Investigating Social Robot Proxies to Support Situationally Impaired Users in Hybrid Meetings

Audrey Balaska
The University of Melbourne
Melbourne, Australia
draudreybalaska@gmail.com

Jens Emil Sloth Grønbæk
Aarhus University & The University of Melbourne
Aarhus, Denmark
jensemil@cs.au.dk

Chuhan Zhang
The University of Melbourne
Melbourne, Australia
chuhan.zhang.professional@outlook.com

Sarah Schömbs
The University of Melbourne
Melbourne, Australia
sschombs@student.unimelb.edu.au

Wafa Johal
The University of Melbourne
Melbourne, Australia
wafa.johal@unimelb.edu.au

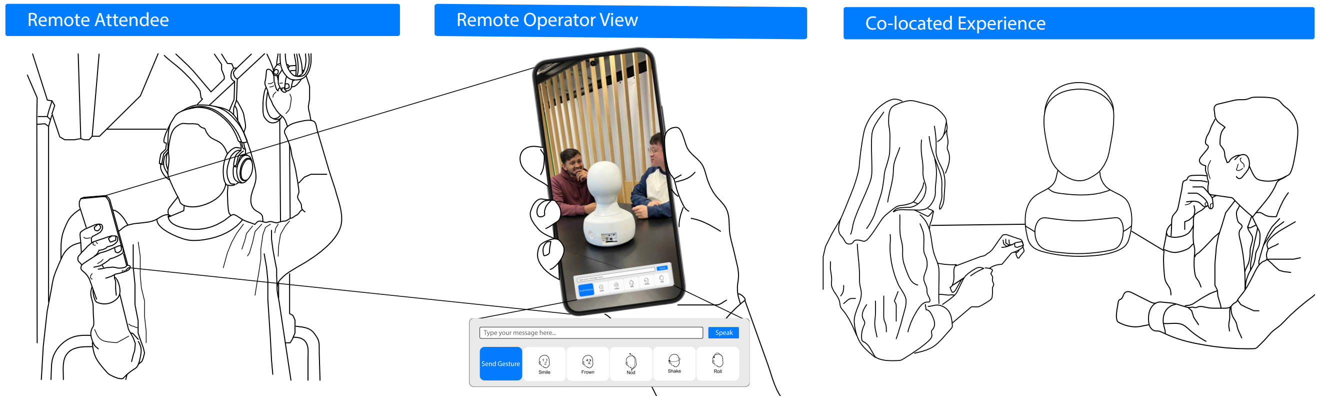


Figure 1: Illustration of scenario where remote participant is experiencing a situational impairment (being on the bus) and uses their handheld device to engage with the in-room participants using the telepresence robot system.

Abstract

Hybrid videoconferencing often results in conversational dominance by physically present participants, particularly when remote members are unable to use their camera or microphone due to situational constraints. This paper invites exploration into the context of remote situational impairment and the potential of re-embodying a remote meeting participant via a robotic surrogate without active user sensing. We integrate a humanoid social robot with videoconferencing software to develop formative insights about this design space. We conducted an exploratory user study involving four triads (each with one remote and two co-located participants). Using thematic analysis of participant interviews, we identify emergent themes that inform design guidelines for synthesized telepresence via a robot proxy. The resulting design guidelines highlight the importance of a robot proxy system's ability to actively capture and sustain the attention of co-located participants and to support shared awareness of the remote user's intent.

AHs 2026, Okinawa, Japan

ACM ISBN 979-8-4007-2351-3/26/03
https://doi.org/10.1145/3795011.3795038

CCS Concepts

• **Human-centered computing → User studies; User interface design.**

Keywords

Human-Robot Interaction, Situational Impairment, Human-Computer Interaction, Social Robot Proxy, Hybrid Meetings, Robot Avatars

ACM Reference Format:
Audrey Balaska, Jens Emil Sloth Grønbæk, Chuhan Zhang, Sarah Schömbs, and Wafa Johal. 2026. Investigating Social Robot Proxies to Support Situationally Impaired Users in Hybrid Meetings. In *The Augmented Humans International Conference 2026 (AHs 2026), March 16–19, 2026, Okinawa, Japan.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3795011.3795038

1 Introduction

Telecommuting, the practice of using technology to facilitate remote and work-from-home opportunities, has become an increasingly common practice since it was introduced as a formal concept in 1973 by Jack Nilles [32, 52], especially in the aftermath of the Covid-19 pandemic [4, 28]. Nowadays, people can remotely join a meeting from anywhere, including from the train, the airport, or from home, while other team members sit co-located together in a meeting room. The additional benefits that these more flexible working arrangements have provided (e.g. reduced financial and environmental burden [54]) have led many experts to conclude that *hybrid meetings* (with both *remote* and *co-located* participants) will now continue as a feature of many work environments [20, 28, 58]. A phenomenon in hybrid meetings which has received considerable attention in prior research is *primary room dominance*, in which remote participants are marginalized compared to their co-located counterparts [24, 39, 66, 67].

However, a less explored aspect of primary room dominance is how to address instances of *situational impairment* for remote meeting attendees. Wobbrock defines *situationally-induced impairments* as "[impairments] caused by situations, contexts, or environments that negatively affect the abilities of people interacting with technology [79]." In particular, some remote meeting attendees encounter situations where they cannot use their computer system's microphone and/or camera, due to reasons such as child privacy concerns [38], environmental challenges with noise[72], or technical issues (e.g. internet connectivity) [77]. These situational impairments can hinder remote users' ability to participate effectively in hybrid collaborative scenarios, with remote participants being relegated to supporting roles and co-located members taking on more influential tasks [10]. Traditional systems often attempt to address this challenge by providing a text-based chat interface and reaction buttons, enabling remote users to communicate with co-located participants through written messages or emoticons. However, this approach restricts remote attendee's expressiveness and social *presence* (a crucial factor for successful collaboration [6]), exacerbating the phenomenon of primary room dominance. The slower pace of typing relative to speaking [30, 65] further contributes to communication imbalances, and the absence of a video feed removes remote participants' nonverbal communication cues entirely, rather than just reducing their effectiveness [67].

One approach to improve remote participants' engagement in hybrid meetings has been to utilize mobile teleoperated robots [11, 13, 14, 21]. Most often, these robots serve as a portable remote conferencing interface, with the robot acting as the vehicle to move around traditional remote conferencing communication tools (e.g. microphone, speaker, camera, screen) [40, 44, 75]. However, they require remote operators to be able to share audio and video, and their largest benefit is their portability rather than enhancing sit-down meetings [75]. Limited work has been conducted regarding hybrid meetings and situational impairment despite the growing occurrence of such scenarios, motivating the exploration of solutions addressing this specific context.

Inspired by research in social robotics and telepresence, we propose a Synthesized Telepresence System where a humanoid robot head is used as a telepresence proxy, transforming the remote user's input into rich verbal and non-verbal behaviors. By offloading the expressive output to a physically embodied agent, the system amplifies the remote user's presence without requiring active sensing of the remote user and allows us to conduct formative work regarding robot appearance and behavior in this unique use case. To explore the potential of such a system, we created a collaborative scenario involving triads (one remote and two co-located users). Following their interaction with the system, we conducted interviews with all participants to explore their perceptions of usability, collaboration, presence, and the robot's appearance.

Results from the thematic analysis of these interviews were used to inform generalized design guidelines for synthesized telepresence via a robot proxy. As part of these guidelines, we identify the importance of attracting attention, communicating intention, mitigating conversational latency, and conveying emotion, leading to new avenues for research in using social human-like robots as proxies for remote collaboration. These guidelines motivate further exploration regarding autonomous system behavior in instances of situational impairment in hybrid meetings.

In summary, this paper presents the following contributions:

- Introduction of a robot proxy system as a solution to specifically address remote situational impairment in hybrid meetings
- Identification of key insights from interviews of both remote and co-located hybrid meeting participants regarding synthesized telepresence in instances of situational impairment, highlighting tensions and overlaps in preferences between the two groups
- Development of generalized design guidelines for synthesized telepresence via a robot proxy

2 Background and Related Work

Our work draws upon hybrid meeting studies, Human-Robot Interaction literature (HRI), and research on situational impairment.

2.1 Situational Impairment and HCI

The specific concerns of situational impairment are under-explored in the context of hybrid meetings. Research regarding Situationally Induced Impairments and Disabilities (SIIDs) focuses on the notion that external circumstances can negatively alter the user experience when interacting with a technology [36, 71]. Designing systems that mitigate situational impairments results in more equitable user experiences and can improve overall system functionality. Most commonly, research efforts have focused on mobile technologies such as cell phones [69, 78], addressing challenges such as cold weather [68], water droplets on the screen [76], motion [47], and encumbrance [51]. Limited work has been conducted regarding situational impairment and videoconferencing or situational impairment and human-robot interaction. Work by Liu et. al. has proposed a general Human-Computer Interaction approach to detecting situational impairments, though the effectiveness of applying this to the context of human-robot interaction has not yet been explored and the work does not focus on resolving these situational impairments [45].

