DT2140 Multimodal Interaction and Interfaces
Royal Institute of Technology KTH

A comparative study of video call, VR and Telepresence robots in group interactions

Amalia Larsson Berglöf
Celine Mileikowsky
Nike Backman Eriksson
Sruti Bhattacharjee

SUPERVISOR
André Tiago Abelho Pereira

January 14, 2021

Contents

Abstract

Remote communication has for many years been used as a means of communication when meeting in real life is difficult. During the covid-19 pandemic, many have been forced to conduct group discussions and meetings using remote solutions. This paper explores how different technologies of remote communication affect group interaction and creativity. To evaluate how the different methods of remote communication affect group interaction and creativity, a comparative qualitative study was performed. Two groups of participants communicated with a remote party using either a Zoom video call, a FurHat Telepresence robot, or in a virtual reality (VR) meeting. The subjects engaged in a conversational task focused on creativity. The results from the study indicate that there are perceived differences in how easy it is to communicate and be creative on different platforms. Using a video call was described as less distracting, but lacking the feeling of the remote participant being physically present. Telepresence robots provided the feeling of physical presence in the room, but the lack of emotion on the robot's face was unnerving. VR-based communication was described as distracting and the discrepancy between the avatar and the real bodies of the participants was disconcerting, however, users found the added movements and gaze useful for group discussion.

A video showing the experiment can be found at `https://vimeo.com/500551973`.

1. Introduction

During the ongoing covid-19 pandemic many aspects of life that previously required physical presence, such as going to work, school, or even spending time with friends, are as a result of social restrictions now conducted fully or partially on a distance. The transition from physical to digital presence poses a challenge for many organizations throughout different realms of society, where they need to implement technical solutions for remote communication that allows for their daily activities to function similarly as to before the outbreak. There are limitations, as well as possibilities, within existing technologies for remote communication. The former vice president of innovation and creativity at Disney, Duncan Wardle, thinks that Zoom kills creativity and suggests new solutions built on video games[16]. The study conducted in this paper will compare three existing platforms for remote communication and how they affect creative work: video conference calls, virtual reality, and a telepresence robot.

In particular, we will focus on how these technologies impact group interaction and creativity, where we define group interaction as the interaction between more than two participants. The most commonly used technologies for group interaction as of today are video conference call platforms (such as Zoom, Microsoft Teams, Google Meet, or Skype). However, these video conference call platforms are missing the essence of physical communication, which at times results in deficient communication. Thus, it is of high interest to evaluate other platforms for remote group interaction and to explore how they compare to the more commonly used video conference call. Our objective is to get a deeper understanding of which communication modalities are important in remote group interaction and creative work, leading to improvements in the field of remote interaction.

2. Background

2.1 Group communication

According to O'Conaill et al. [10] communication is a complex process, where listener feedback (auditory backchannels such as confirming utterances, and visuals such as head nods) plays an important role. Without it, the speaker might feel that the message did not come through and reiterate, and also lead to delayed turn-taking between participants. Turn-taking refers to when the participants change who holds the floor. This is done through intonation, syntactical and non-verbal as well as directive cues, both from speaker and listener. Non-verbal cues can be things such as gaze, facial expressions and postures. During a normal face-to-face conversation, all members in the conversation have equal access to the conversation - they share the same physical space, the transmission isn't delayed. Feedback can be created at the same time as someone is speaking and multiple modalities cocreate the conversation. Remote conversation, on the other hand, often lacks non-verbal cues, making it more difficult to achieve smooth and natural communication. This is particularly prominent in the interaction between more than two participants.

2.2 Remote communication and creativity

Some studies on remote communication, and how it affects creativity, has been made. In the work by Biasutti [2], the author studied how creativity is expressed and supported when music is composed online in an asynchronous virtual environment. The results showed that collaboration and identity were important factors for participation and interest in the work. Also, verbal collaboration was used to democratically develop ideas, and nonverbal collaboration happened during empathetically attunement (such as collective flow) in the participant group. Olson et al. [11]recreated a face-to-face group meeting, but remotely, and studied what mix of video and audio is useful for small groups doing creative work. For example, the screens were positioned so that participants preserved eye contact. Their results showed that when the video was turned on, the quality of the work was as good as that face-to-face, and with audio only the quality suffered small but significant. Without video, the experienced discussion quality was poorer, and the participants could not tell how other group members reacted. To summarize, what seems to be important for remote creative work, is the ability to tell how the other members react to ideas, find a collective flow, but that you're also available to express your identity.

