ProxyBridge: Improving Remote Engagement in Small-Group Hybrid Meetings

Anisa Callis, Steve Chen, Sidney Grabosky, Stephanie Patterson

ABSTRACT

Hybrid meetings often exacerbate social asymmetries, leaving remote participants feeling overlooked and disengaged. We present ProxyBridge, a hybrid conferencing system designed to enhance remote presence and interaction through an embodied tabletop device, mobile interface, and interactive desktop platform. ProxyBridge supports whisper interactions, gaze cues, emotion signaling, and speaker indication via light and haptic feedback. Grounded in literature on social presence, accessibility, and proxemics, our multimethod study included interviews, surveys, focus groups, and Wizard-of-Oz evaluations to identify user needs and assess usability. The results indicate that ProxyBridge promotes impromptu interaction and improves remote visibility, but it also highlights conflicts related to onboarding, privacy, and customization. This study offers empirical insights and a new hybrid interface design for promoting fair participation in remote meetings.

INTRODUCTION

Hybrid meetings—those combining in-person and remote participants—have become widespread across professional, academic, and social settings. A systematic review by Neumayr et al. (2021) notes the surge in hybrid collaboration following the shift to remote work during the COVID-19 pandemic [19], a trend also reflected in academia during pandemic-related disruptions [18].

While hybrid formats expand flexibility and accessibility—especially for those with impairments or across regions—they also pose challenges. Studies report communication breakdowns, disruptions, and reduced engagement compared to in-person meetings [21, 33]. Yet, hybrid systems remain vital for broadening access and participation [18, 21].

Remote attendees often report diminished presence, fewer spontaneous interactions, and difficulty keeping pace. These asymmetries cause exclusion, cognitive strain, and reduced participation. Hybrid meetings often lack immersion, contributing to disconnection [32]. Research shows remote users feel marginalized and struggle to contribute equally [15, 35], with impacts on inclusion, agency, and conversational flow [20].

This project introduces ProxyBridge, a system using physical proxies to enhance remote presence in small-group hybrid meetings. Through a tabletop device, mobile app, and desktop interface, it enables haptic and visual cues, gaze direction, emotion display, and whisper-style interactions.

Building on work in social presence, proxemics, and accessibility, this research uses interviews, surveys, focus groups, and Wizard-of-Oz testing to identify remote user needs and assess ProxyBridge's effectiveness.

This work makes the following contributions:

- We contribute a novel hybrid meeting system, ProxyBridge, which leverages physical embodiment, emotion signaling, and whisper-style communication to support visibility and interpersonal expressivity for remote participants.
- We present a multi-stage user-centered design process, incorporating interviews, surveys, a focus group, and Wizard-of-Oz usability testing to surface pain points and design opportunities in hybrid meeting contexts.
- We provide empirical evidence that embodied proxies can mitigate social asymmetries in hybrid meetings, enhancing spatial presence, engagement, and user agency.
- We offer design implications for future hybrid meeting technologies that emphasize accessibility, user control, and equitable interaction across physical and virtual divides.

LITERATURE REVIEW

Historical Context: Foundations of Hybrid Communication

Hybrid meetings—combining in-person and remote participants—have been explored since early video-mediated systems like DOLPHIN [27] and VideoWindow [7], which used shared displays and ambient cues to support remote presence. Though limited by technology, they established goals around informal communication and co-presence.

Media richness theory [4] shaped later systems such as SideBar [6], designed to foster social engagement via ambient interaction. Follow-up work advanced presence through simulated gaze [34], turn-taking cues [9], and immersive technologies like VR [17] and holographic conferencing [14].

Proxemics and Spatial Mappings

There are numerous ways that spatial and proxemic relationships interact with everyday social cues, and many of them are lost in virtual and hybrid meetings. *Proxemics* is a field concerned with how space and distance impacts human communication and culture [8]. During routine in-person interactions, people naturally utilize proximity and spatial arrangement to facilitate communication or signify social and cultural contexts and cues.

During co-located meetings, seating arrangements shape interaction dynamics and reflect culture and interpersonal bonds. Research on group settings shows that physical proximity influences communication flow and relationship formation. People tend to sit next to or across from teammates or friends to ease interaction [5]. People will often choose to sit

next to specific people, or reserve a seat for another attendee who has yet to arrive [11].

Proximity affords many non-disruptive peer-to-peer communication methods, even during the middle of a meeting or presentation. Common interactions include leaning in, closing the distance even further, and whispering to others at a low volume for quick comments, questions, and sidebar conversations [26].

Subtle nonverbal clues, such as tapping someone to get their attention or speaking quietly to people nearby, are essential for controlling the conversation's flow [23, 11] and preserving spatial links in face-to-face interactions. Participants can more easily negotiate group dynamics and prevent needlessly alerting those who are not directly involved by using these behaviors.

However, studies show that existing teleconferencing platforms lack many of these organic, conversation-enhancing connections, which causes distant participants to lose the richness of their nonverbal communication and their geographical awareness. The limitations of video and audio channels in remote meetings restrict the full spectrum of nonverbal cues, making it more challenging to replicate the fluidity and subtlety of in-person group communication [13, 25].

These emergent social dynamics based on proximity are encapsulated in Halls' proxemics theory through the concepts of public, social, personal, and intimate space [8]. In a larger conference room, some people may be within public space, and thus too far away for social interaction. In a hybrid meeting, *all* remote attendees are arguably within public space by default. In smaller co-situated settings, all attendees are within social space. When whispering discreetly to a neighbor, the boundary is crossed into personal space, while touching somebody to get their attention enters intimate space. Moving to closer circles of distance can result in more personal, less formal social interactions, as well as increased privacy [8, 26].

Social Cues and Signals

Gaze direction and body orientation play a key role in in-person communication. People naturally look toward speakers or visual aids, using gaze as a social signal to focus attention and integrate information [31]. Gaze also helps identify speakers and regulate conversation, making it easy to assess attentiveness or distraction [22].

Speakers may receive passive signals that they need to repeat themselves, or that they are droning on and should accelerate their presentation. Alternatively, they may choose to elaborate more if they see rapt attention [11]. Attention and spatial relationships are further entwined, and they broadly impact the flow of conversations. It is far easier for somebody situated proximally nearby to the active speaker to interject than it is for somebody who is on the other side of the room [11]. Peripheral vision can catch signifiers such as opening the mouth, or a body posture indicating an intent to speak.

With current technology in hybrid meetings, remote participants, in effect, are always the "other side of the room" and "outside of peripheral vision," so they must typically attempt to speak or respond from a "cold start" without having the benefit of peripheral attention. [11, 26]. The need to attract attention before speaking, and the risk of multiple participants speaking over each other at once, is heightened for remote attendees, causing a disjointed conversational flow. This awkward social dynamic results in remote participants frequently opting to vocalize or contribute less than they otherwise would in person [26].

