McIlvenny 2020a and references therein), the Nord360 sessions provided an

[2] We use the term '2D' to refer to video data collected with traditional, approx. 60°–120° field-of-view video cameras.

[3] We also organised a workshop at the *6th Nordic Interdisciplinary Conference on Discourse and Interaction*, which took place as an online conference, to demonstrate our various data sets and hold a demonstration data session in CAVA360VR: Anna Vatanen, Heidi Spets, Maarit Siromaa, Mirka Rauniomaa, Paul McIlvenny, Tiina Keisanen & Jacob Davidsen: *New methods in video-based qualitative research: remote collaborative analysis of 360° video in virtual reality*, 19 November 2021.

important forum for testing, using and experiencing immersive qualitative analytics (Davidsen and McIlvenny 2022) and remote collaboration.

The diversity of our data helps us to illustrate various aspects of data collection. The use of a 360° camera facilitates the aim of covering as much as possible of what is going on in each situation (see Raudaskoski accepted), increasing the chances to follow the EMCA principle of attending to "the demonstrable indigenous import of the events and of their context *for their participants*" (Schegloff 1997: 184; emphasis in original). When multiple 2D, or even 360°, cameras are positioned so as to cover fully or maximally a particular space, the participants and their activities in that space are necessarily covered, too. When participants are mobile, however, even several roving cameras may not provide continuous footage of all the spaces that the participants move in and through; in such a case, it is reasonable to focus on (some of) the participants and their main (mobile) activities or on the general view of the overall event that the recordings may provide.

By recounting our experiences in using 360° video and VR, we also wish to shed light on the reliability of qualitative research conducted on recorded data and transcripts (reliability referring to "the degree to which the finding is independent of accidental circumstances of the research"; Kirk and Miller 1986: 20). As Peräkylä (1997: 206) points out, selection of what is recorded – especially, how much is recorded – as well as the technical quality of the recordings are crucial aspects of reliability in this type of research. He emphasises the importance of paying enough attention to the quality of the equipment and the recording arrangements, such as the quality of the sound and the inclusiveness of the video image (Peräkylä 1997: 206). Goodwin (1993) provides an early demonstration of technical setups and other details of how social interaction data is best recorded. With the current paper, we wish to contribute to this body of literature from the present-day perspective.

2. Collecting 360° Video Data and Analysing Data in VR

In this section, we describe in some detail the technical and spatial setups that we have used in collecting video data and highlight features of the CAVA360VR software that make it particularly apt for analysing such data together. Table 1 provides an overview of the characteristics of each data set as well as of the recording equipment used in collecting the data. The division between "one main" and "multiple", for the number of participation frameworks and activities, is certainly not clearcut but intended here to give a general sense of typical features of each setting.

Data set / Features	Berry	Break	Choir	Drone	Home	Home office	VR
Participation frameworks:							
- one main			X	X		X	X
- multiple	X	X			X		
Activities:							
- one main	X	X	X	X			X
- multiple					X	X	
Mobility:							
- static		X	X		X	X	X
- mobile	X			X			
Recording equipment:							
360° cam	X	X	X	X	X	X	
2D cam(s)	X	X	X	X	X	X	X
audio rec(s)		X			X		
screen rec						X	X

Table 1 – features and recording setups of the data sets examined in RReCo

In our data sets, the number of participants varies from dyads to large groups, and the number of possible participation frameworks varies accordingly. Nevertheless, even a large group of participants may be engaged in one joint main activity, and a small group of participants may be involved in multiple activities from the outset. Furthermore, some of the activities in our data sets are characteristically static and others more mobile in nature, and the activities occur indoors, outdoors or in a virtual environment. Depending on the setting and the availability of recording equipment, we have used a 360° camera, one or more 2D camcorders or action cameras, one to two audio recorders, and screen capture software.

Before any of the data collection commenced, the participants were provided with a privacy notice and relevant information about the project in question, and they gave their informed consent on the use of the data for research purposes. As indicated in the terms of the consent, some images shown in this section have been retouched in different ways for purposes of pseudonymisation.

2.1. One main activity with multiple participants

Let us begin by introducing two sets of data from settings where relatively large groups of participants form multiple participation frameworks while engaging in one main shared activity. During the activity, the participants remain more or less stationary: 1) in workplace breaks, participants are typically seated around a table when not preparing a hot beverage, for example, and 2) in choir rehearsals, most of the participants are seated or stand in assigned positions. In recording these settings, our aim has been to capture, at the minimum, the shared main activity (that is, having a break as a work community and rehearsing as a choir, respectively). This has allowed us to focus, in relevant detail, on how the activity unfolds and how the participants orient to, on the one hand, carrying out their joint activity and, on the other hand, dealing with possible other ongoing involvements.

In collecting data from workplace breakrooms in the *Break* project[4], a 360° camera was placed on a monopod to hang from the ceiling or on a small tripod at the centre of a table, and high-definition camcorders or action cameras on monopods or tripods were placed in those corners of the room where they did not obstruct participants' movement. Additionally, one or two audio recorders were placed on tables. Figure 1 illustrates the layout of a breakroom in which a 360° camera, two camcorders, one action camera and two audio recorders were used to collect data. The breakrooms in which recordings were made generally accommodate from a dozen up to about thirty participants around sets of tables and seating or by the kitchen counter. Participants may come at different times, carry out various break-taking activities and leave individually or in groups of different sizes. The recordings show participants entering the breakroom, engaging in the break-taking activity and leaving the breakroom (see, for example, Helisten and Siromaa 2022; Siitonen and Siromaa 2020).

In this setup, the cameras captured in relevant detail the overall break-taking activity as well as the emergence of individual participants' involvements. Whenever possible, namely in breakrooms where there was only one large table, the 360° camera was placed on the table between the participants and, in this way, in the middle of the break-taking activity. However, when the 360° camera was placed in a breakroom with multiple table sets, some participants are only captured from above or from the back (see Figure 2). In such a case, one or more 2D cameras were positioned so that they provided complementary views (see Figure 3, which shows the view from the action camera positioned by a couch, as indicated in Figure 1). In the setup described in Figure 1, then, one 360° video camera was used to record groups of participants at three sets of tables, and three 2D cameras were used to record each of the groups from another angle.

[4] The project *Interactional organisation of break-taking activities and social support in a changing workplace environment* was funded by the Eudaimonia Institute at the University of Oulu, Finland, in 2018–2022.