Important to note, however, is that many papers touch upon the remoteness and the changes it means for the work-life balance, which can have a positive contribution to creativity. Workers can keep their career despite moving and have improved work-life balance [2, 7], and due to this their creativity is enforced. This paper is not directed at researching these kinds of effects.

2.3 Video conference calls

In a study by Johson et al.[8], turn-taking and the quality of doing interviews in-person, by telephone, or via Skype was researched. Turn-taking was defined as the statements made by interviewer and interviewee. The results showed that the Skype interviews lacked richness in information from the participant, even though the length of the interview was the same as in the other conditions. The results also showed that in-person interviews produce a larger conversation turn number than the other conditions.

There are already some studies on using Zoom as a qualitative research interview tool. The study by Archibald et al.[3] found that Zoom was rated higher than other interview tools among nurses, such as face-to-face or telephone. They found that Zoom had the advantages of rapport, convenience, and user-friendliness. Facial expressions can be picked up and built upon during interviews, giving richer results. But the participants also reported that the technical difficulties sometimes disrupted the conversation, that the sound or video was lagging, or that they had bigger disruptions, up to several minutes, in the call. Another study by Gray et al. [5] researched what attributes Zoom has that contribute to the depth that qualitative interviews need. Things that the participants liked was that they could see the interviewer. However, it's noted as a disadvantage that the interviews are not held in the same physical space; resulting in missing cues such as body language. In the work by O'Conaill et al.[10] where two different video call systems are evaluated concerning face-to-face conversation, they address the fact that simultaneous speech is something that can occur both in face-to-face communication, as well as video calls. Simultaneous speech can overlap as well as interruptions. Where interruptions are more directed at taking the turn from the current speakers, overlaps happen when there's a reason for the person to take the turn, such as helping someone complete a sentence. O'Conaill et al. also hypothesize that users will, in video calls where the transmission is lagging, make more explicit handovers to communicate that they are done speaking. The authors define conversational turns as attempts by a speaker to gain the conversation floor, and they hypothesized that during transmission delays and lower quality, the turns will be longer. The results showed that the two video call platforms had lesser backchannels and less total interruptions than the face-to-face condition, and also that turn-taking becomes more of a slow processor directed by handing over the turn.

2.4 FurHat and telepresence robots

For Telepresence in this experiment, the robot FurHat is used[1]. FurHat is a social robot, which can emulate facial expressions, including lip movement, using a mask backlit by a projector and motors moving its head in 3 degrees of freedom. As a telepresence robot, the FurHat can experimentally be controlled by an Oculus Quest. Previous research [13] has shown that co-presence is increased with the use of telepresence robots compared to video call and that communication tends to be more efficient. The use of telepresence robots also decreases the amount of cognitive load put on the physically present participant, whereas video calls put less of a cognitive load on the participant accessing remotely [13]. It has also been shown that the ability to decide the field of vision is important for the feeling of co-presence and immersion for a remote party[17]. In a study by Keller et al.[9], students were given a tour of a university either in person or via a telepresence robot. The results of

the study indicated that there was no difference in how much affinity the physically present guide felt towards the participants, but that the participants accessing the tour through the telepresence robot felt a higher degree of affinity for the guide than the participants physically present. One study[14] showed that using telepresence robots that displayed body motions can enhance the perception of social presence, but that the use of a robot that did not mimic the remote participant's appearance had a negative effect compared to the use of a virtual avatar with the appearance of the remote party in the aspect of. This resulted in a similar performance in the aspect of social presence for virtual avatars and telepresence robots[14].

The work by Gudmandsen[6] researched the difference between the usage of the Furhat robot for telepresence and a video call. This was evaluated through areas such as who was being spoken to and detecting your turn to speak. The Furhat robot was operated by participants. The results for the participants not directing the call/robot showed that using the FurHat made it easier to understand whom the operator was speaking to and gave an increased perception of eye-contact with the operator. However, using FurHat produced lower scores for understanding the operators' facial expressions. For the operating participant, FurHat received a higher score in the perceived presence and for the feeling that the participant was aware of whom the operator was talking to. With these findings in mind, Gudmandsen argued that the robot improves presence.