Looking at the broader literature regarding hybrid meetings, particularly in the face of their increased prevalence during and after the Covid-19 pandemic [26], some situational impairments specific to videoconferencing and hybrid meetings have been identified, though not necessarily classified as such. This includes concerns about privacy of recording individuals at home [53] and internet connectivity challenges [22].

However, little work has been conducted regarding remote users' situational impairments in hybrid meetings. This is despite the timely need for solutions to support calling from public locations where using camera and microphone is not possible or speaking violates the social norms of their environment [72]. This paper addresses this gap by facilitating synthesized telepresence via a robot proxy in cases where remote meeting attendees cannot use their camera or microphone.

2.2 Hybrid Meetings and Addressing Primary Room Dominance

Different works have explored how to mitigate the challenges of hybrid meetings through improved video- and audio-presence of remote attendees, e.g., via adaptable video windows [24], spatialized audio [33], or integration of 2D and 3D interfaces [74]. Moreover, other studies [16, 56] emphasize the importance of providing a spatially integrated presence in hybrid environments [80]. The ongoing research being conducted in this space highlights that this is a timely and relevant – yet unresolved – issue.

The concept of using teleoperated robots to support remote human-human communication has been applied to contexts with distributed teamwork, facilitating collaboration among geographically dispersed teams [75]. Unlike traditional video conferencing, telepresence robots offer remote participants the ability to move within the physical space, enabling them to communicate not just verbally but also through spatial positioning and non-verbal cues [11–13, 21].

However, even for robotic solutions, these approaches often require that remote users are in an environment where it is situationally appropriate for them to film themselves and speak, in order to effectively exert a presence through audio and video channels. To overcome this fundamental problem, we investigate the ability for a physical robotic proxy to enhance situationally impaired remote participants' presence in the co-located space, enabling them to engage more actively in interactions and reduce primary room dominance—without relying on traditional audiovisual tools.

2.3 Non-Verbal Communication in Hybrid Meetings

Embodiment in telepresence refers to the extent to which a remote user can experience a physical presence in a distant location. This concept extends beyond mere physical movement and includes the ability to engage in social interactions in a way that feels natural and integrated within the remote environment. Research has shown that users perceive telepresence robots as more socially engaging compared to static video displays because of their ability to embody users' movements and gaze direction [55]. Theoretical discussions on embodiment emphasize the importance of control, affordances, and the seamless mapping of users' expressive capabilities to the robotic platform. For example, smooth control of the robot's movements and the ability to navigate the environment autonomously are crucial in enhancing the naturalness of interaction and the user's sense of presence [21, 41–43, 46, 70]. Some investigations regarding embodiment have focused on specific elements of physical presence, such as the effect of robot height on conversational dominance [61] or the combination of a robot arm with a video screen as a means to increase presence [55].

In this study we focus on embodiment in telepresence using the humanoid Furhat robot system [5], utilizing its abilities to change head orientation and mimic human facial expressions. The anthropomorphic design of the robot head allows presence and nonverbal communication to be presented in a natural way, in line with recommendations from telepresence robotics experts [1]. Additionally, users are more engaged when interacting with robots exhibiting complex motion [1]; for this reason, we incorporate simulated autonomy with the robot's head gaze being directed at whichever co-located meeting attendee is speaking, with the intention to incorporate additional autonomous features based on this study's results. Other studies incorporating the Furhat robot system for telepresence have focused on solutions that require active sensing of the remote operator [3, 25]. Unlike other work, this study focuses on the context of situational impairment where such sensing is not possible and motivates exploration of alternative control options [57].

From the perspective of the remote operator, telepresence robots provide a sense of physical agency in a remote environment, which can lead to increased social presence and engagement. However, from the perspective of co-located teammates, the introduction of a telepresence robot can sometimes lead to disruptions in group dynamics, especially in scenarios where several individuals are physically present and only one is remote [21]. Such situations exacerbate issues of situational impairment, as the remote participant may struggle to fully integrate into the social and spatial context of the meeting [35]. This imbalance highlights the ongoing challenge of achieving seamless collaboration in hybrid settings, motivating continued exploration of robot proxy solutions. By collecting insights from remote and co-located meeting attendees attending the same meeting, we hope to facilitate future synthesized telepresence systems to be designed with awareness of how design decisions may impact the whole group.

By using a humanoid robot head, we believe we can address the important context of situational impairment in hybrid meetings, as this modality will allow a remote meeting attendee to have non-verbal presence without relying on audio, video, or real-time tracking. The evaluation study allows us to investigate the potential of this solution, motivate more extensive exploration in this area, and provide initial guidelines to support early-stage prototyping.

3 Prototype System Design

Our synthesized telepresence prototype system consists of two main components: the co-located, "in-person" robot proxy and the "remote" online interface. The Furhat [5] robot functions as the co-located representation of the remote user, giving them a physical presence that facilitates bidirectional verbal and nonverbal communication. The remote user controls the robot with a web interface, including a video stream of the co-located space.

3.1 Re-Embodiment: Visual Appearance

The Furhat robot has a consistent physical face structure, created using a plastic semi-transparent mask. The robot head is able to rotate in the pitch, roll, and yaw dimensions; allowing the robot to make gestures that look like nodding, shaking of the head, and tilting the head to the side.

More refined details are provided by the internal projector, facilitating additional gestures such as winking, blinking, smiling, and talking animations. These projections also allow for variations in appearance, including changing the eye color, skin tone, and inclusion of makeup. Though the physical mask does not change, changes in the shading and coloring of the projected face can give the illusion of different bone structures, though the variation is limited in comparison to the true amount of variation seen with human faces [34, 62]. An example of one of the Furhat character avatars can be seen in Figure 3c.

With this initial iteration of the synthesized telepresence system, we have started to investigate the concept of *re-embodiment*, where the user is represented with a robotic presence reminiscent of their own physicality. In this study, this was achieved by either taking a photo of the remote user and mapping it onto the robot face, or allowing users to select their preferred pre-made face from a selection of face options. This is a relatively low-fidelity prototype to serve as a proof-of-concept and determine if further work regarding robot personalization and re-embodiment is justified. The user selected a text-to-speech voice from a selection of audio clips, though vocal embodiment was not a principal focus of the study.

3.2 Communication Modalities

With our system, the remote participant can communicate both verbally and nonverbally with the co-located participants. For the remote participant, the verbal and nonverbal sensing and communication are conducted with the co-located robot proxy acting as an intermediary agent and the user acting as an off-site processor and controller.

For our prototype system, we utilize some of the existing Furhat robot capabilities. Specifically, we use inbuilt gesture commands and the built in text-to-speech processor. The portion of the interface used to control the robot is shown in Figure 2. To have the robot speak, the user must type text into the text box, then select the “Speak” button. This text-to-speech feature is not the same pace as real-time speech, but provides some form of audio communication without requiring real-time sensing of the remote meeting attendee. To have the robot complete a pre-determined gesture, the user selects one of the included buttons (“nod”, “shake”, “roll”, “smile”, “frown”). The remote user interface including this control interface and the co-located video feed is included in Figure 3b. In this prototype, a Wizard-of-Oz [63] configuration is used to change the gaze of the robot to either one of the co-located participants or straight ahead in response to who is speaking, mimicking robot audio localization [7].

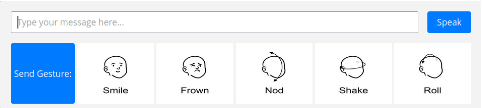


Figure 2: Robot operation interface.

3.3 Co-located Video and Audio Streaming

The interface includes audio and video streaming of the co-located space. The microphone is located between the co-located meeting participants and the camera is located behind the robot facing the co-located participants (see Figure 3c). The camera view is intended to mimic third-person perspective in video games [2], allowing for the remote meeting participant to see both their fellow meeting attendees as well as some of the positions/actions of the robot in the room.

4 User Study

In this study, we investigate the following research questions via a formative study using our Synthesized Telepresence System:

1. To what degree does a situationally impaired remote participant feel that a "re-embodied" robot proxy can represent their presence in a hybrid meeting?
2. How do design characteristics and control features of a telepresence robot proxy affect hybrid meeting experiences of situationally impaired remote participants and their co-located peers?

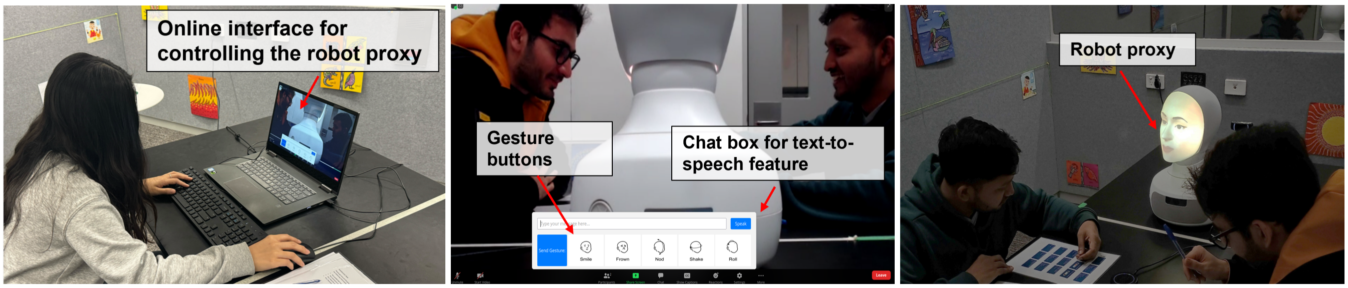
To help address our research questions, we orchestrated a discussion-focused hybrid meeting scenario. In the scenario, there were three people participating in a meeting; two of these participants were co-located “in-person”, and the third participant was both “remote” and unable to use their computer’s microphone or camera. The Furhat robot proxy is located in the co-located meeting room, with the remote participant operating it in a separate room. The remote operator uses the online interface to both control the robot and have a view of the robot co-located meeting participants (Figure 3b). Figure 3 shows staged photos highlighting how participants were situated.