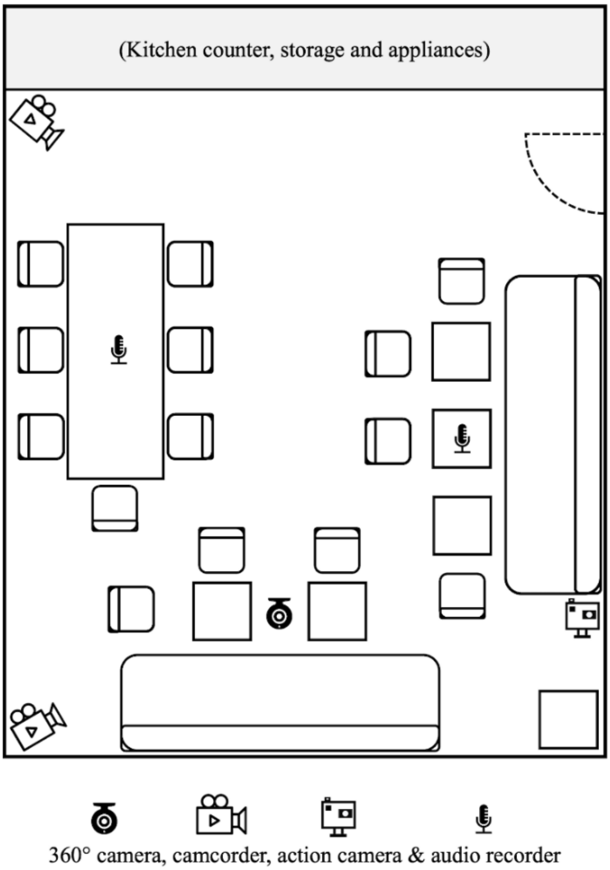

Figure 1 – the layout of a breakroom and placement of recording equipment

Figure 2 – cropped part of an equirectangular projection of 360° footage from a breakroom

Figure 3 – view from an action camera mounted on a monopod in the breakroom

As the layout presented in Figure 1 suggests, in the breakroom setting, where multiple participation frameworks may emerge in and across different parts of the room, it was useful to record sound not only with microphones of the cameras but also with separate audio recorders. In the instance depicted in Figures 2 and 3, for example, an audio recorder had been placed on the table between the three seated participants, and that recording was selected as the main source of audio when preparing the clip for a data session in CAVA360VR. In analysing the data in CAVA360VR, we then used the possibility of embedding a 2D video in the virtual environment, so as to gain a view of all participants in a particular participation framework (for example, the video shown in Figure 3 was embedded in video shown in Figure 2). The 360° and 2D videos are synchronised so that playback starts and stops on both at the same time (see also Figure 16).

Another setting examined in the RReCo pilot project was choir rehearsals, in which a large group of participants is engaged in one main shared activity. These data were originally collected within the project *iTask: Linguistic and embodied features of interactional multitasking*.[5] The data were recorded using a 360° camera that was positioned in the middle of the space between the conductor and the choir (approx. 50 members, seated in two or three rows), and, in some of the rehearsals, using an additional wide-angle action camera that was positioned in the back of the space (see Figures 4–7).

[5] The project has been funded by the Eudaimonia Institute of the University of Oulu and the Academy of Finland (project number 287219).

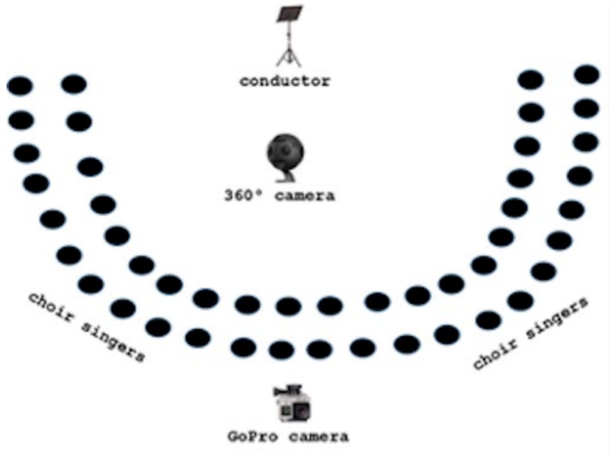

Figure 4 – the choir setting

Figure 5 – crystal ball view from the 360° camera in the choir setting

Figure 6 – equirectangular projection of 360° footage from the choir setting

Figure 7 – view from the action camera in the choir setting

Because of the positioning of the 360° and action cameras, all participants are visible in both views most of the time. The conductor's and, in general, the choir members' actions are thus available for examination, as they engage in their joint main activity in the overall participation framework or momentarily divide into several smaller ones. A separate audio recorder was not used, and the audio quality of these data is relatively poor, especially when the participants sing: the microphones of the cameras were not able to handle the intensity of the volume of the music rehearsal. A high-quality spatial microphone would have worked better for this, and it would also have enabled the researcher analysing the recordings, especially in VR, to sense the spatiality of the sound environment (for example, where each sound comes from in the room).

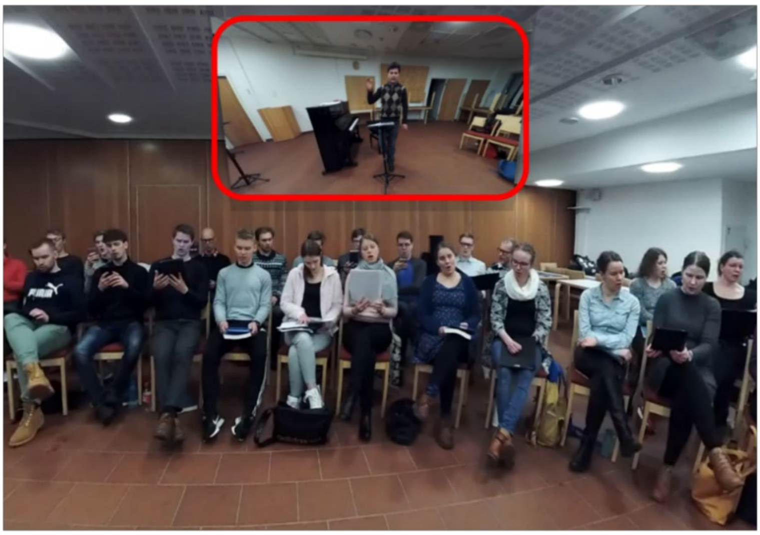

Figure 8 – cropped view of the CAVA360VR software, with a mirror camera view (circled red) included

For the choir data, we found one tool in CAVA360VR especially useful: the incorporated, synchronised mirror camera (see Figure 8), which is in effect very similar to the rear-view mirror in cars. With the help of the mirror camera tool, it is possible to see in two directions simultaneously in the 360° video: the singers (or at least most of them) and the conductor, who would otherwise be behind the viewer's back. The mirror camera tool thus allows analysing the whole participation framework in a synchronised view, without being limited to two separate frames. The mirror camera tool was crucial for analysing the main activity in these data, the rehearsing, as a complete analysis of the rehearsing essentially involves simultaneous monitoring of both the conductor and the singers. The participants themselves, of course, observe the situation in a different way, as they belong to one of the two parties, being either the conductor or a member of the choir. The analyst, however, necessarily needs to observe all participants to be able to make sense of the setting and the activities in it as a whole. A submitted manuscript (Vatanen submitted) investigates the interplay of the multiple participation frameworks in this setting.

2.2. Multiactivities with multiple participants

In addition to examining data sets in which multiple participation frameworks may emerge as multiple participants engage in one main joint activity, during the RReCo pilot project, we collected new data from settings that characteristically involve some form of multiactivity. We focused specifically on families which had children of different ages and in which at least one parent worked remotely at home. The data facilitate the examination of how family members coordinate the various activities that they engage in during an ordinary day, including work, play, domestic chores, and caregiving.