2.5 Virtual Reality

Villani et al.[15] simulated a job interview and compared the feeling of presence in VR compared to a real setting. "Presence" was in the paper defined as the "disappearance of mediation" and the results showed that the experienced presence was higher during the virtual interview. Campbell et al.[4] did a case study where virtual reality meetings were compared to videoconferencing, where two participants at a time met either using virtual reality or Skype. The results showed that overall, meeting in virtual reality environments improved the feeling of presence, closeness, and arousal, even though some notable differences between male and female participants were found.

3. Method

Table 3.1 presents the communication modalities that are supported through each platform - by the operator as well as the participant.

3.1 Recruitment

A total of four participants, divided into groups of two, were recruited. Initially, six participants were planned to partake, but when two participants canceled because of sickness it was deemed appropriate due to the circumstances surrounding the covid-19 pandemic that no new participants should be recruited to take their place. The participants consisted of two men and two women, ranging in ages from 22 to 27 years old. Participants in one group were previously acquainted with each other, while those in the other group had never met before the experiment.

3.2 Set-up

Three platforms for remote communication were prepared for the participants to complete the experimental task within. The conditions and equipment for each platform are described in detail below.

3.2.1 Video conference call

Two participants were placed in chairs approximately 1,5 meters between each other. On a table before them, approximately 1,5 meters, a table with a laptop was placed. In another room, an experimenter was seated, who joined the two participants on a video conference call on Zoom. The two participants could see and hear the experimenter through the laptop placed before them, and the experimenter could see and hear both of the participants. Thus, the experimenter and the participants could communicate through the use of upper body movements, facial expressions, and speech.

Two other experimenters were placed in the room with the participants, where one gave instructions for the experimental task, and one documented the process in pictures and video. Additionally, a camera was placed on a tripod to record the experiment.

3.2.2 Telepresence robot

The participants were placed next to each other in the same manner as in the video conference call setting. On the table, before them, the robot FurHat was placed. An experimenter in

Figure 3.1: Modalities on different platforms.

vTime XR (Virtual Reality)		FurHat (telepresence)		Zoom (Video conference)	
Operator	**Participants**	**Operator**	**Participants**	**Operator**	**Participants**
Upper body movements, including partial hand gestures	Upper body movements, including partial hand gestures	Head movements	Upper body movements, including full hand gestures	Upper body movements, including full hand gestures	Upper body movements, including full hand gestures
-	-	-	Facial expressions	Facial expressions	Facial expressions
Shared virtual environment	Shared virtual environment	Shared physical environment	Shared physical environment	-	Shared physical environment (partial)
Synchronised audio	Synchronised audio	Synchronised audio	Synchronised audio	Synchronised audio	Synchronised audio
Gaze (partial)	Gaze (partial)	Gaze (partial)	Gaze	-	Gaze

another room joined the participants by teleoperating FurHat. This was achieved through the use of a VR-headset with which the experimenter could see the participants from the robot's point of view through the camera on FurHats head. The VR-headset captured the movements of the experimenter's head, as well as the lip movements of the experimenter's speech. The participants could hear the experimenter through a speaker placed next to FurHat, and the experimenter could hear the participants through headphones connected to a microphone in the experiment room. To summarize, the head movements, sound, and lip movements from the experimenter were embodied through FurHat, who was physically present in the same room as the participants.

In the first out of the two experiments, there was an issue with the audio, so that the sound from the operator did not transfer into the experiment room as intended. To pursue the experiment the sound of the operator was transferred through a phone call, where a receiving phone with the speaker on was placed on the table next to FurHat. This did, however, lead to the audio and lip movement synchronization on FurHat stopped working midway through the experiment.

3.2.3 Virtual reality

The participants and one experimenter were seated in three different rooms, where they could not physically see or hear each other. They were all given a VR-headset, as well as handheld controls. Through the use of the social platform vTime XR the participants and the experimenter met in a virtual room. vTime XR was selected over other virtual reality social platforms because it offers easy-to-use interaction where participants are seated down so that the setting is similar to the other experiments. In the virtual room, the participants could see and talk to each other as 3D-animated avatars. The avatars were, however, not customized

after the person using them, and therefore did not visually resemble them. The avatars were seated down, and while they were not mobile in the sense that they could move about in the virtual room, they could use the upper body and hand movements. The first experiment took place in a virtual environment where participants were placed on chairs "floating" in space. The participants were placed so that they faced each other, but they could look out at their surrounding environment, which was filled with stars, comets, spaceships, and other space-related objects. They could also hear environmental sounds from events going on, like a spaceship passing by. The second experiment took place in a virtual office space, in what looked like a meeting room. Participants were placed in chairs around a table, facing each other, but just as in space, they could look out on and listen to ongoing events. The meeting room was placed high up in a building with a view over a virtual city, with helicopters flying by.