The participants performed three discussion-based tasks. A thematic analysis of post-discussion interview responses comprised the bulk of the data collection and processing. This study focused on open-ended interviews as a means of data collection due to the unique context of the system. The sample size is informed by other interview-based studies in HCI literature [17].

4.1 Study Design

Within the group of participants, one was assigned as the “remote attendee”, while the other two were “co-located”. Before completing any tasks, the remote participant selected an avatar and voice from a prepared set of Furhat characters and voice settings. After this was done, the remote participant went into one room (see Figure 3a) and the two co-located participants went to another (Figure 3c). After completing three tasks as a team - with the remote attendee using the online interface (Figure 3b) to control the robot and view their fellow participants - each team member was independently interviewed by a member of the research team about the experience. During the tasks, autonomous gaze behavior was simulated with a Wizard-of-Oz approach (see Section 3.2).

4.1.1 Introduction Task. First, participants were guided by the facilitator to introduce themselves, including their name, preferred pronouns, favorite color, and favorite book or movie. This is a turn-taking task where there is roughly equal amounts of engagement from all participants. With a task like this, in-room dominance is minimized due to the more structured nature of the conversation. Besides introducing participants to each other, this also introduces the participants to the system. Other discussion-based Human-Computer Interaction (HCI) research has also included guided introductions as a way to transition into more intensive discussion



(a) "Remote" room **(b) Online interface** **(c) "Co-located" room**

Figure 3: Staged photos of experimental setup.

tasks [24, 74]. Given the concise, structured nature of the task, there was no set time to complete the introductions.

4.1.2 *NASA Survival on the Moon Task.*

In this task, the group is given instructions describing a space mission scenario, where a space ship has landed on the moon far away from the lunar base that is their final destination. They are given a list of supplies, and asked to rank the importance of the various materials.

Our task is a variation of the one introduced in 1963 as part of an interpersonal relations study [27]. Similar to other studies [60][74][60][49], we made slight alterations. The list was shortened to 8 items, in order to reduce the amount of time needed to reach a consensus. Participants were given five minutes to discuss their solution.

For the purposes of this study, we were not concerned with participants reaching a "best" answer, focusing instead on how participants communicated as they worked to reach a consensus. Each participant was provided with the task instructions and a worksheet to be filled in with their ordered list; the group was instructed that they should all have the same answers on their sheets at the end of the five minutes, and that "It is most important for all team members to have the same things written, even if that means you do not have all 8 items written down."

4.1.3 *Event-Planning Task.*

This task was similar to the NASA Survival on the Moon task, in that the priority was for participants to reach a consensus, rather than reaching a correct answer. This task required the participants to choose from a "menu" of items for the event, while staying within a budget. This included beverage and dinner offerings, color scheme, furniture, and bonus decorations. The "cost" of items was varied enough that participants could not choose the most expensive options in every category and stay under budget, motivating discussion. Planning tasks such as this one have been utilized in similar HCI studies [24][64], though for this task we did not replicate a specific study design. Like the NASA Survival on the Moon Task, participants had 5 minutes for deliberation and were instructed to prioritize agreement rather than completeness when filling out the worksheet.

4.1.4 *Post-Study Interview.*

After the meeting concluded, each participant was interviewed individually by a member of the research team. They were all asked the same initial set of questions, with additional questions being asked based on their responses and/or events that occurred during the activities. The same baseline questions were asked to both the remote and in-person participants, so that the perception of the system from both perspectives could be considered in the evaluation of the prototype. Interviews were audio-recorded and transcribed for thematic analysis.

5 Results

A total of six sessions were conducted, though two were discarded due to technical difficulties with the recording equipment. For the remaining four sessions, each participant was interviewed individually, creating a total of 4 interviews of the remote participants and 8 interviews of the co-located participants. Participants were recruited from the local university community; as part of the individual recruitment process, interested participants shared their availability and were assigned to a session by the facilitator in order to reduce the likelihood of familiarity among participants. This study was approved by our institution's Ethics Committee, and all participants consented to the study.

A thematic analysis [15] was conducted regarding the responses to the 12 semi-structured interviews. First, one member of the research team used the transcripts to create an initial set of codes regarding situational impairment and remote collaboration. Then, three researchers from this project refined the codes to identify themes and sub-themes. After a rigorous familiarisation with the data, the researchers coded the interview transcripts with this refined code list, with each interview transcript being coded by at least two members of the research team. The codes for each interview were reconciled across the research team to create a final collection of codes, themes, and sub-themes.

The results of this analysis are included in the appendices, highlighting elements of the remote and co-located interviews, respectively. In the following sections, we refer to co-located participants with a 'CL' and remote with an 'R'. Additionally, in the labeling, the 'S' refers to their session, and the numbers '1' and '2' are used to distinguish between the co-located participants.

5.1 Interaction and Communication

5.1.1 *Physical Presence and Engagement.*

The presence of a 3D, interactive robot co-located with the in-room participants influenced their level of engagement with the system. Participants expressed sentiments such as the *robot's physical presence is helpful* (S3-R, S4-R, S3-CL1, S5-CL1, S5-CL2, S6-CL2), *robot gestures are helpful* (S3-R, S3-CL2, S4-CL2, S6-CL2, S5-CL1), *robot movement is helpful* (S5-R, S4-CL2, S5-CL1, S5-CL2, S6-CL1, S6-CL2), and *robot gaze behavior improves engagement* (S4-CL2). For instance, S3-CL1 felt that, "The sight that it's actually there...you're consistently aware of this, that you have to ask for [their] opinion". For S4-CL2, "it [the robot] feels like it has eye contact and it's willing to listen to you and willing to participate in the conversation." These insights suggest that the robot's physicality, combined with its movement and gaze behaviors, fostered a greater sense of presence and inclusion for remote participants.

5.1.2 Interaction Confirmation. While remote users could control the robot's facial expression through GUI buttons, the absence of proprioceptive feedback made it difficult for them to confirm whether the gesture was executed correctly and when. One remote participant expressed a desire to have a live *display of the robot's current facial expression* on their screen. They explained that this would provide continuous reminder of their presence, stating, "I wanted to get a continuous reminder that I was not just a block, a piece of plastic somewhere. I was actually being represented as something with a face and with eyebrows or lips" (S5-R).

Additionally, participants mentioned *utilizing the in-room video feed* for interaction confirmation – whether it be to ensure the robot was responding to commands (S5-R) or to check if in-room participants were paying attention to them (S4-R, S5-R). This highlights the importance of proprioceptive cues for the remote participant to ensure effective communication.

5.1.3 Explicit Intent Expression. Both co-located and remote participants highlighted the importance of the robot's ability to effectively *signal when it intended to speak* (S3-CL1, S4-CL1, S5-CL1, S5-CL2, S6-CL2, S5-R). In particular, there was a strong desire for a *typing indicator* to inform in-room participants when the remote participant was typing (S4-CL1, S5-CL2, S6-CL2, S5-R). For in-room participants, this need was often driven by the sentiment that it was *hard to predict what the robot was going to do* (S3-CL1, S5-CL1, S6-CL2), making it difficult for the co-located participants to gauge when the remote participant intended to contribute to the conversation. Additionally, participants suggested the need for a mechanism *to communicate system error* to the remote participants; this recommendation arose particularly from the group that experienced the technical issue of an internet disconnection (S3-R, S3-CL1, S3-CL2). During these moments, co-located participants were unsure whether the remote participant was simply quiet or encountering a technical problem, highlighting the importance of transparent communication such as active listen cues or back-channeling.

5.2 Appearance and Use Cases

5.2.1 Embodied or Communication Tool. For the remote participants, there was a strong emphasis on the robot embodying the human operator. S5-R described the *robot as a representation/extension of the user*: "Towards the end of the experiment, I stopped displaying those non-verbal cues in this room and I started displaying them solely through the robot itself." There was also a *desire for the robot to look like self* (S5-R, S6-R). For participants who chose a built-in robot face, their selection often reflected their personal "identity" (S6-R). Beyond resemblance, some participants also expressed an interest in the robot having a *friendly face and voice* (S5-R, S6-R, S4-CL2).

In-person participants expressed mixed opinions about how much the robot should embody a human. Some felt that a *human face is better than robotic face* (S4-CL1), suggesting to add more human-like features by *adding more "body parts" to robot* (S5-CL1, S5-CL2) or *adding "hair" to robot* (S5-CL1, S6-CL1). There was also interest in *making the voice more human-like* (S5-CL1, S6-CL1).