The first set of data was collected in the home of a family of two parents and their two children (aged 7 and 9 years), in which one of the parents worked remotely. The data consist of recordings from a 360° camera and screen recordings of the computer of the parent working from home (Figures 9–12). The 360° camera was not used to its full potential in this setup: the recording shows the shelf that the camera was on by the back wall of the room (Figure 10), but sections of the 360° footage do provide a good view of the working space, doorways to adjoining rooms and parts of the staircase (Figure 11).

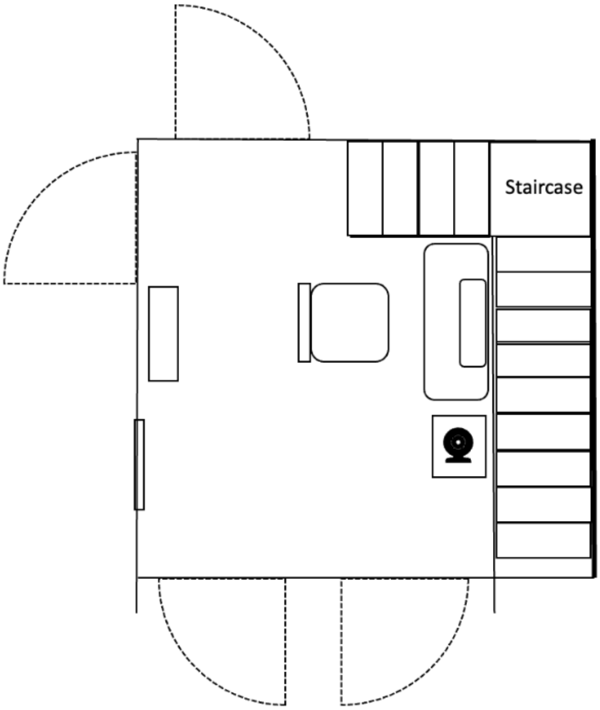

Figure 9 – layout of remote-working space and placement of recording equipment

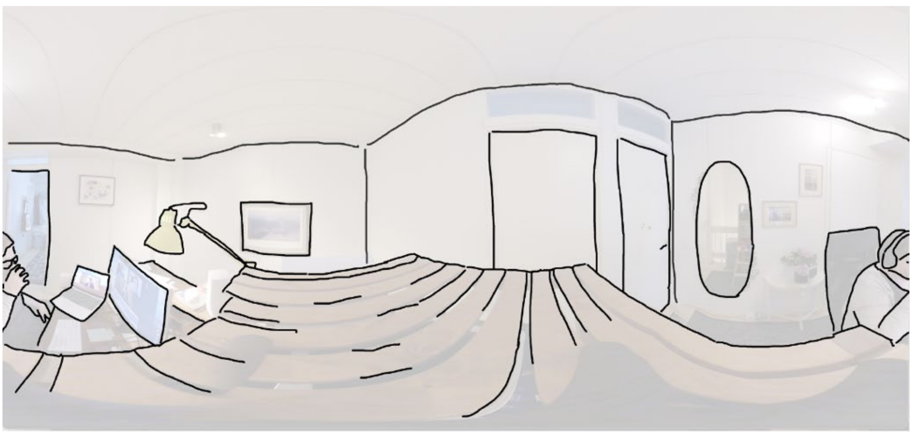

Figure 10 – equirectangular projection of 360° footage from a family home

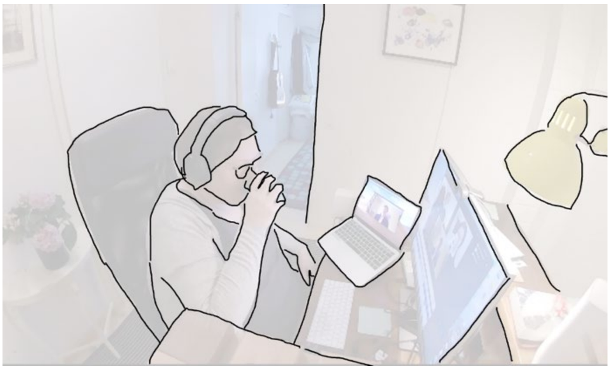

Figure 11 – cropped view of the remote-working space from the 360° camera

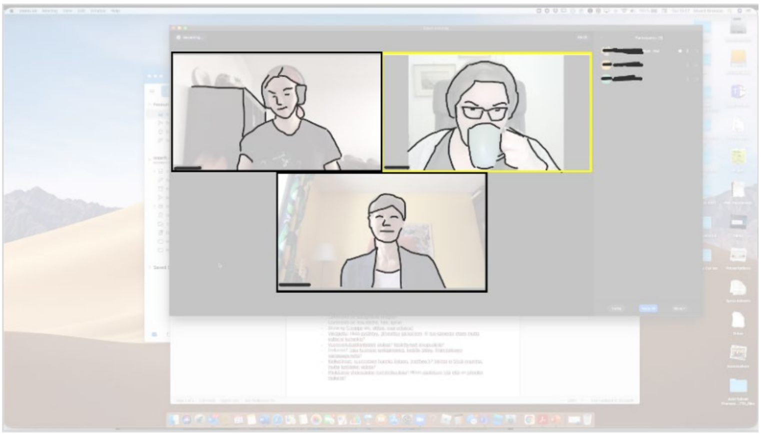

Figure 12 – capture of the screen recording of the remote worker's computer, with a video-mediated meeting going on

The home office data give us access to activities at the interface between work and family life. In particular, in analysing the data together, we focused on how the remote-working parent organises her professional and domestic tasks, by managing transitions between different activities or dealing with potential multiactivity situations. For instance, the data show the children of the family crossing the working space on their way from an adjoining room to the staircase and the mother at those moments suspending her work activities to inquire about the children's day at school or in day care. The data also allow us to examine the role that technology may play in transitions and multiactivity situations and, for instance, to explore interaction in a shared virtual space from the point of view of one participant's local space (see, for example, Holmström, Rauniomaa and Siromaa 2022; Oittinen 2018).

The second set of data was collected in the home of two parents and their five children (aged 1–7 years). During data collection, the father worked remotely from home, and the mother stayed at home taking care of the children. In addition to the 360° camera, we used two separate audio recorders as well as three 2D video cameras: the 360° camera was mounted on a small tripod on an island between the kitchen and the dining- and living-room space (see Figures 13 and 14), one 2D camera was mounted on a tripod in a corner of the living room and two 2D cameras were placed on a drawer in the hallway, one pointing towards the front door and the other towards the dining- and living-room space (see Figure 15 for a compilation of these views). The recordings were saved on a laptop as an automated multi-screen video (all equipment provided by the LeaF infrastructure).