3.3 Experimental task

The experimental task consisted of a storytelling exercise, where the participants were given several parameters to create a story around. In each exercise, they were given a setting, three characters, three objects. They were also given a scenario, which they were prompted to explain how it came about. A total of six exercises were created, with a wide range of settings, characters, and objects. They were designed to stimulate creativity within participants.

Example exercise:
Setting: a busy day on a street in New York
Characters: a middle-aged man, the man's boss, and a talking cat
Objects: a large bag of unidentified, white powder, a birthday cake, and a pair of bloody scissors
Scenario: the middle-aged man finds himself sitting in the middle of the busy street. He is wearing a suit and tie, which are covered in white powder. What has happened?

3.4 Procedure

In each one of the three settings described in section 3.2, the participants and the experimenter joining them were given instructions for the experimental task. One experimenter acted as an instructor and briefed the participants of the purpose of the task: that they were partaking in a group storytelling exercise, where one person (an experimenter) was joining the group remotely through the use of different technologies. The experimenter who joined the group remotely was an active participant in the task and partook in the exercise on the same terms as the participants. The instructor described the exercise at hand (see section 3.3) and asked the participants to engage in discussion. Participants were given an approximate time frame of 10 minutes, but once the conversation seemed to simmer down, the instructor informed them that the task was completed. This occurred in approximately 5-10 minutes for all exercises.

A total of four participants partook, as well as two experimenters. There were two sessions, each with two participants and one experimenter. The order of the exercises in the two ses-

sions was the following:

Session 1: Zoom, FurHat, VR
Session 2: FurHat, Zoom, VR

Originally, the task order was supposed to be the same as in session 1 in both sessions. This order was selected due to that it was deemed to progress from platforms that the participants were likely familiar with, into platforms that were further away from commonly used distance communication tools as of today. However, due to previous technical issues with FurHat's audio, the order was changed in the second session, to ensure that the audio would work properly.

3.5 Evaluation

After each completed exercise, participants were asked to fill out a questionnaire about their experience. First, there was an initial questionnaire with questions about the participants age and previous experience with the three different technologies used in the experiment. For each technology used, there were ten questions about how participants experienced completing the experimental task while being aided by that specific technology. The questions were answered on a seven-point Likert scale, and they touched upon topics such as the naturalness of the conversation, engagement, perceived presence, and creativity.

Once the final questionnaire was filled out, the participants partook in a semi-structured interview where they were asked questions about the experience as a whole. They were asked to reflect upon what they liked and did not like in the different settings, if there was a setting they preferred over the others, how the interaction between participants compared between settings, and how creative they felt like they could be. For the full interview questions, see Appendix A. Note that since the interviews were semi-structured, they did not necessarily follow the interview questions entirely, and some questions may have been left out or altered.

4. Results

4.1 Interviews

4.1.1 VR

Participants generally thought that VR was the most fun to use. VR was also described as being "the most difficult to focus and actually listen in" and by many all of the participants as being distracting. One participant, who felt that VR was the least suitable for creative tasks commented " there was so much going on. [...] because you hear the sounds [from events in the surrounding environment], you intuitively look [..] then I missed the discussion and what was going on". One male participant noted about the design of the VR avatars that "I had a bosom" and that in part due to this "it was really distracting that [the avatar] wasn't like my body", and that the VR environment itself was distracting from the conversation. There were other comments on the fact that the avatars differed from the real people, according to one participant "Everyone looked kind of similar in VR" and "that's not you, [...], I hear your voices, but it doesn't feel like I'm actually talking to you".