Conversely, other participants believed that a *robotic face was better than a human face* (S4-CL1) and considered the *robot as a tool for human to communicate* (S6-CL2). One participant explained their preference by saying, "I don't feel like project[ing] someone's face onto a robot is really presenting himself. I think robot and human [are] different" (S6-CL2). There were also *concerns about uncanniness* (S3-CL2, S5-CL1, S6-CL1) along with *concerns about privacy with re-embodied face/voice* (S6-CL2) when the robot is too human-like.

These differing views were reflected in the discussion about the current robot face, which for one participant was a projected photo of themselves and for the others was selected from pre-made robot character packs. Some felt that the *current face is good* (S3-CL2, S6-CL1, S6-CL2). Others were already experiencing effects of the uncanny valley, believing that the *current face is creepy* (S4-CL1).

One participant highlighted that the ideal appearance may depend on the context, noting that the *ability to customize appearance is important* (S5-CL2) to better suit different situations or preferences.

5.2.2 Privacy vs. Transparency. Participants had differing perspectives regarding the level of privacy the system provided the remote participant. For in-person participants, this sometimes created a sense of *uncertainty about remote participant*, where they "don't know what's going on" (S4-CL1) when the robot is speaking. Due to the fact that the *robot is not emotive*, it was also difficult for co-located participants to fully interpret the tone or emotion behind the remote participant's typed responses (S5-CL1,S6-CL1). Participant S6-CL1 stated, "... he's saying the sentence - and it might be he's neutral, it might be he's happy, he might be sad and you don't know" (S6-CL1).

Conversely, remote participants *appreciated the privacy* provided by the system. S4-R noted, "The best thing I like about this robot is that I can do anything I want without being seen by other people." This suggests that the lack of emotional transparency, while challenging for the in-person participants, offered remote users a sense of autonomy and comfort, allowing them to act freely without being observed.

5.2.3 Possible Applications. When asked about possible applications for the telepresence robot system, one participant discussed demographics of interest. In particular, they mentioned possibilities for *people with disabilities* and *people with technology challenges* to benefit from the system (S5-CL1).

5.3 Group Dynamics

5.3.1 Relative Engagement. Though the robot proxy was positively viewed by participants, instances of clear in-room dominance still occurred. From the perspective of the remote participants, their contributions were sometimes "quite less" (S3-R). There were also moments where the *remote participant felt ignored*, as reflected by S6-R, who noted "they're not listening to me". Interestingly, some co-located participants also acknowledged the uneven communication dynamic with *remote participant contribut[ing] less*: "in-person participants feel more like group setting" (S3-CL1).

However, there were moments where the *remote participant felt listened to*, such as with participant S4-R, who said, "I feel that I contribute a lot to the conversation because my team member listened to my idea" (S4-R). There was even the sentiment that *all participants were equal* across both the remote and co-located participants. When describing their remote experience, S4-R said, "I think I can participate in the meeting with other team member in the same way as in person." Another remote participant, when asked how much they felt they contributed to their team's solutions, classified it "as equally as the other two, even though it meant that I...had to contribute insights...once they had moved on from the topic a little bit" (S5-R). When asked if anyone contributed significantly more or less than their fellow participants, one of the co-located participants described the situation as "equally engaging" (S4-CL1).

In line with these other observations, the co-located participants often *felt listened to in-room*, feeling that their fellow participants "can listen to my idea and my reason for my decision" (S4-CL2).

When asked to make a direct comparison of the robot telepresence system to a chat-only hybrid zoom setup, multiple participants across both rooms expressed that a video call with *chat-only would be worse* than the robot telepresence system, as they "prefer [to interact] with the robot" (S4-CL2). While S6-CL1 shared that the robot proxy or *chat-only would be the same*, S5-CL1 expanded on the fact that the physical present of the robot helps in making sure remote participants are not forgotten in hybrid meetings:

> Compared to this? Yeah, definitely. I think if it's just a - basically - a screen with a chat, that would impact how the discussion [would] go. I think the physical person would- I think not purposely - but they get too deep into the physical interaction and they will just forget about the chat. Because that's happened to me a lot before (S5-CL1).

5.3.2 Attracting Attention.

We identified the ability to *capture attention* as a dominant theme in our analysis. Remote participants need to be able to convey their intent to contribute to a discussion and/or to become the focus point of attention. Our data shows both instances in which the remote participants were able to attract attention through the robot and instances in which they failed. We identified the robot's *voice* as a successful tool to attract attention, while *gestures* were found to be less effective. S5-R noted, "If I [made] a little gesture that my robot is displaying, they sometimes missed it because they just weren't looking at the robot. But in terms of everything that was verbalised, they did pay attention and take it into consideration." This highlights the critical role of voice in ensuring remote participants' contributions are noticed, whereas non-verbal cues were often overlooked by co-located participants.

5.3.3 Effects on Collaboration.

Both remote and co-located participants noted that successful hybrid collaboration depended on the *intentional inclusion of remote participant* by the two co-located participants. For some in-room participants, this was a clear, deliberate effort ("We're trying to understand and we're trying to include the other [remote] person" (S5-CL1)), while for others it was reflected in behavioral adjustments ("I kind of felt myself wanting to wait for her to type" (S5-CL2)).

Empathy for remote participant also played a role in the collaboration dynamic. Participant S5-CL1 shared, "I just felt bad for the person in the robot, because I felt that they weren't heard a lot." However, multiple co-located participants commented on the fact that they sometimes only included the remote participants to seek agreement (S3-CL1, S3-CL2). For example, S3-CL1 observed, "we are just confirming whether they [the person behind the robot] agree or not."

5.3.4 Conversation Flow and Timing.

Both co-located and remote participants identified that the *conversation slows down due to the system* compared to the conversational pacing of just co-located meeting participants. In this mixed-modality setup, *typing speed influences conversation flow*; S4-R argued "I think it's not as effective as in-person because I need more time to type on the keyboard", while S6-R felt that "typing...through text... it's very rudimentary, like it doesn't feel fast enough". It was also identified that *text not deleting caused repetition*, where when the text remained in the text box, the participant would sometimes accidentally send the same message again (S5-R).

The remote participants reflected on the sorts of conversational effects that occurred due to this slower pace. Specifically, it was identified that it was difficult to smoothly interject into the in-person conversation, *interrupting co-located speakers* or *needing to backtrack* the conversation when those in-person had already moved on to another topic.

> ...I had to butt in at times ... by the time I was done typing out the thing I wanted to say they probably already had moved on a little bit. So I think I had to dial the conversation back to the previous topic, but I think once I butted myself into that conversation, it was our conversation then. (S5-R)

These effects were also felt by the in-room participants, identifying both that there was a *robot repetition error* and *robot communication delay* ("the delay and some of the robot repeats" (S5-CL1)). Some co-located participants did note that they were *able to check for agreement* from the remote participant with the head gestures (S3-CL1, S3-CL2).

6 Discussion

Our work introduced synthesized telepresence via a robot proxy as a novel approach to tackle situational impairments of remote participants in hybrid meetings, leveraging its agency and physical presence to enhance inclusivity. Interviews with participants showed an overall appreciation for the system and provided interesting insights for guiding the design of robots as proxies for collaborative hybrid interactions. By examining the similarities and differences between the remote and co-located participants (see Figure 4), we formulate design guidelines that consider the experience of all meeting attendees in the case of remote situational impairment addressed using a robot proxy.

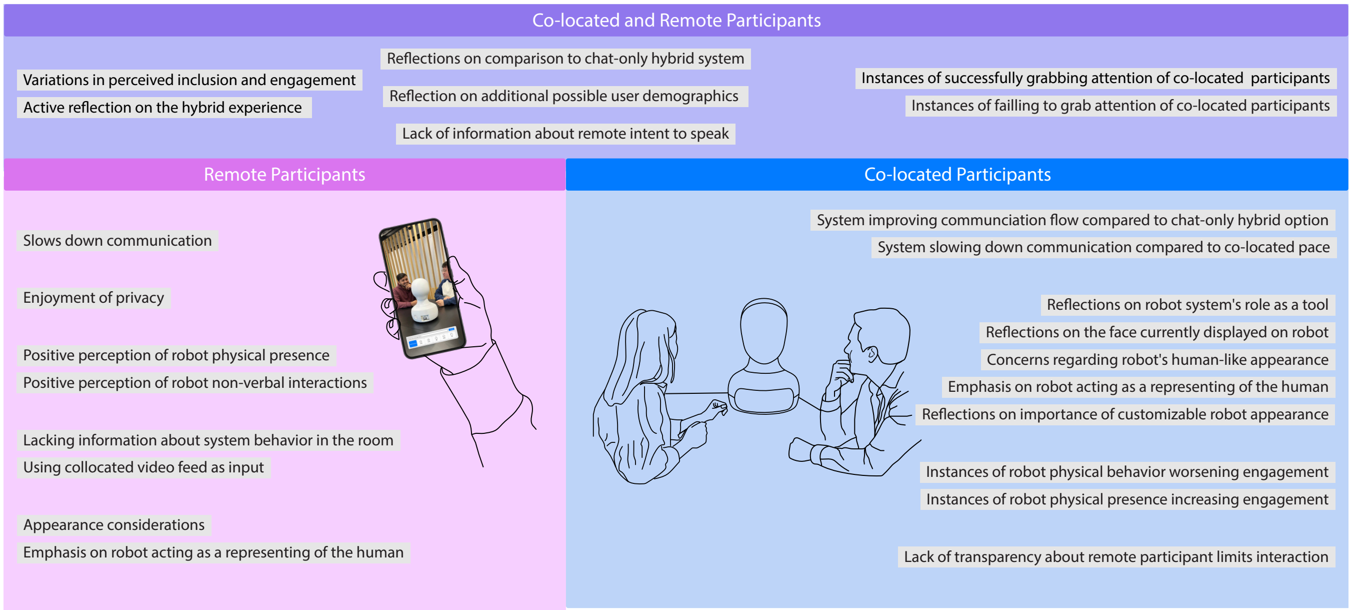


Figure 4: This figure organizes the sub-themes from our thematic analysis into three categories: insights from only remote participants (left/pink), only co-located participants (right/blue), and both remote and co-located participants (top/purple). Sub-themes grouped close together share a theme (the relationship between themes and sub-themes can also be seen in the appendices.

