

The Balancing Act of Social Audio Facilitators: When Self-Promotion Overshadows Community Care

NAZANIN SABRI, University of California, San Diego, USA

YUTONG CHEN, Carnegie Mellon University, USA

MARISSA LEE, University of California, San Diego, USA

STEVEN P. DOW, University of California, San Diego, USA

KRISTEN VACCARO, University of California, San Diego, USA

MAI ELSHERIEF, Northeastern University, USA

Voice-based social media platforms that enable attendees to have real-time, ephemeral interactions with each other—such as X-Spaces, Discord, and Clubhouse—have seen considerable growth in recent years. While prior research on these spaces has predominantly focused on moderating harms, our work seeks to understand emergent practices employed by hosts to proactively shape their discussion space—focusing on the *facilitation* aspect of moderation duties. Drawing on facilitation strategies, we study these practices through three comprehensive studies using mixed-methods: survey of social-audio users, co-design interviews, and analyzing training sessions for hosts. Our findings reveal insights into the issues faced by hosts and attendees, current facilitation practices, opinions on technological solutions, and factors that could be responsible for some of the identified issues such as the available training for hosts. We found that hosts themselves are often significant sources of issues due to practices such as focusing more on self-promotion than facilitating discussions. In addition, host training sessions seem to encourage behaviors that contribute to the negative perception of hosts. We draw on outcomes from co-design interviews to guide the design of future tools to support hosts in facilitating social-audio spaces. Our findings provide insights that could help create a more positive experience for both hosts and attendees.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **HCI design and evaluation methods**; • **Information systems** → **Social networks**.

Additional Key Words and Phrases: Social Audio, Ephemeral Social Media, Content Moderation

ACM Reference Format:

Nazanin Sabri, Yutong Chen, Marissa Lee, Steven P. Dow, Kristen Vaccaro, and Mai ElSherief. 2025. The Balancing Act of Social Audio Facilitators: When Self-Promotion Overshadows Community Care. *Proc. ACM Hum.-Comput. Interact.* 9, 2, Article CSCW083 (April 2025), 34 pages. <https://doi.org/10.1145/3710981>

1 Introduction

In recent years there has been a rise in platforms dedicated to supporting voice-based discussions that allow people to directly speak with each other online (e.g., X-Spaces, Discord, Clubhouse). The real-time nature of voice-based platforms and the ephemerality of interactions can introduce complications to existing moderation pipelines, requiring new strategies and practices [56]. Prior work has identified a number of moderation challenges introduced by real-time voice, such as: (i)

Authors' Contact Information: [Nazanin Sabri](mailto:nsabri@ucsd.edu), nsabri@ucsd.edu, University of California, San Diego, USA; [Yutong Chen](mailto:yutongc2@alumni.cmu.edu), yutongc2@alumni.cmu.edu, Carnegie Mellon University, USA; [Marissa Lee](mailto:mal039@ucsd.edu), mal039@ucsd.edu, University of California, San Diego, USA; [Steven P. Dow](mailto:spdown@ucsd.edu), spdown@ucsd.edu, University of California, San Diego, USA; [Kristen Vaccaro](mailto:kvaccaro@ucsd.edu), kvaccaro@ucsd.edu, University of California, San Diego, USA; [Mai ElSherief](mailto:m.elsherif@northeastern.edu), m.elsherif@northeastern.edu, Northeastern University, USA.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2025 Copyright held by the owner/author(s).

ACM 2573-0142/2025/4-ARTCSCW083

<https://doi.org/10.1145/3710981>

new norms (e.g., avoid interruptions), (ii) new ways to break existing rules, (iii) lack of evidence, and (iv) lack of moderation capabilities typically available in text-based communities (e.g., organizing or editing content) [56].

While addressing harms is of great importance, recent studies of moderation have advocated for proactive practices [63, 94, 103, 110], instead of reactive strategies. Moderators of various social media platforms (e.g., Twitch, Facebook, Reddit) have expressed the importance of pro-active measures [98, 99]. These type of strategies become particularly important on social-audio platforms, as conversations occur in real-time between transient strangers and cannot be easily edited or retracted, requiring moderators to actively ensure the quality of discussions as they happen. While prior work has identified a number of proactive strategies that fall under the umbrella of facilitation, we bring together these aspects and examine their use in ephemeral voice-based interactions.

The quality of a discussion is often dependent on how the host manages the session. In various settings, employing facilitation best practices have been shown to improve the atmosphere of an event and discussion dynamics [7]. For participants to have a positive experience, hosts need to manage a variety of tasks, including giving attendees a fair opportunity to participate, managing the agenda, and preventing conflicts from getting out of hand, among others [7]. These activities have been studied in the facilitation literature to enable productive group interactions in various settings (e.g., teaching, work meetings, group discussions) [7, 13, 82, 120]. However, less is known about the use of facilitation strategies on social-audio platforms. The online social context may create new challenges: for instance, unlike classrooms or work meetings, social-audio discussions are often held between strangers who may or may not return for future discussions. Additionally, participation is pseudonymous, which may lead individuals to feel less accountable for their statements. Such differences motivated us to conduct a dedicated investigation of the current state of facilitation on social audio platforms. In this study, we explore the following research questions:

- RQ1 : What issues do users report having with the current state of facilitation in social-audio spaces?
- RQ2 : How do audience members and hosts feel about current practices and potential technological solutions?
- RQ3 : What are hosts in social-audio spaces being taught, and how do these teachings contribute to the current issues?

We investigate these questions in three studies. In Study 1, we aimed to prioritize the broad range of facilitation issues that users reported on social-audio platforms. To investigate users' opinions of the difficulty of each facilitation task, we surveyed 337 users. Surveyed users (which included listeners, speakers, and hosts) were asked to identify the facilitation tasks where they had observed the most problems, helping us focus attention on the most negative experiences. We found that hosts need the most support for: (1) encouraging participation, (2) helping users become comfortable using the platform, and (3) helping users resolve conflicts.

With an understanding of key challenges with facilitation in social-audio interactions, we then conducted co-design interviews in Study 2. We recruited 14 individuals who had used social-audio platforms in some capacity (either listener, speaker, or host) to view and react to storyboards [23]. One of our goals in Study 2 was to understand the existing practices that led to the tensions and dissatisfaction observed in our survey responses. Additionally, we aimed to use design-based methods to develop ideas for future tools for social-audio spaces, with a special focus on the high-priority challenges identified in Study 1. These interviews allowed us to better understand the issues that occur during online voice-based discussions, as well as current facilitation practices. Our results indicated that there is little consensus on the objectives of or management style for these online audio discussions. Despite these disagreements, we collected design suggestions and feedback on

tools that could help facilitate sessions. These design recommendations could help solve the myriad of facilitation problems currently experienced on these platforms.

Finally, we explored to what extent the host training materials may or may not have contributed to these issues on voice-based social media platforms. In Study 3, we analyzed a corpus of 16 recorded training sessions using a qualitative coding approach. From the training material, we extracted lessons that directly contributed to identified issues, those that were unique to social-audio platforms, as well as lessons that conflicted with facilitation practices. Our results indicate that hosts are being encouraged to exhibit a sense of ownership over sessions, prioritizing control over fostering collaborative discussions. Hosts were also taught to manipulate specific features of the platform to achieve higher attendee and speaker counts.

