The Configuration of Space: Probing the Way Engagement, Empathy, and Awareness are affected by Task-specific Spatial Representations in Online Video Communication

ANONYMOUS AUTHOR(S)

Fig. 1. Study set-up for four conditions. (a) Social support - Room condition, people sit in a circle in the background of one classroom. (b) Social Support - Zoom condition: people talk in a block-like Zoom-style interface; (c) Topic Discussion - Room condition, people sit in two different lines in the background of one classroom. People in each line hold the same opinion. (d) Topic Discussion - Zoom condition: people talk in a block-like Zoom-style interface

Humans live and act in 3D space, but often work and communicate on 2D surfaces. The prevalence of online communication on 2D screens raises the issue of whether human spatial metaphors affect our capabilities, social perception, and behaviors when interacting with others in 2D video chat. How do locations, arrangements, and contexts subtly influence how we communicate online, in for example, social support and topic discussion purposes? Using Ohyay.co as a platform, we compared a block-like Zoom-type interface with a scene-based Room-type interface where participants are located in circular arrangement on screen in a social support task, and found that participants allocated attention to the group as a whole, and had greater self-awareness in the Room format. We then chose a two-sided topic for discussion in the Zoom interface and the Room interface where participants on each team face-off against each other, and found that they utilized spatial references to orient their allegiances, finding greater engagement with those farther away in digital space and greater empathy with those closer, in the Room over the Zoom format. We found spatial effects in the way participants hide from the spotlight, in perspective-taking, and in their use of expressive gestures in time on the screen. This work highlights the need for considering spatial configuration in 2D in the design of collaborative communication platforms to optimize for psychological needs for particular tasks.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**; *Empirical studies in HCI*.

Additional Key Words and Phrases: video conferencing, spatial influence, video analysis, seating arrangement

ACM Reference Format:
Anonymous Author(s). 2022. The Configuration of Space: Probing the Way Engagement, Empathy, and Awareness are affected by Task-specific Spatial Representations in Online Video Communication. *Proc. ACM Hum.-Comput. Interact.* 1, 1 (January 2022), 24 pages. https://doi.org/XXXXXXX.XXXXXXX

2573-0142/2022/1-ART $15.00
https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

When interacting with other people in face-to-face communication, space plays an important role in providing environment cues that help human perceive the surroundings. The environment cues can be the layout of objects seating arrangement which affects the way people perceive each other and effectively communicate with each other [36, 45, 54]. The cues can also come from proxemics (the space around an individual) which studies how the different distances between people [21] bring different types of interpersonal communications.

Moving from 3D physical space to the 2D screen and shifting from face-to-face communication to computer-mediated communication (CMC), the influence of space and social dynamics also change. The affordances of space changed since the video only provide limited environment cues of the remote participants, and only presents part of one's immediate environment on the screen [17]. Social interaction changed since the way people send and receive cues from conversations differently in the virtual environment. Researchers hold different views on how might these cues be picked up differently in CMC compared with physical space, including a "filtered-cue" view arguing that many non-verbal cues are filtered out in CMC [47], and another perspective arguing non-verbal cues like body language can actually help convey social information in CMC [33]. That being said, space and social interactions are recreated and reshaped in the 2D surfaces of the computer screen.

The prevalence of video communication tools during the pandemic immersed people in screen-based communication much more than before. On one hand, people utilized video call to recreate in-person connection [25]. On the other hand, the excessive use of video communication tools also caused exhaustion with the screen-based video communication, such as the "Zoom fatigue" [52]. This raises the issue of whether human spatial metaphors affect our capabilities, social perception, and behaviors when interacting with others in 2D video chat.

To better understand how does the spatial metaphor in online video communication influence the social interaction, we conducted a mixed-method study(Figure 1) including a video chat study with a post-study survey (N=34), and interviews (N=24) with the selected study participants to investigate how the spatial configuration of a video chat space influence people's perception and behaviors with others engaged in different tasks. Every group includes two conditions – social support group (N=18) and topic discussion group (N=16). We set people in Room condition and Zoom condition with an equal number of people. The difference between Room and Zoom is the layout of the video frame and the chat room's background: in Zoom condition, we use a grid layout displaying videos in equal size on a plain background which simulate the space of Zoom; in Room condition, we use a classroom with chairs in it and the video frame "sit" in each chair. For the Social Support, the background of the Room condition is a classroom with a circle chair arrangement; for the Topic Discussion, the background of the Room condition is a classroom with an angular chair arrangement. The post-study survey evaluated participants' social presence, spatial presence, and self-awareness during the group video chat. The interviews are semi-structured investigating participants' experiences during the group video chat study, their perception of self, others, and space, as well as participants' experiences with video chat tools such as Zoom.

Our analysis with the data from the survey, video, and interview revealed that the engagement and empathy level in Room are higher than Zoom under particular arrangement conditions. Our work also found that space can provide references for participants in visual search and identifying their allegiances in the Room condition. In addition, this environment cue helped enhance team awareness for participants in Room with a topic discussion task. Some other patterns and behaviors observed from the study are also discussed, including an intention to hide from the spotlight, and hand gestures using spatial information. Besides, this work highlights design implications of 2D video chat space and considerations for the spatial configuration in satisfying cognitive needs

for tasks with different purposes. Overall, our work contributes to a better understanding of how the configuration of space in online video communication influences the interaction visually and conceptually under different contexts.

2 BACKGROUND AND RELATED WORK

2.1 Video Chat and Video Conferencing from Early Stage to the Pandemic Society

2.1.1 Non-verbal cues in video conferencing. Studies of video communication or computer-mediated communication (CMC) in the early stage investigated how non-verbal cues are conveyed and received compared to face-to-face communication and what are the possible differences in information produced from communication. One argument is the "filtered-cue" model which proposes that a lot of cues that face-to-face communication are filtered out due to the limitation of technology [47]. However, other studies also showed that this non-verbal cues like paralanguage can also help communicate social information in CMC, depending on group size and composition [33]. Some other studies found that non-verbal cues like facial expression or gaze did not influence people's communication because people can use their talk style to compensate for the losing message[8], while others showed that compared to an audio-only setting, people who communicate through a low frame-rate video show more careful articulation [6]. There also researches showing that compared with face-to-face communication, the varying levels of social interaction in computer-mediated communication did not produce significant differences in the quality or quantity of information [51].

2.1.2 Video communication in the pandemic. As video chat (e.g. Skype) becomes popular around the 2010s, it is used on different occasions, ranging from work, family chats to major life events [2, 35]. The COVID-19 pandemic leads to another flourish of video communication technologies in daily life. During this period, people used video chat to communicate with family and friends over distance [25], as well as using video conferencing tools like Zoom for work and class. The excessive use of video communication tools also brought negative effects, including "Zoom fatigue" which is a type of exhaustion with virtual meeting [52]. This brought a critical view on video conferencing tools, raising the awareness on the mental influence of using the video conferencing tool in addition to its functionality. Issues like cognitive load and close-distance eye contact with Zoom are discussed in this context, inspiring further design iterations of video chat [4] .

2.1.3 Space and spatial configuration of video communication tool. Early studies on the affordances of media space shows that the "space" in video communication tool mediated by technology is significantly different from the physical space [17]. Video communication tools only convey a "limited subset of visual and auditory information" compared to the physical space due to the limitation of the medium such as limited field of view and movement possibilities [17]. Harrison and Dourish's work visited the notion of "space" on a more conceptual level by proposing a distinction between and "place" [22] . The original paper held a dualistic view on space and place, focusing on the geometric features of space and the social meaning of place. Later the development of technology in the video communication tool brought opportunities for creating more high-fidelity, immersive space. Later works therefore not only view space on its physical attributes but also as a product of social practices [7, 12] .

Many previous studies have investigated how to improve the configuration of videoconferencing in order to make the experience as similar to the face to face meeting as possible. Approaches taken included achieving mutual gaze [43], capturing spatial cues [9], and enhancing spatial information using multiple cameras with a spatially distorted screen [40, 42].

While these researches focused on how to configure a system in a physical space for supporting the usage of spatial cues, another type of spatial configuration of video chat is implemented or investigated through video chat interface layout. For example, some researchers utilize 360 photo to enhance the contextual awareness [20], and studied the role of video feedback in video-focused and content-focused layout [29]. Some features in recent video conferencing tools such as virtual background are also studied in terms of how it might influence the perception of a person [26].

Several alternatives of the conventional video conferencing tools that emerged during the pandemic also bring new possibilities for the spatial configuration of video chat. These include ohyay.co [32] which is an online platform that allows users to build different video chat rooms, and gather.town, another online platform that combines avatar control in a 2D environment with video chat [16].

2.2 Space and its influence on perception

2.2.1 Physical space's influence on interpersonal interaction. In a physical space, the perception of a social relationship is shaped by interpersonal distance. Hall's theory on proxemics suggests that the distance surrounding a person forms a space [21] . The space with different distances is used for interactions with different levels of social closeness. For example, an interaction at a close distance is for intimate friends, and a further distance is for social with acquaintances. This suggests that in a physical space, people may have different feelings when an individual interacts with them at a different distance.

The spatial configuration of a physical space can also influence interactions among people. The focus of our study is the seating arrangement [36, 45]. Studies show that different shapes of seating arrangement can result in different dynamics in the classroom, for example, changing from row-based arrangement to circle-shape arrangement brought an increase in classroom engagement [15], people who sit in chairs arranged in circular style reveal preferences on family-oriented content compared to angular shape [54]. These studies suggest that in the physical space, the shape of the seating arrangement can be picked up by people and bring subtle influences on their perception and behaviors.

