**ABSTRACT**

Backchannel communication, like whispering or instant messaging, is common in meetings and holds significant value. However, it remains largely unexplored in the context of hybrid meetings with co-located and remote participants. To address this gap, we derived unique challenges of backchannel communication in hybrid meetings through an interview study. These challenges inspired a new voice-based backchannel communication system, WhisperChannel, aimed to be inclusive and low effort. WhisperChannel enables users to whisper remotely to anyone in the meeting—similar to whispering to co-located seat neighbors. In a user study with three groups, each having two sessions of hybrid meetings, we investigated the inclusiveness and effort of WhisperChannel for both co-located and remote participants. We provide insights into the benefits and limitations of using remote whispering. We show that WhisperChannel helped remote and co-located participants feel more included while requiring low effort to use, however, also introduced new challenges in backchannel communication.

**CCS CONCEPTS**

- Human-centered computing → Collaborative and social computing: Interactive systems and tools.

**KEYWORDS**

backchannel communication, hybrid meetings, virtual meetings, whispering
1 INTRODUCTION

Hybrid meetings, where both co-located and remote participants engage in synchronous participation, have gathered increasing attention from researchers since the pandemic [43, 44]. Most remote and hybrid work systems have been designed to support formal interaction during meetings, but not the full social experience that occurs during face-to-face meetings. In face-to-face meetings, it is not uncommon to engage in discreet communication, such as leaning over to the person in closest proximity and whispering inquiries, or utilizing mobile devices to communicate through instant messaging during a presentation [12, 26, 49, 53]. This phenomenon is known as backchannel communication [3]. A backchannel is a line of communication to connect with others inside or outside the room, with or without the knowledge of the speaker at the front of the room, running parallel to the primary frontchannel discourse [3, 4]. Various methods exist for conducting backchannel communication, e.g., physical interactions such as verbally whispering or passing paper notes during face-to-face meetings, or digital interactions based on text-based platforms such as instant messaging, email, Internet Relay Chat, live blogging, and chat rooms [2, 3, 12].

Despite the existing research, a significant gap remains in understanding backchannel communication in hybrid meetings. For instance, consider a hybrid group meeting in which a co-located speaker gives a presentation to co-located participants, while someone joins remotely (see Figure 1). Most studies about backchannel communication focus on its content and function [4, 13, 35, 38, 55], but do not explore the unique challenges that arise when remote and co-located participants engage synchronously in backchannel communications during hybrid meetings.

We believe that investigating backchannel communication in hybrid meetings holds significant value. Unlike face-to-face or entirely distributed meetings where everyone is in a relatively equal context, the asymmetric relationship [6] between remote and co-located people highlights challenges when it comes to backchannel communication. Such backchannel communication is often dispersed across multiple channels and modalities (e.g., speech, text, etc.), and sometimes completely inaccessible to remote participants. Researchers have demonstrated that utilizing backchannel communication can equalize the social relationship between meeting participants [27, 39] by connecting them through shared discussions. Nevertheless, in current systems, this form of communication is limited and primarily restricted to text-based interactions for remote participants. Additionally, as hybrid modes of work have become more predominant following the pandemic, it is important to explore hybrid meetings [14, 43].

To bridge this research gap, we identified two unique challenges related to backchannel communication in hybrid meetings through an interview study. These challenges served as the inspiration for us to build a novel, exploratory voice-based backchannel communication system called WhisperChannel. Aiming to tackle the two challenges, WhisperChannel supports inclusive and low-effort backchannel communications. Participants can whisper to anyone in the meeting, including remote participants. Our system design aims to make this as easy as whispering to co-located meeting participants in close proximity. Finally, we conducted a qualitative user study to gain insights into how our system addressed the two challenges identified in the interview study. We formed three groups, each participating in two sessions of simulated hybrid meetings. Our findings show that WhisperChannel led to both remote, and co-located participants sitting far away from others, feeling more included. WhisperChannel required relatively low effort in certain aspects but added extra effort in others, and it could not yet reach the same level of ease as face-to-face whispering. The paper, thus, contributes the following:

- An interview study that derives challenges of backchannel communication during hybrid meetings.
- A novel voice-based backchannel communication system, WhisperChannel, motivated by these challenges.
- A user study identifying how WhisperChannel tackles these challenges.

2 RELATED WORK

Prior research has investigated challenges in hybrid meetings, and backchannel communication in remote meetings and co-located meetings, but there is a gap in research as regards backchannel communication in hybrid meetings.

2.1 Hybrid Meetings

In the aftermath of the COVID-19 pandemic, the most common work arrangement for knowledge workers is now some form of a hybrid model, a type of flexible working where an employee splits their time between the workplace and remote working [23]. One outcome of this work model is an increase in volume of hybrid meetings — audio- or video-based meetings characterized by the simultaneous presence of both co-located and remote participants [43, 44]. Duckert et al. [14] established the concept of hybrid meeting arrangements as involving collaborative settings, where a minimum of three participants rely on each other’s work while being situated in a number of geographical sites that is less than the total number of participants.

When analyzing the implications of hybrid work, it becomes apparent that hybrid meetings not only inherit challenges from both co-located and remote meetings, but also possess unique attributes that amplify the difficulties linked to collaborative work within hybrid work settings [20, 43, 44]. In hybrid work settings, all participants have only partial access to the hybrid setup, with no single person granted complete access to the entire configuration. These inconsistent technical frameworks make an invisible “wall” in collaboration and cause hybrid work situations to be inherently challenged by asymmetric relationships [6].

Both collocated, remote and hybrid meetings are examples of cooperative work arrangements [45] where people are mutually dependent on one another to get work done. Hence, the setting up and carrying out of hybrid meetings include articulation work to divide,
allocate, coordinate, schedule, mesh, interrelate, etc. distributed individual activities of the participants [45].

Christensen et al. [11] pointed out that collocated relation work is spontaneous, but relation work in a distributed setting is semi-spontaneous, technology-mediated, and requires extra efforts. In cooperative settings, communication assumes a substantial role in relation work [47]. These factors may contribute to the fact that collocated people have a lot of side discussion, storytelling, and interjections, whereas these activities are rare in remote settings [25, 31, 37], emphasizing the possible added effort of articulation work for remote participants engaged in backchannel communication.

One feature of collocation that is missing in remote work is awareness of the state of other people and their activities. This awareness is an important part of common ground [18], and when teams are fully collocated, it is relatively easy to establish such common ground since participants share awareness of who is doing what at the moment. Similarly, in fully distributed teams, participants utilize the same video or audio conference platform and thus have identical means of awareness. However, there are always challenges in the remote situation that make it qualitatively different than collocation [37] and even further accentuated in hybrid situations [6, 14].

In hybrid meetings, there is a potential tension of communication with audio, particularly when it comes to interruptions. This happens because there is only one audio channel from the remote location to the co-located place, with technical constraints, sometimes when someone says something and another person tries to interrupt, there can be problems like network delays and audio glitches in the video conferencing platform, which may result in both parties losing their audio [6].

Since there is a research gap regarding hybrid meetings specifically with backchannel communication, we look further at the particular challenges of hybrid meetings with the technologies deployed backchannel communication for these situations.