Interestingly, while media richness theory [4] suggests that communication benefits from having multiple channels of rich communication available, the symmetry of this arrangement is also important. Simply adding more channels of rich communication from one side to another can actually exacerbate feelings of exclusion and isolation [23], heightening the social presence asymmetry. This suggests it's not sufficient to make remote meeting attendees aware of what's happening in the co-located space; a socially equitable hybrid teleconferencing system must also convey rich media channels and social cues from the remote attendee back to the in-person attendees.

Hybrid Meetings as an Accessibility Tool

Hybrid meetings often act as an accessibility tool by allowing individuals to participate in meetings despite physical barriers or disabilities. Tools like Zoom and Microsoft Teams, which have become ubiquitous since the COVID-19 pandemic, offer a way to participate in a meeting without physically being there, which is particularly helpful for individuals with motor disabilities who may deal with barriers in getting to various physical locations because of inaccessible routes or locations [2].

Hybrid meetings also offer opportunities for accessibility for individuals who are Deaf or Hard of Hearing (DHH). Online meeting technologies like Zoom typically have automatic captions that remote participants can enable, making spoken language more accessible to a DHH meeting attendee. McDonnell et al. conducted 15 semi-structured interviews with DHH individuals exploring their perceptions of current captioning features in videoconferencing software, finding that users wanted features allowing them to give and receive nonverbal feedback related to the captions. Additionally, "five people independently suggested displaying captions next to each speaker's online video feed" [16]. Another study conducted by Seita et al., in which mixed-ability groups designed small-group captioning features, revealed preferences for features allowing meeting attendees to correct automatic captions in real time, and ask the speaker to slow down or speak up [24].

In addition to these benefits, online meetings also introduce new challenges related to social cues and inclusion, especially for neurodivergent individuals. In a recent study featuring interviews with 21 disabled professionals, Alharbi et al. found that while hybrid meetings serve an important accessibility function, they can also introduce access conflicts or tensions that can impact the experience of some participants [2]. In the case of hybrid or remote meetings, teleconferencing tools

have helpful accessibility tools, but they also introduce social tensions that especially impact individuals who are neurodivergent; individuals with autism may struggle more with understanding social cues and focusing for long periods of time, which is exacerbated by joining a meeting remotely [36]. Therefore, it is important that ProxyBridge's design decisions works to make teleconferencing more accessible to neurodivergent individuals, while also maintaining and enhancing the existing accessibility benefits.

Commercial Platforms and Their Limitations

Mainstream hybrid platforms like Zoom, Teams, Webex, and Meet have added features like live captions and reactions [3, 30], yet accessibility and social presence remain limited. Speech-to-text has improved, but spatial feedback, emotional cues, and turn-taking tools are still lacking [3]. Studies highlight persistent issues with rapport, turn-taking, and subtle cues in hybrid contexts [10].

Emerging tools like BlueJeans, Gather, and SpatialChat use avatars and spatial layouts to enhance engagement but often prioritize novelty over accessibility. These formats may improve representation but still fall short on equitable participation [12, 28], continuing to marginalize remote users as static video tiles [28, 29].

IDENTIFYING USER NEEDS AND CONTEXTS OF USE

User Interviews: Findings and Design Implications

We conducted seven semi-structured Zoom interviews with individuals active in hybrid academic, professional, or organizational settings. Participants, recruited via word-of-mouth, discussed role differences and unmet needs around presence and inclusion. Sessions (30–45 min) were transcribed and thematically analyzed. Common themes—feeling "out of sync," difficulty asserting turns, and remote invisibility—informed early design goals such as ambient cues, lightweight signaling, and equitable turn-taking.

Engagement and Focus

Many participants (AC-P1, SC-P2, Sid-P1, Sid-P2) reported that remote participants often feel less accountable for sustained engagement due to their limited visual presence in hybrid settings. The absence of embodied cues, such as posture or facial expressions, made it difficult for co-located attendees to assess attentiveness. This lack of mutual visibility contributes to multitasking and disengagement among remote users, especially when facilitators are not actively managing participation.

Meanwhile, there is additional tension with privacy and social norms; several participants (Sid-P1, Sid-P2, Sid-P3) expressed disdain for the practice of always-on video streams, but without them, even more social cues are lost. Sid-P2 had extreme dislike for compulsory video camera use, Sid-P1 said that it feels "unnatural" and preferred it off, while Sid-P3 remarked that many engineers *"just never turn on their cameras."*

Design Implication: The ProxyBridge prototype uses haptic and ambient light cues to indicate remote users' attentional states in order to solve this problem. These cues seek to re-establish subtle social accountability and awareness, while giving users alternatives to video streams.

Communication and Social Asymmetry

The capacity of distant individuals to have sincere interactions varied significantly. Multiple participants (AC-P1, Sid-P2, Sid-P3) felt excluded from frequent un-moderated side discussions between co-located participants, which are inaccessible to those attending remotely. Sid-P2 who had shifted from a mostly in-person to a mostly-remote role reflected on how useful whispering to people in meetings was in person, while a remote/hybrid setup doesn't facilitate this interaction. Sid-P1 and Sid-P3, both in management roles, believed that sidebar conversations in hybrid contexts were harmful to team cohesion and communication efficacy as they excluded remote attendees. Sid-P1 did not witness exclusion, as he had prohibited sidebar conversations to avoid this problem. Sid-P3 similarly often enforced a "one conversation at a time" rule in meetings he presided over, noting "*The remote people [...] they seem like they're left out of this dynamic.*", and adding that he personally found it *"annoying"* to be left out of side-talk when he is remote.

AC-P1 in addition to pointing out challenges with speaking turn detection and backchannel interactions. The remote participant's confidence in participating in discussions was further hampered by the absence of real-time social cues.

Design Implication: As a result, ProxyBridge has features that support private communication, like whisper-style channels and subliminal emotional feedback indicators. These features aim to support social signaling without interrupting the primary meeting flow, and without excluding remote participants, thereby improving conversational equity.

Meeting Facilitation and Participation Structure

All but one participant emphasized the importance of strong facilitation in hybrid meetings. Meetings in which facilitators explicitly managed turn-taking and enforced agenda-based structures were perceived as more inclusive. However, excessive mediation was described as potentially slowing conversational dynamics or feeling overly hierarchical. Sid-P3 alludes to needing to mute chronic interrupters to keep meetings flowing. Many participants speak of the difficulty of knowing when to talk after questions are posed; Sid-P1 speaks of moments of *"dead silence"* as attendees try to navigate this without natural cues, while Sid-P2 mentions people being constantly talked over.

Design Implication: To reduce reliance on verbal facilitation, the system integrates LED-based signaling to indicate a remote participant's desire to speak. This visual notification is intended to assist both facilitators and co-located participants in recognizing remote contributions within the natural flow of the meeting.