Through these three studies, we identified facilitation practices and challenges, explored potential technological solutions for some of these challenges, and investigated how the available training material could be contributing to the existence of observed issues. We conclude this work by providing design recommendations for improved experiences for both hosts and attendees.

2 Related Work

This project connects two areas of related work. We first discuss prior work on content moderation, before reviewing the literature on facilitation, discussing facilitation aspects and roles. In doing so, we explore how voice-based social media spaces affect existing practices in each space.

2.1 Moderation: A Move From Reactive Strategies to Facilitation

With over 4.9 billion people using social media platforms worldwide [8], these platforms and their governance strategies have garnered significant research [37–39, 69, 77] and public interest [3, 84, 86, 111]. Researchers in this area have studied moderation from various perspectives including the effectiveness of moderation on future behavior [55, 105], the dangers moderation poses to moderators [4, 88], the inequities in moderation decisions particularly when it comes to marginalized communities [26, 27, 44], as well as contestability and appeal strategies and user’s views of them [114, 115].

A number of definitions have been offered for content moderation throughout the years. One common definition of moderation views it as “the organized practice of screening user-generated content (UGC) posted to Internet sites, social media, and other online outlets” [87]. This definition primarily frames moderation as an issue to be addressed reactively. While a portion of the body of work on content moderation focuses on reactive strategies, recent work has discussed and advocated for more proactive strategies when it comes to moderation [90, 99]. One definition of content moderation that is more aligned with this shift is as “governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse” [42]. Both aforementioned definitions approach moderation from a *platform and policy* perspective; however, a large body of work has adopted a *community* perspective on moderation [33, 67, 97, 117]. These studies focus on mechanisms of self-moderation through volunteer moderation. One definition inline with this perspective states that volunteer moderators resemble “community leaders” and “are responsible for building communities, often from their inception, and guiding them toward positive cultures of social interaction” [97]. Our work follows this line of research, focusing on proactive practices of volunteer moderators on social-audio platforms.

A number of studies have investigated proactive moderation strategies [95, 99–101]. Closely related to our focus on the facilitation aspect of content moderation is Seering et. al.’s study of how volunteer moderators envision their roles [99]. This work found that moderators used the metaphors of “facilitator” and “mediator” to describe their roles [99]. Mediators attempted to resolve disputes, stepping in when discussions became heated. The authors also report that facilitators

“described themselves as relaxed about moderation, frequently referencing the effectiveness of community self-regulation” [99]. Another study of proactive moderation investigated effects of norm-setting in Twitch chat messages, showing how users imitated behaviors, with some users (i.e., those with authority) having more influence on others’ behaviors [100]. Screening users prior to joining, and monitoring conversations for potential conflict have also been investigated in the context of Reddit, Facebook, and Twitch moderation as proactive measures [101]. Prior work has also identified major challenges community moderators face, including encouraging contributions and dealing with new-comers [62, 97]. Some of the other proactive strategies that have been investigated in the contexts of post-based platforms, such as Reddit, include: maintaining values of community members [29], managing new comers through guiding or limiting strategies (e.g., not allowing a user to post if they are a new-comer or pairing them with an experienced user) [54, 76], responding to questions [36], and engaging with comment threads [72]. Prior work has also investigated similar strategies on other ephemeral content such as Twitch streams [14, 121].

While strategies such as community self-regulation, norm-setting, encouraging contributions, and helping resolve conflicts are all facilitation strategies, in this work we expand on the existing literature by bringing together these various strategies from the moderation and facilitation literature and investigating multiple aspects of facilitation in the context of social-audio content moderation. This focus on facilitation is also necessitated by the growth of ephemeral social media spaces, where content disappears immediately or shortly after it is shared, such as social-audio platforms or virtual reality. As real-time conflict management and resolution are much more challenging in these spaces, many moderators in ephemeral spaces are attempting to implement facilitation strategies to prevent harms rather than having to address them once they happen [90].

Another factor that further complicates content moderation on these ephemeral spaces is the multiple additional roles moderators have to take on. Moderators in ephemeral spaces such as voice-based platforms have to perform as hosts and content creators. This shared responsibility was rarely the case in post-based spaces such as Reddit or YouTube video uploads. Even though a YouTube content creator might check the comments left on their videos after posting, the mix of roles (i.e., content creation and moderation) often did not occur at the same time. With live-streaming services such as YouTube live or Twitch these simultaneous roles are becoming more common. Social audio platforms follow this trend with moderators needing to manage multiple responsibilities of moderation and content creation at the same time. The desire to create content that is captivating while ensuring it adheres to the guidelines could create a conflict of interest. It is worth noting that while moderators on platforms such as Twitch are often individuals other than the streamer, prior work has shown that the conflict between the viewers and streamers needs is an issue Twitch moderators face [121]. As a result of these complications, assisting moderators in implementing facilitation practices could be beneficial in voice-based spaces.

2.2 Facilitation: Roles and Strategies

Facilitation is a set of “technique[s] by which one person makes things easier for others” [47] and facilitators are identified as “people who make things easier, help others towards achieving particular goals, encourage others, and promote action” [65]. The term facilitation is used in numerous contexts including physiological [12], educational [40, 48], counseling/psychotherapy [71], social, and theological contexts [20], with slightly different definitions in each setting. Two of the closest settings to voice-based social media platforms are classroom and workplace discussions.

When it comes to the facilitation of online classroom discussions, existing literature has identified various roles facilitators take on [49, 73], one categorization fits these roles into four groups: (1) intellectual, (2) social, (3) managerial, and (4) technical [9]. The intellectual role helps “group members achieve predetermined learning objectives, such as understanding critical concepts or

building subject knowledge" [9, 41, 119]. The social role of facilitators is meant to "create and maintain a friendly, interactive environment in which participants feel safe and comfortable to interact with one another" [2]. In their *managerial* role, facilitators "aim at setting the agenda (e.g., timeline, norms, and rules) so that online discussions can go on smoothly" [119]. Finally, the technical role "help[s] group members get familiar and become comfortable with the discussion programs so that they can participate without technical difficulties" [119].

Organizations have also been turning to active-facilitation approaches in the workplace to help improve performance [6]. In active facilitation, the facilitator "has a strong attachment to the group and its task" but takes on an "unbiased service role" [6]. Additionally, they "do not advocate one particular process, but advocate whatever process the group has agreed to use" [6]. The role of facilitators in this approach is often broken down based on the meeting timeline, with strategies being discussed for before, during, and after the meeting [1].

Several techniques have been developed to help facilitators perform their roles. The International Association of Facilitators and SessionLab, for instance, have compiled over 1000 techniques for facilitation [50]. Even though these techniques are useful, they often focus on the use of specific tools (e.g., whiteboards, post-it notes). Additionally, multiple techniques are often offered to resolve the same issue. Due to the large number of techniques and their overlap, we focus on higher-level roles rather than specific techniques for our analysis.