2.2 Backchannel Communication

The term backchannel can be defined differently based on the specific discipline one is working in [12]. In our research, we draw upon Goffman’s concepts of frontstage and backstage behavior to explain what backchannel means [2, 16, 48]. Frontstage behavior occurs in the presence of an audience adhering to expected roles, and actors adjust their actions to align with an idealized image. Backstage behavior, on the other hand, involves interactions behind the scene among teammates.

Backchannel communication has diverse interpretations, with two primary references. In linguistics, backchannel often involves simple affirmations like “uh-huh” used by listeners to indicate their attentiveness or agreement with the speaker [12]. Our focus lies on the aspect where backchannel communication as a discreet secondary channel running parallel to the primary frontchannel discourse [2, 12, 33, 52, 55]. This frontchannel discourse typically involves official and formal communication, e.g., speakers delivering content in various settings, with the backchannel serves as a behind-the-scene space where additional interaction takes place [4]. Different studies have employed varied terminologies in the same meaning as backchannel, such as out-of-band communication [54], parallel conversation [31], side conversation [2, 31] and multicomunication [10, 36, 49].

Similar to a double-edged sword, backchannel communication can serve a constructive purpose by augmenting and expanding valuable information and relationships, offering instant feedback to the speaker [3, 12, 31, 32] and facilitating more dynamic group conversations [24]. It could even alter the collaborative decision-making processes and may impact overall outcomes [2, 9]. Conversely, backchannel communication can be destructive due to its visibility and potential audibility to other attendees and create distraction [3, 32].

Previous research has examined backchannel communication either in physically co-located settings [2, 31, 35], or in distributed online meetings [2, 12, 24, 31, 33]. However, none of these studies have specifically investigated backchannel communication in the context of hybrid meetings that involve both co-located and remote participants simultaneously. Researchers found that backchannel communication in remote meetings is sometimes lost because the majority of such meetings are characterized by formality and rigidity [31].

2.2.1 The Content of Backchannel Communication. People use backchannel communication channels for various purposes. The categorization of Cogdill et al. [12] includes process-oriented (focuses on the environment and generally attempts to work with group process), content-oriented (reflects upon the content of the frontchannel discussion), participation-enabling (provides technical and social assistance so participants can engage more fully in the public conversation), tangential (briefly diverges from the frontchannel discussion), and independent discourse (being unassociated with the frontchannel). Alan R. Dennis [2] developed a set of six somewhat overlapping sub-genres: directing the meeting, providing task support, seeking clarification, providing social support, participating in a subgroup meeting, and managing extra-meeting activities.

These two categorizations can be summarized as: facilitating the process, addressing technical or social aspects, providing support and helping participants engage effectively; involving reflections on or discussions related to the frontchannel discussion, providing additional insights or analysis; and unrelated communication with the frontchannel, such as greetings, farewells, non-work-related discussions, and humor. We use these insights in conducting our user study, when we assign participants to tasks (see also in section 6).

2.2.2 Types of Backchannel Communication. Voice-based backchannel communication encompasses both physical and digital forms of interaction. In physical forms, backchannel communication occurs without any reliance on digital tools, often involving whispered conversations. Conversely, digital forms involve the use of digital platforms to exchange voice notes or similar. However, both forms have been relatively understudied in the past within the context of meetings. Text-based backchannel communication, which has evolved from handwritten notes to predominantly digital tools like instant messaging, has undergone comprehensive research and been widespread used [22].

Backchannels in physical conversations are usually private conversations like whispering in the context of a larger meeting. They
are limited to people who are co-located and in physical proximity [31], and even co-located people sitting apart can not whisper to each other. Furthermore, whispering is synchronous similar to face-to-face communication. Due to the social pressure to adhere to established norms, whispering is a common communication method people prefer in order to not interrupt the frontchannel [2]. Whispering is a lack of vibration of the vocal cords, an excessive private backchannel systems, such as Backstage [35] as well as Twitter [1, 13, 22, 38]. Additionally, designs have emphasized the potential to foster collaboration and enable participants to share ideas, clarify concepts, and resolve issues more effectively.

While whispering was imagined for video conferencing by Mantie et al. [34] in 1991, only very few systems have actually experimented with or implemented verbal whispering features in meetings, and none in hybrid meetings. The Psst add-on in the Argo system allows one-way whispering between meeting participants, facilitating discreet communication during distributed meetings [5]. However, the one-way nature of Psst differs from real-world interactions (two people cannot talk at the same time in Psst). Meeting Central offers an add-on tool with a whisper function [54], where a short audio chime alerts recipients to an incoming voice chat request. To prevent disruptions, recipients must explicitly accept the incoming voice chat, and the main meeting volume is lowered when accepted. FluidMeet [24] incorporates a private calling feature during breakout-room conversations; ongoing private conversations are indicated by a halo on the participants’ videos. However, the potential benefits of these systems for hybrid settings are unknown.

The insights and discoveries obtained from these designs suggest that this direction of research has potential for hybrid situations: For instance, the proposals of Psst and FluidMeet emphasize the potential of effortless audio-only communication. Meeting Central’s analysis, contrasting the functions of voice-based and text-based communication, points to the advantages and limitations of both approaches. Their findings regarding preferences for replicating face-to-face dynamics in virtual meetings have inspired our design.

Various forms of text-based backchannel communication, including handwritten notes as well as platforms like Internet Relay Chat (IRC), live blogging, instant messages, email, and personalized chat rooms [3] have all been used for backchannel communication. Digital platforms enable backchannel communication to transcend physical proximity constraints and reach any location. They can be used asynchronously, meaning that people can read and reply to messages at any time. Studies have explored both private and public modes of communication, yielding insights into the dynamics of text-based backchannel communication during meetings [2, 9, 12, 22, 31, 35, 55]. They demonstrate that backchannel communication is a valuable aspect of creative teamwork, has the potential to foster collaboration and enable participants to share ideas, clarify concepts, and resolve issues more effectively.

Much of the research has been directed towards public text-based backchannel communication, which is visible to other participants, e.g., during conferences [3, 22, 40] and classes [13, 15, 55], IRC channels [35] as well as Twitter [1, 13, 22, 38]. Additionally, designs have been explored to enhance the functionality and user experience of private backchannel systems, such as Backstage [38] and Meeting Central [54]. Those studies found some users prefer private over public chats since private chats allow users to share opinions and ideas with others in a confidential manner, fostering a sense of social support. Furthermore, users view private communication as a stepping stone to eventually participating in public discussions, helping them gain confidence and build the necessary skills for more open communication.

Since there is a lack of research on backchannel communication in hybrid meetings, especially when it involves voice-based backchannel communication, we explore the specific challenges posed by hybrid scenarios, and the voice-based technologies used in these situations in a private manner.

3 INTERVIEW STUDY
To define challenges of backchannel communication during hybrid meetings, and to gain deeper insights into their origins and different contexts, we conducted four different studies, using different methods like workshops, focus groups, and interviews with individuals.