Environmental and Interface Factors

The ability to perceive the spatial dynamics of a room was associated with improved situational awareness. AC-P1 expressed that seeing the full meeting room view—rather than individual video tiles—provided useful context regarding group engagement and mood. Text-based backchanneling via

Table 1: Interview Participant Demographics and Roles

#	Title	Manager	Age	Gender	Industry
Sid-P1	Senior Staff Engineer	Technical Lead	35–44	M	Tech
Sid-P2	Senior Software Engineer	No	25–34	M	Tech
Sid-P3	Senior Manager Digital ASIC Development	Yes, Director	55+	M	Tech
SC-P1	Student	N/A	18–24	F	Creative
SC-P2	Sales Director	Yes	55+	F	Manufacturing
SC-P3	Clinical Professor in Foundations	No	25–34	M	Education
AC-P1	Equal Employment Opportunity Director	Yes	55+	F	Government

chat or messaging platforms was also described as a valuable secondary layer of interaction, especially for asynchronous contributions.

SC-P2 and Sid-P3 stated that video streams create difficulty in interpreting one's social signals; they are more likely to misread attendees' state through video feed than when in person. Potential technological difficulties add more uncertainty to this confusing situation. Sid-P3 further added that larger meeting rooms often present difficulty for remote attendees, since microphones often struggle to pick up audio from co-located attendees in far-off corners of the room.

Design Implication: The physical embodiment of the Proxy-Bridge device was developed to provide an embodied, spatial representation of remote attendees. The gadget, which is positioned inside the conference area, supports awareness for both co-located and remote users by communicating speaker concentration and attention states using position-based, haptic, and ambient light signals. Even in far-off corners, sounds will be recorded thanks to the placement of several devices throughout the space.

Survey

In order to gather a wider range of user perspectives, we also conducted an online survey using Google Forms. The survey had 11 total responses; nine of the survey participants were between the ages of 18 and 24, two were between 25 and 34, and one was above 55. Most of the participants were either employed, students, or both.

The survey results reinforced many of the findings from our interviews. Although all 11 respondents rated themselves as either Familiar or Very Familiar with teleconferencing tools, 8 reported experiencing exclusion or disengagement while being an online participant in a hybrid teleconferencing meeting. This suggests that the disconnect in hybrid meetings is not a result of users being unsure how to engage using teleconferencing software, but rather a limitation of the teleconferencing software itself.

When asked about missing teleconferencing features that would make it easier to collaborate between remote and in-person meeting participants, survey respondents expressed a desire for features that simulate in-person interactions, often noting the ways that current teleconferencing tools are insufficient. Below are some key open-ended survey responses that show users' desired features in teleconferencing tools:

- S1: "Allow for side conversations that don't interrupt the meeting."
- S3: "During hybrid meetings, in-person attendees are often looking at each other instead of at their screens, so the hand-raising feature is often not noticed right away when used by remote attendees."
- S7: "Most teleconferencing software is fairly feature-rich, but accessing those features is often non-obvious, especially to those who are less comfortable with using technology."
- S8: "I think having more subtle communication would help, such as if the hand raise feature also counted who raised their hand first, second, etc. Including subtitles can help understand online participants and being able to react to what people are saying through emotes would also be more engaging."

The need for improved interaction between remote and in-person participants in a hybrid meeting is highlighted by these remarks from survey respondents. This can be achieved by including features that mimic in-person interactions, such as side conversations, emotional reactions, and improved hand-raising. The comment from S7 was particularly interesting, as this participant noted the importance of teleconferencing features being easy and "obvious" to access, especially for meeting attendees who are less tech-savvy.

RESULTS

Findings

Results from surveys and semi-structured interviews showed that remote individuals consistently exhibited patterns of marginalization. The structural restrictions in the hybrid system design were the main cause of engagement barriers rather than a lack of experience with teleconferencing tools.

Theme 1: Visual Presence Without Agency

Despite being visible on screen, participants lacked embodied cues—like gaze, gesture, or posture—to signal attention or intent. This reduced their agency in discussions. One participant described feeling like "a floating head on the screen" as in-person conversation continued. In the survey, 8 of 11 respondents—though "Familiar" or "Very Familiar" with teleconferencing—reported feeling excluded or overlooked.

Theme 2: Richness Versus Clarity
Chat, reactions, and hand-raising aimed to support remote users, but many found them ambiguous in practice. One respondent noted hand-raises "often go unnoticed" since in-person participants focus on each other, not the screen. Others mentioned that chat inputs and reactions lacked consistent interpretation, delaying responses and causing misalignment.

Theme 3: Structure Versus Spontaneity
Participants valued structured support (e.g., round-robin speaking), but also noted that rigid formats can suppress organic interaction—especially in small-group settings. Several respondents emphasized the importance of "side conversations" or private backchannels for engaging without interrupting the main discussion.

Design Implications
Hybrid meeting systems should prioritize legible, expressive remote presence over mimicking in-person interaction. Ambient cues (e.g., lights, haptics), intuitive design, and lightweight side channels (e.g., whispers, nudges) enable spontaneous yet unobtrusive participation. These principles shaped ProxyBridge, which uses subtle embodiment and ambient signaling to support equitable, natural remote engagement.

INITIAL PROTOTYPE

ProxyBridge Device Interface: In-Person Perspective
The need to rethink how remote participants are represented, engaged, and perceived in hybrid meetings led to the creation of ProxyBridge. Early user studies highlighted recurring issues of imbalance, disengagement, and exclusion—rooted not just in technology but in the lack of physical presence and limited communication affordances. A Figma-based prototype was developed as a speculative interface to explore these challenges and propose solutions.

Serving as both a design artifact and ideation tool, the prototype illustrated how remote users might claim presence via a proxy device and interact through ambient, non-verbal, and expressive features. Key components included:

- Whisper and Nudge Buttons for initiating discreet one-on-one interactions without disrupting the main discussion.
- Emotional Reaction Emojis, allowing remote users to express affective states.
- Speaking Queue, to ensure turn-taking fairness and minimize interruptions.
- Shared View, a collaborative interface for indicating attention or reference points on shared content.

To support spatial orientation, the prototype featured a seat selection flow enabling remote users to claim a physical proxy. Designed to promote equity and presence, early testing (Figure 1) uncovered usability issues such as low contrast, small icons, and unclear claim status. Participants raised concerns about professionalism and privacy, noting that public emotional reactions and visible whisper cues—though conceptually helpful—could be inappropriate in formal settings.