3 An Introduction to Social Audio Applications

In this section, we will provide a brief overview of social audio applications and their affordances. We will focus on the social audio applications of (a) Clubhouse, (b) X-Spaces, and (c) Discord audio. All three applications enable users to hold real-time audio-only conversations. Live audio events on Clubhouse are referred to as "rooms" by the application's users. We use the terminology of *room*, *session*, and *event* interchangeably in this work to refer to live audio-only discussions. Events could be public or private—we focus on public events for our studies.

Users within a Clubhouse event are split into two groups with regards to speaking privileges: those with speaking privileges (i.e., users who are "on the stage") and those who are not able to speak (i.e., users who are "in the audience"). Once a user has been moved to the stage, they can begin or stop speaking (i.e., mute or unmute themselves) at any time during the conversation. Both groups of users can send messages in the public chat dedicated to the event and use emoji reactions which are displayed on the side of their avatars for a limited amount of time. X-Spaces offers similar functionality with minor differences. For instance, only up to 13 people can be on stage at any point in X-Spaces¹, while Clubhouse allows for much larger online gatherings [122].² Another difference is that X-Spaces doesn't have a dedicated "chat" section and text-based messages for the event appear as new X posts. Clubhouse and X interfaces are displayed in Figure 1.

Discord, on the other hand, has a unique characteristic. While Clubhouse and X-Spaces automatically assign "host" duties to the creator of every event, Discord audio spaces don't require a host to be present for an audio event to begin. This difference means that audio conversations on X-Spaces and Clubhouse are time-limited, beginning and ending when the hosting team is available. Discord conversations, on the other hand, can be joined at any time without the need for a member of the hosting team to be present, making them continuously open audio "channels" rather than one-time "events". However, it is worth noting that while hosts do not have to be present, a large number of

¹This limit is only placed on the number of speakers in the room, and not on the number of people who can attend the event.

²We were not able to find an exact number for Clubhouse. However, members of our team have seen events on the application with hundreds of users "on stage".

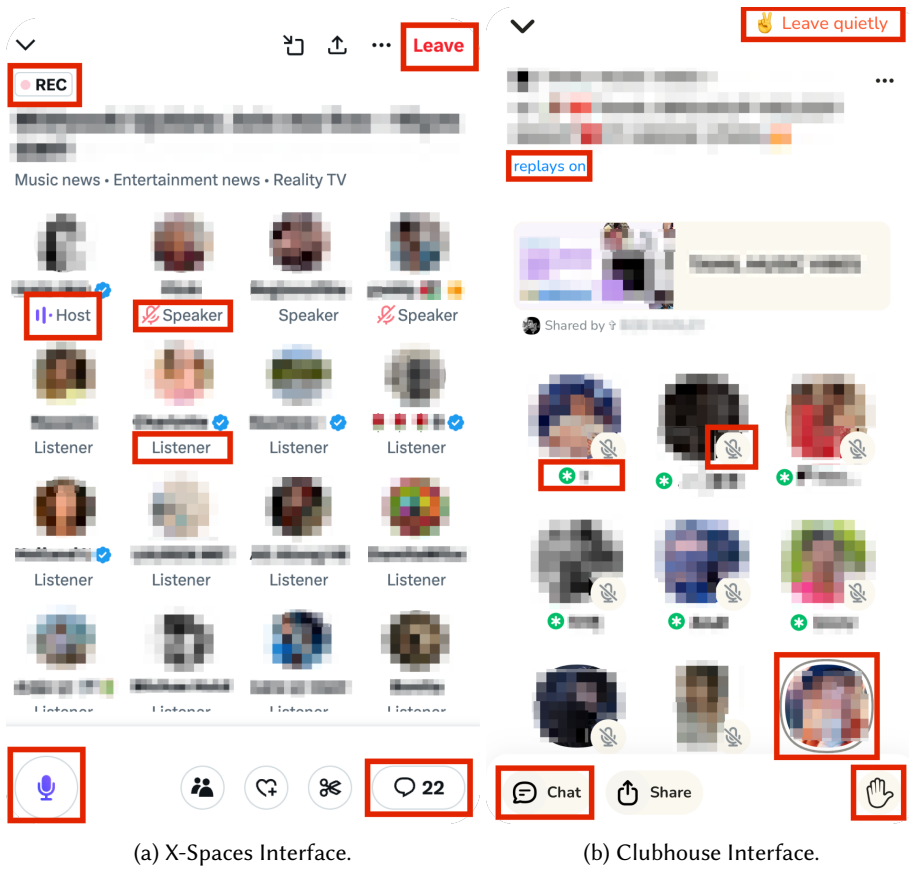


Fig. 1. Screenshots of events in Clubhouse and X-spaces. Usernames, images, and event titles have been blurred to preserve the privacy of the users. In both interfaces, the event title is shown on top of the room. Users hosting or attending the event are shown using their images with their names written underneath the circular representation. The green icons next to the user's name on Clubhouse indicate that the user is one of the hosts. X-Spaces, on the other hand, uses text descriptions to indicate each user's role. Both platforms allow users to send chat messages and request to speak (through the hand-raise icon in Clubhouse and the microphone icon in X-Spaces). Hosts can also choose to record their session in which case attendees will be notified through a "Rec." or "replays on" indicator shown on top of the page. The current speaker is also pointed out on both platforms, either through a gray circle around the user's avatar on Clubhouse (e.g., the avatar on the third row on the right) or through wave signs next to the user's name on X-Spaces (e.g., the avatar on the first row on the left).

hosts (i.e., server owners) schedule specific audio events they plan to hold and are present for the duration of these scheduled events.

X-Spaces assigns a "host" badge to the user who has started a session. Clubhouse refers to the creator of a space with the terminology of "moderator". However, in both cases, their tasks are more inline with those of a host and facilitator, rather than a moderator. Namely, these users are responsible for planning the event, setting and re-iterating the rules and norms, managing the progress of the event, and responding if or when conflicts arise, among other things. All these tasks are in line with facilitation duties that have been discussed in prior work [7]. We will use the terminology of "host" in the remainder of this work to refer to users who conduct such facilitation

tasks. To perform these tasks hosts are provided with several capabilities within the social-audio application, such as (i) the ability to mute or remove users, (ii) the ability to select speakers, and (iii) the ability to remove chat messages. The host who created an event, can appoint co-hosts who will then have the same capabilities (including the ability to appoint additional co-hosts).

Having discussed the differences and similarities of these platforms, we believe that they all share several fundamental characteristics concerning the host's functions that justify investigating our research questions within the context of all three platforms rather than limiting our work to a single one. We are interested in how hosts perform their tasks and interact with the group of people who have come together for an audio event. Since these platforms create similar dynamics between the hosts and other users within the event, investigating our research questions within the context of them all is meaningful.

4 Study 1: Surveying How Users Prioritize Facilitation Challenges

Prior work on social-audio applications (i.e., Discord Audio) [56] has highlighted challenges with the current state of moderation on these platforms. These challenges were identified based on the host's recollections of specific negative events and with a focus on harm. However, hosts within social-audio events are expected to perform a variety of other tasks, such as opening the session, engaging participants, and conflict resolution, which go beyond the management of harm. These types of responsibilities fit into the scope of tasks the facilitation literature explores [7]. As a result, in this study, we conducted a survey in which users were presented with a high-level overview of aspects of facilitation and asked to indicate areas where they believed current practices were inadequate. We believe the outcomes of this survey provide more clear insights into the facilitation challenges within social audio environments.