Given these study findings on physical spaces' influence on social interactions and behaviors, when the environment changes from 3D space to 2D computer screen, how the human spatial metaphors shifted and whether the human spatial metaphors similar effects is worth investigating.

2.2.2 Perception of Space. Epstein's work on scene perception provides a possible explanation for understanding how human beings gather and process information from the environment and form their perception of space [14]. According to this theory, when looking at an image depicting a scene, the observer extracts information from the scene properties and categorizes it on different levels. This includes low-level concrete properties (e.g. edges), mid-level elements (e.g. layout), and high-level abstract properties (e.g. category). These information can be used in many cognitive functions including searching an object in the scene and recognizing the context of the scene.

The perception of the virtual space of video chat can be different from perceiving the scene of an image. Firstly, in a video chat, each participant has a different background, which means a person in video chat not only processes information from one scene, but instead processes multiple scenes at the same time. This can bring more cognitive load for people in the video chat [26]. Different spatial configurations of the video chat space also bring in different types of information for people to process. For example, in a Zoom-style layout where the video feeds of each participant are displayed on a plain background, the environment cues mainly come from the background of participants. But in a Room-style configuration where is a background image that presents a scene in the shared space, the scene perception for people in the video chat becomes more complicated as they are

going to process two different types of scenes. Besides, in the Room configuration, the participants are located in the scene in a way that is different from several people standing in a physical space.

2.3 Social Interaction and Self Perception in Virtual Space

The social interaction in virtual space is often studied in the lens of presence. There are three dimensions of presence: spatial presence, social presence, and self presence[50]. Spatial presence measures the extent to which an individual feels present in a media environment compared to the physical space [48, 53], which is closely related to how they perceive the environmental and spatial cues of the current media environment. Social presence measures the level of feeling "being together with others" in the virtual space which can be used as an indicator of the social interactions [5]. The measurement of social presence can be divided into different dimensions, including psychological involvement and behavioral engagement [5, 27].

Social presence and spatial presence are often measured in video game which, compared to video chat, gives the user more agency and control and provide different viewpoints [11, 18, 24, 27]. But on the other hand, video chat also shares the social aspect that video game provides and create a shared space for participants to communicate and interact with each other. Therefore, in other studies of video communication tools, a sense of presence is used to evaluate the social experiences and team collaboration [30, 39]. The evaluation matrices in the related fields suggest that presence can be a useful measurement for spatial awareness and social interactions in video chat.

The dimension of self-presence is related to one's perception of the virtual self versus the physical self [3]. This is not very relevant to the context of video chat since participants are presented with a "talking head" which is a video capturing their physical self instead of using a virtual body [39]. The self-perception is still valuable to the exploration of spatial influence in video chat, therefore we turn to other measurements that capture the self-awareness, such as the situation self-awareness scale (SSAS) [19] which uses a self-report scale to evaluate self-awareness on three dimensions: private, public, and the immediate surroundings.

3 METHOD

We used a mixed-methods methodology to probe user interactions in video conferencing environments in a block-like Zoom-type interface and in a spatially formatted Room-type interface, and collected survey and semi-structured interview data after the intervention.

3.1 Participants

Participants (N=34) (Table 1) were college students recruited through social media and posts on campus, and participated in the Social Support (N=18, 13 females, 5 males) and Topic Discussion (N=16, 12 females, 4 males). All participants signed an online consent form and were given a prompt for their particular condition. They were also provided with an honorarium for completing the study.

3.2 Study Set-up

The user study was conducted in ohyay.co, which is a video platform that allows users to build customized video chat room layouts according to research needs.

To generalize our study of spatial configuration's influence on the video chat, we designed two different tasks (Figure 1) –Social Support, in which people sit together, share personal experience and support each other; and Topic Discussion, in which people hold opposite opinions about one topic sit and discuss together.

In the Social Support task, participants (N=18) discussed challenges and experiences during Covid-19, with the purpose to support other participants. The prompt given by the researcher is:

Table 1. Summary demographics of the participants

Social Support					Topic Discussion				
ID	Condition	Age group	Gender	Interviewed	ID	Condition	Age group	Gender	Interviewed
P1	Room	18~24	F	✓	P19	Room	18~24	F	
P2	Room	18~24	F	✓	P20	Room	18~24	M	✓
P3	Room	25~31	M	✓	P21	Room	18~24	M	✓
P4	Room	18~24	F	✓	P22	Room	18~24	F	✓
P5	Room	18~24	F		P23	Room	25~31	F	✓
P6	Room	18~24	F		P24	Room	18~24	F	✓
P7	Room	25~31	M		P25	Room	18~24	F	
P8	Room	18~24	F	✓	P26	Room	18~24	F	✓
P9	Room	25~31	M	✓	P27	Zoom	18~24	M	✓
P10	Zoom	25~31	F	✓	P28	Zoom	18~24	F	✓
P11	Zoom	18~24	F		P29	Zoom	18~24	F	✓
P12	Zoom	25~31	F		P30	Zoom	18~24	F	
P13	Zoom	18~24	F		P31	Zoom	18~24	F	✓
P14	Zoom	18~24	F	✓	P32	Zoom	18~24	F	✓
P15	Zoom	18~24	M	✓	P33	Zoom	18~24	M	✓
P16	Zoom	18~24	F	✓	P34	Zoom	18~24	F	✓
P17	Zoom	32~38	F						
P18	Zoom	18~24	M	✓					

"The uncertainty of the pandemic has brought stressful and anxious moments to people all over the world. Today we're here to support everyone in combating the upset feeling toward this issue. Everything in this chat will be kept confidential." All the participants were equally divided into two 9-people groups with two different spatial configurations. One group called Zoom is a grid layout for simulating the interface of Zoom, which displays the video of each participant in equal size. Another spatial configuration is Room using a background image of a classroom which is a photo taken by the researchers. In the photo, 9 chairs were placed in a circle. The video frame of each participant was positioned based on the location of each chair in the background. Based on the position, the perspective of the video frame is adjusted using the rotation and skew parameters on each axis. Except for the layout of each video chat room, we kept the rest of the settings the same in different conditions. This conversation for each group lasts around 30 minutes.

16 people (N=16) participated in the Topic Discussion task. Based on their background, the task topic is about whether the university should launch a compulsive policy that undergraduate students need to finish a full-time internship in one term. All the participants were equally divided into two 8-people groups. One group called Zoom is a grid layout for simulating the interface of Zoom, which displays the video of each participant in equal size. Another spatial configuration is Room, using a background image of a classroom which is a photo taken by the researchers. In the photo, 8 chairs were placed in two lines. The video frame of each participant was positioned based on the location of each chair in the background. Based on the position, the perspective of the video frame is adjusted using the rotation and skew parameters on each axis. Except for the layout of each video chat room, we kept the rest of the settings the same in different conditions. This conversation for each group lasts around 30 minutes. For both Room and Zoom, Participants are randomly and equally divided into two groups holding opposite views. They discussed their standpoint rationally.

3.2.1 Procedure. The user study is conducted on an online video chat platform, ohyay.co. Recruited participants are randomly assigned to different conditions, including 9 people in Room with a social support task, 9 people in Zoom with a social support task, 8 people in Room with a topic

discussion task, and 8 people in Zoom with a topic discussion task. Two researchers monitored the study, one as note-taker and one as the facilitator. In the beginning, one researcher facilitated a short self-introduction session for participants to get familiar with each other. Then the researcher gave a topic for them to discuss. Participants in Room and Zoom condition with the same task followed the same procedure. The researchers did not make interventions during the 30-minute discussion. The process of video chat is recorded in video form for data analysis. At the end of the study, participants were asked to finish an online survey and then leave the room. We also conducted post-study interviews with selected participants after they completed the survey. Details of the interview are elaborated on in a later session.

3.3 Data Collection: An Online Survey after the Group Chat

To understand the dynamic interaction and people's spatial perception during the group chat, an online survey was developed with mainly closed-form and a few open-ended questions. The survey took approximately 10-20 minutes for participants to complete after the group video chat. The closed-form questions in the survey are adopted from the following scale measures.

Social Presence Gaming Questionnaire (SPGQ) [10]. This scale measures the level of social presence during the video chat. The original scale includes three subscales: (1) Psychological involvement – Empathy, (2) Psychological Involvement – Negative feelings, and (3) Behavioural involvement. In our survey, we used items from the first subscale and the third subscale to measure the level of empathy and behavioral engagement. Items in each subscale are rated on a 7-point Likert scale from "Strongly Disagree" (1) to "Strongly Agree" (7).

Since the scale was used in a pair of two, we adjusted the definition of "the other" in the original scale. Each participant was asked to self-report their feelings toward two individuals who have a specific spatial relationship with them - "the person who sit next to me" and "the person who sits far from me". Participants were given a figure showing who they should rate when answering the survey questions. In different task conditions, the definition of these two rated individuals has adjusted accordingly. In the Room condition with social support task, "the person who sits next to me" is defined as the person who sits on one's left side, "the person who sits far from me" is defined as the fifth person on the right (counted anti-clockwise). For an individual in Zoom of social support task, "close" is defined as the person whose video is positioned this participant, and the definition of "far" is defined as this figure (Figure 2). In the Room condition with topic discussion task, "the person who sits far from me" is defined as the person who sits opposite to one participant. In the Zoom condition with the topic discussion task, the definitions of "close" and "far" are similar to the Zoom condition with social support task (Figure 3).