3.1 Participants and Procedure
As part of a large Danish research project exploring the future of work and future work technologies, HCI and CSCW researchers from four different universities collaborate with twelve different industrial partners. The partners encompass a spectrum from large-scale to smaller enterprises, spanning industries ranging from food to IT. In total this large research project has carried out 13 research activities, using workshops, observation studies, interviews, and other methods, all of which were based in Denmark. Some of the studies involved all of the researchers, others only one or two. The main focus is exploring and re-thinking hybrid work situations with focus on collaborative technologies. As part of this we also build and explore technologies of which the current example is one.

Over the period from June 2022 to March 2023, we worked with 18 participants across several partners. Specifically, we conducted four distinct empirical sessions, as shown in Table 1. In Session 1, we invited eight participants (S1P1 to S1P8), one of them joined remotely. The focus group centered on exploring the participants’ experiences and challenges associated with hybrid work. Session 2 consisted of a semi-structured interview with one participant (S2P1), to gain deep insights into the specific challenges encountered in the particular hybrid work setting. For Session 3, we visited BigFoodCorp and conducted a workshop with two participants (S3P1 and S3P2), to understand their present challenges and the vision of the future of hybrid work. In Session 4, we invited fourteen participants (S4P1 to S4P14), four of them joined remotely. The workshop included different group works, the one led by the first author with a theme on informal communication in the hybrid work, including focus group discussion about challenges encountered within this framework.

The subject of backchannel communication was identified as part of these investigations. By exploring the challenges of hybrid work, we recognized the significance and absence of backchannel communication. As we delved deeper, we uncovered the substantial challenges associated with backchannel communication, particularly synchronous communication in hybrid meetings.
Industrial Partners | No. of Participants | Participant IDs | Duration | Method
---|---|---|---|---
Session 1 | Multiple Partners | 8 | S1P1–S1P8 | 3 hours | Focus Group (Hybrid)
Session 2 | NationalFin | 1 | S2P1 | 2 hours | Interviews
Session 3 | BigFoodCorp | 2 | S3P1, S3P2 | 2 hours | Workshop
Session 4 | Multiple Partners | 14 | S4P1–S4P14 | 4 hours | Workshop (Hybrid)

Table 1: Overview of the participants of the interview study. The names of industrial partners are anonymized with invented names that point to their sizes, reach, and domain.

### 3.2 Data Collection and Analysis
In total we recorded 11.5 hours of audio and 7.5 hours of video from the four sessions, as shown in Table 1. We analyzed the transcribed contents from the video and audio recordings and our notes. The first author employed NVivo to code the data and three authors collaboratively identified themes related to challenges of backchannel communications in hybrid meetings.

### 4 CHALLENGES IN BACKCHANNEL COMMUNICATIONS DURING HYBRID MEETINGS
After analyzing the data from interview study, we identified two fundamental challenges for backchannel communication during hybrid meetings. These challenges are C1 — Backchannel communication faces limitations in fostering inclusiveness among all meeting participants, especially remote participants, potentially hindering equal opportunities and awareness for engagement; and C2 — Engaging in backchannel communication may require significant effort in terms of added articulation work with technologies such as texting apps on smartphones.

#### 4.1 C1: Unequal Access to Backchannel Communication Limits Inclusiveness
In our research, we use the definition of *inclusiveness* provided by Shore et al. [46]. This points to ensuring equal access and opportunities for participation, regardless of individuals' background or abilities, including accommodating diverse communication preferences to help all participants engage in the meeting. Furthermore, inclusiveness relates to participants feeling valued and genuinely included in the meeting. It fosters a sense of belonging and psychological safety, where individuals feel comfortable expressing their thoughts and ideas without fear of judgment or exclusion. The lack of sufficient inclusiveness leads to participants feeling unequal and isolated [46].

Overall, when backchannel communication is going on, anyone who becomes aware of it, either in the same space or distributed space, while not being part of it, may perceive it as exclusive [3] and themselves as being excluded. This is in particular a problem among remote participants, because they lack the ability for face-to-face whispering. This implies that co-located and remote participants do not have equal access to the backchannels. Furthermore, remote participants being aware of whispers among co-located participants have difficulties figuring out the contents of these whispered conversations.

When we engage in hybrid meetings remotely, we have repeatedly encountered a common challenge related to sound. Specifically, when two people engage in whispered conversations [in-person], it becomes increasingly difficult to figure out their dialogue [remotely] because the sound equipment does not keep up. (S1P7)

However, it is interesting to note that this feeling of being left out is not limited to remote participants. In certain situations, co-located participants also experience a sense of exclusion:

> When we organize hybrid events, we provide an online platform for participants to watch the presentation and engage with each other through a chatbox. However, for those physically present in the room, they are unable to communicate with each other using the chatbox. It’s quite fascinating to observe this division between the online and physical community, when the chatbox is on fire and the physical attendees have to keep silent. (S4P5)

Unequal access to the backchannel in hybrid meetings also relates to unequal awareness. Awareness in face-to-face interactions partly results from visual cues that enable participants to identify with whom to have backchannel communication, and see who is interrupt or talk. It’s a little bit more different when... (S1P8)

#### 4.2 C2: Articulation Work of Technology Use Requires Excessive Effort
The second challenge relates to *effort*, specifically in terms of the added articulation work for participants, when having a backchannel conversation. We are interested in the additional work or effort it takes to make hybrid meetings, and in particular, backchannel communication work. In line with, e.g., Bowers [8], we focus on the overhead and added effort introduced with the technology.
In the fourth session of interview study, to make the hybrid workshop work, we used in a set of artifacts [6], including seven laptops (four for remote participants) with internet access, external cameras, headphones, a speakerphone, cell phones for texting, and an assortment of wires and cables. To provide a smooth experience — particularly for the remote participants — investigators utilized various backchannel communication platforms, including email, instant messaging services like Slack. To foster stronger communication among the investigators and with the participants, one of the seven laptops, which was dedicated to running Microsoft Teams, served as a backchannel communication hub, sending information to remote participants who faced difficulties, such as issues with joining different group work. However, despite our preparations, some remote participants experienced connectivity problems with the backchannels. This led to delayed messages and communication breakdowns, and the cluttered state of the backchannel platform occasionally added to the confusion, etc. Hence, some remote participants were unable to fully engage, and they experienced difficulties navigating between different group work.

In addition to the instances of articulation work required to establish backchannel communication in hybrid meeting, we also gained valuable insights into the extra effort needed to make this articulation work facilitate backchannel communication. Bowers focuses on the extra work added by technology and advises not to “dismiss it as merely ‘oiling the wheels’” [8]. Whereas face-to-face verbal whispering typically requires minimal effort, text-based backchannel communication demands more effort, due to its complexity and asynchronous nature. Our findings include:

- **Verifying the intended recipients.** Participants expressed concerns about sending messages in the video conferencing platform’s chatbox to the right receiver. It is crucial to see up-front whether the message is intended for a particular individual or the entire group.

- **Revising messages for clarity.** Messages are pre-composed with words, structure, sentences to convey the intended meaning. In contrast, speaking verbally is considered less planned and more effortless [54] as stated by S4P3: “It takes much more time when we tried to describe something through a team’s chat. It’s much easier and more efficient to simply make a phone call.”