The participants overall felt like it was clear to tell who was being addressed, "It was very clear [where attention was directed] because that person turned their body back and forth. But it was still difficult, because you lose facial expressions. Even if someone has their body directed towards you, they could still be focusing their attention on something that's over your head, and that's two very different things"'. Another participant commented on the seating of the avatars, and how that made it easier to understand the handovers and turn-taking; "I was seated on one side of the table, and the two others were placed on the opposite side. So for me it was very clear if they looked at me, but maybe it was more difficult for them to interpret who I was talking with."

The difference between the real world and the VR environment was also noted by one of the participants as they commented on the fact that in the VR environment they had a table in front of them, but it was not present in the real world and the participant said they would have "felt more present during the meeting" if there was not this discrepancy.

Another thing that one participant noted was how they wanted to communicate, such as using facial expressions, but that they were halted by the headset being too intrusive, and also the knowledge that their expressions wouldn't be shown to the others. "... we have this headset that kind of weighs down your whole face, so that even if you want to use facial expressions just as being part of you talking for yourself, you're haltered by that you have something in your face [...] and then also it isn't shown to the other participants".

There were also comments saying that the difficulties might have been due to the fact that the participants weren't acquainted with the specific VR-environment where they were placed: "... it was the first time you were in that environment... and that I saw the characters for the first time and things like that". Some participants believed that the results would have

been better if they had gotten some more time to get used to the environment, or if they got to do the task again in the same environment: "...if I had got to put on [the VR headset] and see everything before maybe, then I might have thought it was less exciting [...] I think that I would like to try VR again, that if you were more familiar with the setting you were in... I think it would have gone better then."

4.1.2 Zoom

All of the participants had used Zoom previously. One participant noted that communication via Zoom was "a skill [I] already have" from using it almost every day. This participant also answered they would prefer to do this kind of creative meeting over Zoom, with the reasons being it was the platform they felt most engaged in and there were no distractions. When asked which platform they felt most creative in, this participant answered Zoom, but gave the reasoning that "not sure if it was because it made me feel creative, rather [...] I felt like here it's not as much that distracts my train of thought". Another participant noted that even though they had a slow start on Zoom as it was their first test case, it was easier for them to "focus a lot more on the actual task" due to the lack of distractions.

Other participants noted about Zoom that they feel like they are "not given any physical presence by that screen" due to the fact that "everyone's webcams are so bad and you don't get the perception that people are there in the room". This made the participant "think that I directed all my talking towards [participant 1] when we were sitting in the Zoom meeting. It was almost like I didn't notice that [operator] was in the room". One other participant had the opposite opinion, they felt like they could sense the presence of the remote party in the room. Two of the participants noted that they seemed to talk more to each other than the remote party either regardless of the platform used, and that it might be hard to hold a conversation due to this, but that they felt it was easy to listen to someone talk over Zoom. One of them put it like " I think that having a combination of certain people physically present in a room, and some on Zoom or on Furhat, is not a good idea, because naturally, I would rather talk to the person physically present in the room".

One participant noted there is lag and delay when using Zoom, which they do not like. They also noted that in contrast to the other platforms, in Zoom "If there are many people in a Zoom meeting it is harder to address a particular person".

One participant reflected on that if the remote participant directed their attention at something else, they felt a need to know what they were looking at: "In Zoom everyone looks at their own screen, and it's difficult to achieve shared attention in that sense [...] you're kind of like, okay, I can see that you're looking at something over there, what are you looking at?"

4.1.3 FurHat

Two of the participants preferred FurHat for the discussion. One of them described it as "the discussion was more free flowing with FurHat I think". The other noted that the discussion was better and that it was the "easiest" to use.

One participant noted the head movements improved the sense of copresence: "with Furhat, it was still that [operator] turned her head, and then you felt that she was there", and "you feel more that you sit in a physical room", and thanks to the head movements "you felt more

physically addressed like when [FurHat] turned to you". Another participant said "the mouth is moving [...] the head is turning, but you don't get any emotional expressions from the eyes", and that this was something they did not like with FurHat. Important to note, which the participant also reflected on herself, is the prior knowledge of the person operating the FurHat. "Because we know [the operator] right, I could like hear her voice and like feel the feelings around [what the operator said].... But if I were to talk to someone that I didn't know via FurHat it would probably be more stiff." The same participant noted that they talked less to the remote participant because of this, as "it felt like [it was not really] a human that was sitting there so I always turned to [Participant 2]". Another participant in the other experiment expressed a similar thing. That the robot was just a head and not a full-body, and also was light up in a blue color, made him think the following. "I want to direct my attention there, but at the same time it's kind of, this uncanney valley issue, like do I actually interpret this as a real person?".