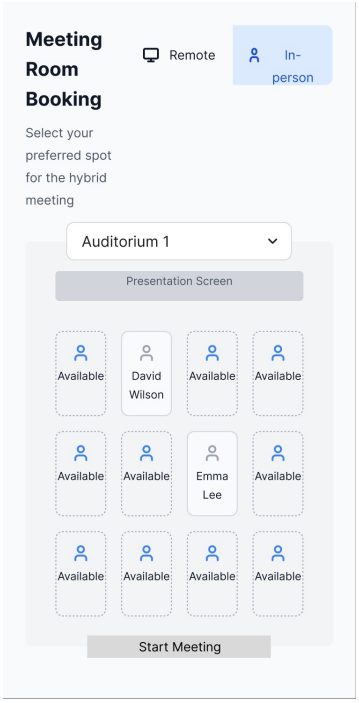


Figure 1: Seat claiming interface in the Figma prototype. Remote users can select and claim a physical seat in the meeting space, supporting spatial presence and orientation.

This informed later iterations, which explored haptic or auditory whisper cues and more context-sensitive avatars. The final prototype screen illustrated the live meeting interface, integrating expressive modalities to support continuous presence and subtle interaction for remote users.

The initial Figma prototype acted as a boundary object linking HCI, accessibility, and social computing. It translated abstract user needs into tangible concepts and guided ProxyBridge's development by envisioning ambient, inclusive interaction beyond traditional teleconferencing.

Desktop Interface: Remote Perspective
Alongside the in-person proxy device, we developed a Figma mock-up of the remote user interface (Figure 3). This desktop interface mirrors familiar teleconferencing platforms like Zoom to support first-time usability through external consistency—featuring a bottom toolbar for mic/camera control and modular participant tiles. Distinctive features that set ProxyBridge apart from conventional platforms include:

- Whisper and Nudge buttons by each participant to match the ProxyBridge device's functionality. The speech bubble shows what a Whisper message might look like to a remote participant.
- Emotional Reactions, displayed in the corner of each participant's video feed. The system recognizes each user's emotional state based on their facial expression and asks the user before sending emotional reactions. The user also

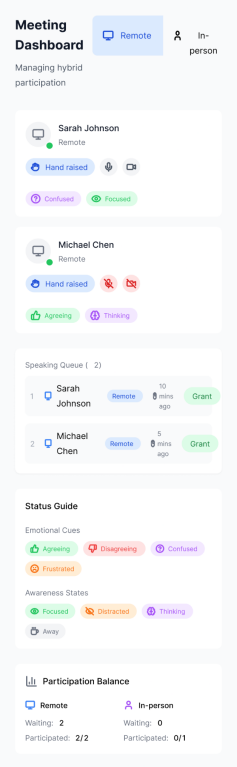


Figure 2: Active meeting interface. Remote participants can use whisper, nudge, and emotional reaction features to communicate subtle intent and expressivity without interrupting the main conversation.

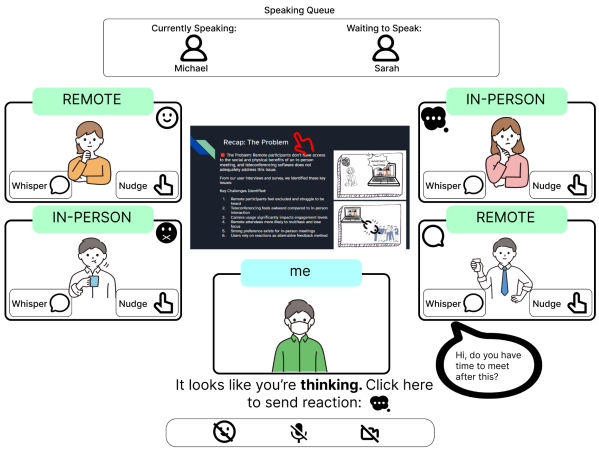


Figure 3: Desktop interface for the remote participant. This mock-up was created in Figma and features emotional reactions, a speaking queue, a shared view, and the video feeds of each participant with a Whisper and Nudge button.

has the option to turn off all emotional reactions using the left button of the bottom toolbar.

- Speaking Queue on the top, showing the current speaker as well as individuals who are waiting to speak.

- Shared View in the middle, where everyone can see a presentation or shared workspace. The red pointer shows where the speaker is pointing.

The interface includes options to approve or disable emotional reactions, addressing privacy concerns around sharing affective data. By consistently seeking user consent and providing transparency, ProxyBridge fosters trust and helps users feel more secure when sharing personal cues in meetings.

User Feedback Session 1: Focus Group

To evaluate the initial prototype, a remote focus group was held with three HCI graduate students experienced in hybrid collaboration. Recruited via university channels, participants completed scripted tasks using the Figma prototype in a simulated meeting. Sessions were recorded, transcribed, and inductively coded to identify usability, spatial mapping, and communication challenges, which shaped later design iterations.

Focus Group Findings

Spatial Orientation and Seat Claiming

Participants began by "claiming" a seat using a layout inspired by classroom/theater metaphors, which helped visually orient them; one noted it "helped organize the meeting." However, usability issues arose around distinguishing available seats, unclear claim status, and low text visibility. One participant noted, "it wasn't obvious which seats were taken," while another suggested, "claimed seats should be grayed out more prominently." In response, the interface was updated with larger icons, clearer claim indicators, and higher-contrast fonts to improve glanceability.

Expressivity and Emotional Cues

The meeting dashboard used emoji-style indicators to show participants' emotional states and engagement. While participants appreciated the concept, they noted ambiguous iconography and high cognitive load—e.g., "thinking" and "confused" were hard to distinguish, and low-contrast colors (e.g., green on green) reduced readability. These concerns prompted revisions to icon design and color palette to improve clarity and accessibility across environments.

Whisper and Gaze-Based Interactions

Participants reviewed storyboarded features such as whisper interaction and gaze cues. Participants proposed haptic or auditory choices since they found the flashing light bothersome and "not discreet enough," even if whispering was preferred for facilitating private communication. Reactions to gaze-based attention tracking were varied; some people thought it was helpful, while others thought it was "too intrusive," comparing it to surveillance. These responses highlight the issues of user comfort and embodiment in accessibility.

Implications for Design

The focus group highlighted the need to balance social presence with user boundaries. Participants valued features that improved awareness and inclusion but stressed the importance of transparency and customization in attention and emotion cues. Their feedback guided design updates in visual hierarchy, discreet signaling, and ethical framing of embodiment.

REFINEMENT OF PROTOTYPE

ProxyBridge Embodied Avatar Device

We produced a hardware prototype artifact, an embodied computing device called the ”ProxyBridge” (or affectionately, ”The Orb”), to serve as the embodied representation of remote meeting participants in the central meeting room location.

Implementation

The main body of the device is a translucent, textured glass orb. Our design had long called for a relatively small device in a spherical or cylindrical shape; this allows us to have social cue indicators that can be signaled in 360°. The size of the device makes it noticeable, but still portable. The translucency of the orb allows the light from an embedded twelve-segment LED ring to shine through, which is the primary modality used to indicate various social cues and communication functions. The device also includes a haptic vibration motor, adding audio and touch modality to select interactions.