Besides, in data analysis, we differentiated the perspective of different questions in empathy. Questions, like "When who sits next to me were happy, I was happy" or "I found it enjoyable to be with people who sit far from me", were marked as "self" in the variable of perspective, since it stated how "I" felt; Questions, like "When I was happy, people who sit next to me were happy" or "People who sit next to me paid close attention to me", were marked as "other" in the variable of perspective as it stated how "other people" felt. Similarly, the engagement also included the perspective of "self" and "other". Questions like "My actions depended on people who sit next to me" or "I paid close attention to people who sit next to me", were marked as "self". Questions, like "The actions of people who sit next to me were dependent on my actions" or "People who sit next to me paid close attention to me". The definition of close and far for each condition is shown in Figure 2 and Figure 3.

Situational Self-Awareness Scale (SSAS) [19]. This scale measures the level of self-focus of the participants during the video chat. It has three subscales measuring private self-awareness (the awareness level of one's internal and personal feeling), public self-awareness (the awareness level

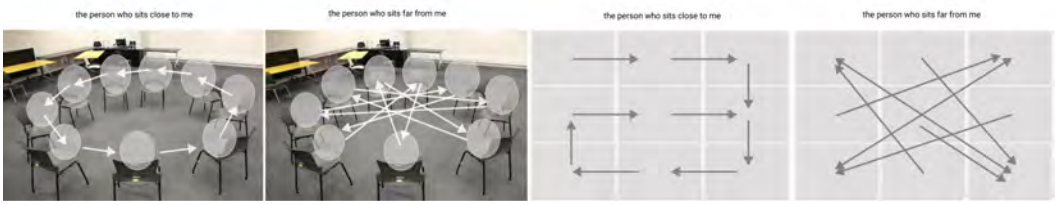

Fig. 2. Each participant in the video chat was asked to report their engagement and empathy level on "the person who sits next to me" and "the person who sits far from me". This figure shows how "close" and "far" in survey and data analysis are defined in Room - Social Support (left) and Zoom - Social Support (right). For each arrow, the start point means the participant who is going to give the rating, and the endpoint is the participant who will be rated.

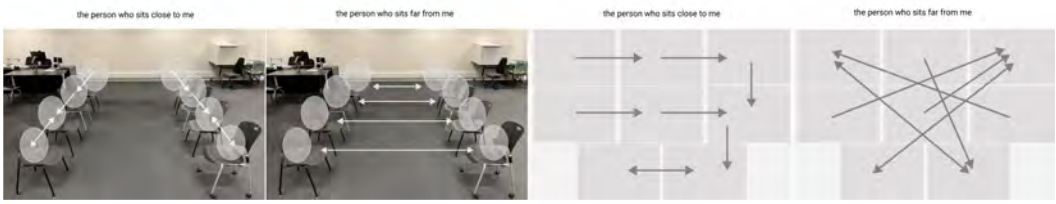

Fig. 3. Each participant in the video chat was asked to report their engagement and empathy level on "the person who sits next to me" and "the person who sits far from me". This figure shows how "close" and "far" in survey and data analysis are defined in Room - Topic Discussion (left) and Zoom - Topic Discussion (right). For each arrow, the start point means the participant who is going to give the rating, and the end point is the participant who will be rated.

of how oneself is presented to others), and awareness of immediate surroundings (the awareness level of the physical environment around the person). We defined these category as "context" and as an independent variables in our data analysis. Each subscale includes three items rated on a 7-point Likert scale.

The Spatial Presence Experience Scale (SPES) [23]. This scale measures the level of spatial presence which is the sense of "being there" in the media space. This scale consists of two dimensions: self-location and perceived possible actions in the media environment. Since the participants in video chat don't have the freedom on moving around or interacting with virtual objects, we used and modified the items in the self-location session for measurement. Five items were rated on a 7-point Likert scale.

3.4 Data Collection: Post-survey Interview

To garner a deeper understanding of the survey responses and trends that were identified in the data, follow-up interviews were conducted with selected participants who participated in the group video chats. The selected participants cover lurker and active participants in the video chat, and people who sit on special spots, such as in the center or on the corner. 24 participants (N=24) in total had the interview, including 12 participants from the Room condition groups, and 12 participants from the Zoom condition groups.

The interviews were conducted in participants' native language in less than 12 hours after the group chats to make sure they could remember and elaborate the details of video chat vividly. The interviews were conducted using video or voice calls. Each interview lasted about 30 minutes.

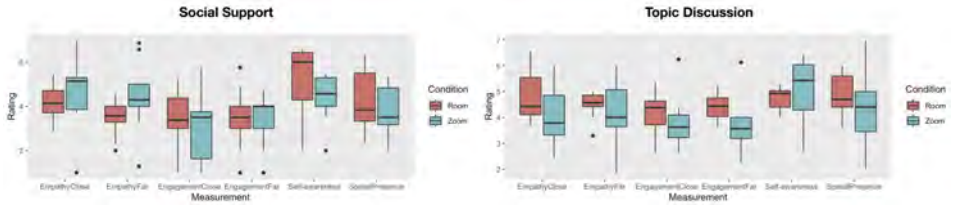

Fig. 4. Descriptive statistics of the measurements ratings, including the social support group and the topic discussion group

The interviews were semi-structured, with questions probing participants' experiences during the group chat. After recalling people's feelings in that process, we specifically asked them questions related to "spatial perception"– how they feel about the "space" in the video chat room and whether they were influenced by the spatial perception. Moreover, we asked some questions about how they thought about the gap between group chat online and in the physical space. With approval from interviewees, all 24 interviews were audio-recorded and transcribed.

3.5 Data Analysis: Video and Interview

The video recordings and interview transcriptions were analyzed using an open coding method. For interviews, researchers coded segments of the interview scripts by themselves, then came together to discuss their codes and reach an agreement for a codebook. Each author then categorized the rest of the transcriptions using the codebook. All the codes were translated into English.

For video recordings, the code theme is based on the summary of interaction features that emerged in the pilot test and also refers to methodologies on analyzing the group interactions in communication, such as the Discussion Coding System (DCS) developed by Schermuly [46]. Among the code theme, facial expressions–smiles, measured people's attitudes to the chat content and how the social dynamic changes in different conditions. The body language – "handclap", "nod head" and "look other places outside the screen", can evaluate the participants' engagement in different conditions, which is related to the level of connection with others and belonging in a team. Other body languages–"turn heads" and "point to someone" are usually used to locate other's location in the physical space and we are supposed to measure whether people take actions which related to their spatial perception.

The videos were analyzed in time series. Researchers coded the video every 20 seconds. This time length is neither too short to generate many empathy codes nor too long to miss many interaction details. Participants' performance was coded individually, as everyone had different experiences and responses in this chat room. At last, we plot all the events of everyone and also sum all the events in every condition.

4 QUANTITATIVE RESULTS

In this study, all participants (N=34) responded to the survey (25 females, 9 males). The descriptive statistics of the measurements using R (1.4.1717) studio (Figure 4).

To test whether there's a difference in the measurements including empathy, engagement, spatial presence, and self-awareness between the Zoom and Room conditions, we first checked whether they are normal distribution or not by using Shapiro-Wilk. The results showed that they were not non-normal distribution. This allowed us to use a t-test and two-way ANOVA to tease out the contributions of each factor.

Table 2. Statistics of Two-way Interaction Term of Condition * Perspective

		Social Support		Topic Discussion	
Measurement	**Variable**	**F**	**P-value**	**F**	**P-value**
EmpathyClose	Condition	2.286	0.140	4.418	**0.045** *
	Perspective	0.693	0.411	1.261	0.271
	Condition*Perspective	0.000	1.000	0.604	0.443
EmpathyFar	Condition	0.339	0.565	1.640	0.211
	Perspective	0.045	0.833	2.646	0.115
	Condition*Perspective	0.001	0.982	0.634	0.433
EngagementClose	Condition	0.950	0.337	0.658	0.424
	Perspective	0.002	0.961	0.164	0.688
	Condition*Perspective	0.214	0.647	0.105	0.748
EngagementFar	Condition	3.347	0.077	4.589	**0.041** *
	Perspective	0.008	0.929	0.111	0.741
	Condition*Perspective	0.004	0.950	0.111	0.741

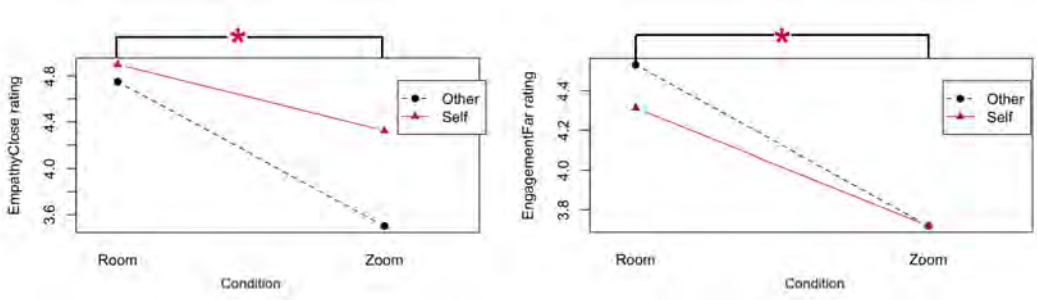

Fig. 5. EmpathyClose and EngagementFar in Topic Discussion: the interaction results of conditions and perspective in two-way ANOVA

Spatial presence. First, we look at spatial presence, it measures whether participants had a stronger spatial perception when entering the Room compared to the Zoom. By using a t-test, no significance was found between Room and Zoom, either in the social support group ($p = 0.4503$) or in the topic discussion group ($p = 0.3379$).