- **Keeping up with the constant stream of messages.** Since some information is time-sensitive, not responding promptly may make it difficult to carry forward or continue, as stated by S4P5: “I sometimes find it difficult to keep on top of my texts. So I end up not replying to people for extended periods, and I admit it’s not ideal.” Additionally, as some instant messages come with a “read”-indicator, partners may think they have missed each other in the conversation [10]. Lastly, upon entering the meeting, some individuals described how they keep their phones in silent mode. As a result, they would not receive and respond to messages.

### 4.3 Relationship Between C1 and C2

Returning to the findings made in C1 (remote participants feel excluded in backchannel communication), we argue that C2 (involving excessive effort) might also be a contributing factor to C1. As backchannel communication between remote and co-located participants is only possible through technology, the articulation work required to use this technology involves extra effort, such as following and revising messages. Thus in some contexts, both remote individuals and co-located participants may experience a reluctance to engage in backchannel communication. For example, co-located individuals may prefer less effort approaches, which do not require technology when sharing information without specific targets, such as choosing to whisper to participants sitting close by, but not texting remote participants. Consequently, co-located participants have less chance of engaging in backchannel communication with remote participants, thus leading to C1.

These insights have inspired us to develop a backchannel communication system that requires low-effort, similar to whispering, while enhancing the inclusiveness for all participants in during hybrid meetings.

### 5 WHISPERCHANNEL

After identifying the aforementioned challenges, we started designing and prototyping the system WhisperChannel to address them. To tackle the first challenge (C1) and improve inclusiveness, we incorporated a layout overview and shared backchannel dynamics, which were equally accessible to all users. To address the second challenge (C2) and reduce the effort required, users only need press once to open a bidirectional voice channel, to hear and be heard by others. In this section, we provide an overview of the capabilities of WhisperChannel including its different design dimensions, and briefly summarize its implementation.

#### 5.1 System Overview and Dimensions

WhisperChannel is a web application which is used on a mobile phone that is connected to earphones (see Figure 2a). The user interface (see Figure 2b and 2c) shows an overview of the meeting, its participants, and current active backchannels. The interaction follows a “walkie-talkie-style” communication that requires only a single type of interaction. In the following sections, we introduce different dimensions of the system and how they connect to the challenges in the interview study.

- **Layout Overview.** The user interface of WhisperChannel shows a top-down view of the meeting room and its participants. Participants are illustrated as circles with a name. Furniture or any other objects can be added as images and moved around freely in the interface. When not in a backchannel conversation, one’s own avatar is colored in orange and has a house mark (see participant Charlie in Figure 2b) and avatars of other participants are colored grey.

The layout of the avatars mimics the physical layout of participants and furniture in the main room, while positioning remote participants at the bottom of the screen. In our current prototype, however, this layout has to be prepared before the meeting by the meeting host. This is done using a desktop version of the system that allows to add more avatars or images to the layout or remove objects. Also avatars need to be added manually to the layout beforehand and are not automatically added or removed by the system during a meeting.

We chose this spatial layout of participants over a simple participant list or avatars placed in a grid, because related work shows that digitally compositing users into a virtual meeting scene can help create a more inclusive and equitable meeting space [50] (C1).
Whispering Through Walls: Towards Inclusive Backchannel Communication in Hybrid Meetings

WhisperChannel is used with a regular phone and earphones with microphone.

(a) WhisperChannel is used with a regular phone and earphones with microphone.
(b) The WhisperChannel GUI from the perspective of Charlie.
(c) The WhisperChannel GUI from the perspective of Alice.

Figure 2: The WhisperChannel system. In the GUI, the circles are avatars representing users and the other rectangular objects represent furniture in the room.

Arranging the virtual meeting room in WhisperChannel similar to the physical meeting room enables remote participants to get a better picture of the layout of the meeting. Thus, building common ground and situational awareness through a shared view of the meeting environment could help remote people feeling more included in the hybrid meeting [20].

5.1.2 Shared Backchannel Dynamics. Open backchannels are indicated by the changing the color of avatars and a line between them: In Figure 2b, for example, the blue color of the participants Alice and Bob indicates that they are currently in a backchannel communication. From the perspective of the participant who is in a backchannel conversation, the avatars of themselves and the targeted participant turn green (see Figure 2c). The aim of this dynamic overview is to raise participants’ awareness of the backchannel dynamics within a meeting. On the one hand, remote participants can be more aware of backchannel conversations in the main room. On the other hand, co-located participants have higher awareness of remote participants and their backchannel communications.

Showing the backchannel dynamics in this way aims at increasing the inclusiveness of participants (C1). We considered showing only whether a person is in a backchannel (blue color) without disclosing with whom (the connecting line), however, this would exclude remote participants who cannot see participants in the room who have backchannel communication with each other.

5.1.3 Simple “Walkie-Talkie-Style” Interaction. Opening a backchannel works similar to a walkie-talkie: as long as a participant taps on another participant’s avatar in the application, a one-to-one backchannel to that participant is opened. Once the avatar is released, the channel closes again. Users can only be in one backchannel at a time: a user that is currently in a backchannel cannot be selected and, instead, the user has to wait for the other to finish the backchannel conversation, thus an ongoing backchannel could not be interrupted by other participants.

This dimension of WhisperChannel is strongly motivated by C2: To simplify and reduce the effort of the system, we decided against requiring a confirmation of the targeted participant to open a backchannel, similar to when receiving a phone call, and instead open it directly, like starting a conversation face-to-face. This removes an extra step to open a backchannel.

5.1.4 Immediate One-to-one Voice-Based Communication. While a backchannel is open, a participant can speak to and hear the targeted participant using their earphones with a microphone. Contrary to a walkie-talkie, however, the backchannel is opened in both directions and both participants can talk to and hear each other at the same time. Backchannels in WhisperChannel are always one-to-one — a user cannot select multiple users to talk to. Still, there can be any number of one-to-one backchannels within a meeting.

With WhisperChannel, we aim to create a low-effort (C2) way of having backchannel communications, similar to whispering to one’s neighbor. The related study FluidMeet revealed that private calls is highly effective for facilitating quick and effortless conversations, also fosters a closer connection between users compared to text messages [24]. Hence, we decided to use synchronous speech as the modality over text messages or voice notes. For the same reason, we decided to keep backchannels one-to-one and do not allow group backchannels. We created a version of WhisperChannel
with support for group backchannels, however, it increased the complexity of the interface considerably, and added extra steps of actions.

5.2 Implementation

*WhisperChannel* is built with web technologies and can be accessed using a regular web browser. For our study we used Android phones but *WhisperChannel* also works on, e.g., laptops or tablets. After accessing the URL of *WhisperChannel*, a user can select their user name and connect to the meeting, activating their microphone. The system is then ready and can be used either with the phone microphone and speaker or connected earphones. *WhisperChannel* also comes with a desktop admin interface, which allows to create users, their avatars in the meeting overview, and arrange avatars in the meeting. This process has to be done manually and for the user study we prepared the users and avatars for the participants.

We adapted the **Mirrorverse** [19] prototype to implement *WhisperChannel*. **Mirrorverse** is a platform for tailorable video conferencing software. We adapted the functionality of Mirrorverse by removing not required components such as most tools or content types, and used the audio routing to control which clients can hear which other clients depending on whether their avatar is tapped.