An Arduino Uno microcontroller was used to control the electronic components of the prototype, which was programmed with default behaviors, as well as the ability to trigger different interaction sequences over a serial connection via USB. The FastLED library [1] was utilized to program a series of custom animation sequences in the C++ programming language for the LED ring.

The first sequence is ”moveToPosition”, which animates three white LEDs from a current starting position to an ending position sequentially along the radial path using a specified duration, intended to indicate the gaze direction in which the remote participant is looking.

The second is a ”speaking” animation, which lights up the entire orb in a moderate-luminosity green color to indicate that they are currently speaking; the animation subtly pulses or flickers to add a sense of vitality and to subtly increase the prominence of the speaker.

The third is an ”upNext” animation, used to indicate that the represented participant has indicated that they wish to speak, and that they are the next up in the speaking queue; this sequence animates one blue LED around in a loop reminiscent of a progress indicator, and after it concludes it segues directly into the green ”speaking” sequence while the user speaks.

The fourth is a ”receiveWhisper” sequence, this involves three pulses of purple light over the span of one second on the side of the orb facing the in-person participant, indicating that they are receiving a direct whisper communication; each pulse is accompanied by a haptic vibration to draw the attention of the in-person participant who is receiving the communication.

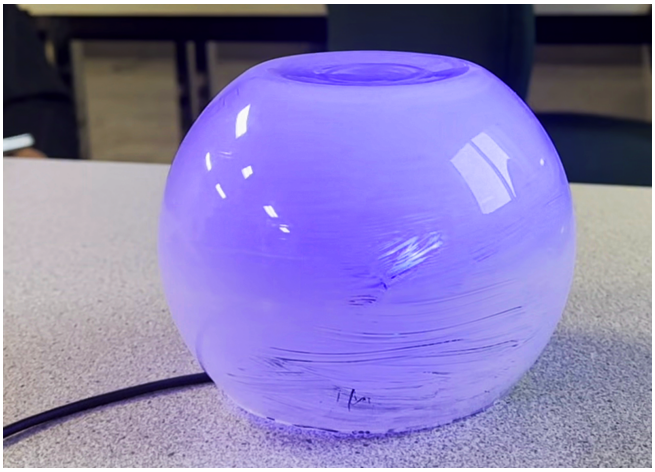


Figure 4: ProxyBridge Device, seen glowing blue in a radial animation during the ”upNext” sequence.

The fifth is the closely related ”nudge” sequence, which is the same as the whisper, but only a single pulse; this cue simply indicates to the participant that somebody is trying to get their attention.

Wizard of Oz Methods

For the purposes of the subsequent user interview sessions, ”Wizard of Oz” methods were used to simulate the actions of the remote participant who was being represented by the ProxyBridge device. While the device and its interactions were high-fidelity and real, the system to gather the social cues of the remote participants was simulated.

To represent the gaze social cue, a ”wandering gaze” default sequence was implemented: this sequence triggered the *moveToPosition* sequence, moving the gaze to a random new position quickly or slowly (duration between a half-second and five seconds), and then pausing in that new position for between zero and ten seconds before triggering again.

To clarify the meaning of the gaze indicator, one portion of the test has a video of remote participants' *whisper* appear; a custom animation sequence was programmed and initiated at the start of this video to move the gaze indicator to predefined positions that map to where she is looking over the course of the pre-recorded video.

Each of the sequences, aside from the default *wanderingGaze* sequence and the automatic segue between *upNext* and *speaking*, were triggered from a command line on a connected laptop, sending that command over a serial connection to the Arduino microcontroller and triggering the appropriate sequence in the test or demo.

ProxyBridge Device Interface: In-Person Perspective

Figma Updates

Following the first focus group, we implemented several minor revisions to the Figma-based prototype to address usability concerns. These adjustments aimed to improve clarity, accessibility, and overall user experience, particularly for remote participants in hybrid meetings.

Enhanced Seat Claiming Interface
Because to the lack of iconography and low text contrast, participants had trouble telling the difference between claimed and open seats. Consequently, the following functionalities were added to the seat claiming interface:

- Improved Visual Indicators: Claimed seats are now prominently grayed out, and available seats are highlighted with higher contrast colors to enhance visibility.
- Larger Icons and Labels: Seat icons and accompanying text labels were enlarged to improve readability and ease of selection.
- Clear Claim Status: Visual cues were added to clearly indicate the claim status of each seat, reducing ambiguity.

Refined Emotional Reaction Emojis
The initial design used visual signals that resembled emojis to represent the participants' emotional states. Feedback indicated that several emoticons were too similar to one another, making them confusing. The following improvements were made:

- Distinct Iconography: Emojis were redesigned to have more distinct shapes and expressions, reducing visual similarity between different emotional states.
- Accessible Color Palette: Colors were adjusted to ensure sufficient contrast, making the emojis more distinguishable for users with visual impairments.

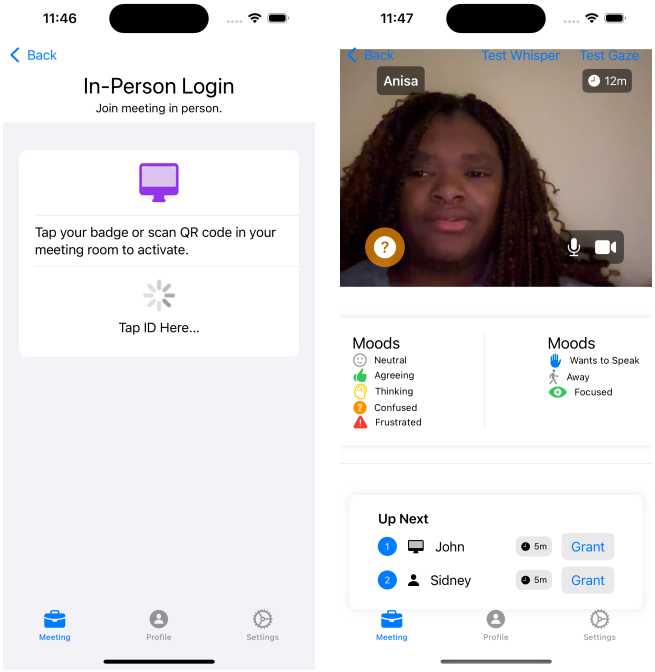
iOS Application

A native iOS (iPhone) application prototype was developed using the Swift programming language and SwiftUI GUI framework. The application is intended to visualize meeting state and facilitate certain forms of interactivity. The device running the app is mounted to the ProxyBridge device. It displays important meeting metadata such as the number of participants and the queue of people who wish to speak. It provides the interface by which meeting attendees can easily authenticate a device as their own during a meeting (Figure 5a), this authentication enables spatial communication methods and peer-to-peer communication via *whispers* and *nudges*.

When a device is representing a remote user, this user interface is capable of optionally displaying a video stream, and/or displaying the remote participants' emotional state, focus level, and active/idle status, as seen in Figure 5b.