Empathy and engagement. Then we checked the measurements of empathy and engagement. We used "perspective" as an independent variable for each measurement. Two-way ANOVA with independent variables–Condition and Perspective– was used here (see the Table 2, Figure 5). Results showed that no significance was found in the Social Support group. But in the Topic Discussion, We found the empathy rating of participants who were close to the informants by condition were significantly higher in Room condition compared to zoom ($F=4.418$, $p < 0.05$), though the interaction between conditions and perspective was not significant. Similar to empathy, we found a significant difference in Engagement of participants who were far from the informants by the condition($F=4.589$, $P< 0.05$), and interaction between conditions and perspective was not significant.

Self-awareness. Applying two-way ANOVA, in Social Support groups of Room and Zoom, self-awareness rating in Room is significantly higher compared to it reported by participants in Zoom

Table 3. Statistics of Two-way Interaction Term of Condition * Context

		Social Support		Topic Discussion	
Measurement	**Variable**	**F**	**P-value**	**F**	**P-value**
Self-awareness	Condition	4.216	**0.046** *	0.807	0.374
	Context	0.292	0.748	0.749	0.479
	Condition * Context	0.532	0.591	0.626	0.540

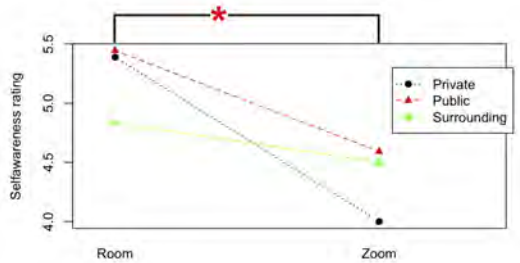

Fig. 6. Self-awareness in Social Support : The interaction result of conditions and context in two-way ANOVA

(F=4.216, p<0.05), though the interaction between these terms was not significant (Table 3, Figure 6).

4.1 Video Data Results

The video data allows us to dissect how each session occurred in terms of the way people talked to each other, and in terms of the nonverbal gestures that people make with each other. The conversation flow amongst participants does not appear to differ between Zoom and Room conditions, in either Social Support or Topic Discussion (Figure 7).

In the video data, we observed different behavior patterns between the group of Social Support and Topic Discussion. First, we found no difference in the conversation flow is shown (Figure 7) between different conditions in each group. The conversation flow amongst participants does not appear to differ between Zoom and Room, either in Social Support or in Topic Discussion. It seems that the space configuration may be one of the variables that influence participants' behavior and experience in the whole process.

Except for conversation flow, we found only "number of total smile" is one of the obvious and common behaviors in the Social Support group(see Figure 8).

5 QUALITATIVE FINDINGS

As discussed in the method session, we aim to investigate how does the spatial configuration influence the social dynamics of a group chat under different tasks. Our analysis reveals patterns around how the participants gathered and processed the spatial information of video chat using visual cues, and how they form the perception of self and others in group video chats with different spatial configurations.

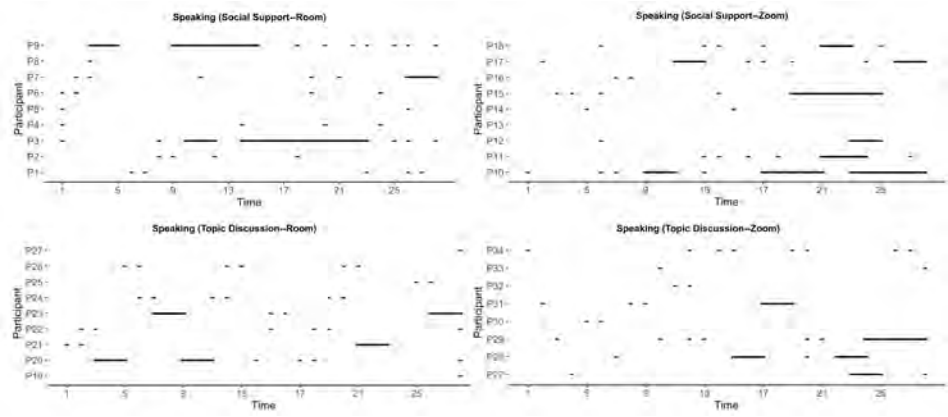

Fig. 7. 30 minutes conversation: Who was speaking

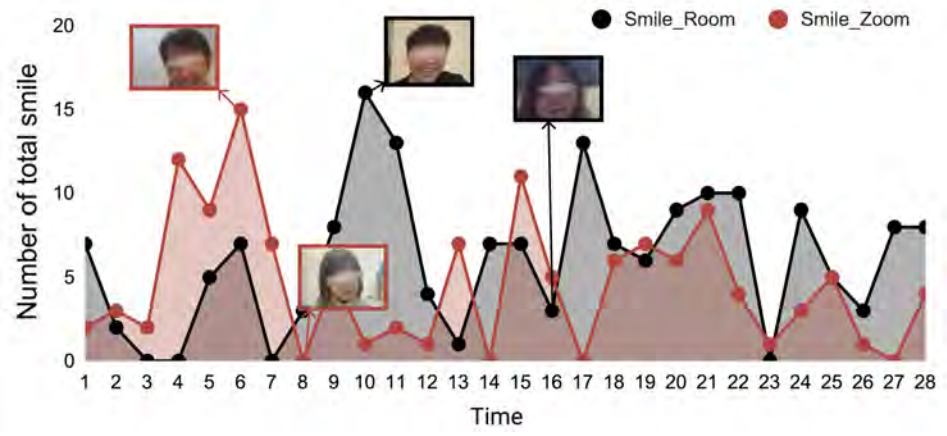

Fig. 8. Plot of smile in Social Support task: measures in the validate 30 minutes who smiled in different periods

5.1 Gathering and processing spatial information using visual cues

Our study investigates the information collecting and processing in the context of video chat. From interview analysis, we conclude emerging themes about how the participants pick up these visual cues from both the conversation and the environment of the video chat, and how does it relate to their behaviors. There are two main aspects when a participant gathers information through observations: tagging the participants with a spatial location and glancing through the group in a certain order.

5.1.1 Tagging the participants with a spatial location. One theme that emerged from the analysis is about searching and locating other individuals in the group video chat. In the real world, a scene can provide useful guidance for visual search with its properties [14], such as finding the front door of a house. Our analysis find similar behaviors happening in the virtual space of the group video chat under the Room condition - participants were able to locate other participants with a

virtual "seat" and glance in a certain order using the chairs in the background image. Details of this finding are elaborated next.

Instead of finding different objects in a scene, the "objects" of the video chat are the videos of participants. The need of locating a specific participant during the group video chat occurred more often in the condition with the topic discussion task where people need to identify who is on their side or on the other side. In the Room layout, by putting the pro and con team separately on the left and right side of the room, participants were able to "label" other participants based on their standpoint thus saving the time on repeating this process, e.g. *"It's easier to locate other participants using this position in the space so that I don't need to find them every time [...] If this is on Zoom it will be much difficult to find person one by one"(P26, Room - Topic Discussion).*

This is considered as an advantage of Room compared with Zoom where every participant is given a random position. Participants in Zoom with the same topic discussion task needed more mental efforts on identifying team members. *"One boy seems confused which side he was on, so I have no impression of him at the beginning. Yes, I was also confused, I thought he was on the opposite side, so I didn't pay attention to his talking" (P27, Zoom - Topic Discussion).* This struggle, as a result, elicited the urge for a participant to rearrange the arrangement based on task needs, *"I want people on my side to be put together rather than scattered on the top left or bottom right corner. [...] I want my teammates to sit closer to me, and also their video to be positioned closer to me" (P34, Zoom - Topic Discussion).*

Therefore, the use of chairs in the scene attaches a spatial attribute for a participant, which can be used by other participants in memorizing and referring to an individual, such as *"the girl opposite to me" (P23, Room - Topic Discussion).* The way of referencing an individual in Room is similar to the way we used in a physical space. But in Zoom one person is rarely referred to by the "position" (e.g. the person sitting opposite to me) because the display of one's video is randomized, and on each participant's screen, the order of the video display may not be the same.

However, this visual search guidance also has its limitations. Although the chairs in the scene helped participants establish a connection between an individual and their location on the scene, this information did not help them process time-sensitive information in the group chat such as which individual is currently talking in the group chat, e.g. *"I think compared to Zoom the discussion in this room is more chaotic because I keep finding who is talking" (P21, Room - Topic Discussion).*

Therefore, whether this visual guidance is considered useful also depends on the extent to which an individual relies on visual information, and whether they care more about short-term information such as who is talking or long-term information such as the location. Some participants who are visual-focused may find this useful, while some content-focused participants may choose to filter out the visual information, e.g. *"I'll choose not to look at the screen, and take notes which helped me be more focused [...] those visual things may be a distraction for me" (P34, Zoom - Topic Discussion)*, or memorize another individual based on their opinions rather than using the spatial information, e.g. *"I mainly focus on what their content, and then realize whether they have the same opinion as me or not [...] did not remember who is in the same group with me" (P32, Zoom - Topic Discussion).*

5.1.2 The order of a glance. In addition to using a chair as the spatial reference to locate one participant, the chairs as a whole also allowed participants to scan through the group members in sequential order. According to the scanning behavior reported by the participants during interviews, the scan order varies based on the specific visual cues in the virtual space. In the context of this study, it is the arrangement of chairs in the background image (in a circle vs on opposite sides of a table) that influences the way participants scan through video frames in the video chat room. For instance, in the topic discussion room where the video frame are positioned in two opposite rows,

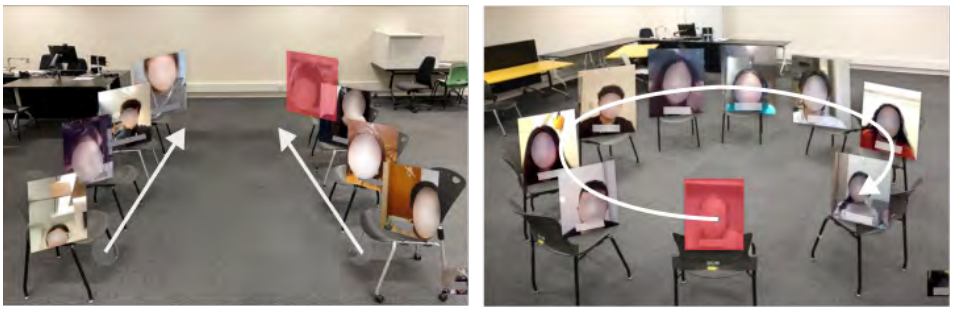

Fig. 9. Glance behaviors reported by Room participants. The participant who reported this glance behavior is marked with a red rectangle. The arrow illustrates the direction of the glance.