4.1 Method

4.1.1 Survey Design. As the goal of our survey was to identify aspects of facilitation where issues arise, the main question asked participants to select up to *three* areas of facilitation from a provided list where they had observed hosts struggling. The list of issues was compiled by combining the facilitation responsibilities presented in prior literature [6, 119]. Wang focuses on facilitation in online discussions[119]. While the paper studies educational settings, it focuses on online sessions which is closely related to voice-based social media platforms. Another resource we used was a handbook released by the Management Consulting Associates focusing on facilitation in workplace settings [6]. While this setting is further from social audio, they offer a different perspective on facilitation roles by separating the tasks based on the timeline of the session rather than roles offered in the previous paper [6]³. Combining both perspectives offers a more holistic view of facilitation responsibilities.

The combination of these responsibilities created a list of 17 aspects that were presented to the users in our survey. These aspects are displayed in Table 1. To ensure that the first few options were not disproportionately chosen due to their prominent positioning, each respondent was presented with the options in a randomized order. Respondents were shown only the issues, not the higher-level categories displayed in the table. It is also worth noting that the displayed "Issue Categories" are just one way of grouping facilitation tasks, and other prior work have offered different ways of categorizing facilitation functions, such as the five clusters introduced by Den Hengst and Adkins [25]. Other questions on the survey queried the demographics of respondents.

4.1.2 Participant Recruitment. A short description of the study, the compensation information, and a link to the survey were posted in the following locations to recruit participants: (1) on four

³These facilitation roles are discussed in our related work section (see Section 2.2), and displayed in Table 1.

subreddits, the topics of which were related to voice-based social media platforms (i.e., r/Clubhouse-guide, r/ClubhouseInvites, r/ClubhouseApp, r/twitterspaces), (2) as public chat messages in rooms that were being held in Clubhouse, (3) as private direct messages to users on X who tweeted about X-Spaces or private messages to users on Clubhouse who were active during the recruitment period, and (4) advertisement through X posts by two members of our research team. Users were recruited through various channels to capture the opinions of individuals who use a variety of social audio platforms such as Clubhouse, X-spaces, and Discord (Audio). While each platform could have distinct norms, the facilitation tasks included in the survey are broad tasks identified by prior work to be useful in group settings [6, 119].

We did not restrict participation to specific roles (e.g., only hosts). As a result, hosts, listeners, and speakers could all respond to the survey. While each role has different capabilities, including multiple perspectives is valuable because it ensures a comprehensive understanding of the platform's dynamics and user experience. Hosts can offer insights into the challenges of managing discussions, speakers can share their experiences in engaging with the audience, and listeners can highlight how the interactions impact their overall experience. This holistic approach allows for a better understanding of challenges on social-audio platforms. Including multiple stakeholders in the need-finding and design process is a fundamental aspect of Human-Computer Interaction (HCI) to ensure the development of user-centered and effective solutions [18, 34, 92]. The same strategy with regards to the inclusion of multiple stakeholders and platforms is applied to all studies within this work.

Our method of compensation was to provide five 20\$ gift cards to five individuals who completed the survey at random. This study was reviewed by our institution's IRB.

4.1.3 Validating Survey Responses. We received a total of 1,643 responses to our survey. To validate collected data, first any responses submitted from duplicate IP addresses were removed (i.e., only unique IP addresses were kept). This step reduced the response count to 404. Next, duplicate email addresses were removed, leaving 366 responses. We then removed any response that was completed in less than one minute, resulting in 358 remaining responses. Lastly, any responses from outside of the United States were removed (due to IRB limitations) resulting in a total of 337 responses. After the aforementioned cleaning steps, we find that it took users an average of 4.6 minutes to fill in the survey with a minimum of 1.03 minutes and a maximum of 39.9 minutes.

4.1.4 Demographics of Respondents. The majority (N=179, 53%) of respondents were between the ages of 25 and 34, followed by 18-24 year olds (N=64, 19%) and 35-44 year olds (N=63, 19%). Additionally, 82% were White, followed by 9% Black/African American respondents. Male users submitted 51% of responses. Through a multiple-select question, 224 (66%) respondents indicated using Clubhouse, 275 (82%) mentioned using X Spaces, and 199 (59%) selected Discord (Voice)⁴.

4.2 Findings

In this section, we discuss the results of our survey on the aspects of facilitation in need of improvement. Table 1 displays the number of votes in our survey for each aspect of facilitation hosts struggle with. We found the following four issues to be reported most frequently: (1) encouraging participation or inviting responses (N=101), (2) helping group members become comfortable with using the platform (N=95), (3) helping the group resolve differences or conflicts (N=86), and (4) setting up rules/norms (N=81).

⁴Users were instructed to select all the applications they had experience with. This is why the percentages do not add up to 100%.

Issue Category	Issue	# Votes	%
Social	Encouraging participation or inviting responses	101	29.9%
Technical	Helping group members become comfortable with using the platform	95	28.1%
Social	Helping the group resolve differences or conflicts	86	25.5%
Managerial	Setting up rules/norms	81	24%
Intellectual	Making connections between relevant pieces of information	73	21.6%
Managerial	Keeping discussions focused	72	21.3%
Managerial	Demonstrating leadership	70	20.7%
Intellectual	Asking questions to clarify	66	19.5%
Intellectual	Developing an understanding of the issues	62	18.3%
Social	Acknowledging contributions	61	18.1%
Managerial	Managing the agenda and progress	60	17.8%
Managerial	Opening the session	59	17.5%
Intellectual	Summarizing key points	54	16%
Intellectual	Initiating questions	54	16%
Managerial	Closing the session	54	16%
Social	Setting the tone	51	15.1%
Managerial	Preparing and circulating an agenda	49	14%

Table 1. Survey responses highlight the aspects of facilitation most challenging for hosts in online voice-based spaces. The horizontal line separates the top four issues which are further examined in Study 2 (Section 5). Higher-level categories (social, technical, managerial, intellectual) are drawn from facilitation literature [9] and discussed in Section 2. These “issue categories” were not shown to survey respondents and are included here for reader clarity.

Encouraging participation or inviting responses was the main aspect of facilitation where the respondents observed difficulties. Lack of participation could slow group progress or create a boring session [7]. As a result, the facilitation literature has studied circumstances in which small or large portions of participants do not participate in the discussion, offering techniques to improve these situations [7]. Previous studies exploring facilitators’ perspectives on challenges have reported that encouraging participation is the second most difficult task for facilitators both experienced (that is, 5 years or more: 66.7%⁵) and less experienced (that is, less than 5 years: 71. 9%) [108].

Making participants feel comfortable using the platform was the second most voted-for issue. This issue indicates that even if hosts are personally comfortable using the technology, they are not successful at creating the same feeling for participants. Through a survey of facilitators with different levels of experience, Tarmizi et. al. had also found that creating comfort with the environment and tool to be considerably difficult for facilitators [108].

Conflict resolution (#3) is the topic of many studies on facilitation [17, 96]. Conflicts can be very challenging to resolve as they are heavily influenced by emotions, making it difficult for people to act rationally [113]. This inherent difficulty could be exacerbated by the anonymity of involved parties in voice-based spaces, which could affect the discussants’ willingness to adhere to prosocial behaviors. These factors could explain why *helping the group resolve differences or conflicts* is our third most voted-for issue.