The **physical presence** of the robot emerged as a critical factor in including participants in the collaboration. By **reducing the in-room dominance—inherent to hybrid meetings and exacerbated by the situational impairment of remote participants**—the robot was perceived as a good alternative to prior experiences with solely screen-based solutions. This insight aligns with work by Biehl et al. [9], which found that embodied video interfaces enhanced co-located participants' awareness of their remote collaborators. Beyond that, our robot proxy system also improved opinions of the remote participants regarding the system and their own social presence, possibly because it allowed for directly transferable nonverbal communication, unlike other work. We hypothesize that the camera view for the remote participant, showing some of the ways in which the robot they controlled was interacting within the room, improved their own perceived engagement.

Additionally, having a **text-to-speech feature** was identified as a crucial element of the system, particularly as a means of **attracting attention of the co-located participants** when they are not already directing attention to the robot proxy, and subsequently the remote meeting participant. As the auditory text-to-speech consistently succeeded in grabbing the attention of the co-located participants, it motivates investigation into additional, non-verbal auditory cues as a means of attracting attention [37].

The two groups perceived the robot gestures differently. For co-located participants, the **gestures were one of the highlights of the system, facilitating an easy way to check for agreement and helping them stay engaged with the remote participant**. For the remote participant, the gestures were still useful, but there was higher awareness of **situations where the robot was making a gesture and the co-located participants missed it,** which could be observed via the camera feed and increased their perception of primary room dominance. Combining gestures with lighting or audio cues may be necessary to help attract attention [8].

Additionally, there was significant interest in creating a shared awareness of the current conversation state amongst all meeting members. In particular, co-located participants desired **increased situational awareness regarding the remote attendee**, whether that be via a **typing indicator** (so that co-located participants will wait to progress the conversation to another topic) or conveying the **emotional tone** of the remote person's responses. There was also desire amongst co-located participants to know more about the **state of the telepresence robot proxy**, so that inactive robot behavior could be accurately attributed to the remote participant or to the system's failure. Remote participants also wanted to be aware of the system state, relying on their video feed to get confirmation regarding the gestures that they sent. It could be valuable to explore how an adjustable view of the meeting space influences the interaction (reminiscent of the work by Spittle and Panda et. al. [74]). It seems likely that if this shared awareness was improved, the conversation flow would also improve.

Continuing on the theme of giving the remote participants knowledge on what is happening in the co-located space, the system would benefit from a **robot proxy self-view** showing the actions of the robot, particularly those that cannot be picked up by the camera. This reflects preferences of people regarding their own camera self-view in traditional videoconferencing [18], demonstrating interesting similarities in monitoring of embodiment for video recording and robot proxy representation. It is important to note that remote participants who desired more information specifically said they would like an additional window to show the robot's face, similar to existing self-view features in video conferencing platforms, rather than requesting the webcam for the system be moved to another location. This highlights the importance of maintaining both the existing camera angle and adding in a robot proxy self-view, and again supports the idea of an adjustable view of the meeting space. Furthermore, additional cues such as pointing or cues assisting co-attention could be introduced to enhance collaboration.

Regarding the robot's appearance, there was not a clear consensus on how much the robot should represent the operator. Remote participants conveyed more interest in the robot looking similar to whomever was operating it. However, the co-located participants that had to look directly at it expressed mixed opinions about the robot appearance, with some expressing concerns in-line with existing literature regarding the uncanny valley [50]. Though a remote participant may prefer to have a hyper-realistic version of their face on the robot, if co-located participants find that "creepy", it may be worth reducing the level of human-likeness. Alternatively, repeated exposure to the system may reduce the uncanny valley effect [81], and could be incorporated into the system deployment. As such, it seems that the **level of accurate representation through the robot would need to be customizable depending on both the user and the situation** (in line with similar social virtual reality embodiment literature [23]).

There were additional interesting cases of differences between the remote and co-located participants. **For co-located participants, the lack of live mapping of the remote participant's actions and emotions was perceived as a barrier to successful communication, while remote participants sometimes enjoyed it as a useful privacy perk of the system.** A nuanced approach to tackle this could be to enhance the robot proxy's expressivity via autonomously generated behaviors to suit the conversation; enhancing emotive communication while keeping the privacy of the remote user.

Continuing the discussion of robot autonomy in this system, the robot head-gaze behavior led co-located participants to feel that the robot was listening to them based on when the robot was looking at them. Notably, since this was a simulated case of autonomous auditory localization, it did not reflect whether the remote participant was paying attention. This motivates discussion about **robots mimicking human autonomy and how that influences perception of the remote participant**, reminiscent of discussions regarding traditional HRI Wizard-of-Oz studies and how human operators influence participants' perceptions of robot autonomy [59]. When is it misleading or deceptive to allow autonomous robot motion or actions, if ever?

6.1 Design Guidelines

Using the results of this thematic analysis, design guidelines for a robot proxy system to embody a situationally impaired remote meeting participant in a hybrid meeting are proposed. The guidelines are summarized below:

- **Attracting attention** The system must enable the remote participant to grab the attention of the co-located participants. In particular, auditory cues are necessary to attract immediate attention, while physical gestures help maintain sustained awareness of the remote participant throughout the interaction.
- **Turn-taking and communicating intention** The system must address the differences in conversation mediums, particularly that the remote participant does not have reflexive body language to communicate intent-to-speak nor the ability to automatically communicate nonverbally.
- **Mitigating conversational latency** In particular, it is important to provide methods for real-time communication from the remote participant, as typing is never going to be at the same pace as speaking. Both increasing typing speed and providing alternative means of communication (ex: nonverbal gestures) help improve conversation flow.
- **Communicating system state** Both remote and co-located participants depend on awareness of the system state, relying on knowledge that the system is working, or being alerted if there is an error.
- **Customizing appearance** There is no consistent preference regarding the robot's appearance - customization allows the appearance of the robot to be personalized to the remote operator's preference while also allowing consideration for the meeting participants that are co-located with it. There is an ongoing tradeoff between considerations regarding embodiment and the uncanny valley.
- **Conveying emotion** By implementing vocal tone and emotive facial expressions, the communication of the remote meeting attendee's intent and sentiments is improved, allowing them to be better represented in the meeting.
- **Leveraging system autonomy** Autonomous features increase presence of the remote attendee while maintaining privacy. With autonomous embodiment, the possibility of inaccurately representing the remote attendee's intent must be considered.

Based on the results of this study, we are motivated to improve upon our robotic telepresence prototype using our own design guidelines and then run more in-depth studies. In practice, this includes adding a light-based typing indicator to the base of the robot, adding a robot proxy self-view for the remote participant, and increasing the emotive abilities of the system. Additionally, future work should investigate methods for increasing shared awareness, explore options for customization of appearance, and work to identify methods and implications of autonomous robot proxy behavior when it's re-embodying a situationally impaired human. Investigations regarding different levels of autonomy could provide additional insights regarding the benefits and challenges of autonomous proxy behaviour; an example of this may include allowing the remote operator to directly control the head gaze of the robot, rather than having this be an autonomous process. This work also motivates a larger-scale study to directly compare this system to a chat-only or solely text-to-speech interface using quantifiable metrics such as eye-gaze[19], turn-taking [73], or speaking-time[48].

Alternative metrics such as task completeness should also be considered, as increased social presence does not necessarily correlate to task metrics[29]. Eventually, deploying this system "in the wild" could provide unique insights regarding the possible feasibility of large-scale adoption of such a system; especially given the known barriers to commercial social robot adoption [31].

6.2 Limitations

We acknowledge some limitations of this study. As noted, there were originally six groups, though two were removed from the thematic analysis due to data collection issues. The participants from Session 3 had their experience influenced by the robot system experiencing connectivity issues partway through the Survival on the Moon Task; this ended up being a benefit and informing useful insights regarding the system. These results are not generalizable, due to the small sample size, though they invite useful initial discussion and help motivate studies with greater participant sample size in the future.