On a lower level, **Mirrorverse** builds on the Webstrates [28] and Varv [7] software stack for live reprogrammable, collaborative software. The audio signal is transferred using the WebRTC API. Building on this software stack and, in extension, web technologies, enabled us to run *WhisperChannel* on most modern off-the-shelf devices such as regular phones, tablets, or laptops, independently of their operating system.

6 USER STUDY

We conducted a qualitative user study with three groups with four participants each. The main goal of the study was to gain deep insights into how *WhisperChannel* can make backchannel communications in hybrid meetings more inclusive while keeping the required added effort low. To gain qualitative feedback from participants, we designed a simulated hybrid meeting situation in which participants used the system during two sessions and experienced different seating positions within the meeting.

6.1 Participants

Twelve people (P1–P12) participated in the user study and formed three groups. Three participants described their gender as female, nine as male. Seven participants were in the age bracket 20–30 years old and five in the age bracket 30–40 years old. Participants described their occupation as the following: Master’s student (1), Ph.D student (8), postdoc (1), assistant/associate/full professor (1), and research assistant (1). All participants had attended hybrid meetings before.

We recruited participants from our local university in Aarhus, Denmark, and authors invited them directly to participate in the study. Since we aimed to let participants engage more with our system, we selected participants who work in academia, so that groups would share similar backgrounds and academic knowledge.

6.2 Procedure

The study was conducted at our university in Aarhus, Denmark, and was spread over two rooms: the main room and the remote room (see Figure 3a). The main room was a large meeting room that we used for joint activities like the introduction. During the task sessions, three participants were co-located in the main room. Since we were also intrigued by how *WhisperChannel* would have an effect in a co-located setting, two participants in the main room were seated close to each other (P_{C1}, P_{C2}) and one far away (P_{R}). The remote room was a neighboring room to the main room that was used by one remote participant (P_{R}) during the task sessions. Participants with an uneven ID (P1, P3, etc.) experienced positions P_{C1} and P_{R}, and participants with an even ID (P2, P4, etc.) experienced P_{C2} and P_{R}.

Each participant was provided with a phone (Motorola moto g5 5G) and wired earphones. The phones were running *WhisperChannel*, which we set up for the participants (i.e., opening the prototype in the web browser and setting up the correct user for each phone), with their real name on the avatar Figure 3c. Each study lasted two hours and consisted of an introduction with a pre-questionnaire and focus group, a training session, two task sessions, and a concluding session with a post-questionnaire and focus group.

6.2.1 Introduction with Questionnaire and Focus Group (30 min).

Each study started with participants filling out a consent form and an introduction of the topic and domain of our research, including explanations of what hybrid meetings and backchannel communications are. After the introduction, participants were asked to fill out a questionnaire\(^1\) with demographic questions and questions about their past hybrid meeting and backchannel communication experiences, and challenges with either. Following the questionnaire, we conducted a short focus group interview to discuss these past experiences with participants.

6.2.2 System Introduction and Training Session (10 min). Next, we introduced *WhisperChannel* to the participants and explained its functionality and use. In a short training session, participants then had time to try out the system and use it while co-located in the main room. Participants were also meant to try whispering during this phase to familiarize themselves with the quality of the microphone and earphones of the system.

6.2.3 Task Sessions (50 min).

The task session of the study consisted of two sessions with identical tasks, each lasting 15 minutes. In between the two sessions we scheduled a short break where participants switched seats: P_{C1} switched with P_{R} and P_{C2} switched with P_{R}. This allowed each participant to experience the system as a participant with a direct seat neighbor and as a participant without. Participants was not mandatory, it was encouraged based on relevant research, as it is easier for people to talk with people they know rather than with strangers [25, 32, 37]. Thus, when forming groups, we tried to cluster participants who were already familiar with each other, in the same group.

\(^1\)WebRTC API: https://developer.mozilla.org/en-US/docs/Web/API/WebRTC_API (Retrieved: August 30, 2023)

\(^2\)The questionnaires and lists of questions of the focus groups can be found in the supplementary material of this article.
sitting alone (either further away in the main room or remote). In each session, we simulated a hybrid meeting with a speaker that gives a presentation in the main room. In order to guarantee the same attending experience for all three groups, instead of having an actual person as a speaker, we played a video of a presentation on a big screen in the main room (see Figure 3b) and streamed it to a laptop in the remote room using Zoom.

Our goal was that participants use the system as freely as possible. Hence, participants were tasked with using our system to talk to each of the other participants at least once during each session — even when sitting directly besides each other (positions PC1 and PC2). As for the content of the backchannel communication, participants could talk either about topics related to the presentation, e.g., How does this presentation relate to your research?, or small talk, e.g., Where shall we go for lunch?. However, at least one backchannel communication per session should relate to the presentation, so that participants do not just use the tool while ignoring the presentation. Apart from these requirements, participants were free to decide when and how they wanted to use the system.

6.2.4 Debriefing Questionnaire and Focus Group (30 min). After the participants were done with the two task sessions, they were asked to fill out another questionnaire with questions about their experience using WhisperChannel and how they would rate its effort compared to texting or whispering. Lastly, we conducted a semi-structured focus group interview in which we asked participants about their opinions focusing on the topics we provided, including inclusiveness and effort identified in a previous interview study, as well as an additional topic concerning future use.

6.3 Data Collection and Analysis
We collected the handwritten results of the two questionnaires and recorded the two focus group interviews in audio and video. During the task sessions, we recorded each participant with cameras and recorded their screens and backchannel communications using a computer that was connected to the WhisperChannel.

For the analysis, we first digitized the handwritten replies of the questionnaires and transcribed the recordings of the focus group interviews. Next, we examined how participants addressed the topics we provided them, and also employed affinity diagramming to identify and explore other points that participants discussed, which were relevant and interesting for our study. Utterances from participants were sorted into these guiding themes and afterwards clustered into sub-themes. The most interesting and relevant of these sub-themes form the content of the next section.

7 FINDINGS
We structure the findings around the two main challenges we addressed with WhisperChannel, inclusiveness (C1) and effort (C2), with an additional topic of future use. For each of these themes we present a number of sub-themes emerging from the responses from the focus groups.

7.1 Inclusiveness
The first theme relates to the first challenge of a lack of inclusiveness in backchannel communication in hybrid meetings (C1). The challenge surfaced in a tension between the increased awareness of backchannels that others are having, and privacy concerns about exposing one’s own backchannel communication. Further, while WhisperChannel increased inclusiveness for both remote and co-located participants, it still did not remove all barriers of inclusion.

7.1.1 Increased Backchannel Awareness Compromises Privacy. WhisperChannel increases the awareness of backchannel communications by changing the color of avatars that are currently in a backchannel to blue and connecting them with a line. Most participants found the blue color of avatars was useful to know which participants in a meeting were unavailable at the moment. They
saw it as an indicator like in applications such as Teams that shows whether a person is busy — or as P12 put it: “I think for me it was only important because that channel is blocked out, so I know that I cannot communicate with that person.” This was especially useful for remote participants who could not observe other co-located participants to see whether they were chatting. P10 also compared this to how they would not attempt to whisper to a co-located participant while they are already leaning over to another participant — a similar indicator of a person being in a backchannel conversation already.