Finally, *setting up rules/norms* (#4) was reported to be the fourth most challenging aspect hosts struggle with. Prior work has also indicated that facilitators find atmosphere management to be a very demanding task [24]. While difficult, this aspect of facilitation is extremely important as setting appropriate ground rules can help ensure meetings are more productive [82]. These rules can also help enforce effective group interactions.

⁵Percentage of experienced facilitators who reported encouraging participation to be a difficult task.

4.3 Study 1 Discussion

While we focused on the four aspects of facilitation that received the most votes, the distribution of responses is relatively flat, ranging from 29% to 14%. Furthermore, even the issue with the fewest votes still had a significant number of participants ($N = 49$) identifying it as a problem. This suggests that facilitation in voice-based spaces may be underperforming across all aspects and requires substantial improvement.

One reason why users report negative experiences with all aspects of facilitation could be the host appointment mechanism. Currently, any user on social-audio platforms can create and host or co-host events, regardless of their duration of experience with these applications. Additionally, users interested in hosting are not required to complete any form of training. As a result, hosts are often unfamiliar with facilitation techniques. Moreover, since there are no limits on how long hosts have been using the platform, they could be unfamiliar with social-audio norms. Including mandatory training or requiring hosts to have experience with the application (e.g., having used the platform for a minimum number of days or having attended a minimum number of hours of other events) might be helpful. Only being able to use voice (i.e., not being able to have visual or physical contact with participants) could be another reason why facilitation on social-audio platforms is so challenging. Prior work has found that some facilitation functions become more demanding to perform when there is a lack of proximity to the participants [24].

Facilitators' inability to know who their participants are before each session prevents them from fully preparing. Prior work recommends understanding the audience and tailoring content to meet the needs and interests of attendees [6]. However, this is not entirely possible for social-audio hosts. Most social audio platforms allow hosts to schedule and announce their upcoming events, but to the best of our knowledge, they do not provide information to hosts about the users who "set reminders" for themselves to join the events. Providing summary demographics or a list of expectations—such as topics of interest or the anticipated level of participation (e.g., speaking or listening)—could help hosts plan more effective sessions.

Events not being part of a recurring long-term series or conversation, might be another reason why hosts might have less incentive to support users in contexts that aren't directly related to their event. For instance, helping the user become comfortable with the platform might not be a priority for the host as they might not interact with the same group of users after that one event.

A number of the most observed issues could also be connected with one another. For instance, prior work has discussed how one way of encouraging participation is introducing new members to the group's norms [108] as without knowing these norms new users might feel unequipped to participate. However, since setting rules and norms is an area social-audio facilitators struggle with, the effects propagate to encouraging participation as well.

Having identified aspects of facilitation where users express a need for improvement, we conducted more in-depth co-design interviews in our next study. Through these interviews, we asked participants about positive and negative experiences they had when using social audio platforms. Additionally, focusing on the top issues reported through the survey, we discussed users' perspectives on potential technological solutions for these issues. Through these interviews we explore current practices that might explain the reason behind this level of dissatisfaction with host's implementation of facilitation tasks.

5 Study 2: Co-Design Interviews to Understand Facilitation Practices

Having identified significant facilitation challenges on social-audio platforms in our previous study, our objective was to delve into the practices employed by hosts when participating in online audio discussions. Additionally, we aimed to investigate user reactions, perspectives, and suggestions for

potential technological solutions to address facilitation challenges. To achieve these goals, inspired by the methodology presented in [23], we conducted co-design interviews with 14 individuals who had engaged with social audio platforms in various capacities (either as listeners, speakers, or hosts).

5.1 Method

5.1.1 Developing Storyboards as Co-Design Prompts. Storyboards are often used in qualitative research [91] to generate richer data and prompt more meaningful participation [21]. To prompt ideas and inputs from participants with respect to facilitation tools, we used storyboards in our experiments [23]. Each storyboard was made up of five panels with the first two describing the problem, and the next three proposing a potential solution. Visuals were intentionally designed to be rough sketches to ensure participants felt comfortable criticizing and suggesting modifications.

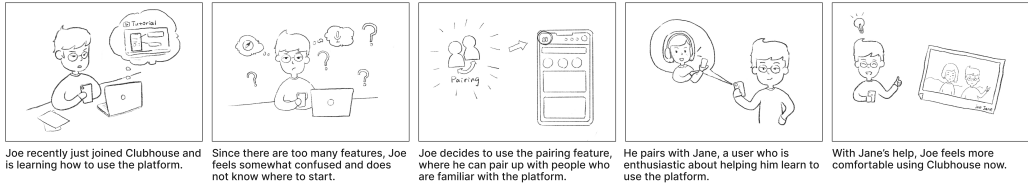
A total of five storyboards were created. Four of these storyboards showed potential solutions for the four most common issues identified in our survey study (discussed in Section 4). We focused on the top four issues identified as the most challenging by social-audio users in our co-design interviews. Limiting the scope to these issues helped prevent participants from feeling overwhelmed by too many divergent topics while allowing for in-depth discussion of each issue without requiring excessively long interviews [68, 107]. However, we acknowledge that other facilitation functions displayed in Table 1 are also of value and should be considered in future work. It is important to note that while our storyboards focused on the issues with the highest number of votes, participants were encouraged to discuss any challenges or negative experiences they had in social-audio events. Consequently, additional challenges were identified and are addressed in our findings.

A fifth storyboard targeting facilitator collaborations was also added (Figure 2e) to help explore issues with collaboration in these voice-based discussions. Facilitator collaboration was not one of the issues shown to users on the survey as it was not indicated as a major task or issues by the two sources used. Recent studies on moderation, however, have discussed the importance of teamwork when it comes to content moderation [15]. As a result, we include another storyboard to investigate issues and a potential solution for host collaboration in social-audio spaces. All storyboards are shown in Figure 2.

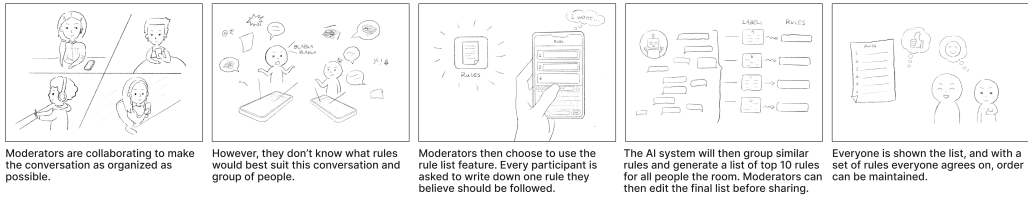
Solutions presented for each issue were developed through multiple brainstorming sessions between the members of our research team (N=5). Storyboards were designed to convey the core concept behind the solution without focusing on the specific design elements of any existing social audio application interfaces. The goal of the presented designs was not to provide optimal solutions, but to probe participants for discussion.

5.1.2 Co-Design Interviews. We designed an interview protocol that asked users of voice-based platforms (e.g., listeners, speakers, and hosts) about the current issues they faced on the platform, and how they thought these issues could be solved. Participants were first asked to recall a positive and a negative experience they had on voice-based platforms. Within the exploration of each experience, facilitation practices and strategies were also explored.