Additionally, some participants noted that they felt pressure to complete the tasks within five minutes, which influenced how they behaved in the conversation. Though, given that meetings often have a set cut-off time, this is arguably reflective of typical hybrid meeting parameters. Some participants emphasized that a camera and/or microphone would improve the interaction, demonstrating that they did not fully understand the context of our situational impairment scenario. We do not believe this devalues their insights, but future work will better emphasize situational impairment. Though these results in the controlled lab environment are useful for initial discussion, implementing this study in a more natural situational impairment setting (such as public transportation, as highlighted in the example in Figure 1) will be necessary to comprehensively evaluate the effectiveness of synthesized telepresence in contexts of situational impairment. The system could also be evaluated in more diverse meeting configurations, where there may be more co-located participants in the room or even multiple remote attendees.

7 Conclusion

Overall, our robot telepresence system shows high potential for improving situationally impaired hybrid meeting scenarios, and our interview study provides unique insights into the re-embodiment of remote participants in this context. This paper provides a novel outlook into a hybrid meeting scenario by having all meeting attendees be interviewed, to allow for a comparison of the co-located and remote perspectives. This study highlights the benefits of having a physical embodiment of the remote participant and suggests ways to improve the conversation flow within the mixed-modality communication system. Future work will evaluate the success of these suggestions and investigate the benefits and implications of incorporating autonomy into a robot proxy system when it is representing a remote meeting attendee that cannot be directly observed.

Acknowledgments

This research was partially supported by the Independent Research Fund Denmark (DFF) (grant ID: 10.46540/3104-00008B) and by the Australian Research Council (Grants: FT250100459 and CE260100108)

References

[1] Lorenza Abbate and Claudio Germak. 2024. Human-Robot-Human: The Natural Dimension of the Telepresence Robotics Design. In *For Nature/With Nature: New Sustainable Design Scenarios*, Claudio Gambardella (Ed.). Springer Nature Switzerland, Cham, 637–656. https://doi.org/10.1007/978-3-031-53122-4_38

[2] Ernest Adams. 2014. *Fundamentals of Game Design* (3rd ed.). New Riders Publishing, USA.

[3] Priyanshu Agarwal, Samer Al Moubayed, Alexander Alspach, Joohyung Kim, Elizabeth J. Carter, Jill Fain Lehman, and Katsu Yamane. 2016. Imitating human movement with teleoperated robotic head. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 630–637. https://doi.org/10.1109/ROMAN.2016.7745184 ISSN: 1944-9437.

[4] Cevat Giray Aksoy, Jose Maria Barrero, Nicholas Bloom, Steven J. Davis, Mathias Dolls, and Pablo Zarate. 2022. Working from Home Around the World. https://doi.org/10.3386/w30446

[5] Samer Al Moubayed, Jonas Beskow, Gabriel Skantze, and Björn Granström. 2012. Furhat: A Back-Projected Human-Like Robot Head for Multiparty Human-Machine Interaction. In *Cognitive Behavioural Systems*, Anna Esposito, Antonietta M. Esposito, Alessandro Vinciarelli, Rüdiger Hoffmann, and Vincent C. Müller (Eds.). Springer, Berlin, Heidelberg, 114–130. https://doi.org/10.1007/978-3-642-34584-5_9

[6] Luis Almeida, Paulo Menezes, and Jorge Dias. 2022. Telepresence Social Robotics towards Co-Presence: A Review. *Applied Sciences* 12, 11 (Jan. 2022), 5557. https://doi.org/10.3390/app12115557 Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

[7] S. Argentieri, P. Danès, and P. Souères. 2015. A survey on sound source localization in robotics: From binaural to array processing methods. *Computer Speech & Language* 34, 1 (Nov. 2015), 87–112. https://doi.org/10.1016/j.csl.2015.03.003

[8] Thibault Asselborn, Wafa Johal, and Pierre Dillenbourg. 2017. Keep on moving! Exploring anthropomorphic effects of motion during idle moments. In *2017 26th IEEE international symposium on robot and human interactive communication (RO-MAN)*. IEEE, 897–902. https://doi.org/10.1109/ROMAN.2017.8172409

[9] Jacob T. Biehl, Daniel Avrahami, and Anthony Dunnigan. 2015. Not Really There: Understanding Embodied Communication Affordances in Team Perception and Participation. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1567–1575. https://doi.org/10.1145/2675133.2675220

[10] Nathan D. Bos, Ayse Buyuktur, Judith S. Olson, Gary M. Olson, and Amy Voida. 2010. Shared identity helps partially distributed teams, but distance still matters. In *Proceedings of the 2010 ACM International Conference on Supporting Group Work (GROUP '10)*. Association for Computing Machinery, New York, NY, USA, 89–96. https://doi.org/10.1145/1880071.1880086

[11] Andriana Boudouraki, Houda Elmimouni, Marta Orduna, Pablo Perez, Ester Gonzalez-Sosa, Pablo Cesar, JesÚS GutiÉRrez, Taffeta Wood, Verónica Ahumada-Newhart, and Joel E Fischer. 2023. Emerging Telepresence Technologies for Hybrid Meetings: an Interactive Workshop. In *Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing* (Minneapolis, MN, USA) *(CSCW '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 547–552. https://doi.org/10.1145/3584931.3611283

[12] Andriana Boudouraki, Joel E. Fischer, Stuart Reeves, and Sean Rintel. 2021. "I can't get round": Recruiting Assistance in Mobile Robotic Telepresence. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3 (Jan. 2021), 248:1–248:21. https://doi.org/10.1145/3432947

[13] Andriana Boudouraki, Joel E. Fischer, Stuart Reeves, and Sean Rintel. 2023. "Being in on the Action" in Mobile Robotic Telepresence: Rethinking Presence in Hybrid Participation. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (Stockholm, Sweden) *(HRI '23)*. Association for Computing Machinery, New York, NY, USA, 63–71. https://doi.org/10.1145/3568162.3576961

[14] Andriana Boudouraki, Stuart Reeves, Joel Fischer, and Sean Rintel. 2023. "There is a bit of grace missing": Understanding non-use of mobile robotic telepresence in a global technology company. In *Proceedings of the First International Symposium on Trustworthy Autonomous Systems (TAS '23)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3597512.3599710

[15] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1191/1478088706qp063oa.

[16] Bill Buxton. 2009. Mediaspace – Meaningspace – Meetingspace. In *Media Space 20 + Years of Mediated Life*, Steve Harrison (Ed.). Springer, London, 217–231. https://doi.org/10.1007/978-1-84882-483-6_13

[17] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 981–992. https://doi.org/10.1145/2858036.2858498

[18] Jose Eurico de Vasconcelos Filho, Kori M. Inkpen, and Mary Czerwinski. 2009. Image, appearance and vanity in the use of media spaces and video conference systems. In *Proceedings of the 2009 ACM International Conference on Supporting Group Work (GROUP '09)*. Association for Computing Machinery, New York, NY, USA, 253–262. https://doi.org/10.1145/1531674.1531712

[19] Ziedune Degutyte and Arlene Astell. 2021. The Role of Eye Gaze in Regulating Turn Taking in Conversations: A Systematized Review of Methods and Findings. *Frontiers in Psychology* 12 (April 2021). https://doi.org/10.3389/fpsyg.2021.616471 Publisher: Frontiers.

[20] Nick Dua, Mattias Fyrenius, Deborah L. Johnson, and Walter H. Moos. 2021. Are in-person scientific conferences dead or alive? *FASEB BioAdvances* 3, 6 (Feb. 2021), 420–427. https://doi.org/10.1096/fba.2020-00139

[21] Maja Dybboe, Johannes Ellemose, Alexander Langagergaard Vastrup, Andriana Boudouraki, Sean Rintel, Marianne Graves Petersen, Jens Emil Sloth Grønbæk, and Clemens Nylandsted Klokmose. 2024. TableBot: Getting a Handle on Hybrid Collaboration by Negotiating Control of a Tabletop Telepresence Robot. In *Proceedings of the 13th Nordic Conference on Human-Computer Interaction (NordiCHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3679318.3685350

[22] Ricky Ellis, Tim Goodacre, Neil Mortensen, Rachel S Oeppen, and Peter A Brennan. 2022. Application of human factors at hybrid meetings: facilitating productivity and inclusivity. *British Journal of Oral and Maxillofacial Surgery* 60, 6 (July 2022), 740–745. https://doi.org/10.1016/j.bjoms.2021.12.055

[23] Guo Freeman and Divine Maloney. 2021. Body, Avatar, and Me: The Presentation and Perception of Self in Social Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3 (Jan. 2021), 239:1–239:27. https://doi.org/10.1145/3432938

[24] Jens Emil Grønbæk, Banu Saatçi, Carla F. Griggio, and Clemens Nylandsted Klokmose. 2021. MirrorBlender: Supporting Hybrid Meetings with a Malleable Video-Conferencing System. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3411764.3445698

[25] Magnus Gudmandsen. 2015. *Using a robot head with a 3D face mask as a communication medium for telepresence*. Ph. D. Dissertation. KTH ROYAL INSTITUTE OF TECHNOLOGY. https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-171402