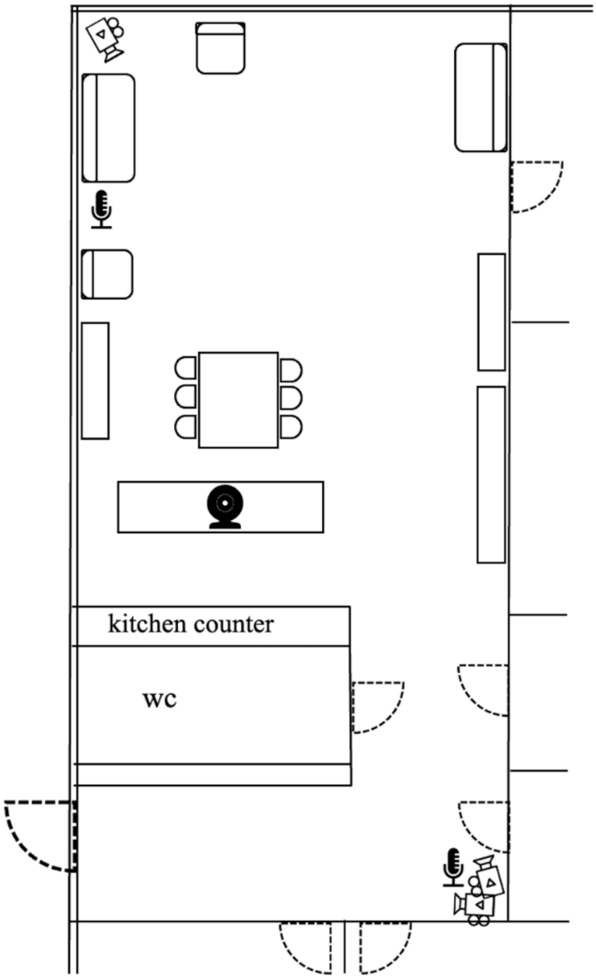

Figure 13 – the layout of the family home and placement of recording equipment

Figure 14 – equirectangular projection of 360° footage from another family home

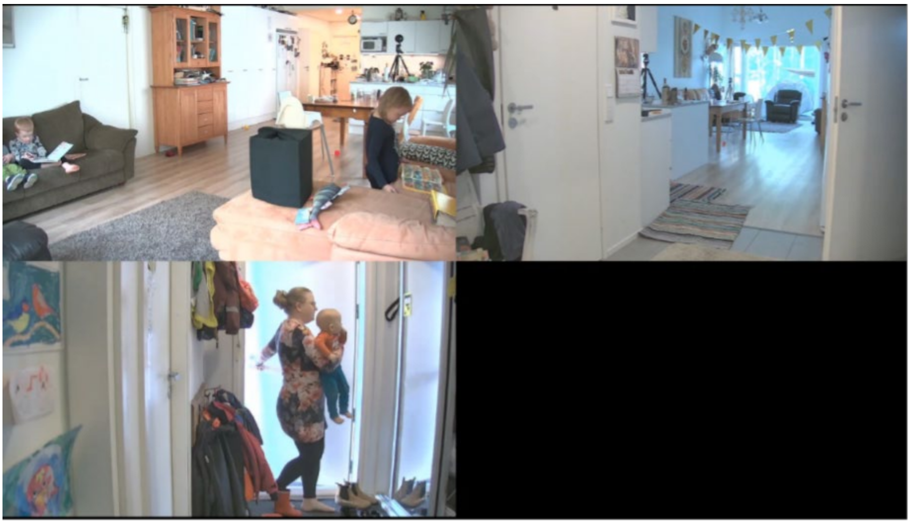

Figure 15 – views from the three 2D cameras in the family home

In this family, similarly to the previous family, we have paid attention to how a parent – in this case the father – moves between family and work, or between the family space and his remote-working space (i.e., a car parked by the front entrance to the house). We analysed this phenomenon in the demonstration data session held as a part of the NORDISCO conference workshop that we organised in 2021 (Vatanen et al. 2021), as well as in other venues (Vatanen and Rauniomaa 2021, in prep.).

Additionally, in these data, we have focused on what the mother and the children do at home and how they engage in multiple activities simultaneously (see, for example, Vatanen and Haddington 2021). Analysing the data in CAVA360VR, we used both the mirror camera tool (see Figure 8) and footage from a separate 2D camera (Figure 16).

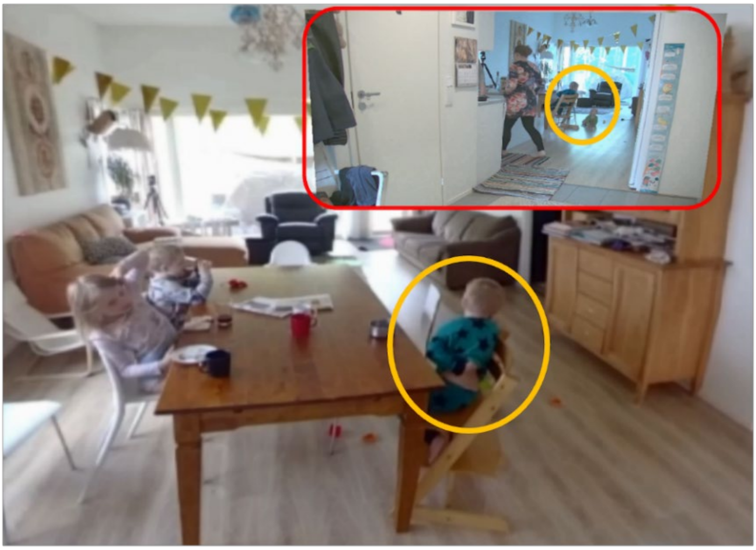

Figure 16 – cropped view of the CAVA360VR software, with synced footage from another camera included (circled red)

Figure 16 illustrates the advantages of viewing 2D video within 360° video from a setting with multiple participants: from the synchronised, incorporated 2D video we acquire another camera angle, which here allows us to see the F-formation between the two children, circled in yellow in the Figure. In the main 360° footage, the younger child is behind a chair and his actions would thus not be available for analysis. The multiple synced camera angles provide us as analysts with access to the participants' perspectives as they engage in various activities in their everyday life.

2.3. Mobile activities with participants as data collectors

In contrast with recording in (indoor) settings where researchers may simply set up the recording equipment and leave the scene, recording in (outdoor) settings where participants engage in mobile activities requires more attention on the presence of either researchers or participants to take care of following the action and moving the recording equipment along with the participants. Some of the data that we examined in the RReCo pilot project arise originally from the project *Human Activities in Natural Settings* (HANS)[6], which documented participants' interactions and mobile activities in nature. One of the stated aims of the HANS project was to explore the possibilities new technologies offer for collecting data in outdoor settings and of mobile activities. The development of technology during the HANS project in 2015–2019 was fast: at the beginning of the project, small action cameras were popular on the consumer market and, by the end of the project, affordable, easy-to-carry 360° video cameras were becoming widely available. Furthermore, much of the data

[6] The project was financed by the Academy of Finland in 2015–2019 (project number 285393).

collection in the HANS project was done in collaboration with the participants, and below we provide two examples of collecting and analysing such data from mobile settings.