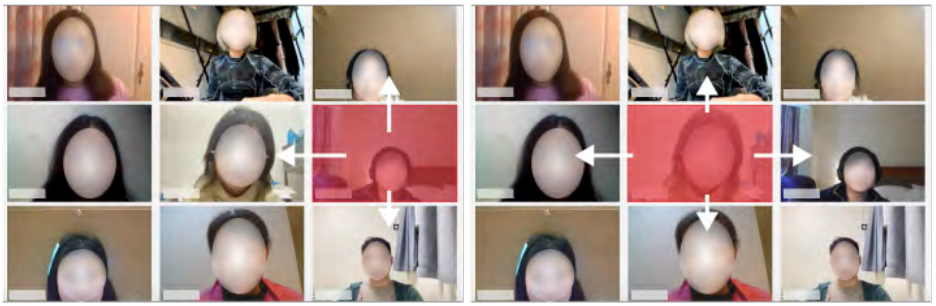

Fig. 10. Glance behaviors reported by Zoom participants. The participant who reported this glance behavior is marked with a red rectangle. The arrow illustrates the direction of the glance.

participants said they *"scanned from closer to the screen to further from the screen (P26, Room - Topic Discussion)"*.

While in the social support room where chairs are arranged in a circle, a participant "glanced through the participants in the room in the clockwise order" and another participant also "look around" (P1, P9, Room - Social Support). This behavior in both conditions (circle vs line) is a one-way scan that has a start point which is either the front and center seat of the circle or the front seat of a line that is closest to the screen (Figure 9). In both conditions, the computer screen is used as the frame of reference, instead of the video of participants.

This glancing behavior in Room is different from that in Zoom in terms of the center point. A few Zoom participants (P14, P15, Zoom - Social Support, P31, Zoom - Topic Discussion) noted that they glanced through the video around their own video, such as allocating the attention "mainly based on a circle using my video as the center" (P15, Zoom - Social Support) (Figure 10). This is reported by people whose videos are on different locations of the grid layout in zoom condition, including the center and the right side on the second row. This may indicate two different types of viewpoints adopted by participants during the video chat, which we'll elaborate on further in the next section.

5.2 Forming the Perception of Self, Others, and the Space

Continuing on 5.1 in which we talk about findings on how participants gather information from the space, in this following section we are going to discuss our findings on how participants further process the visual cues in forming the perception of others. An effect of space on enhancing team awareness was reported by participants in Room condition with the topic discussion task. Besides, we find an intention of "hiding from the spotlight" which is mentioned by both Room and Zoom participants. This offers us an understanding of how the participants perceive themselves in the video chat.

5.2.1 Perception of others: Sense of belonging and connection shaped by the perception of the space. In two groups with different tasks, we found the sense of belonging and connection is perceived differently in Room. For the topic discussion task, the sense of belonging is related to how one participant identifies their relationship with people who hold the same opinion toward the topic, which are often described as "my team" or "my side". We found a difference between the level of belonging and connection perceived in Room and Zoom - participants in Room reported higher level of team awareness compared to Zoom. For example, *"It gave me a stronger sense of debating and confrontation since the layout puts two parts separately and makes the boundary obvious. For example, then I feel the person sitting opposite to me and I are on different teams, and people sitting on the same side with me are on my team" (P24, Room - Topic Discussion).*

A sense of belonging to one's own team is also emphasized by the boundary between two teams, which put participants closer to their teammates and further to people who held different opinions with them, *"Putting the affirmative team together and the negative team together, from this format I feel this boundary conceptually influence my perception, and I'll recognize who are with me or against me" (P23, Room - Topic Discussion).* From another perspective, when such a team grouping did not happen in Zoom condition, a negative feeling arose, e.g. *"I will feel disappointed if people next to me are not on the same side as me. I want my teammates to sit closer to me" (P34, Zoom - Topic Discussion).*

However, the background image of the video chat only conveys one scenario of the team segment. When the individual's behavior matches the information given by the scene, the visual information enhances the perception. But when the behavior mismatches the visual information, the influence of the space seems to be surpassed by the behavior, according to the interviews from our study.

Specifically, in our case, people on the left and right sides hold different opinions. This means a participant holds a different opinion with people who sit opposite to them and shares the standpoint with people sitting on their side. As the discussion goes, at one point in our topic discussion task, a person on the left side seemed to lean toward the opinion on the right side, which created a mismatch between his "location" and his behavior. This also changed another participant's perception of him and the team identification, as this participant put it,

> *"I consider people on my side as a team, and also people on the other side as another team. But there is a person on the opposite who seems to change his standpoint and show some agreement on our opinions as the discussion goes. So then I view people on another team separately. It feels like dragging that person to our team." (P24, Room - Topic Discussion).*

This shows that there are two different ways for people in Room condition to perceive others in the same video chat. One is identifying the team as a whole based on their standpoint. Another is identifying each individual in a team based on what they say and act. When an individual's behavior matches his or her role in the team, these two ways may act similarly, such as a participant on the right side of Room arguing for the team's opinion, as this participant is also identified as a member of the team. When an individual's behavior mismatches the team's mission, this individual

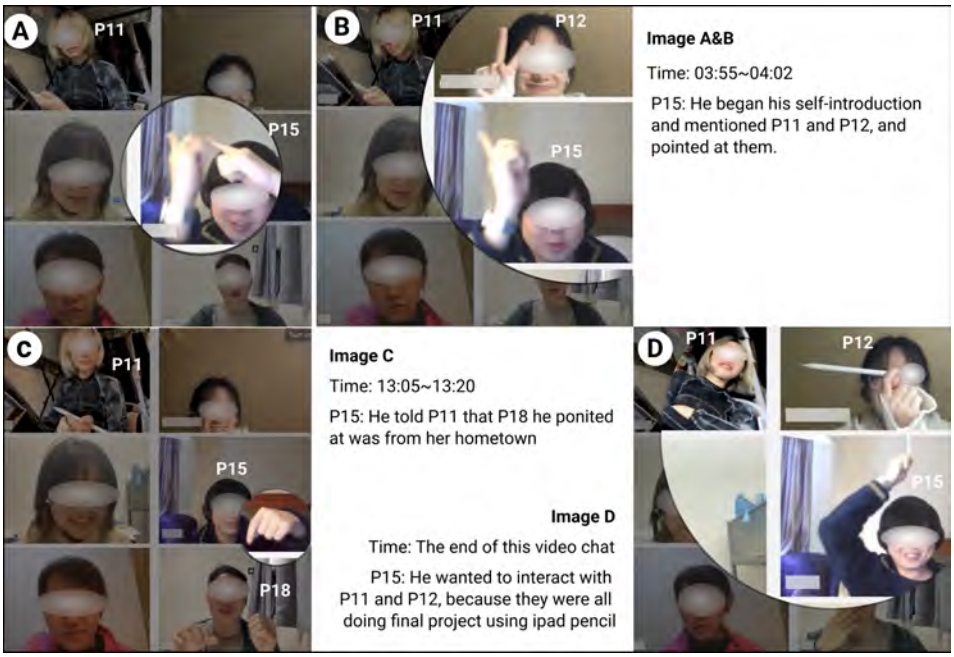

Fig. 11. Social Support - Zoom: Participant utilized spatial information to interact with others

may be excluded from the side and perceived separately. The visual environment cues provided by the space works in the first scenario in enhancing the feeling. But in the second scenario where the environment cues mismatch the individual's behavior, the influence may be diminished.

In addition, there could be a mismatch between the context provided by the space and the activity actual happening in the video chat, which may also break the immersion in video chat. For example, although the Room - topic discussion condition helped some participants to form team awareness, it is also questioned by participants, "Since the goal is not to rather than fighting against the people on the other side, I'm not sure whether this layout is too confrontational for a group discussion (P24, Room - Topic Discussion)."

In the Zoom condition, although participants didn't mention any effect on how the grid layout affects their sense of belonging, we observed from the Zoom with the social support group that some participants take advantage of the spatial configuration to enhance the sense of connection. For example, one participant used a hand gesture to point another participant whose video is positioned above. "This is cooler, and probably also help to warm up the atmosphere. Pointing with the finger will send clearer signal than just saying the name" (P15, Zoom - Social Support) (Figure11) This action also inspired another participant to use a hand gesture later. This action did seem to bring more sense of connection and warm up the atmosphere as the participant expected, given the fact that many participants smiled or laugh when the hand gesture are used.