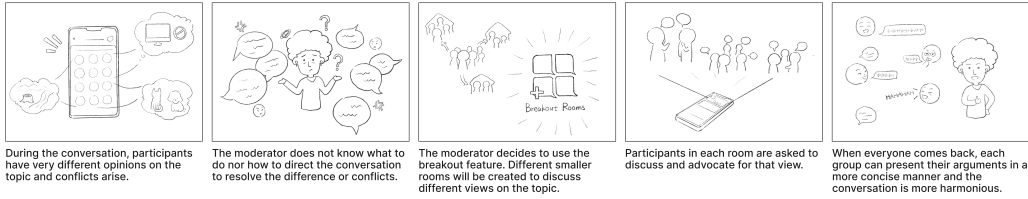
Once those experiences were explored, to help participants with ideation on potential solutions, we showed each participant 2–4 storyboards. The variation in the number of storyboards presented was influenced by the progress of the interview (i.e., time limitations) and whether the participant resonated with the issue being presented. If a participant did not view the scenario as a problem, we would not discuss the solution with them. For each storyboard, we initially showed only the first two panels, depicting the situation and problem. We then asked if they related to the issue and what they thought could be done to address it. If the issue resonated with them, and having heard their potential solutions, we shared the full storyboard. We asked for their feedback on the



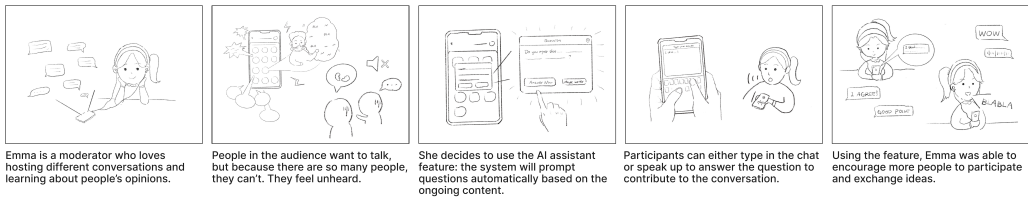
(a) Helping group members become comfortable with using the platform



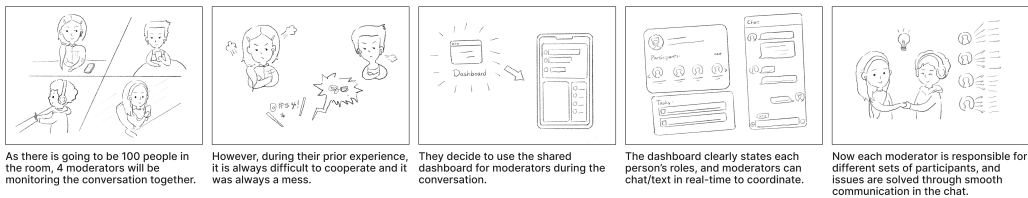
(b) Setting up rules/norms



(c) Helping the group resolve differences or conflicts



(d) Encouraging participation or inviting responses



(e) Helping moderators collaborate more effectively.

Fig. 2. The 5 storyboards presented to participants in interviews.

solution and any shortcomings they identified. Finally, we asked how they thought the solution could be improved or replaced with other methods.

Table 2 displays a sample of the questions participants were asked. Follow-up questions were asked in each section to clarify or expand on points. Participants were offered 15\$/hr (in line with our state's minimum wage regulations) for their time and the interview took one hour on average (maximum length = 1 hour & 27 minutes, minimum length = 38 minutes).

Section	Example Questions
Opening	What role do you think moderators should play in voice-based spaces?
Understanding Experiences	Think of a positive/negative experience you've had on [voice-based platform]: - What happened? - How did the host respond? - What do you think should have been done?
Ideation Over Storyboard	Please read the storyboard: - Do you resonate with the scenario that is depicted? - Do you think this is a problem? - What do you think can be done in this situation to solve the problem?

Table 2. A sample of the questions asked within our co-design interview sessions. Followup questions were introduced depending on the response to encourage discussion.

5.1.3 Participant Recruitment & Demographics. Participants were recruited for the interview through three methods: (1) reaching out to the survey respondents⁶ who had consented to be contacted for a follow-up interview (N=9), (2) students in our institute who had experience using social audio platforms (N=2), and (3) referral by previously interviewed participants (N=3). Interviews were conducted with a total of 14 participants over Zoom.

Table 3 displays the characteristics of the individuals who participated in our co-design interview study. The storyboards each participant was shown is also presented in table. The majority of participants were African American (53.8%). Women also made up the majority (61.5%) of our participant pool. Variations in the number of storyboards shown to participants is explained by time restrictions, and whether they resonated with the issues depicted in the storyboards or not. The variation in the number of responses to each storyboard is due to our interview strategy; when a participant brought up an issue before we reached the storyboard segment of the interview, we made an effort to present the relevant storyboard to gather their viewpoint.

5.1.4 Analysis. We conducted inductive coding [31] on the entirety of transcripts, extracting: facilitation practices, expectations of voice-based spaces, and reactions to specific technological solutions presented in the storyboards. A total of eight aspect of facilitation that participants had polar opinions on were identified. These findings are presented in Section 5.2.

5.2 Findings

In this section, we present our findings along the aspects of facilitation and running sessions that participants had differing strategies for or face the most challenges when trying to implement. The participant that discussed each point is referenced using the notation of (*P* + participant #).

5.2.1 Podcast vs. Conversation: Differing Opinions on Levels of Participation and Interactivity. It is often assumed that hosts would like to have users engage in the discussion, however, both the hosts and speakers/listeners whom we interviewed had conflicting views on the level of involvement they liked to have. Within interviews with participants, one group emphasized the importance of getting users involved in the discussion and having an interactive conversation:

⁶Who were recruited through social-audio related spaces.

⁷These users participated in sessions as listeners or speakers, without possessing moderation capabilities.

ID	Age	Applications Used	Duration of Use	Most Common Role	Storyboards
P1	18-24	Clubhouse, Discord	4-6 Months	Listener/Speaker	2a, 2c
P2	25-34	Clubhouse, Twitter Spaces	Less than 1 Month	Listener/Speaker	2d, 2b
P3	55-64	Discord	1 Year or more	Host/Moderator	2c, 2d
P4	25-34	Twitter Spaces, Discord	1 Year or more	Listener/Speaker	2a, 2b, 2e, 2c
P5	18-24	Twitter Spaces, Discord	6 Months - 1 Year	Host/Moderator	2d, 2c, 2e
P6	18-24	Twitter Spaces, Discord	1 Year or more	Host/Moderator	2a, 2b, 2e
P7	35-44	Clubhouse, Twitter Spaces	1 Year or more	Listener/Speaker	2a, 2b, 2e
P8	25-34	Twitter Spaces, Discord	1 Year or more	Host/Moderator	2c, 2d, 2e
P9	25-34	Twitter Spaces, Discord	1 Year or more	Host/Moderator	2a, 2b, 2c
P10	25-34	Twitter Spaces	6 Months - 1 Year	Listener/Speaker	2c, 2e
P11	18-24	Twitter Spaces, Discord	1 Year or more	Host/Moderator	2d, 2a, 2b
P12	25-34	Clubhouse, Twitter Spaces, Discord	1 Year or more	Host/Moderator	2b, 2d
P13	55-64	Clubhouse, Twitter Spaces	6 Months - 1 Year	Listener/Speaker	2b, 2c
P14	25-34	Twitter Spaces	6 Months - 1 Year	Listener/Speaker	2c, 2b, 2e

Table 3. Demographics & experience level of our interview participants. We interviewed 7 listeners/speakers⁷ and 7 hosts/moderators. The number of times each storyboard was shown to participants is as follows: (a) = 6, (b) = 9, (c) = 9, (d) = 6, (e) = 7. While we have discussed how Discord audio channels do not require a host to be present, P3, who exclusively used Discord audio, clarified that they hosted events on Discord—meaning they were present for the entire duration, and performed hosting tasks.