The lines connecting avatars, however, were received with a mixed response. While some participants liked the increased awareness of the backchannel communication in the meeting, others disliked the feature as it compromised their privacy: “Usually when you’re whispering, you don’t want other people to know that you’re whispering” (P12). Especially for meeting scenarios in which participants do not know each other well, participants disliked the idea of sharing such information. As P10 remarked:

“It works fine for this setting, because I kind of like trust the people here. But say you were at a conference with unknown people, it would cross a border for me, definitely.” (P10)

There is a trade-off that has to be made when enabling more or less awareness and that may influence how people use the system. For instance, P4 voiced that they “would use the system less because everybody else can see that [they’re] talking to [someone].” This trade-off is both situational and personal, as participants would tolerate more awareness in meetings with familiar audiences, and as some participants liked the idea while it went too far for others. P3, for instance, found it useful to know who another participant was talking to, which is usually not possible when people are in text-based backchannel communication during a meeting.

### 7.1.2 Increasing Inclusiveness for Remote and Co-Located Participants

A key goal of WhisperChannel was to make backchannel communication more inclusive for both remote and co-located meeting participants. Indeed, our system had the effect of making communication more inclusive for both remote and co-located participants.

For co-located participants that sit further away (position P3), WhisperChannel supported them as well in being more included into the backchannel. For instance, P1 mentioned feeling more included into the backchannel of the two participants sitting close to each other (positions P\textsubscript{C1} and P\textsubscript{C2}) while sitting in position P\textsubscript{F}. Also P3 felt like it “reduced this distance” between the participant sitting further away of the other two co-located ones. Some participants enjoyed the spatial layout of the avatars, as “you had a sense of being co-present, even if you weren’t really” (P9) and it helped them to “quickly match the name and the seat” (P11) when talking to other participants they did not know.

Even P9, who did not experience the remote position, felt it was nice that “even though [the remote participant] is not here, I know that he’s still included”, when seeing remote people was talking with other participant from the backchannel dynamic of the interface. This hints towards co-located participants putting value into including remote participants even though it makes whispering for co-located participants more effortful than simply whispering.

### 7.2 Effort

The second theme focuses on the challenge of excessive effort required for articulation work due to technology (C2). We found that while participants agreed that the effort of our system being low, it still did not reach the same level of low effort as purely whispering. We also uncovered a trade-off particularly in its synchronous and voice-only features. E.g., while it reduced the effort of receiving quick responses, it simultaneously raised the effort needed to provide immediate responses. The “press and talk” simplicity of the system, also introduced the potential extra effort due to the absence of facial and gesture expressions in remote communication.

#### 7.2.1 System Interaction Is Not as Low-Effort as Whispering Yet

Since WhisperChannel users only needed to press the button to talk, in general, most of the participants agreed that WhisperChannel was easy to use, and “It’s a very intuitive interface as well. It was really easy to get started with” (P9).

However, participants revealed some additional effort with technology, making it less effortless than whispering. E.g., users had to visually ensure that they tapped the correct avatar and needed to check their channel status. Additionally, two participants sometimes tapped on the same avatar at the same time during a conversation, causing the slowest one to start talking while only noticing afterwards that the backchannel was not opened. The adjustability of volume also added extra effort for participants as they had to consider their speaking volume so that the other participant would understand them. Furthermore, technical constraints in form of a weak microphone required users to hold the microphone with one hand (see also subsection 8.4). Overall the extra effort of interaction could be summarized as what the participant addressed:

While WhisperChannel fulfills the goal of feeling included in the backchannels, it cannot resolve the “feeling of being remote” for P12. Yet, it is unclear if the latter is something that can be resolved at all with conventional hardware like phones and laptops. P2 also liked including remote participants “into the whole social aspect of being part of the presentation, when you can whisper to them also when they’re remote.”
You need to be aware of what’s going on, and you need to press the phone, and you need to make sure that it was actually heard. (P12)

7.2.2 Synchronous One-to-one Communication Has Benefits and Drawbacks. In WhisperChannel, backchannels are always one-to-one, allowing both participants to talk and listen simultaneously when the channel is open, leading to different perceptions of effort in synchronous and asynchronous systems in various contexts. On one hand, our system’s effort appeared low compared to asynchronous methods like instant messaging and voicemail, as users did not need to worry that other person noticed the conversation, which enabled quick answers. One participant preferred to use WhisperChannel over voicemail when joining remotely: “When you’re remote then I would prefer such a system, because then I don’t have to send WhatsApp voicemails” (P6). P7 also mentioned that the habit of muting the phone in meetings and ignoring the vibrations, made traditional backchannel communication like instant messages even more inaccessible:

The most difficult part is to get the notifications. I usually mute my phone during the lecture, and it’s difficult to get the notifications all the time when somebody is texting to me. Even though it’s vibrating, often I can ignore the vibration. So I have to really focus on the smartphone continuously to see the notification from the others. (P7)

P10 shared a previous experience of waiting for a reply for hours, and having the receiver of the message not understanding it, because it was referring to the context of the meeting. Therefore, participants believed the immediacy of synchronous communication in WhisperChannel was valuable in certain scenarios, such as when they needed assistance but did not want to gather too much attention.

On the other hand, immediate replies also increased effort, as the person had to reply even if the timing was not ideal. E.g., P2 remarked: “Whereas when we talk, then you kind of have to reply within, I don’t know, a second or so, otherwise it’s awkward.” One-to-one communication could lead to conflicts when someone called the moment a participant intended to speak to someone else. In addition, sometimes users had to wait until the person finished the communication. The absence of an accept button before initiating a conversation meant that some participants worried about whether someone else would suddenly connect with them or open their channel when they talked to themselves.

7.2.3 Voice-only Backchannels Decrease Effort. The communication of WhisperChannel is solely voice-based. Participants discussed the advantages of using voice over texting, for instance, that there is no record of the conversation, that it is low effort, and the ability for remote individuals to speak naturally. One of the participants who experienced the remote position, expressed WhisperChannel was preferable compared to texting:

It was better to talk instead of texting or something like that. Because I was faster and lower effort, for sure, because I don’t have to type the entire time. It was like natural speaking. (P6)

However, the lack of texting or video functionality resulted in varying perceptions of effort among participants. Many participants viewed WhisperChannel as a low-effort system: it eliminated the need for touching or eye contact before a conversation, only allowing for a simple press-and-talk interaction. Additionally during conversations, unlike whispering with occasional eye contact, and texting with focus solely on the screen, WhisperChannel only required checking the phone at the start of a conversation. Using WhisperChannel permitted users to sit still when whispering and did not require them to turn their heads or move around. P12 noticed that while using WhisperChannel they could even simultaneously watch the presentation: “If I’m speaking through WhisperChannel I can still look and listen, but I can still speak as well.”