5.2.2 Perception of self: "Hiding from the spotlight" in the video chat as sitting on the back rows of the classroom. A theme that emerged from the analysis is an intention to "hide from the spotlight", or in other words, avoid receiving too much attention from others during the video chat. This point is raised by participants from both Zoom and Room conditions. A Room participant noted that this condition produced a lower level of anxiety compared to the speaker view in Zoom which highlights the current speaker,

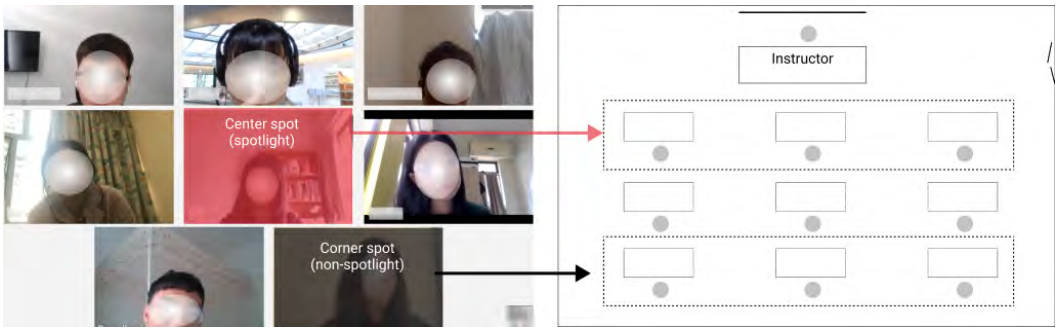

Fig. 12. A spatial metaphor made by the participants: "sitting at the corner of Zoom is like sitting at the back rows of the classroom".

> *"In Zoom, when someone is talking, their video will be put in the middle with a larger size, and everyone knows who you are and looks at you. From that point, I feel more comfortable on this platform" (P23, Room - Topic Discussion).*

Another Room participant noted a different feeling compared with previous experience in Zoom,

> *"I feel more awkward in this room because everyone's video is displayed in a rectangle shape, and sometimes I will look at every one of the videos, which makes me feel stressed. [...] When having a face to face communication I feel a lot of people are watching me, which made me feel uncomfortable" (P4, Room - Social Support)*

The source of the anxious and awkward feeling is the attention from people with who they are not very familiar. This finding aligns with the researches investigating the level of social anxiety level and video chat [29], suggesting a higher level of anxiety when seeing one's video presented in a large size in the video chat.

In addition to the quote suggesting Room eased the nervous feeling, we also found from a Zoom participant that the feeling of being "put in the spotlight" is related to the anxiety of video chat. The intention of "hiding from the spotlight" may help mitigate the anxiety. Furthermore, as some participants mentioned, this intention of hiding is similar to sitting at the back rows of the classroom in physical space (Figure 12). The perceived level of attention from other participants is the factor that contributes to this anxiety. For example, a participant who is put on the bottom right corner of the Zoom grid layout reported that,

> *"this corner spot brought me a safe and comfortable feeling just like sitting at the back row of a classroom [...] probably because it does not receive much attention from other people [...] As long as I'm not in the center spot which is like the first row of a classroom" (P34, Zoom - Topic Discussion).*

This quote reveals a thought of feeling less anxious in a seat that receives less attention from other participants and feeling more anxious in a seat that receives too much attention from other participants. Based on the perceived level of attention of a seat, the spots on the video chat space are "spotlight" and "non-spotlight." Since every individual perceived the attention in different ways, the "spotlight" seat can be defined by a visual focus (e.g. the center spot of the Zoom grid is considered as the seat receiving the most attention) or the seat around a person with a special role (e.g. host of an event or lecturer of a class).

If the participants have an agency to the seat they want, the intention of hiding may also influence the behavior in choosing a seat of a video chat. As a participant noted, *"When going to an in-person*

class, a lot of people will prefer the back row seats. Therefore the front rows are often left empty. Similarly, if we sit in a circle, there is often a facilitator, and the seat next to the facilitator may also be left empty (if I can choose the seat)" (P3, Room - Social Support). This also enhances the potential connection between the intention of "hiding from the spotlight" and "sitting on the back row of the classroom". This spatial metaphor of video position may provide another possible understanding of Zoom's interface from a spatial perspective.

6 DISCUSSION

The survey, video code and interview findings uncovered nuances of spatial perception in video chat in block-like (Zoom) vs. physically mapped (Room) settings. In what follows, we reflect and highlight the opportunities and challenges that video chat presents, situate our results within other video chat or video conferencing research, and discuss insights and future avenues of research.

6.1 Spatial influence in video chat with a Topic Discussion task

Comparing the survey data from Room and Zoom condition with a topic discussion task, we found the level of empathy towards the person who sits next to oneself is significantly higher in Room than Zoom. The level of engagement towards the person who sits opposite to a participant in Room is also higher than that in Zoom condition. From the interview data, we also had two findings that are only reported by participants in Room with a topic discussion task. The first one is the use of chairs in Room as visual guidance to locate people who held the same standpoint or the opposite. Secondly, participants in Room with topic discussion task also revealed a higher level of team awareness, indicating the positioning of chairs may help enhance this feeling.

The results from both survey and interview suggested some possible influences of the spatial configuration in Room. Since these findings only occurred in Room with the topic discussion task, the influence of space may be task-specific. There are several reasons for that.

Firstly, compared to the social support task in which every participant in the room are considered as on the whole group, in the Topic Discussion task there exists sub-groups within the video chat room. Identifying those sub-groups may require extra cognitive bandwidth. The conventional block-like Zoom-style layout may not be sufficient for this requirement since it only provided visual cues on who is currently talking. The Room-style interface performed better in this task because the chair in the background provide additional environment cues for the participants to identify the sub-groups in the chat by creating sub-spaces within the video chat space.

This is in line with the word "boundary" used by participants. What they mean about "the boundary helps separate two sides and identify myself with my team" (P26, Room - Topic Discussion), in the framework of scene perception [14], is that the chair in the background image is a scene property which helps participants' cognitive function - recognizing their teammates and their allegiances. Meanwhile, according to researches on social presence [41], the background of the Room condition created a shared space for the participants in the group which may help participants develop a stronger sense of being together. This also aligns with our finding with the survey data that shows the level of empathy and engagement in Room condition (topic discussion) is significantly higher than that in Zoom condition for two particular scenarios: empathy with the person who sits close and engagement with the person who sits on the opposite side.

6.2 Spatial influence in video chat with a Social Support task

The survey data of participants in Room condition with a social support task shows that the self-awareness in Room is significantly higher than that in Zoom. Among the three dimensions of self-awareness which are private self-awareness, public self-awareness, and the awareness of

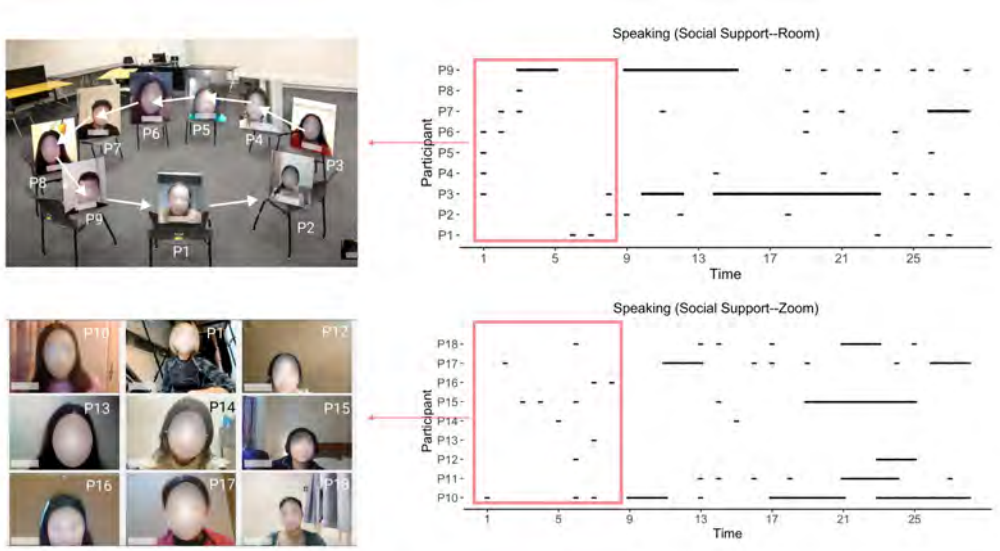

Fig. 13. The order of speaking at the beginning of conversation in Room vs Zoom condition with a social support task. Participants in Room talked in an order based on their location, while participants in Zoom talk in a random order.

immediate surroundings, private self-awareness in Room is most significantly higher than that in Zoom.

This finding leads to a further discussion on the relationship between the social dynamics of a group chat and the way space influences perception. Our assumption for this difference in self-awareness among different tasks is that these two different tasks require different levels of self-disclosure [13, 37] - in social support task, people need to share their personal experiences while in the topic discussion task, the focus is more on the topic itself. Revealing more personal stories in the chat leads to a higher level of self-disclosure to the group, which could make the participant more aware of their inner feelings [1]. The higher level of self-awareness shown in the survey data may also suggest the potential influence of the space on this social support activity. The environment cues of the circle-shaped chairs from the scene subtly enhanced the atmosphere of sharing and helping on the perception level and made this online communication closer to the social support happening in a physical space.

We also found from interviews that the circular shape of chairs in Room with the social support task may help participants build a sense of order. Since the order of a circle is either clockwise or anti-clockwise, this visual hint can be very helpful in tasks that require participants to interact following sequential order such as self-introduction, ice-breaking activities, or the mafia game in which players often sit around a round table and talk one by one. The layout in Zoom, by contrast, doesn't satisfy this need since in the grid layout there could be multiple ways to decide the order, and there is no visual cue that facilitates this. Therefore, at the beginning of the conversation for a social support purpose, we saw participants in the Room condition speak in an anti-clockwise order based on their position, while participants from Zoom condition with both tasks gave self-introduction in a random order (Figure 13).