The first case illustrates the recording of opportune situations and events, which require flexibility and planning on the fly. One such opportunity was provided by a one-day drone-flying training, which was aimed primarily at private forest owners and forestry professionals. The training included hands-on practice in small groups around a park area (Figure 17). One of us participated in the training in a group that had an instructor and three other trainees (two of whom are present in the situation depicted in Figures 17–19). We discussed data collection with the organiser in advance and prepared to record something, but the possibility to collect data was only confirmed on site and potential participants were recruited once the training was already ongoing. Therefore, decisions about which pieces of equipment to use and how had to be made in haste: the researcher who participated in the training ended up wearing a chest-mounted action camera on a harness, which captured the drone controls when the researcher was holding the device (Figure 18), and a 360° camera was mounted on a short, handheld tripod and passed on from one participant to another, depending on who was the most convenient person to hold it at any given time (Figure 19).

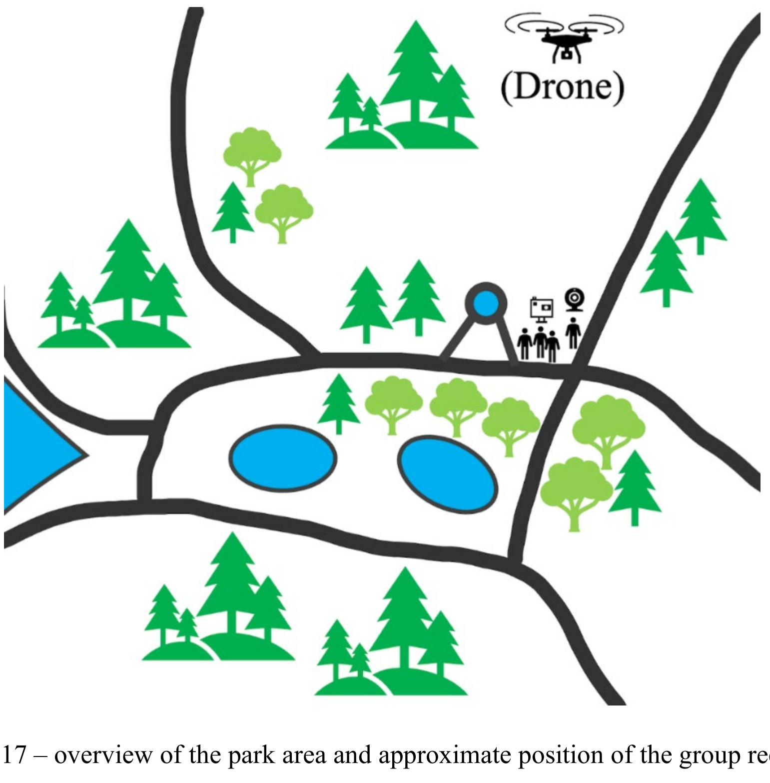

Figure 17 – overview of the park area and approximate position of the group recorded

Figure 18 – footage from chest-mounted action camera, showing drone controls

Figure 19 – equirectangular projection of 360° footage from drone-flying training, with camera held by a participant on a short tripod

Because the 360° camera was held in hand and passed on from one participant to another, rather than mounted on a static tripod, it does not always provide continuous footage of particular action trajectories or of participants whose conduct the analyst might be interested in. Nevertheless, the 360° video in this case has allowed us to follow, on the one hand, the participants' orientation to the drone-flying, which spans across large open spaces and some twenty metres in the air, and, on the other hand, their engagement in the training, which relies on the use of various vocal, embodied, material and spatial resources (for example, the instructor's verbal instructions and the trainees' noticings, the participants' body orientations, gestures and gaze, as well as the drone controls).

Foraging activities were the main nature-related mobile activities focused on in the HANS project, and several occasions of families or groups of friends engaged in picking berries or mushrooms were recorded. Most of these events were captured with small action cameras, but towards the end of the project we also collected 360° video data on a group of family members picking wild berries in the woods (see Figure 20 for an overview of the area). To enable the recording of the mobile activity, one small, lightweight 360° camera was integrated into the cap worn by the mother of the family (Figure 21). Additionally, the grandfather, who is visible in Figure 21 with the two children of the family, had a chest-mounted action camera on a harness.

Figure 20 – overview of the forest area and approximate position of the participants

Figure 21 – placement of the 360° camera and a cropped view of the 360° footage

When the participants were moving and searching for an appropriate place to pick berries, the mother wore the camera on her cap, but when a good berry-picking spot was found, the camera was placed on a stand. Such a situation, in which the person doing the recording also participates in the activity, may require some balancing between the roles of data collector and participant (see Pehkonen, Rauniomaa and Siitonen 2021, and other papers in the same special issue). These different participant positions are socially negotiated and flexibly adjusted. As observed by Raudaskoski (accepted), 360° footage makes the camerapersons' analytical work in these different participant positions also available for analysis.

In this case, placing the 360° camera on a stand stabilises the video and enables the person making the recording to participate in the actual foraging activity. The 360° video provides an overall view on the participants' ongoing activities, much more so than a single or even multiple 2D cameras would. However, mounting an additional action camera on an individual participant's chest makes it possible to capture, close-up, their manual activities. In analysing such activities as foraging, this setup provides the possibility, for example, to examine the coordinated use of language, participant's gestures and other embodied actions in how they direct others to look at a target or to do something with or to it (Siitonen, Rauniomaa and Keisanen 2021). Altogether, the use of 360° cameras and additional 2D cameras increases the reliability of data collection and the analysis of interactions in mobile settings.

More generally, in studying mobile settings, we have been interested in how the different nature-based activities and their related mobility practices, such as wayfinding, are organised, and how participants assist another's movement or draw each other's attention to events and objects in their changing surroundings. A 360° camera has been useful in capturing emergent phenomena, such as wildlife that appear from unforeseen directions or scenery that becomes visible as participants move along in nature. Thus, analysing these data collaboratively in CAVA360VR, we found it helpful to utilise the avatar's hand pointing gestures to draw each other's attention to specific objects (see Figure 22 for an illustration). In avatar-based interaction in CAVA360VR, the ability to be able to point and draw attention to specific features in the environment is an essential resource for enabling joint action and attention among the data session participants. Furthermore, a basic resource in any conversation-analytic data session is the transcript of the event that is being examined together. Immersive qualitative analytics provide nearly first-hand access to the embodied features of interaction, but the details of talk are still best represented in a transcript. It is possible to include a transcript in CAVA360VR and each participant can individually have it appear and disappear, as well as move it around as necessary (see Figure 22). This facilitates analysis as the position of the transcript may be changed, to align with the changing position of the mobile participants in the video(s).

Figure 22 – cropped view of the CAVA360VR software, showing a transcript, the pointing gesture by the user's avatar (hands and green line on the right) as well as another avatar ("Mirka" on the left)

As mentioned, participants engaged in mobile activities continuously move and change position. Thus, another tool in CAVA360VR that we have found useful is the drawing tool: data session participants in CAVA360VR can draw on the video, which helps others to pay attention to relevant features in it. In analysing the foraging data, for example, we used this tool to draw arrows to highlight the direction of the participants' gaze or movement and to write letters to identify the participants (see drawings in blue in Figure 23).