Two participants answered they felt it was easiest to be creative in the FurHat condition. One said it might have to do with this being their first task, and the other that FurHat "was enough to be creative" without being as distracting as the other conditions.

Due to some technical difficulties, the sound for the FurHat condition was broadcasted to participants via a phone put on speakers in the first experiment. This affected the sound quality negatively, and the problems also resulted in the lip-sync of the robot being turned off halfway during the experiment. The participants both mentioned this during their interview, noting that it was "easier to interpret where the head was looking when it was also talking" and that they experienced a "clear connection" of where to divert attention when the sound coincided with the lip movements of the robot.

4.2 Questionnaire

Four participants with an average age of 23.5 rated the three technologies out of 7. All participants had used VR and Zoom before, and 2 out of the 4 participants had seen or used FurHat before. The data collected in the questionnaire can be found in figure 4.1.

Figure 4.1: Results from the questionnaire.

Questionnaire	Zoom	Furhat	VR
How easy did you find it to be creative during the task?	4,75	6,25	5,5
How much did it seem as if you and the people in your group were together physically in the same place?	2,5	5	3,75
During the conversation how well were you able to observe the body language of the other participants?	4	5,25	4,625
How well did you understand who the other participants were addressing?	5.75	5	5,375
How well were you able to determine the mood of the remote participants?	5.5	5	5,25
How natural did the interaction with the other participants feel?	4	5,25	4,625
How close did you feel to the other participants?	4	4	4
To what extent did you feel the technology used was distracting you from the conversation at hand?	3,5	2,5	3
To what extent did you feel that you could express yourself and take part in the conversation as you would if it were a face-to-face conversation?	3,5	5,75	4,625
To what extent did you feel as though all participants could equally contribute to the conversation?	4	6	5

5. Discussion

Although the overall results don't point to a clear preference for any single platform in terms of creativity, there are still interesting findings, which will be discussed below.

5.1 Prior knowledge and Distraction

The participants' thoughts on and preferences for the different platforms seems to partly have relied on two important factors: prior knowledge of the technology at hand, and how distracting they deemed them to be.

All of the participants had used Zoom prior to the experiment and were very acquainted with it. One participant expressed clearly that they felt the most creative when using Zoom, and one other participant agreed that it was the setting where they felt the most focused. These conclusions did, however, not seem to stem from the communication modalities within the platform itself: on the contrary, most participants took note of issues when communicating with the remote participant, such as lack of their perceived presence and difficulties to address them. Instead, their preference mainly lied in their prior knowledge and experience with using the tool. This allowed participants to keep focused on the given task, and in that sense, to be more creative. However, the Zoom call wasn't entirely free from distractions. One participant commented that they felt as it was distracting not knowing where the remote person was directing their attention, or what they were looking at. Not being able to see what goes on in the remote parties' surrounding environment takes the focus from the conversation at hand. This can be seen as a consequence of the lack of a shared physical space, a concept of interest for further research within technologies for remote communication.

As for VR, the participants' experience was quite the opposite. While they generally enjoyed the experience and thought that the platform was fun to use, there was consensus among the participants that they were very distracted and therefore had difficulties in focusing on the task at hand. All participants made comments saying that they were distracted by events going on in the surroundings of the virtual environment, which led to less focus being directed towards the conversational task. They looked at, listened to, and turned their virtual avatars towards objects and events in their surrounding environment. Apart from being distracted yourself, observing this sort of body language in others indicates that there's something more important happening than the conversation at hand. It seems like, in this case, that accessing certain communication modalities (upper body and hand movements) was obstructing creativity rather than promoting it. It is possible that the results would have been different in other virtual environments with fewer distractions in the surrounding environment: one participant even commented that they would probably have performed better in a more "boring" environment. Also important to note, is that the software used for the VR did not support any convenient way to quickly let 3 participants join the same room so that the participants had to find the correct user and connect to them before they could

join the room. Since this was the first time the participants used this particular software, this took some time and effort. This, in combination with the distractions participants noted due to their inexperience with the system, emphasizes the need to learn the system to use it efficiently, which is an important issue if the software is not used often enough to warrant learning.