Further researches can be done on investigating the influence of the object layout as an environmental cue. For example, how might the glance of order be different in a video chat with a

seating arrangement in a square shape, rows and columns, or a circle. Methodologies including eye-tracking can be used for gaining more quantitative findings [28, 31, 44]. Researches can also further compare the spatial influence of different seating arrangements with the same task, e.g. having the topic discussion task in Room condition with different arrangements of chairs (circle vs in two sides of a table).

6.3 Design implications for video chat space

6.3.1 Use of Visual Cues for Video Chats with a Different Group Structures. Our findings regarding the influence of a Room-style video chat interface compared to Zoom offers useful insights on the design of video chat space. As we observed from participants, Room-style configuration can provide additional visual cues which have the potential of influencing perceived engagement and empathy with others. The cases of chairs helping participants search and locate others, enhance team awareness and chairs in a circle helping participants' turn taking activities indicate a pathway about how Room participants collect visual cues (chairs and their arrangement) as a property of the scene and use that on forming recognition of the space. The limitation of the additional visual cues offered by the Room condition is that such types of visual cues can not change dynamically during the video chat. Therefore, people who configure such a video chat room can only make an assumption about the possible team dynamics beforehand and use that as a preset. When the assumptions mismatch what actually happens in the video chat room, such intervention may fail to bring its expected influence. Instead, this mismatch between a visual cue and the immediate situation may even bring the risk of causing more cognitive load for participants [49].

Based on what have been discussed above, the design implication of future video chat interfaces is that there can be additional visual cues that help fulfill needs for particular tasks but we should also be aware of the risk of cognitive overload provided by mismatched cues. Therefore, such visual cues also need to be adjustable and flexible to accommodate changes happening during the video chat. The lesson learned from how chairs' arrangement facilitates turning taking in a group can also be applied in video chat interface design, and be combined with the aspect of visual cues. One design solution using spatial interventions proposed by video chat platforms Gather Town [16] is building a game-like scene for the video chat and giving users control of a small virtual avatar that can move in the scene. Meanwhile, the video feed of participants can also be displayed on the screen. The spatial configuration inside the scenes, therefore, functioned as the visual cues, and the control of avatars allows user to change their spatial location during the video chat. This spatial intervention may help fulfill the needs of team-based tasks by providing more environment cues.

6.3.2 Anxiety and the spotlight area in video chat. Our findings of the intention to "hide from the spotlight" lead to a further discussion on how does a spatial metaphor of "seat" and "spotlight" in the video chat related to the feeling of anxiety. In the block-like Zoom style interface, participants refer to "seat" as the position of their video in the grid layout, e.g. the bottom right corner or the center. In the scene-based Room-style interface, since the video of each participant is positioned with a chair, the location of a chair in the scene is referred to as the "seat".

From some participants from the Zoom condition, we found an intention of avoiding the seat that receive the most attention, such as the center spot of the Zoom condition. And some Zoom participants noted that they feel less anxious in a corner spot compared to the center spot. From Room participants, such concern is not noted. But instead, Room participants felt less nervous compared to the dynamic view of Zoom which highlights the current speaker.

The spatial metaphor of "seat" and the intention of hiding is also related to a social context - seat choice in the classroom. Depending on the seating arrangement of the classroom, the spotlight area one try to avoid has different types of spatial metaphor. In a row-column seating arrangement,

putting on a corner spot is compared with the similar feeling of sitting in the back rows, while a center spot is compared with the seats on the first row, depending on the layout of the classroom. In a circle, the spotlight area is defined as the area near the host of the event. What is shared in common among these spatial metaphors and the feeling of anxiety related to different spots is the perceived level of attention. In general, the spotlight area is considered by participants as spots that received more attention from the whole group. And it is the awareness of being watched by too many people that makes them feel anxious.

These findings from Room and Zoom in the sense of anxiety and the intention of hiding may suggest a potential "hiding" difference in different spatial configurations of video chat. For example, video conferencing tools like Zoom provide different view options for the users including a possibility of hiding one's video feed. In our Room condition, however, the video is still displayed during the video chat, which means it's harder for participants to hide. The engagement (with people who sit opposite) and the empathy (with people who sit close) which are higher in Room over Zoom are also aligned with a lower perceived sense of hiding.

This spatial metaphor provides another way of thinking about the layout design of a video chat interface. Even in a layout where every video is displayed in equal size, there can still be a "spotlight" area where receive more attention compared to other spots. The level of anxiety in the video chat may also be related to a spotlight spot.

Besides, another facet of this spatial metaphor which is the empty seats in the classroom inspired the thoughts on the "blank space" in video chat. For example, how might the portion of background area play a role in mitigating anxiety and whether that will create new spatial relationships or power dynamics are questions worth exploring in the future.

The design implication for this is a goal of balancing some participants' need to avoid the spotlight area while also ensuring their engagement in the video chat. Given Miller's research on the protective role of video feedback for people with social anxiety, [29], simply hiding the video of oneself may not be a good solution to mitigate the anxiety. More spatial configurations of the video chat space can be explored in this context.

6.4 Limitations

Our study is limited in that the participants only cover the age group from 18 to 38. Future researches should explore more demographics, such as the elderly people who have different digital literacy and perceived telepresence than the younger generation [34], and gender-specific influence of the digital space.

Besides, since our study is based on the interventions that combine the classroom scene and the spatial seating arrangement in a 2D image, thus for a specific understanding of different elements' role in the configuration of video chat, future works should separately each factor. Other features of the video chat layout such as the video frame can also affect people's self-awareness [38]. In our study in the Room condition for either the Social Support group or Topic Discussion group, the video frame twisted according to the seating arrangement. The spatial distortion of the frame may also influence participants' perception of the space. We can discuss the frame's influence in the configuration of video chat in the next step.

7 CONCLUSION

Our study contributes to an understanding of how might the 2D screen-based interface of the video chat influence people's perception and behavior in online communication for social support and topic discussion purposes. The results show a higher engagement with people who sit opposite and higher empathy with people who sit close in Room vs Zoom with the topic discussion task. Qualitative results also show more details about how participants in Room utilize spatial representations

as visual guidance for them to identify their team members. These findings provide inspiration for designing a video communication tool that can configure a 2D space for collaboration, discussion, and sharing experiences.

REFERENCES

[1] Irwin Altman and Dalmas A Taylor. 1973. *Social penetration: The development of interpersonal relationships.* Holt, Rinehart & Winston.

[2] Morgan G Ames, Janet Go, Joseph'Jofish' Kaye, and Mirjana Spasojevic. 2010. Making love in the network closet: the benefits and work of family videochat. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 145–154.

[3] Laura Aymerich-Franch, Cody Karutz, Jeremy N Bailenson, et al. 2012. Effects of facial and voice similarity on presence in a public speaking virtual environment. In *Proceedings of the International Society for Presence Research Annual Conference*. 24–26.

[4] Jeremy N. Bailenson. 2021. Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior* 2, 1 (Feb. 2021). https://doi.org/10.1037/tmb0000030

[5] Frank Biocca, Chad Harms, and Jenn Gregg. 2001. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual international workshop on presence, Philadelphia, PA*. 1–9.

[6] A. Blokland and A. Anderson. 1998. Effect of low frame-rate video on intelligibility of speech. *Speech Commun.* (1998). https://doi.org/10.1016/S0167-6393(98)00053-3

[7] Johanna Brewer and Paul Dourish. 2008. Storied spaces: Cultural accounts of mobility, technology, and environmental knowing. *International Journal of Human-Computer Studies* 66, 12 (Dec. 2008), 963–976. https://doi.org/10.1016/j.ijhcs.2008.03.003

[8] Vicki Bruce. 1996. The role of the face in communication: Implications for videophone design. *Interacting with Computers* 8, 2 (June 1996), 166–176. https://doi.org/10.1016/0953-5438(96)01026-0

[9] William Buxton. 1992. Telepresence: Integrating shared task and person spaces. In *Proceedings of graphics interface*, Vol. 92. Citeseer, 123–129.

[10] Yvonne De Kort, Wijnand Ijsselsteijn, and Karolien Poels. 2007. Digital Games as Social Presence Technology: Development of the Social Presence in Gaming Questionnaire (SPGQ).

[11] Alena Denisova and Paul Cairns. 2015. First Person vs. Third Person Perspective in Digital Games: Do Player Preferences Affect Immersion?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 145–148. https://doi.org/10.1145/2702123.2702256

[12] Paul Dourish. 2006. Re-space-ing place: "place" and "space" ten years on. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work (CSCW '06)*. Association for Computing Machinery, New York, NY, USA, 299–308. https://doi.org/10.1145/1180875.1180921

[13] Frederick G Elias, Mark E Johnson, and Jay B Fortman. 1989. Task-focused self-disclosure: Effects on group cohesiveness, commitment to task, and productivity. *Small Group Behavior* 20, 1 (1989), 87–96.