Interactive Desktop Interface: Remote Perspective

For the refined prototype of a remote user's perspective, the desktop interface for ProxyBridge was redesigned with interactivity and updated to reflect user feedback. When joining the meeting in the desktop view, the remote participant has



(a) In-person authentication (b) "Whisper" P2P message received

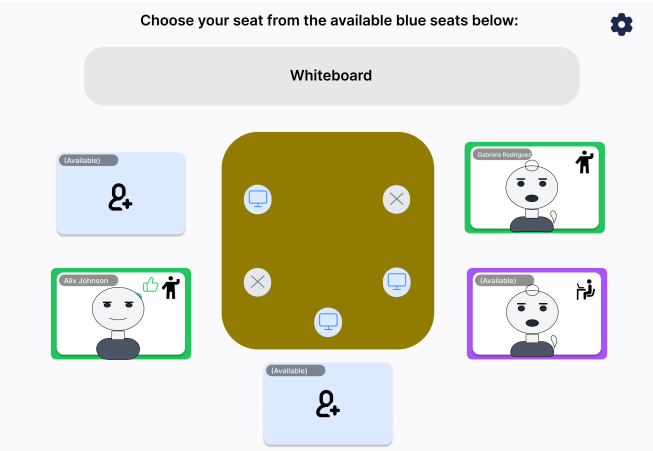
Figure 5: iOS application GUI

the benefit of seeing the spatial layout of the physical meeting space as well as where their device is in the space, as shown in Figure 6a.

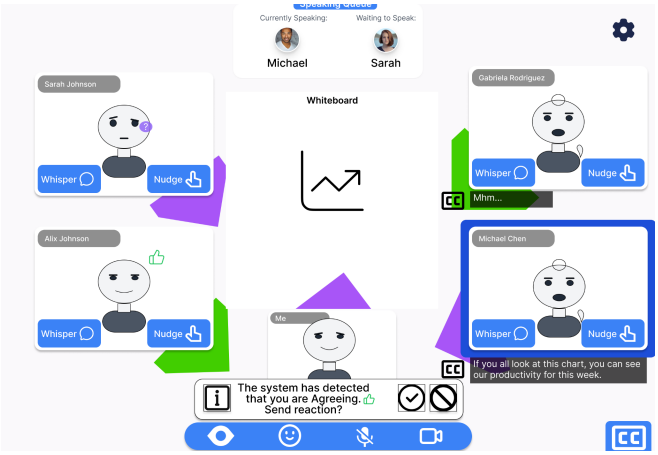
The overall layout of the main screen, shown in Figure 6b, remained the same from the initial prototype, but a few key changes were made:

- Aesthetic Improvements: The visual design of the prototype was improved, which involved creating a consistent color scheme and cleaning up the layout.
- Gaze Indicators: This interface incorporates arrows showing the direction that each meeting attendee is currently looking; the purple arrows are the gaze of remote participants, and the green arrows are the gaze of in-person participants. This color difference is not essential information, but it may serve as a reminder of who is in-person and who is remote.
- Spatial Captioning: As shown under the two speakers on the right side of Figure 6b, captions are shown directly under the speaker. This design decision draws from previous research indicating that this layout would be preferred by some DHH individuals [16]. This could also reduce the visual burden of glancing at captions at the bottom of the screen for all users.

The gaze indicators and emotional reactions from other participants provide a replacement for missing social cues. The whisper and nudge features allow for one-on-one interactions that remote participants often miss out on.



(a) Spatial view of meeting room. This screen shows the user what the physical meeting space looks like, and allows the user to choose where they would like to be in the room.



(b) Refined remote desktop interface. This is the main screen that a remote meeting attendee interacts with during the meeting. There are Whisper and Nudge buttons for each participant, a Speaking Queue, a bottom toolbar, and a shared view in the middle.

Figure 6: Desktop User Interface

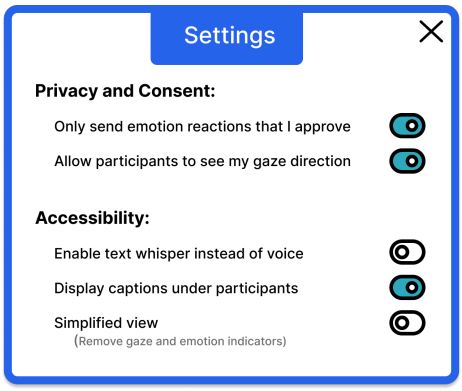


Figure 7: Configurable settings. When the user clicks the gear on the main screen, this pop-up appears with privacy and accessibility settings.

The configurable settings for ProxyBridge are shown in Figure 7, including both privacy and accessibility settings. To prioritize user consent and privacy, it was important to offer control over both gaze and emotion because we found perceptions on these to be very individualized from user feedback; participants often had different opinions about which information was invasive to share, so the user has control over both emotional reactions and gaze independently. The accessibility settings allow the user to set the text whisper as default, display captions under each speaker instead of all at the bottom, and activate a simplified view that removes features that may distract or confuse some users, offering a more familiar experience like Zoom. This simplified view could be particularly helpful for those with cognitive disabilities or individuals who are not tech-savvy.

User Feedback Session 2: Hands-On Interviews

For the second round of interviews, the priority was to gather feedback on the first-use intuitiveness of the ProxyBridge system. Our rationale for this was that a videoconferencing tool should not have a large learning curve or add unnecessary cognitive load because in a real meeting context, the user's main focus should be joining and engaging with the meeting. Further, users often do not have time to test out teleconferencing features before joining a meeting. The format of the hands-on interviews followed this procedure, which was repeated for each major feature in the prototype: (1) Show a feature of the prototype, such as the purple flashing light for a Whisper notification. (2) Ask the user to guess the intended functionality of the feature. (3) Probe the user's mental model of the device by asking follow-up questions. (4) Explain the intended functionality and ask about intuitiveness, as well as their opinions about the design decision.

We interviewed three different participants, all of whom were between the ages of 18 and 24 and familiar with teleconferencing software. Each interview took place in-person at Wallace Library, and the interviews lasted between 30 and 60 minutes. One of the interviewees also participated in User Feedback Session 1: Focus Group.

Interview Findings

During the test, we noticed that the participants initially had difficulty understanding the device behavior, causing confusion. Upon explaining the functionality, participants found the system to be intuitive once they understood. On average, participants gave a rating of 3.67/5 for the overall experience interacting with the system, and 3.67/5 for the intuitiveness. Two participant commented on the learning curve, and how the system becomes more intuitive with use. All three participants agreed that the device improves the remote attendee's presence in the hybrid meeting setting, although some concerns about the device being too loud or distracting were raised. Participants also expressed concerns about their privacy, requesting options to turn off their gaze indicator and emotional status. Additional privacy concerns regarding social interactions, such as Nudge and Whisper requests being

too loud, were largely due to hardware limitations in the prototype's vibrations and lights.