The ability to draw on the video further enables the sharing of the same perspective on the ongoing social interaction when immersed in the data. The achievement and maintenance of the shared perspective has been identified as one of the challenges in VR-based social interaction (see Hindmarsh, Heath and Fraser 2006; Spets submitted), and thereby it presents a potential challenge also for VR-based data sessions. It is also worth noting that the type of activity and the setting may have additional consequences for immersive qualitative analytics: the stability of any video recording of mobile activities may become an issue, but especially data captured with a 360° camera may cause severe nausea when viewed in VR. This can partly be remedied by editing the video feed and stabilising it as far as possible, but the problem may still remain for those prone to motion sickness.

Figure 23 – cropped view of the CAVA360VR software, with materials drawn in CAVA360VR (in blue on the right)

The main observation here regarding making video recordings of mobile activities is that not only it is about the cameras and how they are set up, but also – and in many cases essentially – about the role of the cameraperson in the ongoing activities (see Raudaskoski accepted). The changing setting also requires that the person(s) doing the recording, whether they are researchers or participants, should be able to adapt the recording plan flexibly. Decisions made in the field thus have consequences for data analysis also.

2.4. Interaction in immersive co-present virtual reality

One more set of data recorded during the RReCo pilot project was collected from the virtual environment Rec Room for examining the complexity of interaction in virtual reality. The data were collected at the LeaF infrastructure at the University of Oulu.

The research participants engaged in various activities in the Rec Room virtual space. Rec Room provides its users with a large virtual space full of activities such as Disc Golf, Paintball and various ball games. It is, as the name implies, a virtual recreation room where people can gather to do things together and socialise. The software is available for free on most virtual reality platforms. The participants in our study used Oculus Quest 2 headsets and handheld controllers during the recording sessions (see Figure 25). Two participants were recorded at a time, and there were three pairs in total. The participants were in different rooms (as shown in Figure 24). However, they were separated only by a thin wall and could sometimes hear each other (partially due to Quest 2 using small speakers instead of headphones for the audio feed to the user). There is around one hour of video from each pair of participants.

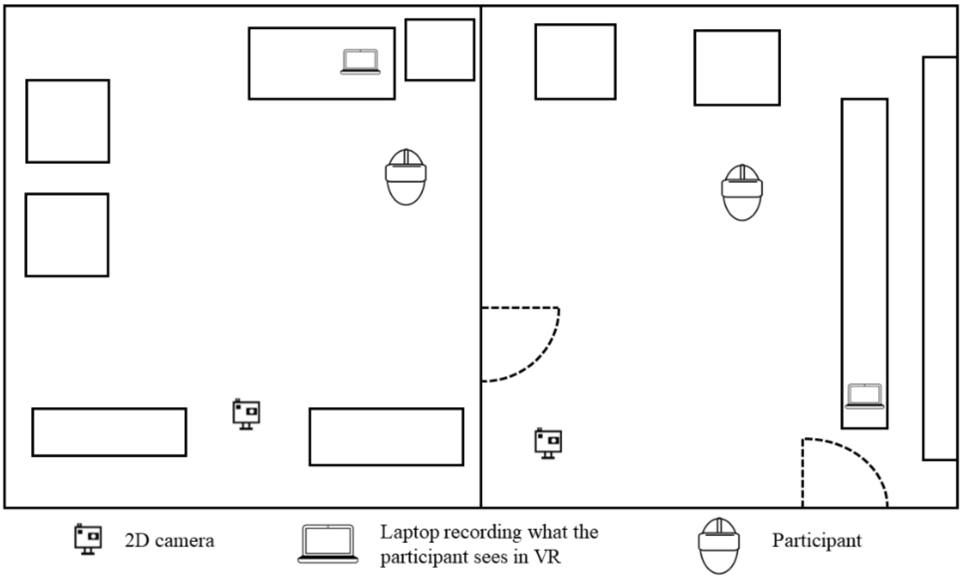

Figure 24 – the layout of the VR recording setting

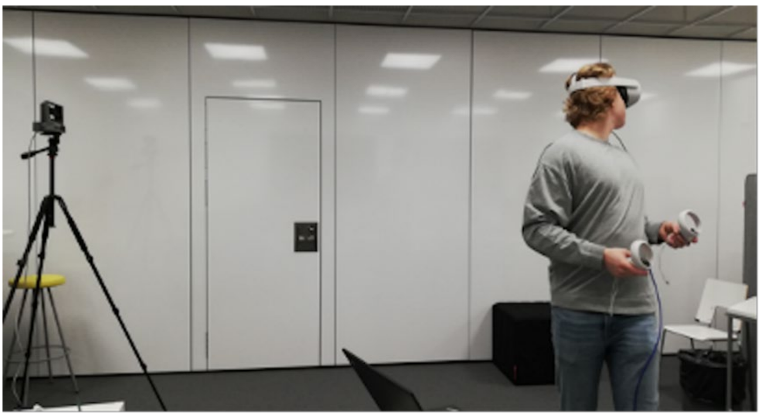

Figure 25 – the recording set-up: camera on tripod (left) recording the participant in the physical space and the participant (right) wearing a VR headset and holding controllers

Each participant was also recorded by a 2D camera in the physical space (see Figures 24 and 25). The camera followed them, keeping them in focus as they moved around in a small area. These cameras were connected to the same computers as the participants' VR headsets. This allowed the recording software (OBS) to compile the in-game footage with the recording from the physical space into a single video file. The result was a video that showed both views in sync without the need for further editing, saving time when preparing the materials. The materials were later edited to include both participants' recordings in one composite video (see Figure 26).

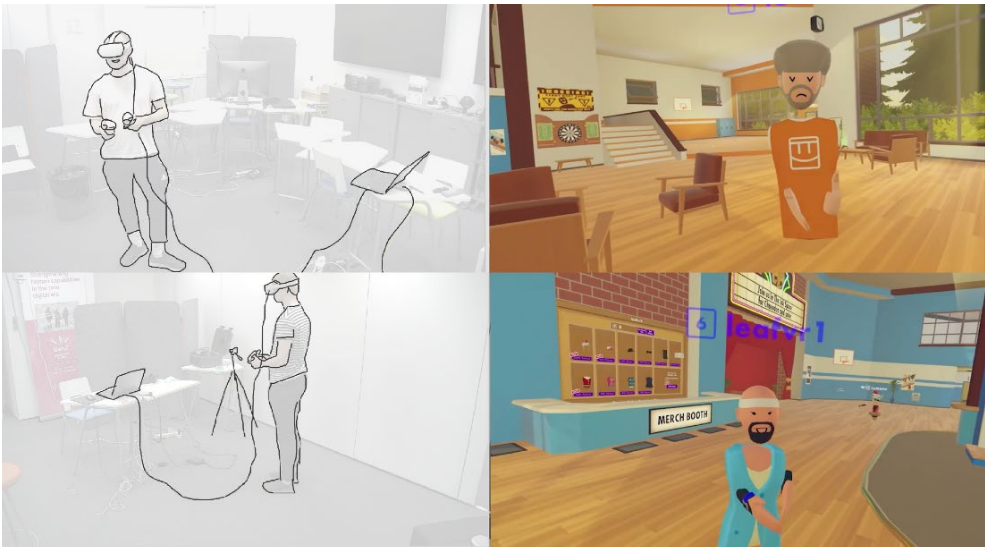

Figure 26 – the edited result showing both participants in virtual reality and the physical space

It is important to record both the physical space and the virtual space when examining interaction in VR (see Kohonen-Aho and Haddington accepted): these parallel videos allow for insights into participants' actions in a way that capturing only the VR space does not. In VR, participants inhabit both their physical and virtual bodies simultaneously. Participants' physical actions are the basis for the movements of their virtual bodies, avatars. However, due to technological reasons, not all their actions are reliably replicated to the avatar body. This can lead to situations in which participants' actions can be either not recognisable in VR or entirely inaccessible to others (Kohonen-Aho and Haddington accepted; Spets submitted). Having access to the physical space aids in the analysis of such situations by providing access to participants' embodied conduct and orientation. However, the use of the parallel video from the physical space should always be considered carefully since it contains more information than what is available to the participants themselves.