[14] Russell A. Epstein and Chris I. Baker. 2019. Scene perception in the human brain. *Annual review of vision science* 5 (Sept. 2019), 373–397. https://doi.org/10.1146/annurev-vision-091718-014809

[15] Joseph Falout. 2014. Circular seating arrangements: Approaching the social crux in language classrooms. *Studies in Second Language Learning and Teaching* 4, 2 (Jan. 2014), 275–300. https://doi.org/10.14746/ssllt.2014.4.2.6

[16] Inc. Gather Presence. 2021. Gather.town. https://www.gather.town/

[17] William W. Gaver. 1992. The affordances of media spaces for collaboration. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work - CSCW '92*. ACM Press, Toronto, Ontario, Canada, 17–24. https://doi.org/10.1145/143457.371596

[18] Geoffrey Gorisse, Olivier Christmann, Etienne Armand Amato, and Simon Richir. 2017. First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and Performance Analysis of Embodied Users. *Frontiers in Robotics and AI* 4 (2017), 33. https://doi.org/10.3389/frobt.2017.00033

[19] John M. Govern and Lisa A. Marsch. 2001. Development and Validation of the Situational Self-Awareness Scale. *Consciousness and Cognition* 10, 3 (2001), 366–378. https://doi.org/10.1006/ccog.2001.0506

[20] Jiajing Guo, Yoyo Tsung-Yu Hou, Harley Mueller, Katherine Tang, and Susan R. Fussell. 2019. As If I Am There: A New Video Chat Interface Design for Richer Contextual Awareness. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3312759

[21] Edward Twitchell Hall. 1966. *The hidden dimension*. Vol. 609. Garden City, NY: Doubleday.

[22] Steve Harrison and Paul Dourish. 1996. Re-place-ing space: the roles of place and space in collaborative systems. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work (CSCW '96)*. Association for Computing

Machinery, New York, NY, USA, 67–76. https://doi.org/10.1145/240080.240193
[23] Tilo Hartmann, Werner Wirth, Holger Schramm, Christoph Klimmt, Peter Vorderer, AndrÃ© Gysbers, Saskia BÃ¶cking, Niklas Ravaja, Jari Laarni, Timo Saari, Feliz Gouveia, and Ana Sacau. 2015. The Spatial Presence Experience Scale (SPES). *Journal of Media Psychology: Theories, Methods, and Applications* 1 (2015), 1–15. https://doi.org/10.1027/1864-1105/a000137
[24] Michael Havranek, Nicolas Langer, Marcus Cheetham, and Lutz Jäncke. 2012. Perspective and agency during video gaming influences spatial presence experience and brain activation patterns. *Behavioral and Brain Functions* 8, 1 (2012), 34. https://doi.org/10.1186/1744-9081-8-34
[25] Yasamin Heshmat and Carman Neustaedter. 2021. Family and Friend Communication over Distance in Canada During the COVID-19 Pandemic. In *Designing Interactive Systems Conference 2021*. ACM, Virtual Event USA, 1–14. https://doi.org/10.1145/3461778.3462022
[26] Angel Hsing-Chi Hwang, Cheng Yao Wang, Yao-Yuan Yang, and Andrea Stevenson Won. 2021. Hide and Seek: Choices of Virtual Backgrounds in Video Chats and Their Effects on Perception. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 303 (oct 2021), 30 pages. https://doi.org/10.1145/3476044
[27] Wijnand A. Ijsselsteijn, Huib de Ridder, Jonathan Freeman, and S. E. Avons. 2000. Presence: Concept, determinants and measurement.
[28] Robert JK Jacob and Keith S Karn. 2003. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The mind's eye*. Elsevier, 573–605.
[29] Matthew K. Miller, Martin Johannes Dechant, and Regan L. Mandryk. 2021. Meeting You, Seeing Me: The Role of Social Anxiety, Visual Feedback, and Interface Layout in a Get-to-Know-You Task via Video Chat.. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3411764.3445664
[30] Seungwon Kim, Sasa Junuzovic, and Kori Inkpen. 2014. The Nomad and the Couch Potato: Enriching Mobile Shared Experiences with Contextual Information. In *Proceedings of the 18th International Conference on Supporting Group Work*. ACM, Sanibel Island Florida USA, 167–177. https://doi.org/10.1145/2660398.2660409
[31] Raymond W Kulhavy, William A Stock, Kristina A Woodard, and Robert C Haygood. 1993. Comparing elaboration and dual coding theories: The case of maps and text. *The American journal of psychology* (1993), 483–498.
[32] Pine Labs. 2021. ohyay.co. https://ohyay.co/
[33] Martin Lea and Russell Spears. 1992. Paralanguage and social perception in computer-mediated communication. *Journal of Organizational Computing and Electronic Commerce* 2, 3-4 (1992), 321–341.
[34] Qian Liu, Yanyun Wang, Qingyang Tang, and Ziwei Liu. 2020. Do You Feel the Same as I Do? Differences in Virtual Reality Technology Experience and Acceptance Between Elderly Adults and College Students. *Frontiers in Psychology* 11 (2020), 2555.
[35] Michael Massimi and Carman Neustaedter. 2014. Moving from talking heads to newlyweds: exploring video chat use during major life events. In *Proceedings of the 2014 conference on Designing interactive systems*. 43–52.
[36] Albert Mehrabian and Shirley G. Diamond. 1971. Seating Arrangement and Conversation. *Sociometry* 34, 2 (1971), 281–289. https://doi.org/10.2307/2786417 Publisher: [American Sociological Association, Sage Publications, Inc.].
[37] Ken G Meleshko and Lynn E Alden. 1993. Anxiety and self-disclosure: toward a motivational model. *Journal of Personality and Social Psychology* 64, 6 (1993), 1000.
[38] Matthew K. Miller, Regan L. Mandryk, Max V. Birk, Ansgar E. Depping, and Tushita Patel. 2017. Through the Looking Glass: The Effects of Feedback on Self-Awareness and Conversational Behaviour during Video Chat. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 5271–5283. https://doi.org/10.1145/3025453.3025548
[39] Bonnie A Nardi, Heinrich Schwarz, Allan Kuchinsky, Robert Leichner, Steve Whittaker, and Robert Sclabassi. 1993. Turning away from talking heads: The use of video-as-data in neurosurgery. In *Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems*. 327–334.
[40] David T. Nguyen and John Canny. 2007. Multiview: improving trust in group video conferencing through spatial faithfulness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. Association for Computing Machinery, New York, NY, USA, 1465–1474. https://doi.org/10.1145/1240624.1240846
[41] Catherine S. Oh, Jeremy N. Bailenson, and Gregory F. Welch. 2018. A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI* 5 (2018), 114. https://doi.org/10.3389/frobt.2018.00114
[42] Kenton O'hara, Jesper Kjeldskov, and Jeni Paay. 2011. Blended interaction spaces for distributed team collaboration. *ACM Transactions on Computer-Human Interaction* 18, 1 (May 2011), 3:1–3:28. https://doi.org/10.1145/1959022.1959025
[43] Ken-Ichi Okada, Fumihiko Maeda, Yusuke Ichikawaa, and Yutaka Matsushita. 1994. Multiparty videoconferencing at virtual social distance: MAJIC design. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. 385–393.
[44] John O'keefe and Lynn Nadel. 1978. *The hippocampus as a cognitive map*. Oxford university press.

[45] J. Curtis Russell, Ira J. Firestone, and Reuben M. Baron. 1980. Seating Arrangement and Social Influence: Moderated by Reinforcement Meaning and Internal-External Control. *Social Psychology Quarterly* 43, 1 (1980), 103–109. https://doi.org/10.2307/3033752 Publisher: [Sage Publications, Inc., American Sociological Association].

[46] Carsten C. Schermuly and Wolfgang Scholl. 2012. The Discussion Coding System (DCS):A New Instrument for Analyzing Communication Processes. *Communication Methods and Measures* 6, 1 (2012), 12–40. https://doi.org/10.1080/19312458.2011.651346 Publisher: Routledge _eprint: https://doi.org/10.1080/19312458.2011.651346.

[47] Lee Sproull and Sara Kiesler. 1986. Reducing social context cues: Electronic mail in organizational communication. *Management science* 32, 11 (1986), 1492–1512.

[48] Jonathan Steuer. 1992. Defining virtual reality: Dimensions determining telepresence. *Journal of communication* 42, 4 (1992), 73–93.

[49] John Sweller, Jeroen JG Van Merrienboer, and Fred GWC Paas. 1998. Cognitive architecture and instructional design. *Educational psychology review* 10, 3 (1998), 251–296.

[50] Ron Tamborini and Nicholas Bowman. 2010. Presence in video games. *Immersed in Media: Telepresence in Everyday Life* (Jan. 2010), 87–109.

[51] Christina Underhill and Murrey G. Olmsted. 2003. An Experimental Comparison of Computer-Mediated and Face-to-Face Focus Groups. *Social Science Computer Review* 21, 4 (Nov. 2003), 506–512. https://doi.org/10.1177/0894439303256541 Publisher: SAGE Publications Inc.

[52] Brenda K Wiederhold. 2020. Connecting through technology during the coronavirus disease 2019 pandemic: Avoiding "Zoom Fatigue". , 437–438 pages.

[53] Werner Wirth, Tilo Hartmann, Saskia Böcking, Peter Vorderer, Christoph Klimmt, Holger Schramm, Timo Saari, Jari Laarni, Niklas Ravaja, Feliz Ribeiro Gouveia, Frank Biocca, Ana Sacau, Lutz Jäncke, Thomas Baumgartner, and Petra Jäncke. 2007. A Process Model of the Formation of Spatial Presence Experiences. *Media Psychology* 9, 3 (May 2007), 493–525. https://doi.org/10.1080/15213260701283079

[54] Rui Zhu and Jennifer J. Argo. 2013. Exploring the Impact of Various Shaped Seating Arrangements on Persuasion. *Journal of Consumer Research* 40, 2 (2013), 336–349. https://doi.org/10.1086/670392 Number: 2.