DISCUSSION

Summary of Findings

The light indicators from the dome can successfully attract attention, which would make it easier for the meeting attendees to notice and respond. In addition, participants felt that ProxyBridge would increase remote engagement with the implementation of social interactions using the dome. Most of the concerns were due to the privacy of certain features. All three participants also asked for more a discrete notification for a whisper request, such as reduced vibration and light intensity to match the privacy level of the interaction. Participants also agreed that both the gaze indicator and emotional status contribute to in-person attendee interaction and involvement with the remote attendee.

All participants found the dome design and the light signals representing social interactions to be intuitive, but hard to grasp initially. One participant has suggested giving users more room to customize their device, as it helps meeting attendees to identify people based on the appearance. The customization includes idle color theme, pattern/logo and profile picture.

Design Implications

The feedback from the second usability study provided additional insight for the prototype. The system needs to allow user more control over what to display regarding their status, specifically gaze indicator of the dome and status display for the screen. Allowing users to freely turn on and off for these two information, addressing the privacy concerns of the users.

Secondly is to enable customization of the physical device, for users to personalize the appearance. Once they login and claim a place the device will take on their personalized appearance. This way it helps attendees to better identify the users based on the device. And it has the advantage of creating a more personable experience for the in-person attendee, as they are not interacting with a plain electronic dome. Throughout the user studies the participants have emphasized on the importance of anthropomorphic design, that let the device become the physical avatar of the remote attendee in the meeting space.

Future Work

Future work could focus on the impact of allowing the user to personalize their ProxyBridge device to better distinguish meeting attendees and make the devices more personable. It would also be helpful add a better onboarding experience to address the initial learning curve participants faced. To verify the functionality of the system's accessibility features, it is crucial to conduct dedicated accessibility testing.

LIMITATIONS

While we initially gathered information from professionals across industries, the sample pool of our usability studies is limited to on-campus students at RIT. In addition, due to the testing being scheduled near finals, we had difficulty recruiting new participants, so we reused one participant for both user feedback sessions.

Furthermore, this study is hindered by the limited time and resources, leading to reduced building time. In the first iteration there was no physical prototype, and we relied on a storyboard and Figma prototypes to convey the idea and goal of ProxyBridge. Storyboards were also used to demonstrate the interaction concepts to participants. The prototype also had some malfunctions during the second usability study, causing confusion and breaking immersion for the participants.

In addition, we only developed one working prototype, so the study was solely relying on one prototype to convey different functions and perspectives of the meeting. During the testing, the participants were instructed to switch the roles of different users mid-testing, which could be confusing and disruptive to demonstrating the user flow.

CONCLUSION

Using today's teleconferencing tools, remote attendees in hybrid meetings often find themselves overlooked and struggle to remain engaged and included. To address this issue, we introduce the ProxyBridge system, including both a physical prototype and a suite of accompanying software. By embodying remote participants with a virtual seat at the table, and by incorporating the missing social cues normally present in human interaction through multi-modal channels, we can rebalance the social asymmetry that prevails today.

In multiple rounds of user feedback sessions, participants responded positively to the prototype, confirming the design choices and the device's ability to increase remote attendees' presence. However, there is room for improvement: while participants eventually found the system intuitive, testing suggests an initial learning curve and the need for an improved onboarding experience. Features like Whisper and Nudge were received positively, but participants agreed that the notifications needed to be more private. Participants believed the devices would greatly improve remote meeting attendees' ability to be noticed in meetings. Significant differences in privacy preferences verifies the design choice to allow for granular customization of social cue channels and features. Participants also suggested additional options for customization and personalization, such as a color representing a certain person.

Overall, it is clear that the ProxyBridge system is a meaningful step towards improving the hybrid meeting experience. By emphasizing rich channels of social cues, proxemic and spatial relationships, and accessibility, the system restores much of the natural, rich human interaction that is too frequently lost when communicating through technology.

REFERENCES

1. 2025. FastLED/FastLED. (May 2025). https://github.com/FastLED/FastLED original-date: 2013-11-10T01:41:14Z.
2. Rahaf Alharbi, John Tang, and Karl Henderson. 2023. Accessibility Barriers, Conflicts, and Repairs:

Understanding the Experience of Professionals with Disabilities in Hybrid Meetings. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 605, 15 pages. DOI:http://dx.doi.org/10.1145/3544548.3581541

3. Valentin Bercaru and Nirvana Popescu. 2024. A Systematic Review of Accessibility Techniques for Online Platforms: Current Trends and Challenges. *Applied Sciences* 14, 22 (2024), 10337.
4. Richard L. Daft and Robert H. Lengel. 1986. Organizational Information Requirements, Media Richness and Structural Design. *Management Science* 32, 5 (1986), 554–571. https://www.jstor.org/stable/2631846 Publisher: INFORMS.
5. Roelof Anne Jelle de Vries, Juliet A. M. Haarman, Emiel Harmsen, Dirk K. J. Heylen, and Hermie J. Hermens. 2020. The Sensory Interactive Table: Exploring the Social Space of Eating. *Proceedings of the 2020 International Conference on Multimodal Interaction* (2020). https://api.semanticscholar.org/CorpusID:224816444
6. Morten Esbensen, Paolo Tell, and Jakob E Bardram. 2014. SideBar: Videoconferencing system supporting social engagement. In *10th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*. IEEE, 358–367.
7. Robert S Fish, Robert E Kraut, and Barbara L Chalfonte. 1990. The VideoWindow system in informal communication. In *Proceedings of the 1990 ACM conference on Computer-supported cooperative work*. 1–11.
8. Edmund T. Hall. 1969. *The hidden dimension*. Doubleday.
9. Komei Hasegawa and Yasushi Nakauchi. 2014. Telepresence robot that exaggerates non-verbal cues for taking turns in multi-party teleconferences. In *Proceedings of the second international conference on Human-agent interaction*. 293–296.
10. Lene Heiselberg and Agnieszka Stepińska. 2022. Transforming Qualitative Interviewing Techniques for Video Conferencing Platforms. *Digital Journalism* 11 (2022), 1353 – 1364. https://api.semanticscholar.org/CorpusID:247729991
11. Erzhen Hu, Jens Emil Sloth Grønbæk, Austin Houck, and Seongkook Heo. 2023. OpenMic: Utilizing Proxemic Metaphors for Conversational Floor Transitions in Multiparty Video Meetings. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–17. DOI: http://dx.doi.org/10.1145/3544548.3581013
12. Nuur Roofiah Kusumawati, Agus Efendi, and Cucuk Wawan Budiyanto. 2024. Comparative Analysis Between Using Google Meet and Gather Town Video Conferencing Platforms Viewed from User Experience Using meCUE Questionnaire on Students. *IJIE (Indonesian Journal of Informatics Education)* (2024). https://api.semanticscholar.org/CorpusID:272148939
13. Jingjin Li, Shaomei Wu, and Gilly Leshed. 2024. Re-envisioning Remote Meetings: Co-designing Inclusive and Empowering Videoconferencing with People Who Stutter. *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (2024). https://api.semanticscholar.org/CorpusID:270822727
14. Nai Li and David Lefevre. 2020. Holographic teaching presence: participant experiences of interactive synchronous seminars delivered via holographic videoconferencing. *Research in learning technology* 28 (2020).
15. Siran Ma, Qingyu Hu, Yanran Chen, Zhilong Zhao, and Houze Li. 2023. Social Bots that Bring a Strong Presence to Remote Participants in Hybrid Meetings. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 853–856.
16. Emma J. McDonnell, Ping Liu, Steven M. Goodman, Raja Kushalnagar, Jon E. Froehlich, and Leah Findlater. 2021. Social, Environmental, and Technical: Factors at Play in the Current Use and Future Design of Small-Group Captioning. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–25. DOI:http://dx.doi.org/10.1145/3479578
17. Joshua McVeigh-Schultz and Katherine Isbister. 2022. A "beyond being there" for VR meetings: envisioning the future of remote work. *Human–Computer Interaction* 37, 5 (2022), 433–453.
18. Carol Nash. 2020. Report on digital literacy in academic meetings during the 2020 COVID-19 lockdown. *Challenges* 11, 2 (2020), 20.
19. Thomas Neumayr, Banu Saatci, Sean Rintel, Clemens Nylandsted Klokmose, and Mirjam Augstein. 2021. What was hybrid? a systematic review of hybrid collaboration and meetings research. *arXiv preprint arXiv:2111.06172* (2021).
20. Payod Panda, Lev Tankelevitch, Becky Spittle, Kori Inkpen, John Tang, Sasa Junuzovic, Qianqian Qi, Pat Sweeney, Andrew D Wilson, William AS Buxton, and others. 2024. Hybridge: Bridging Spatiality for Inclusive and Equitable Hybrid Meetings. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–39.
21. Muhammad Haris Patel, Jamshed Akhtar, Syed Muhammad Raees Hussain Taqvi, and Tayyaba Batool. 2021. Analysis of challenges faced and the scientific content of a hybrid pediatric surgical conference arranged during the COVID-19 pandemic. *Annals of Pediatric Surgery* 17, 1 (2021).