During the RReCo pilot project we realised that the VR data that we collected cannot be viewed in CAVA360VR in the same way as our other materials, because the data is not 360°. The VR materials can, nonetheless, be viewed together in Zoom or on another virtual platform that makes remote collaboration possible. However, using such software can create issues of latency due to, for example, the stability of internet connections, which is not an issue in CAVA360VR since each participant has a local copy of the data (see McIlvenny 2020a). Indeed, working with VR data and video-mediated interaction in general, latency needs to be considered (see, e.g., Seuren, Wherton, Greenhalgh and Shaw 2021).

3. Conclusion

Based on our experiences with different types of settings and contexts for collecting video data and with engaging in immersive qualitative analytics, here are some of the most central points to consider in these research activities in the future. A recording from a 360° camera can be used to provide an overall view of the setting or activity, and other (360° or 2D) cameras and audio recorders can complement that in various ways. This allows for analytic attention to be directed to multiple participants and multiple participation frameworks, as appropriate. The use of spatial microphones can further facilitate the analysis of 360° video data by enabling spatial sound experience in VR. However, using technical devices such as 360° cameras, multiple cameras, spatial audio and so forth, the analyst does not go beyond what is called the member's perspective in ethnomethodological and conversation analytic research, but is, instead, able to gain deeper or broader access to it (see Eilittä et al. accepted), at the same time enhancing the analytical potential of the data (Raudaskoski accepted) and ensuring its reliability. The data can thus provide the analyst with a multitude of perspectives to choose from when analysing the recorded situation, and thus a possibility to choose the perspective that is closest to that of a particular participant at a particular moment. Moreover, with multiple views on the same interactional event, it is possible to change and adjust the analytical perspective to follow a participant's perspective in a changing situation. In itself, the qualitative analysis of 360° video data does not differ from the qualitative analysis of 2D video data (for example, the analyst necessarily has to choose where to look, as the human field of vision is not 360°) – except that it provides the possibility to experience 360° video data in VR. It should be noted, however, that the analytical procedures themselves are not in principle different in VR either.

In general, data sessions in CAVA360VR provide a rich sensorial experience – both in terms of being immersed in the data, together with other researchers who are visible as avatars, and in terms of the analytical affordances of the software. CAVA360VR includes several tools that aid the joint analysis of video data: as we have discussed, it is possible to include a transcript or other visual files in the scene, to point to objects of analysis with the avatar's hand and a tracked visual pointer, to draw on the video, to use a mirror camera tool to gain a view in two directions simultaneously, as well as to use a separate synced 2D video that provides an additional view to the situation (for these tools, see also McIlvenny 2019). We have illustrated these tools with examples from our data,[7] and considered how their use has helped us in immersive qualitative analytics. In CAVA360VR data sessions, researchers are able to "inhabit" the data in virtual reality (McIlvenny 2019), experiencing the playback of the data simultaneously and using the tools for interactive, collaborative analysis (Davidsen and McIlvenny 2022). In comparison with traditional data sessions, which may be organised in a hybrid mode but are also

[7] At present, unlike in AVA360VR, in CAVA360VR it is only possible to include one additional video, which is a disadvantage when there are several videos available, such as in our data from the family home. We are, however, aware of the software developers' plans to release a version of CAVA360VR that includes all features available in AVA360VR.

then mostly based on one 2D video and one-directional sound, VR allows for a marked sense of presence, the spatiality of participants' voices and their movement in the data and the ability to zoom in on certain parts of the video canvas. The co-presence of other researchers experiencing the same data provides – what feels like – more direct access to the events on the video.

4. Acknowledgements

The pilot project *Remote Research Collaboration Using VR and 360° Video* (RReCo) received funding from the GenZ profiling project at the University of Oulu, Finland. Part of the work was carried out with the support of the LeaF Research Infrastructure, University of Oulu, Finland. We are grateful to Henna Kaaresto, our research assistant in RReCo, and Pauliina Siitonen, a fellow researcher in RReCo, for their work and contribution to the project. We extend our thanks to Marika Helisten, the two anonymous reviewers and the journal editors for their constructive comments on earlier versions of this paper. Finally, we thank Paul McIlvenny, Jacob Davidsen and others in the BigSoftVideo team for their continued support and guidance, as well as Pirkko Raudaskoski, Christopher Jenks, Tobias Boelt Back, Laura Bang Lindegaard and other Nord360 participants for the exciting collaboration.

References

Boonbrahm, Poonpong, Charlee Kaewrat, and Salin Boonbrahm. 2020. Effective Collaborative Design of Large Virtual 3D Model Using Multiple AR Markers. *Procedia Manufacturing* 42: 387–392. https://doi.org/10.1016/j.promfg.2020.02.058.

Davidsen, Jacob, and Paul McIlvenny. 2022. Towards Collaborative Immersive Qualitative Analysis. In *CSCL Proceedings: 15th International Conference on Computer-Supported Collaborative Learning (CSCL) 2022*, edited by Armin Weinberger, Wenli Chen, Davinia Hernández-Leo and Bodong Chen, 304–307. International Society of the Learning Sciences (ISLS).

Eilittä, Tiina, Pentti Haddington, Antti Kamunen, Laura Kohonen-Aho, Tuire Oittinen, Iira Rautiainen, and Anna Vatanen. Accepted. Ethnomethodological Conversation Analysis in Motion: An introduction. In *Ethnomethodological Conversation Analysis in Motion: Emerging Methods and Technologies*, edited by Pentti Haddington, Tiina Eilittä, Antti Kamunen, Laura Kohonen-Aho, Tuire Oittinen, Iira Rautiainen and Anna Vatanen. London: Routledge.

Goodwin, Charles. 1993. Recording Human Interaction in Natural Settings. *Pragmatics* 3 (2): 181–209. https://doi.org/10.1075/prag.3.2.05goo.

Helisten, Marika, and Maarit Siromaa. 2022. 'Right, Already Brewing There': Arriving to the Staff Breakroom and Beginning Interaction with a Co-Present

Colleague. *Journal of Pragmatics* 193: 139–154. https://doi.org/10.1016/j.pragma.2022.03.016.