In terms of prior knowledge, comments were most neutral when it came to FurHat. Two participants were previously familiar with FurHat, but there were no major differences between their comments compared to those from the participants who weren't previously familiar with FurHat. Unlike both VR and Zoom the comments from participants didn't circle around their previous experience, or inexperience, with the technology at hand. There also were, generally, few comments regarding participants feeling distracted while conversing with FurHat, which is also reflected in the questionnaire rating regarding distraction, where FurHat was rated the least distracting with an average rating of 2.5. Two participants felt as though they could be the most creative when interacting with FurHat. It should, however, be noted that these participants were previously acquainted with the experimenter who operated FurHat.

5.2 Embodiment

The quote from one of the male participants regarding having a body that conflicted with their own and how it distracted them from the discussion is interesting. A study by Rae et al. [12] showed that control of the embodiment is something that can build trust among participants. Further, the previous knowledge of the people you're attending the meeting with might have had an impact; since they had met each other physically before the VR meeting, and the avatars virtually did not represent the other participants, this can have resulted in the feeling "that's not you". This discrepancy between the physical embodiment and the virtual can affect how the participants used the different modalities to communicate, as well as affected the creativity; if you don't feel like you are you, you communicate differently and become distracted by other trails of thoughts.

It's also interesting to discuss one participant's view that a screen or face might not be enough embodiment for the attending operator to feel present, and that in both the FurHat condition and the Zoom condition, it was easier to direct the word towards the other participant. In both of the conditions, the participants saw more of the upper body and face, and the lack of a full-body and full body gesture might lead to you attending that person less. Another participant also talked about this regarding FurHat, that it gave him an uncanny valley feeling; he wanted to direct himself towards the robot but it also felt unnatural.

The lack of physical embodiment in the Zoom conditions resulted in less communication with the remote participant according to one participant, and another noted it is hard to address a specific person when there are several people in the Zoom conference call. With FurHat, the remote participant is embodied through the robot, and the physical movement of FurHat's read gave a good indication of whom the remote participant was talking to. This was true even for the case when the lip-sync lagged, which indicates that the embodiment of the head movements is important for the sense of presence.

5.3 Methodological impacts

Considering the restrictions by the ongoing covid-19 pandemic, as well as the short time frame of the project, it was decided that the number of participants should be limited. Even though we received exhaustive and interesting answers from participants, the results would have been more reliable if there had been a larger number of participants.

Due to both the small number of participants, as well as some technical issues, there was no proper counterbalancing of the order of the experiments. This likely affected the results, as both of the group expressed that they felt the most creative in the first platform where they tried, which they also touched upon in the interviews.

6. Conclusion

Through a qualitative study of how three platforms that support remote communication (Zoom, VR, and FurHat), the platform's impact on creative group communication has been researched. The results spoke clearly to the fact that to feel creative, one must not be distracted. At the same time, the results point towards the need for physical embodiment for equal contribution among participants in remote group communication. Adding modalities that allow for a higher degree of physical embodiment, such as upper body and hand movements, gaze, and facial expressions did not necessarily increase the perceived creativity in group interaction but did however increase the perceived naturalness, flow, and engagement. This poses a challenge for the future use of remote technologies: while they should not be too complicated or perceived as distracting, they should still offer sufficient embodiment of remote participants to create an engaging experience.

The results also emphasize the importance of familiarity with the system used in order to promote creativity. While participants noted several weaknesses caused by the lack of sufficient communication modalities via Zoom, such as the remote person not feeling present and difficulties with addressing them in the conversation, two out of the four still meant this is where they felt the most creative, solely based on their familiarity with the system. This is important to note for enterprises or organizations considering implementing other systems than video conference calls for remote meetings. Time and effort might be needed to learn a new system. In addition, implementing other solutions can also be costly.

The number of participants was smaller than hoped, but this was deemed necessary due to the ongoing covid-19 pandemic. Similar studies on remote group communication and creativity should be conducted, with a larger number of participants. Even though the results of this study are an interesting, introductory exploration of the topic, more data should be gathered. This is necessary to produce a more in-depth analysis of how these platforms can support, or obstruct creative group communication. Future research should continue to explore the possibilities of limitations of remote group communication since this will likely be a recurring phenomenon even in a post-pandemic world.