“that’s the whole point of, you know, social audio, like to actually have a conversation, not just listen to one person” (P7)

Another set of participants, however, revealed that they use social audio platforms to teach certain topics and hold classes. For this group, controlling the number of attendees (through limiting the number of people allowed in the space (P11)), when they could join the session (through disabling the join link after a certain period of time (P5)), and how much involvement in the content they could have, was very important. Citing a need to control attendees, prevent rowdiness, and prevent themselves and the conversation from getting distracted, they would sometimes mute all attendees and disable the chat, and only enable some of these capabilities for short periods of time during which attendees could ask their questions (P5). One example of how they would run a session was as follows: “I had to reduce the number of [attendees]. I disabled the chats. And whenever I wanted your opinion, I would call your name and allow you to talk” (P11).

Hosts who wanted to create an interactive space believed that if “you create that kind of environment where we can laugh at our stupidity, we can laugh at our silliness. It makes people feel safer” (P3), thus encouraging them to participate. Additionally, hosts would verbally encourage members through asking for contributions (P7). This encouragement would at times be done through reaching out to members privately through text messages telling them that they are “welcome to join the stage if [they] wanted, [or had] any questions” (P13). Private messages, however, were thought to be “helpful, although sometimes annoying” (P13) if the participant did not have anything to add at the time.

Among speakers/listeners, some expressed that “the main good point [about voice-based spaces is that] it doesn’t really convey any like pressure [...] and you don’t need to actively participate” (P1). This group enjoyed listening to these events “almost like a podcast going on in the background while I’m like cooking or doing something else” (P2). Others disagreed, expressing that participation is “important because otherwise, it’s one person talking to you. So why have an audio space? Just go listen to a podcast or watch a YouTube video or something, you know. So I think that’s what audio spaces are for, for people to participate” (P7).

When talking about the type of engagement attendees expected, they expressed that “it would have been nice to have a host who was actively asking people to engage by asking specific questions

[...] right now it was very much like question answer sessions and not very interactive or engaging otherwise" (P2).

5.2.2 Panel of Experts vs. All Opinions Welcomed: Composition of Speakers Based on Different Characteristics. Participants also discussed different properties of speakers they believed were important. Overall levels of expertise of speakers was one of the mentioned properties. Participants mentioned enjoying listening to professionals or experts in a topic speak, going as far as praising the hosts if they had good knowledge of the topic, such as:

"the fact that the host was so knowledgeable on the topic. And when I [...] went up on the stage and I asked specifically about my upbringing and my education [...] she took the time and she, like, really explained and went in the background of what happened [...] So, the fact that she, one took the time, and two was extremely knowledgeable of the topic [was important to me]" (P7)

However, they also mentioned feeling intimidated to contribute if all speakers are experts: "[if the panel of speakers are all] so specialized that I don't even feel comfortable participating [...] I'd be more interested in being in something where it's more everybody's opinion is welcome" (P13). They further expressed:

"I like to hear people other than the experts, be able to step in and talk to me, that's the benefit of having a platform like this. Otherwise I will just, you know, turn on the TV and watch some show where they're interviewing somebody that knows about quantum physics, for example. It's the engagement that makes it interesting, it makes it a social platform." (P13)

While the feeling of being uncomfortable participating in rooms where speakers are experts was echoed by other participants, some believed their level of comfort in participation was also greatly influenced by the tone the host had set. In other words:

"[in] some of the [...] conversations, there are people who are [...] really high up or really removed from what I do. So I wouldn't feel comfortable in this space unless the host sets that tone [...] having a host constantly interacting or having a thread throughout that conversation, asking for engagement from people that's low stakes engagement, but still makes you feel part of a larger community in the conversation." (P2)

Diversity of speakers was another characteristic of the speaker group participants cared about, but believed was often not implemented well: "[most of the spaces I've attended] wasn't very diverse or inclusive" (P2). Finally, participants felt like they were not able to participate in discussions due to the large number of people within the space, one person taking up too much time, or because of how speakers were selected: "the few people that are speaking are often the friends of the host" (P4).

5.2.3 Shutting Down Disagreements vs. Exploring Differing Opinions: Exploring Conflict Management Styles. In contrast to workplace conversations or educational environments where achieving consensus may be the objective, social-audio discussions are oriented towards embracing diverse opinions. Participants expressed joining spaces out of curiosity (P1) or to hear "different opinions from different people" (P14). The majority of participants believed that "you do need [a] diversity of perspectives" (P2) on the topics that are being discussed. They expressed that being able to hear differences in opinions is "something that makes a space sweet and great, because as one person is supporting, definitely another group of people [are] opposing" (P4). They acknowledged that the expression of these differences needed to be done in a "concise manner, in a good manner, in a respectful manner" (P4). When that wasn't the case, they believed hosts needed to intervene, with

participants giving-up on the session if the host did not intervene in time: “I feel like the moderator, just let things play out for so long. So I felt [that even if I] speak up, what difference does it make? Probably nobody would even try to listen” (P14).

Hosts implemented different types of intervention. One method was to completely shut-down the discussion:

“I have the power to mute everyone that is speaking. If you guys are all fighting [...] then basically what I’ll do is I’ll mute everyone, then I’ll talk about it, or what you guys are doing is not proper, that you guys should allow each person to express their views. Then I would open the mic again for one person or the other person or the next person. [...] If I notice that, ok, everyone is cool. But if I notice that, OK, you guys are still fighting, the best thing is just to end the space” (P4)

When asked if they thought there were any downsides to their suggested strategy, they did not believe there were any (P4). Another strategy hosts employed to stop the disagreement from continuing, was to acknowledge differences in opinion and “tell participants that their idea also works” (P8). They expressed that this strategy would stop any further disagreements.

Finally, one participant believed that when there is conflict “you first have to address the emotional response to that, [...] you have to understand the psychology and really pay attention to that detail so that you can then direct the conversation in the right way” (P3). They believed that conflicts occur either because (1) the participants are seeking attention or (2) they are truly interested in exploring the topic. If they are seeking attention, the host would “put them in the spotlight, because clearly they need it, then [the host] can turn that spotlight off” (P3). Times of conflict were viewed as occasions when “delicate manipulation” (P3) is useful. However, even if the participant is hoping to explore a topic, the host has to “think about the 40 other people in the room [and if] it benefit[s] them to do [that exploration]” (P3).