Hindmarsh, Jon, Christian Heath, and Mike Fraser. 2006. (Im)materiality, Virtual Reality and Interaction: Grounding the 'Virtual' in Studies of Technology in Action. *The Sociological Review* 54(4): 795–817. https://doi.org/10.1111/j.1467-954X.2006.00672.x.

Holmström, Mari, Mirka Rauniomaa, and Maarit Siromaa. 2022. Zooming in on a Frame: Collectively Focusing on a Co-Participant's Person or Surroundings in Video-Mediated Interaction. In *Shaping the North through Multimodal and Intermedial Interaction*, edited by Juha-Pekka Alarauhio, Tiina Räisänen, Jarkko Toikkanen and Riikka Tumelius, 95–120. Arctic Encounters. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-99104-3_5.

Kirk, Jerome, and Miller, Marc L. 1986. *Reliability and Validity of Qualitative Research*. London: Sage.

Kohonen-Aho, Laura, and Pentti Haddington. Accepted. From Distributed Ecologies to Distributed Bodies in Interaction: Capturing and Analyzing 'Dual Embodiment' in Virtual Environments. In *Ethnomethodological Conversation Analysis in Motion: Emerging Methods and Technologies*, edited by Pentti Haddington, Tiina Eilittä, Antti Kamunen, Laura Kohonen-Aho, Tuire Oittinen, Iira Rautiainen and Anna Vatanen. London: Routledge.

McIlvenny, Paul. 2019. Inhabiting Spatial Video and Audio Data: Towards a Scenographic Turn in the Analysis of Social Interaction. *Social Interaction. Video-Based Studies of Human Sociality* 2 (1). https://doi.org/10.7146/si.v2i1.110409.

McIlvenny, Paul. 2020a. New Technology and Tools to Enhance Collaborative Video Analysis in Live 'Data Sessions'. *QuiViRR: Qualitative Video Research Reports* 1: a0001. https://doi.org/10.5278/ojs.quivirr.v1.2020.a0001.

McIlvenny, Paul. 2020b. The Future of 'Video' in Video-Based Qualitative Research Is Not 'Dumb' Flat Pixels! Exploring Volumetric Performance Capture and Immersive Performative Replay. *Qualitative Research* 20 (6): 800–818. https://doi.org/10.1177/1468794120905460.

McIlvenny, Paul, and Jacob Davidsen. 2017. Big Video Manifesto: Re-sensing Video and Audio. *Nordicom Information* 39 (2): 15–21. https://doi.org/10.5281/zenodo.3972007.

Oittinen, Tuire. 2018. Multimodal Accomplishment of Alignment and Affiliation in the Local Space of Distant Meetings. *Culture and Organization* 24 (1): 31–53. https://doi.org/10.1080/14759551.2017.1386189.

Pehkonen, Samu, Mirka Rauniomaa, and Pauliina Siitonen. 2021. Participating Researcher or Researching Participant? On Possible Positions of the Researcher in the Collection (and Analysis) of Mobile Video Data. *Social Interaction. Video-Based Studies of Human Sociality* 4. https://doi.org/10.7146/si.v4i2.127267.

Peräkylä, Anssi. 1997. Reliability and Validity in Research Based on Tapes and Transcripts. In *Qualitative Research. Theory, Method and Practice*, edited by David Silverman, 201–220. London: Sage.

Raudaskoski, Pirkko. Accepted. 360-cameras Used by a Team Participating in a Mobile Gathering. In *Ethnomethodological Conversation Analysis in Motion: Emerging Methods and Technologies*, edited by Pentti Haddington, Tiina Eilittä, Antti Kamunen, Laura Kohonen-Aho, Tuire Oittinen, Iira Rautiainen and Anna Vatanen. London: Routledge.

Schegloff, Emanuel A. 1997. Whose Text? Whose Context? *Discourse & Society* 8(2): 165–187. http://dx.doi.org/10.1177/0957926597008002002

Seuren, Lucas M., Joseph Wherton, Trisha Greenhalgh, and Sara Shaw. 2021. Whose Turn Is It Anyway? Latency and the Organisation of Turn-Taking in Video-Mediated Interaction. *Journal of Pragmatics* 172: 63–78. https://doi.org/10.1016/j.pragma.2020.11.005.

Siitonen, Pauliina, Mirka Rauniomaa, and Tiina Keisanen. 2021. Language and the Moving Body: Directive Actions with the Finnish *kato* 'look' in Nature-Related Activities. *Frontiers in Psychology* 12: 661784. https://doi.org/10.3389/fpsyg.2021.661784.

Siitonen, Pauliina, and Maarit Siromaa. 2020. Accounting for Leaving the Break Room: Work Obligations as a Resource in Transitions from One Activity to Another at the Workplace. *Journal of Pragmatics* 171: 178–199. https://doi.org/10.1016/j.pragma.2020.09.032.

Spets, Heidi. Submitted. Intersubjective Interaction During the Word Explanation Activity in Social Virtual Reality. In *Complexity and Intersubjectivity: Shared Understanding in Multi-layered Social Interaction*, edited by Pentti Haddington, Tiina Eilittä, Antti Kamunen, Laura Kohonen-Aho, Iira Rautiainen and Anna Vatanen.

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

Vatanen, Anna. Submitted. Embodied Noticings as Repair Initiations: On Multiactivity in Choir Rehearsals. In *Complexity and Intersubjectivity: Shared Understanding in Multi-layered Social Interaction*, edited by Pentti Haddington, Tiina Eilittä, Antti Kamunen, Laura Kohonen-Aho, Iira Rautiainen and Anna Vatanen.

Vatanen, Anna, and Pentti Haddington. 2021. Multiactivity in Adult-Child Interaction: Accounts Resolving Conflicting Courses of Action in Request Sequences. *Text & Talk* (ahead of print). https://doi.org/10.1515/text-2020-0165.

Vatanen, Anna, and Mirka Rauniomaa. 2021. Laadullinen tutkimus työ- ja perhe-elämän yhteensovittamisen käytänteistä [A Qualitative Study on Practices of Reconciling Work and Family Life]. Presentation at *Sosiaalipolitiikan päivät*, Turku, 28 October 2021.

Vatanen, Anna, and Mirka Rauniomaa. In preparation. Practices of Combining Work and Family Life: A Single-case Analysis of Transitioning Between Work and Family Activities During Remote Work at Home.

Vatanen, Anna, Heidi Spets, Maarit Siromaa, Mirka Rauniomaa, Paul McIlvenny, Tiina Keisanen, and Jacob Davidsen. 2021. New Methods in Video-based Qualitative Research: Remote Collaborative Analysis of 360° Video in Virtual Reality. Workshop at *the 6th Nordic Interdisciplinary Conference on Discourse and Interaction* (an online conference), Uppsala, 19 November 2021.

Wienrich, Carolin, K. Schindler, Nina Döllinger, Simon Kock, and Ole Traupe. 2018. Social Presence and Cooperation in Large-Scale Multi-User Virtual Reality – the Relevance of Social Interdependence for Location-Based Environments. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*: 207–214. https://doi.org/10.1109/VR.2018.8446575.

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