22. Viktor Richter, Birte Carlmeyer, Florian Lier, Sebastian Meyer zu Borgsen, David Schlangen, Franz Kummert, Sven Wachsmuth, and Britta Wrede. 2016. Are you talking to me?: Improving the Robustness of Dialogue Systems in a Multi Party HRI Scenario by Incorporating Gaze Direction and Lip Movement of Attendees. *Proceedings of the Fourth International Conference on Human Agent Interaction* (2016). https://api.semanticscholar.org/CorpusID:15092781

23. Camille Sallaberry, Gwenn Englebienne, Jan Van Erp, and Vanessa Evers. 2023. Out of Sight,... How Asymmetry in Video-Conference Affects Social Interaction. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION*. ACM, Paris France, 465–469. DOI: http://dx.doi.org/10.1145/3577190.3614168

24. Matthew Seita, Sooyeon Lee, Sarah Andrew, Kristen Shinohara, and Matt Huenerfauth. 2022. Remotely Co-Designing Features for Communication Applications using Automatic Captioning with Deaf and Hearing Pairs. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–13. DOI: http://dx.doi.org/10.1145/3491102.3501843

25. Ingo Siegert and Oliver Niebuhr. 2021. Case Report: Women, Be Aware that Your Vocal Charisma can Dwindle in Remote Meetings. In *Frontiers in Communication*. https://api.semanticscholar.org/CorpusID:231694381

26. Maurício Sousa, Daniel Mendes, Alfredo Ferreira, João Madeiras Pereira, and Joaquim Jorge. 2015. Eery Space: Facilitating Virtual Meetings Through Remote Proxemics. In *Human-Computer Interaction – INTERACT 2015*, Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler (Eds.). Springer International Publishing, Cham, 622–629. DOI: http://dx.doi.org/10.1007/978-3-319-22698-9_43

27. Norbert A Streitz, Jörg Geißler, Jörg M Haake, and Jeroen Hol. 1994. DOLPHIN: integrated meeting support across local and remote desktop environments and LiveBoards. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. 345–358.

28. Joseph Tu, Arielle Grinberg, Mark Hancock, and Lennart E. Nacke. 2023. Multimedia Showdown: A Comparative Analysis of Audio, Video, and Avatar-Based Communication. *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (2023). https://api.semanticscholar.org/CorpusID:263697141

29. Javlon Tursunov, Gregor Rozinaj, Vivek Dwivedi, and Ivan Minárik. 2024. A Customizable WebRTC-based Video Conferencing System For Real-time Communication. *2024 31st International Conference on Systems, Signals and Image Processing (IWSSIP)* (2024), 1–8. https://api.semanticscholar.org/CorpusID:271939591

30. Mitali Vishwekar, Daniel Injeti, Dr. Rakhi O. Gupta, and Nashrah Gowalker. 2025. SMART SIGN: Live Sign Language Interpretation For Barrier-Free Video Conferencing. *International Journal For Multidisciplinary Research* (2025). https://api.semanticscholar.org/CorpusID:276154075

31. Basil Wahn, Laura Schmitz, Alan Kingstone, and Anne Böckler-Raettig. 2022. When eyes beat lips: speaker gaze affects audiovisual integration in the McGurk illusion. *Psychological Research* 86, 6 (2022), 1930–1943.

32. Cheng Yao Wang, David Saffo, Bill Moriarty, and Blair Maclntyre. 2024. CollabXR: Bridging Realities in Collaborative Workspaces with Dynamic Plugin and Collaborative Tools Integration. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 454–457.

33. Siqin Wang and Qingdu Li. 2023. Remote collaborative process optimization in research and design of industrial manufacturing. *Autonomous Intelligent Systems* 3, 1 (2023), 10.

34. Bin Xu, Jason Ellis, and Thomas Erickson. 2017. Attention from afar: simulating the gazes of remote participants in hybrid meetings. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 101–113.

35. Zhilong Zhao, Yanran Chen, Qingyu Hu, Siran Ma, Houze Li, Yijie Guo, and Haipeng Mi. 2023. Buzzo or Eureka–Robot that Makes Remote Participants Feel More Presence in Hybrid Discussions. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 323–327.

36. Annuska Zolyomi, Andrew Begel, Jennifer Frances Waldern, John Tang, Michael Barnett, Edward Cutrell, Daniel McDuff, Sean Andrist, and Meredith Ringel Morris. 2019. Managing Stress: The Needs of Autistic Adults in Video Calling. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 134 (Nov. 2019), 29 pages. DOI: http://dx.doi.org/10.1145/3359236