[26] Jan Gulliksen, Joakim Lilliesköld, and Stefan Stenbom. 2023. The 'New' New Normal—Digitalization and Hybridization of Work and Education Before, during and after the Covid-19 Pandemic. *Interacting with Computers* 35, 5 (Dec. 2023), 691–706. https://doi.org/10.1093/iwc/iwac034

[27] Ernest James Hall. 1963. *The Rejection of Deviates as a Function of Threat*. Ph. D. Dissertation. The University of Texas. https://apps.dtic.mil/sti/trecms/pdf/AD0421742.pdf

[28] BM Zeeshan Hameed, Yiloren Tanidir, Nithesh Naik, Jeremy Yuen-Chun Teoh, Milap Shah, Marcelo Langer Wroclawski, Afrah Budnar Kunjibettu, Daniele Castellani, Sufyan Ibrahim, Rodrigo Donalisio da Silva, Bhavan Rai, J. J. M. C. H. de la Rosette, Rajeev Tp, Vineet Gauhar, and Bhaskar Somani. 2021. Will "Hybrid" Meetings Replace Face-To-Face Meetings Post COVID-19 Era? Perceptions and Views From The Urological Community. *Urology* 156 (Oct. 2021), 52–57. https://doi.org/10.1016/j.urology.2021.02.001

[29] Jörg Hauber, Holger Regenbrecht, Mark Billinghurst, and Andy Cockburn. 2006. Spatiality in videoconferencing: trade-offs between efficiency and social presence. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work (CSCW '06)*. Association for Computing Machinery, New York, NY, USA, 413–422. https://doi.org/10.1145/1180875.1180937

[30] Alexander G. Hauptmann and Alexander I. Rudnicky. 1990. A Comparison of Speech and Typed Input. In *Speech and Natural Language: Proceedings of a Workshop Held at Hidden Valley, Pennsylvania, June 24-27,1990*. https://aclanthology.org/H90-1045/

[31] Damith C. Herath, Nicole Binks, and Janie Busby Grant. 2020. To Embody or Not: A Cross Human-Robot and Human-Computer Interaction (HRI/HCI) Study on the Efficacy of Physical Embodiment. In *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. 848–853. https://doi.org/10.1109/ICARCV50220.2020.9305520

[32] E. Jeffrey Hill and Kaylene J. Fellows. 2014. Telecommuting. In *Encyclopedia of Quality of Life and Well-Being Research*, Alex C. Michalos (Ed.). Springer Netherlands, Dordrecht, 6599–6600. https://doi.org/10.1007/978-94-007-0753-5_2985

[33] Jeremy Hyrkas, Andrew D Wilson, John Tang, Hannes Gamper, Hong Sodoma, Lev Tankelevitch, Kori Inkpen, Shreya Chappidi, and Brennan Jones. 2023. Spatialized Audio and Hybrid Video Conferencing: Where Should Voices be Positioned for People in the Room and Remote Headset Users?. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3544548.3581085

[34] Daisy Ingle, Nadine Marcus, and Wafa Johal. 2021. The Valley of non-Distraction: Effect of Robot's Human-likeness on Perception Load. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 99–103. https://doi.org/10.1145/3434074.3447137

[35] Baptiste Isabet, Maribel Pino, Manon Lewis, Samuel Benveniste, and Anne-Sophie Rigaud. 2021. Social Telepresence Robots: A Narrative Review of Experiments Involving Older Adults before and during the COVID-19 Pandemic. *International Journal of Environmental Research and Public Health* 18, 7 (Jan. 2021), 3597. https://doi.org/10.3390/ijerph18073597 Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.

[36] Robert J. K. Jacob. 2006. What is the next generation of human-computer interaction?. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems*. ACM, Montréal Québec Canada, 1707–1710. https://doi.org/10.1145/1125451.1125768

[37] Wafa Johal, Gaëlle Calvary, and Sylvie Pesty. 2015. Non-verbal signals in HRI: interference in human perception. In *International conference on social robotics*. Springer, 275–284. https://doi.org/10.1007/978-3-319-25554-5_28

[38] Dima Kagan, Galit Fuhrmann Alpert, and Michael Fire. 2024. Zooming Into Video Conferencing Privacy. *IEEE Transactions on Computational Social Systems* 11, 1 (Feb. 2024), 933–944. https://doi.org/10.1109/TCSS.2022.3231987 Conference Name: IEEE Transactions on Computational Social Systems.

[39] Demetrios Karis, Daniel Wildman, and Amir Mané. 2016. Improving Remote Collaboration With Video Conferencing and Video Portals. *Human–Computer Interaction* 31, 1 (Jan. 2016), 1–58. https://doi.org/10.1080/07370024.2014.921506 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07370024.2014.921506.

[40] Annica Kristoffersson, Silvia Coradeschi, and Amy Loutfi. 2013. A Review of Mobile Robotic Telepresence. *Advances in Human-Computer Interaction* 2013 (April 2013), e902316. https://doi.org/10.1155/2013/902316 Publisher: Hindawi.

[41] Hideaki Kuzuoka, Jun'ichi Kosaka, Keiichi Yamazaki, Yasuko Suga, Akiko Yamazaki, Paul Luff, and Christian Heath. 2004. Mediating dual ecologies. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work (CSCW '04)*. Association for Computing Machinery, New York, NY, USA, 477–486. https://doi.org/10.1145/1031607.1031686

[42] Hideaki Kuzuoka, Shinya Oyama, Keiichi Yamazaki, Kenji Suzuki, and Mamoru Mitsuishi. 2000. GestureMan: a mobile robot that embodies a remote instructor's actions. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work (CSCW '00)*. Association for Computing Machinery, New York, NY, USA, 155–162. https://doi.org/10.1145/358916.358986

[43] Hideaki Kuzuoka, Keiichi Yamazaki, Akiko Yamazaki, Jun'ichi Kosaka, Yasuko Suga, and Christian Heath. 2004. Dual ecologies of robot as communication media: thoughts on coordinating orientations and projectability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vienna Austria, 183–190. https://doi.org/10.1145/985692.985716

[44] Jamy Li. 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies* 77 (May 2015), 23–37. https://doi.org/10.1016/j.ijhcs.2015.01.001

[45] Xingyu Bruce Liu, Jiahao Nick Li, David Kim, Xiang 'Anthony' Chen, and Ruofei Du. 2024. Human I/O: Towards a Unified Approach to Detecting Situational Impairments. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3613904.3642065

[46] Yu Liu, Gelareh Mohammadi, Yang Song, and Wafa Johal. 2021. Speech-based Gesture Generation for Robots and Embodied Agents: A Scoping Review. In *Proceedings of the 9th International Conference on Human-Agent Interaction* (Virtual Event, Japan) *(HAI '21)*. Association for Computing Machinery, New York, NY, USA, 31–38. https://doi.org/10.1145/3472307.3484167

[47] Georgios Marentakis and Arman Balic. 2024. Situational impairment due to walking with conversational versus graphical interfaces. In *Proceedings of the 19th International Audio Mostly Conference: Explorations in Sonic Cultures (AM '24)*. Association for Computing Machinery, New York, NY, USA, 77–85. https://doi.org/10.1145/3678299.3678307

[48] Marianne Schmid Mast. 2002. Dominance as Expressed and Inferred Through Speaking Time. *Human Communication Research* 28, 3 (2002), 420–450. https://doi.org/10.1111/j.1468-2958.2002.tb00814.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2958.2002.tb00814.x.

[49] Nicoleta Meslec and Petru Lucian Curşeu. 2013. Too Close or Too Far Hurts: Cognitive Distance and Group Cognitive Synergy. *Small Group Research* 44, 5 (Oct. 2013), 471–497. https://doi.org/10.1177/1046496413491988 Publisher: SAGE Publications Inc.

[50] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine* 19, 2 (June 2012), 98–100. https://doi.org/10.1109/MRA.2012.2192811 Conference Name: IEEE Robotics & Automation Magazine.

[51] Alexander Ng, Stephen A. Brewster, and John Williamson. 2013. The Impact of Encumbrance on Mobile Interactions. In *Human-Computer Interaction – INTERACT 2013*, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi, Gerhard Weikum, Paula Kotzé, Gary Marsden, Gitte Lindgaard, Janet Wesson, and Marco Winckler (Eds.). Vol. 8119. Springer Berlin Heidelberg, Berlin,

Heidelberg, 92–109. https://doi.org/10.1007/978-3-642-40477-1_6 Series Title: Lecture Notes in Computer Science.

[52] Jack M. Nilles. 1994. *Making Telecommuting Happen: A Guide for Telemanagers and Telecommuters.* Van Nostrand Reinhold. Google-Books-ID: MkgdAQAAMAAJ.

[53] Kenneth Okereafor, Philip Manny, and Mustafa Sabri. 2020. UNDERSTANDING CYBERSECURITY CHALLENGES OF TELECOMMUTING AND VIDEO CONFERENCING APPLICATIONS IN THE COVID-19 PANDEMIC. *SSRN Electronic Journal* 8 (June 2020), 13–23.