7.2.4 A Lack of Facial Expressions Increases Effort. The absence of facial expressions and body language in WhisperChannel seemed to cause more effort, leading to difficulties in certain aspects: If people want to physically reject the backchannel communication in a physical meeting, they can lean away to show they do not want to talk at the moment. However users could not give or get a non-verbal reaction beforehand from WhisperChannel. P9 was leaning forward to indicate that “you are really paying attention” to the presentation after which P10 remarked that this way was “much easier for him to teach that it’s not a good time” compared to using WhisperChannel.

Without seeing the physical reaction, participants felt challenged to gauge if the listener was attentive or not, as P10 mentioned: “When I’m broadcasting to someone here, it’s hard to know whether they are paying attention.” Furthermore, as there were no cues like nodding or other gestures, and, e.g., one participant sitting in P3 responded to P9 by nodding before realizing that they should have answered “yes” instead.

When we talk to people face-to-face, comprehension becomes simpler as part of our understanding relies on observing a speaker’s mouth movements [56]. However, WhisperChannel posed challenges to remote communication. As P7 said: “When I was talking with the remote participant, I think I was more aware of some mistakes and misunderstandings.” In contrast to prolonged phone conversations with specific individuals, WhisperChannel caused participants to lose track of their conversation partners’ identities:

Usually verbal speaking is to somebody right away. Or either it’s to a phone or it’s to actually a person. And I feel like in this case I kind of lost the, at sometimes I kind of forgot who I was talking to because it kind of, you lost that physicality of the talking. (P4)

Moreover, some participants found that using the same modality for both the backchannel and frontchannel in voice-only backchannels was confusing, as P3 stated: “Audio is how this communication is happening in the main channel and you’re also using audio overlaid on that to do it [the backchannel].” Also other participants agreed that using audio modalities for both front- and backchannels could cause more distractions.

7.2.5 Social Acceptability Challenges Can Increase Effort. Participants found some interactions with WhisperChannel to be unconventional. Typically, one would lower ones voice when whispering, but “it’s kind of weird that you can just talk normally as the remote
person and that people like [imitates inaudible whispering]” (P4).

This became even more challenging when using earphones to communicate. For instance, some of the participants felt this action would attract the attention of the presenter as if they wanted to interrupt. And in some instances, whispering to someone who was not sitting closely but, e.g., behind the whisperer, was awkward. P2 felt it was unusual to be using WhisperChannel for whispering to someone behind them: “It was a bit weird to whisper to myself while he was behind me.”

We also found that the social acceptability of distraction could be related to different roles in a meeting. Participants pointed out that in situations when the lecture is dull, unimportant or when participants are merely a passive audience, using WhisperChannel to enhance whispering could be more acceptable. However, when in an important meeting, using WhisperChannel could be problematic, since “if you miss any information from this instruction, you are in trouble” (P7). Furthermore, in the case of a large lecture hall, participants would feel more at ease using the WhisperChannel, because the significant distance from the speaker could make the use less disruptive. However, participants mentioned that when sitting close to the presenter they would feel less inclined to use WhisperChannel, as potential distractions become a more significant consideration.

7.3 Future Use Cases

As part of the concluding focus group of the user study sessions, we also asked participants about their thoughts on which types of meetings a system like WhisperChannel could be useful for.

Most participants preferred using WhisperChannel as a remote participant because they could speak normally without the need to whisper. This naturally led them to the use case of fully remote meetings. In these, participants would not interrupt the speaker when talking and hence it would feel less uncomfortable to use the system (see also subsubsection 7.2.5). P5 talked about including the feature into Zoom and P2 mentioned how it could be useful in breaks between presentations: “Instead of doing the breakout room thing in Zoom, you just hold down and then ‘Hey, how is it going?’”

Another promising scenario is using WhisperChannel in groups of friends or social groups in general. P5 stated that such a tool could be useful for “stable groups of people” such as groups of students that go through classes at university together. And P12 mentioned an experience of visiting a conference with colleagues and found it difficult to stay in touch in the breaks or lunch time between sessions. For such situations, WhisperChannel would allow immediate communication opposed to, e.g., following WhatsApp group chats.

Lastly, participants identified remote meetings with multiple frontchannels as a promising use case for systems like WhisperChannel. P3 elaborated that it would be useful “for a completely remote setting where the focus of this is one-on-one communication,” such as “networking” or “speed dating type setting.” Participants compared with apps like Gather⁴ and suggested that an addition of a video channel would be useful in such a scenario where participants do not know each other. P11, further, described design sprint group meetings as a possible use case, because these include people with “different responsibilities” that need deliverables from different colleagues, where a system like WhisperChannel could allow them to talk to each other in parallel.

8 DISCUSSION

In this section, we discuss the relationship between the effort required by co-located participants when using a backchannel communication system, and the ways in which remote participants might feel included. Further, we discuss the possible negative impact of using WhisperChannel for people who sit next to one another, and whether we should even make backchannel communication easier in the first place. Lastly, we unfold the limitations of our user study, and technical limitations of the system, followed by an outline of directions for future research.

8.1 The Relationship Between Inclusiveness and Effort

Our research revealed that WhisperChannel holds value in enhancing the sense of inclusiveness for remote participants. Yet, the extra effort required by the backchannel technology — whether in WhisperChannel or in other tools — is undeniable. Our overall aim was to reduce the amount of effort required to make remote participants feel more included. The results of this research are promising: As presented in the findings, remote participants felt more included with participants in the co-located space. Some participants attributed this to the low effort required for voice communication compared to texting, others expressed a preference for WhisperChannel over voicemail because of its synchronous features.

More importantly, co-located participants who observed that the remote participant was having backchannel conversations felt positive about seeing them included. This, however, raises the following questions: To what extent are co-located participants willing to invest effort in systems like WhisperChannel to make backchannel communication more inclusive for remote participants? Where is the threshold and what elements contribute to it?

Currently, these questions need further investigation. However, our studies indicate that the answers may diverge significantly depending on individuals, their roles in different situations, the social acceptability in a specific context, their familiarity with the remote people, and the significance of the backchannel content. These factors present interesting opportunities for future research.

8.2 Close Together, Yet Far Apart?

During task sessions, we recommended that participants sitting together (P_C1 and P_C2) would use WhisperChannel for backchannel communication, even when traditional whispering was possible. Even though the participants did not whisper (in the traditional sense), some still used other communication mechanisms when co-located, such as shared glances and looking at each other to see reactions. However, some participants continued to focus on the presentation while using WhisperChannel, appearing to talk to someone in a different location while actually talking to the person next to them.

With whispering no longer being limited to participants in close proximity, we showed that WhisperChannel can reduce the perceived distance between participants sitting further apart. But could it also increase the sense of distance among those sitting closely

⁴Gather: https://www.gather.town/ (Retrieved: August 30, 2023)
together? Is the intimate [32] aspect of whispering then lost? Does this have an impact on the sense of inclusiveness? These questions provide exciting new directions for future research.

8.3 Should We Make Backchannel Communication Easier?
Some of our participants disliked the idea of making backchannel communication easier as they perceive it be undesirable behavior during meetings. For example, they mentioned that sometimes people are not aware of how loud they are when whispering, and lengthy text-based chats require visual attention, which can be perceived as bothersome from the presenter’s perspective. They do have a point since backchannel communication can indeed introduce distractions and be perceived as impolite [3, 32]. However, we should not completely disregard its value: both related research and the experience of some of our participants confirm the positive impact of backchannel communication in meetings. From fostering equal relationships [27, 39] to providing social support [2] and enhancing presence [29].