Bibliography

[1] Furhat robotics. `https://furhatrobotics.com/`. Accessed: 2020-11-30.

[2] Creativity in virtual spaces: Communication modes employed during collaborative online music composition. *Thinking Skills and Creativity*, 17:117 – 129, 2015.

[3] Mandy M. Archibald, Rachel C Ambagtsheer, M. Casey, and M. Lawless. Using zoom videoconferencing for qualitative data collection: Perceptions and experiences of researchers and participants. *International Journal of Qualitative Methods*, 18, 2019.

[4] Abraham Campbell, Thomas Holz, Jonny Cosgrove, Mike Harlick, and Tadhg O'Sullivan. *Uses of Virtual Reality for Communication in Financial Services: A Case Study on Comparing Different Telepresence Interfaces: Virtual Reality Compared to Video Conferencing*, pages 463–481. 01 2020.

[5] Lisa Gray, Gina Wong, Gwen Rempel, and Karen Cook. Expanding qualitative research interviewing strategies: Zoom video communications. *Qualitative Report*, 25:Article 9, 05 2020.

[6] Magnus Gudmandsen. Using a robot head with a 3d face mask as a communication medium for telepresence, 2015.

[7] Philip Hunter. Remote working in research: An increasing usage of flexible work arrangements can improve productivity and creativity. *EMBO reports*, 20(1), January 2019.

[8] David R. Johnson, Christopher P. Scheitle, and Elaine Howard Ecklund. Beyond the in-person interview? how interview quality varies across in-person, telephone, and skype interviews. *Social Science Computer Review*, 0(0):0894439319893612, 0.

[9] Lisa Keller, Kevin Pfeffel, Karsten Huffstadt, and Nicholas Müller. *Telepresence Robots and Their Impact on Human-Human Interaction*, pages 448–463. 07 2020.

[10] Brid O'Conaill, Steve Whittaker, and Sylvia Wilbur. Conversations over video conferences: An evaluation of the spoken aspects of video-mediated communication. *Human-Computer Interaction*, 8:389–428, 12 1993.

[11] Judith S. Olson, Gary M. Olson, and David K. Meader. What mix of video and audio is useful for small groups doing remote real-time design work? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, page 362–368, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[12] Irene Rae and Leila Takayama. In-body experiences : Embodiment , control , and trust in robot-mediated communication. volume 15, 04 2013.

[13] Gabriella Rydenfors. Telepresence and remote communication through virtual reality. Master's thesis, Linköping UniversityLinköping University, Media and Information Technology, The Institute of Technology, 2017.

[14] Kazuaki Tanaka, Hideyuki Nakanishi, and Hiroshi Ishiguro. Comparing video, avatar, and robot mediated communication: Pros and cons of embodiment. In *Collaboration Technologies and Social Computing*, pages 96–110, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg.

[15] Daniela Villani, Claudia Repetto, Pietro Cipresso, and Giuseppe Riva. May i experience more presence in doing the same thing in virtual reality than in reality? an answer from a simulated job interview. *Interacting with Computers*, 24:265–272, 07 2012.

[16] Duncan Wardle. Why Zoom isn't the future of work, url = https://www.ozy.com/the-new-and-the-next/why-zoom-isnt-the-future-of-work/393371/, urldate = 2020-12-20.

[17] J. Young, T. Langlotz, M. Cook, S. Mills, and H. Regenbrecht. Immersive telepresence and remote collaboration using mobile and wearable devices. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1908–1918, 2019.

7. Appendix A

Introduction

Finally, we're going to conduct an interview with you to sum up your participation(s), your thoughts about the study that the questionnaires maybe didn't give place to discuss further, and debrief you.

General comparison

Which setting did you enjoy the most, in terms of the task given to you?

How did you experience interacting as a group in VR, compared to the other settings?

How did you experience interacting as a group with the telepresence robot, compared to the other settings?

How did you experience interacting as a group in a video conference call, compared to the other settings?

If you would do this task again, which setting would you prefer?

Creativity

In which setting did you feel as though you could be most creative when creating the story, and why?

For a creative group task like this, which platform would you say is most fitting?Why?

Which is least fitting? Why?

Communication modalities

Did you notice any changes in how you communicated with the others in the different platforms? If so, how?

Debriefing/cool down

Do you have anything else to add?