[54] Dennis Ong, Tim Moors, and Vijay Sivaraman. 2014. Comparison of the energy, carbon and time costs of videoconferencing and in-person meetings. *Computer Communications* 50 (Sept. 2014), 86–94. https://doi.org/10.1016/j.comcom.2014.02.009

[55] Yuya Onishi, Kazuaki Tanaka, and Hideyuki Nakanishi. 2016. Embodiment of Video-mediated Communication Enhances Social Telepresence. In *Proceedings of the Fourth International Conference on Human Agent Interaction (HAI '16).* Association for Computing Machinery, New York, NY, USA, 171–178. https://doi.org/10.1145/2974804.2974826

[56] Rieks op den Akker, Dennis Hofs, Hendri Hondorp, Harm op den Akker, Job Zwiers, and Anton Nijholt. 2009. Supporting Engagement and Floor Control in Hybrid Meetings. In *Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions*, Anna Esposito and Robert Vích (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 276–290.

[57] Jiahe Pan, Jonathan Eden, Denny Oetomo, and Wafa Johal. 2024. Effects of shared control on cognitive load and trust in teleoperated trajectory tracking. *IEEE Robotics and Automation Letters* 9, 6 (2024), 5863–5870. https://doi.org/10.1109/LRA.2024.3396111

[58] Sharon K Parker, Timothy Ballard, Mark Billinghurst, Catherine Collins, Maureen Dollard, Mark A Griffin, Wafa Johal, Karina Jorritsma, Marek Kowalkiewicz, Eva Kyndt, et al. 2025. Quality work in the future: New directions via a co-evolving sociotechnical systems perspective. *Australian Journal of Management* (2025), 03128962251331813. https://doi.org/10.1177/03128962251331813

[59] Hannah R.M. Pelikan, Fanjun Bu, and Wendy Ju. 2025. The People Behind the Robots: How Wizards Wrangle Robots in Public Deployments. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems.* ACM, Yokohama Japan, 1–21. https://doi.org/10.1145/3706598.3713237

[60] Sarah Jane Pell. 2024. Augmented Astronaut Survival: Updating the 'How to Survive on the Moon' scenario workshop in preparation for an Artemis Edition.. In *Proceedings of the Augmented Humans International Conference 2024 (AHs '24).* Association for Computing Machinery, New York, NY, USA, 331–341. https://doi.org/10.1145/3652920.3654916

[61] Irene Rae, Leila Takayama, and Bilge Mutlu. 2013. The influence of height in robot-mediated communication. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction (HRI '13).* IEEE Press, Tokyo, Japan, 1–8.

[62] Stephen Richmond, Laurence J. Howe, Sarah Lewis, Evie Stergiakouli, and Alexei Zhurov. 2018. Facial Genetics: A Brief Overview. *Frontiers in Genetics* 9 (Oct. 2018). https://doi.org/10.3389/fgene.2018.00462 Publisher: Frontiers.

[63] Laurel D. Riek. 2012. Wizard of Oz studies in HRI: a systematic review and new reporting guidelines. *J. Hum.-Robot Interact.* 1, 1 (July 2012), 119–136. https://doi.org/10.5898/JHRI.1.1.Riek

[64] Loïc Rosset, Hamed Alavi, Sailin Zhong, and Denis Lalanne. 2021. Already It Was Hard to Tell Who's Speaking Over There, and Now Face Masks! Can Binaural Audio Help Remote Participation in Hybrid Meetings?. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems.* ACM, Yokohama Japan, 1–7. https://doi.org/10.1145/3411763.3451802

[65] Sherry Ruan, Jacob O. Wobbrock, Kenny Liou, Andrew Ng, and James A. Landay. 2018. Comparing Speech and Keyboard Text Entry for Short Messages in Two Languages on Touchscreen Phones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4 (Jan. 2018), 159:1–159:23. https://doi.org/10.1145/3161187

[66] Banu Saatçi, Kaya Akyüz, Sean Rintel, and Clemens Nylandsted Klokmose. 2020. (Re)Configuring Hybrid Meetings: Moving from User-Centered Design to Meeting-Centered Design. *Computer Supported Cooperative Work (CSCW)* 29, 6 (Dec. 2020), 769–794. https://doi.org/10.1007/s10606-020-09385-x

[67] Banu Saatçi, Roman Rädle, Sean Rintel, Kenton O'Hara, and Clemens Nylandsted Klokmose. 2019. Hybrid Meetings in the Modern Workplace: Stories of Success and Failure. In *Collaboration Technologies and Social Computing*, Hideyuki Nakanishi, Hironori Egi, Irene-Angelica Chounta, Hideyuki Takada, Satoshi Ichimura, and Ulrich Hoppe (Eds.). Springer International Publishing, Cham, 45–61. https://doi.org/10.1007/978-3-030-28011-6_4

[68] Zhanna Sarsenbayeva, Jorge Goncalves, Juan García, Simon Klakegg, Sirkka Rissanen, Hannu Rintamäki, Jari Hannu, and Vassilis Kostakos. 2016. Situational impairments to mobile interaction in cold environments. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16).* Association for Computing Machinery, New York, NY, USA, 85–96. https://doi.org/10.1145/2971648.2971734

[69] Zhanna Sarsenbayeva, Niels Van Berkel, Chu Luo, Vassilis Kostakos, and Jorge Goncalves. 2017. Challenges of situational impairments during interaction with mobile devices. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction.* ACM, Brisbane Queensland Australia, 477–481. https://doi.org/10.1145/3152771.3156161

[70] Martin Schuessler, Luca Hormann, Raimund Dachselt, Andrew Blake, and Carsten Rother. 2024. Gazing Heads: Investigating Gaze Perception in Video-Mediated Communication. *ACM Transactions on Computer-Human Interaction* (June 2024), 3660343. https://doi.org/10.1145/3660343

[71] Andrew Sears, Min Lin, Julie Jacko, and Yan Xiao. 2003. When Computers Fade … Pervasive Computing and Situationally-Induced Impairments and Disabilities | Request PDF, Vol. 2. 1298–1302. https://www.researchgate.net/publication/255625951_When_Computers_Fade_Pervasive_Computing_and_Situationally-Induced_Impairments_and_Disabilities

[72] Samarth Singhal, Carman Neustaedter, Thecla Schiphorst, Anthony Tang, Abhisekh Patra, and Rui Pan. 2016. You are Being Watched: Bystanders' Perspective on the Use of Camera Devices in Public Spaces. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems.* ACM, San Jose California USA, 3197–3203. https://doi.org/10.1145/2851581.2892522

[73] Gabriel Skantze. 2021. Turn-taking in Conversational Systems and Human-Robot Interaction: A Review. *Computer Speech & Language* 67 (May 2021), 101178. https://doi.org/10.1016/j.csl.2020.101178

[74] Becky Spittle, Payod Panda, Lev Tankelevitch, Kori Inkpen, John Tang, Sasa Junuzovic, Qianqian Qi, Pat Sweeney, Andrew D Wilson, William A.S. Buxton, Abigail Sellen, and Sean Rintel. 2024. Comparing the Agency of Hybrid Meeting Remote Users in 2D and 3D Interfaces of the Hybridge System. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems.* ACM, Honolulu HI USA, 1–12. https://doi.org/10.1145/3613905.3651103

[75] Katherine M. Tsui, Munjal Desai, Holly A. Yanco, and Chris Uhlik. 2011. Exploring use cases for telepresence robots. In *Proceedings of the 6th international conference on Human-robot interaction (HRI '11).* Association for Computing Machinery, New York, NY, USA, 11–18. https://doi.org/10.1145/1957656.1957664

[76] Ying-Chao Tung, Mayank Goel, Isaac Zinda, and Jacob O. Wobbrock. 2018. RainCheck: Overcoming Capacitive Interference Caused by Rainwater on Smartphones. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI '18).* Association for Computing Machinery, New York, NY, USA, 464–471. https://doi.org/10.1145/3242969.3243028

[77] Sebahat Sevgi Uygur and Yasemin Kahyaoğlu Erdoğmuş. 2025. (In)visible students: Investigating why students turn off their cameras during live lessons. *International Journal of Educational Research* 132 (Jan. 2025), 102638. https://doi.org/10.1016/j.ijer.2025.102638

[78] Jacob O. Wobbrock. 2019. Situationally aware mobile devices for overcoming situational impairments. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '19).* Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3319499.3330292

[79] Jacob O. Wobbrock. 2019. Situationally-Induced Impairments and Disabilities. In *Web Accessibility: A Foundation for Research*, Yeliz Yesilada and Simon Harper (Eds.). Springer, London, 59–92. https://doi.org/10.1007/978-1-4471-7440-0_5

[80] Emily Wong, Adélaïde Genay, Jens Emil Sloth Grønbæk, and Eduardo Velloso. 2025. Spatial Heterogeneity in Distributed Mixed Reality Collaboration. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems.* ACM, Yokohama Japan, 1–19. https://doi.org/10.1145/3706598.3714033

[81] Jakub A. Zlotowski, Hidenobu Sumioka, Shuichi Nishio, Dylan F. Glas, Christoph Bartneck, and Hiroshi Ishiguro. 2015. Persistence of the uncanny valley: the influence of repeated interactions and a robot's attitude on its perception. *Frontiers in Psychology* 6 (2015). https://doi.org/10.3389/fpsyg.2015.00883