Furthermore, WhisperChannel is not a definitive solution to address these problems, rather, it is an exploration of an inclusive backchannel communication platform focusing on making backchannel communication easier to use. We do not propose that participants should increase their amount of backchannel communication during a meeting, and it is an open question whether WhisperChannel would increase the frequency of backchannel communication, and whether this is a good thing or not. More research is needed to provide answers to these questions.

8.4 Limitations and Directions for Future Work
We discuss limitations of, on the one hand, our study design and pool of participants, and, on the other hand, technical limitations related to the software and hardware of our WhisperChannel prototype setup.

8.4.1 Study Design Limitations. In our study, participants mentioned that their actions or thoughts might have evolved differently if the study was conducted over a longer time. It is rather obvious that the study setup neither captures participants’ past experiences and routines of collaborating, nor does it help address the learning that happens in longer-term use and the routines developed among participants. This would require other methods and also a different selection of participants, e.g., studies closer to the work of Larsen-Ledet et al. [30] or Griggio et al. [17]. Our current work paves the way for future lab studies, which could involve incorporating a control group without access to the WhisperChannel. Additionally, conducting fieldwork studies with one or two conditions (e.g., with and without the WhisperChannel) during regular meetings over a longer period of time could be considered for future work.

Despite our efforts to group participants who already knew each other, it should be noted that some of the participants in each group did not know each other and had no experience of working together. We asked them to imagine they knew each other and were willing to have conversations with each other in the backchannel. For future use, WhisperChannel needs to be explored in real-life situations where people have worked together before or currently are, in order to understand how and when backchannel communication would be more natural, hence leading to other interesting findings.

Considering that hybrid meeting dynamics are influenced by cultural norms and perspectives [44], it is worth noting that the context of this study, being situated in Denmark, may have influenced certain outcomes of backchannel communication. Also people from different parts of the world may have varying views on whispering. However, in our study, we did not include nationality as a demographic question or into our analysis. An interesting direction for further investigation is how nationality and culture could potentially impact the outcomes of backchannel communication.

Additionally, in real-life situations there are numerous types of hybrid meetings — each with its unique characteristics and dynamics. However, in our study, we focused exclusively on testing one specific type of hybrid meeting, where one remote participant was included in each task session. In a real-world case there could be multiple remote people joining the hybrid meeting together. Furthermore, we only set up one arrangement of co-located individuals, with P_{C1} and P_{C2} positioned in front of P_{F}. Different arrangements, such as the opposite configuration or participants sitting in the same line among others, could potentially lead to different findings, but would also require methods and analyses that focused in greater detail on, e.g., turn-taking [21, 41].

Furthermore, we found that during the second focus group discussion, the topic of social acceptability came up (see subsubsection 7.2.5). Participants were interested in the social factors that made them comfortable using WhisperChannel, and other social factors that discouraged them from doing so. Also some participants discussed that, from the speaker’s point of view, there would be different levels of acceptability of subtle or less subtle backchannel communication. For instance, seeing active backchannel communication in the audience might indicate to the speaker and participants that something would need to be addressed, however, if very obvious, the activity could also become disruptive. Thus for future studies, it is relevant to add social acceptability as a specific topic, as seen from the perspectives of both the audience and the speaker. Lastly, the technical limitations have the potential to impact the data collection of the study and the overall experience of the users.

8.4.2 Technical Limitations. The suboptimal quality of the earphones and network delay of the audio impacted participants’ experiences. Currently, with our existing earphone setup, it is challenging to both speak clearly and quietly, and the network delay caused participants sitting close to each other to hear messages twice. Furthermore, the need for participants to hold the microphone with one hand was not an ideal design. In the future, we plan to address these issues by upgrading to higher-quality wireless earphones that can capture and transmit whispers more effectively. However, it is important to note that the network delay remains a challenge that may not be easily resolved at this time.

Based on the feedback we have received regarding the usage of WhisperChannel, there are several design suggestions that should be considered:

- **Do-not-disturb Feature or Whitelisting**: Some participants preferred the option to set a “Do-not-disturb” mode or whitelist for specific individuals who can communicate with them.
Others outside the whitelist would need to send a request or “knock on the door” before initiating communication.

- Notification Cues: Users suggested adding notification sounds or vibrations to alert them when a backchannel is opened or closed. This would reduce the need to constantly check the screen for updates.

- Simplified Interaction: To minimize the effort required to engage in backchannel communication, the hold-down interaction could change to a simple toggle for activating and deactivating a backchannel.

- Voice-to-text Functionality: To improve comprehension, participants have requested the addition of voice-to-text functionality, which would convert spoken messages into text.

- Broadcast Feature: Some users expressed a desire to communicate with multiple people simultaneously. To address this, adding a “Broadcast” button that allows users to send messages to multiple recipients at once could be helpful.

- Profile Pictures: Adding profile pictures to avatars is suggested to enhance networking capabilities and make it easier for users to identify one another.

These design suggestions will be taken into account in our further explorative development of WhisperChannel. As face-to-face communication often involves multiple modalities like spoken words and non-verbal cues, hybrid meetings also present a wide range of options for backchannel interaction. Stepping back from the specific adjustments, future work should also consider the exploration of not only speech, but also alternative modalities. For example, voice-to-text (and vice versa) functionality, gestures, touch, facial expressions and eye gaze, etc., could be considered to explore and provide insights into the preferred and most expressive modalities in various situations.

9 CONCLUSION

Our empirical analysis of the interview study presented in this paper points to challenges in hybrid meetings, with participants feeling excluded from backchannel communication. This is true for both remote and (to some extent) co-located participants, as the meeting technology often offers asymmetric or unequal possibilities. An additional challenge is the amount of articulation work required, in terms of setting up and carrying out backchannel communication. This extra effort also seems to contribute to the feeling of exclusion.

In order to address these backchannel communication challenges, we developed an exploratory, voice-based backchannel communication system WhisperChannel aiming to promote inclusiveness and low effort through whispering. Studying the use of WhisperChannel in a setting with co-located and remote participants, we demonstrated the potential for voice-based backchannel communication. Our results show that WhisperChannel made both remote participants, and those co-located participants seated at a distance from each other, more involved in the hybrid meeting. While generally WhisperChannel requires little effort to be used, our study points to some issues when it comes to having to reply immediately to synchronous communication; a lack of facial expressions while using the system; and social acceptability concerns that can increase effort.

Future work includes conducting a longer-term study where participants use the system during their actual meetings, as well as iterating on the design of WhisperChannel to overcome its technical limitations. Lastly, the interrelationship between inclusiveness and effort, and possible negative impacts of systems like WhisperChannel for co-located participants, warrants further research.

ACKNOWLEDGMENTS

This work is supported by the Innovation Fund Denmark for the project DIREC (9142-00001B), the Aarhus University Research Foundation (AUFF-E-2022-9-33), and the Independent Research Fund Denmark (DFF) (10.46540/3104-000080).

REFERENCES