When it came to situations where conflict could arise, participants also expressed that they expected hosts to “know the intent of what [a participant] wants to talk about” (P14) instead of just allowing anyone to speak. They believed that the host “must have been listening and noting people who were speaking so from the username, [...] next time when that person wants to speak they should know the angle that person might come from, so [the host] might not allow them to speak at that point in time and let another person speak [...] based on previous opinions people have said in the spaces they should be able to know what kind of things they are going to contribute to the discussion” (P14). While this could help guide the discussion, without the help of any tools, it could be extremely taxing for the host.

5.2.4 Soliciting Community Feedback vs. Appearing Incompetent: Setting and Informing Users of the Rules. The importance of rules was brought up by participants both without any prompting and as a response to our associated storyboard (Figure 2b). The unprompted mention of rules was linked to some participants defining the role of the host as someone who “gives the guidelines [...], rules, and regulation” (P10) or the person who “maintains order by setting rules and regulations” (P14). When it came to how attendees are informed of the rules of each space, participants expressed that while “you can write it in the room description, [...] most people don’t read that” (P7). As a results, hosts would often “state them at the beginning of the room and then as more people come in” (P7), or they would do a “reset” (P7) as part of which they would remind users that “this is this space, we’re talking about this topic, [and] these are our basic rules” (P7). Another way to remind people of the rules was that “when it happens, it’s brought up” (P7). While resets seem to be a common practice used in most events, some expressed issues with the practice as hosts were “repeating [the rules] over and over” (P4) which attendees found “more like a distraction [...] So I think it

should be written down [...] [and you see the rules] before you join in the space, [...] then you can go ahead and join" (P4).

While informing users of the rules is an important aspect, selecting rules is another issue that needs to be addressed. Participants expressed that the selection of rules could be difficult because the hosts "don't know the kind of people who will join the space" (P14) or "they are a novice and they've never [held events] before so they don't know what to expect from the space" (P14). While many agreed that "there are some rules that could apply to almost any room" (P7), they also believed changes need to be made "based on [...] the topics" (P4). When it comes to who is responsible for the selection of rules, one suggestion from prior work is to ask participants to generate ground rules [6] (One implementation of this practice was depicted in Storyboard 2b). Interviewees however, had conflicting views on the matter. Some believed that if they are the host of a space then they "should be able to create the rules and regulation for everyone that will be joining this space" (P4), and if the rules are shown to attendees prior to joining, they would have the option not to join if they did not agree with the rules (P4). Others believed more flexibility is required, pointing out that while "if it's [the host's] room [...] and they're taking ownership and responsibility" (P7) hosts should be selecting the rules, "but at the same time [hosts] should also listen to the members [...] if the members collectively feel that there's a better rule, or one rule shouldn't be used" (P7). Participants also thought that while the need for some rules might not be clear to hosts from the beginning, they should be vigilant that if something they don't think is appropriate "keeps happening, [it] should be a rule now" (P7).

When it came to asking for participant input participants expressed concerns about attendees not being able to "see [the prompt to send in a rule] on time" (P4) to be able to provide input, and doubting that they "would have the right expertise to be making good recommendations" (P2). Another participant expressed that:

"when you do this, your audience might feel that maybe you don't know what to do with the space you're trying to hold. They kind of view you as maybe incompetent because if you are coming to host a space. Then definitely, you should know what to do and what not to do. So why are you asking them to make the rules for you?" (P14)

The need to present one-self as competent and in-charge rather than trying to get the community's input is an interesting aspect of facilitation in these spaces to consider.

5.2.5 Enforcing Rules vs. Allowing Issues to Self-Resolve: Handling Harms in Voice-Based Events.

Deviating from the main topic of discussion and interruptions were some of the main harms users mentioned when discussing negative experiences: while some expressed that they sometimes enjoyed breaking the rule of *staying on topic* and believed that they should be allowed to do so if the majority of participants agreed with the change in the direction of the discussion (P1), others believed that "there are some free-for-all rooms but a lot of the rooms are structured to a specific topic and [...] the host should be trying to keep everyone on topic as possible" (P7). This contradiction around addressing harms when they occurred, was an additional source of frustration for some participants. Reflecting on the hosts' lack of action when it came to individuals who went off topic, one participant expressed:

"[the host] is simply just not putting any kind of a control. And they're just allowing people that are willing to, spout off [...] and they're just letting them go" (P13)

In this scenario, having an event in which participants are willing to talk and contribute to the discussion, instead of a quiet meeting, might have been the driving factor behind why the host did not take action. It is also worth noting that the facilitation literature discusses different strategies for dealing with problems that arise [7]. One of these strategies is to do nothing in cases where the

problem does not seem to be too serious and is likely to disappear or be handled by other members if left alone [7]. As a result, the lack of action by hosts could be thought of as the implementation of this strategy. Our participants however, had very negative reactions to this lack of action.

When discussing conflict resolution, participants also brought up the importance of listening and being attentive to prevent any mistakes by the hosts themselves:

“one of the hosts said that, well, what I think she’s trying to say, which is my pet peeve, when people try to put words in my mouth, and then she got my points wrong and I was like, no, no, that was not my point” (P7)

While this participant acknowledged that the host “tried to help” (P7) they still felt unsatisfied because “one, it was after a while but two, she still got the point I was trying to make wrong [...] I felt I wasn’t being listened to” (P7).

5.2.6 Support Systems vs. Sources of Conflict: Navigating Co-Hosting Challenges. While hosts expressed that managing sessions is a task that could be undertaken by one person, they often believed having at least two people available was a good idea. One motivation for having more than one person was to be able to split tasks and focus on specific tasks (P5). Another reason was to ensure that someone could take over and run the session in case of technical problems, such as losing internet connection or experiencing audio quality issues. This preference for multiple hosts requires additional steps to ensure that these individuals collaborate effectively.

When describing sessions with multiple hosts, participants described a hierarchical structure with the event creator being in charge of major decisions (e.g., rules, and expectations) and the management of other hosts (P7). Participants described holding preparation sessions prior to events, as well as getting together after events if issues occurred (P7). The plan for the event was also revealed to be discussed in the private sessions co-hosts held. However, in contrast with the suggestion to share the agenda with participants [6], hosts often kept the agenda to themselves and used it internally. While not knowing the participants beforehand may contribute to hosts not sharing the agenda prior to the event, they also chose not to share it after the event had started. Hosts explained that they believed the agenda would not be useful to the participants (P5). Participants however, expressed an interest in “see[ing] what’s the whole purpose of this talk so that they can decide whether or not [...] it is aligned with their purpose” (P1).

Even though there is a hierarchical structure between hosts, one participant recalled a clear breakdown of co-hosting efforts that went unnoticed by the lead host for far longer than it should have. The conflict was understood to be due to personal conflicts between the co-hosts:

“they mute each other’s microphones, they get into arguments over certain topics [...] one closes the room even though the other one didn’t want to” (P7)

5.2.7 The Use of AI vs. Authenticity: Perspectives on Technological Solutions. When discussing the use of technologies in general, but also AI in particular, the majority of users did not have specific hesitations or only questioned the ability of the systems to perform well or correctly. A subset of participants, however, believed that the use of such systems need to be carefully considered:

“AI or any kind of assist feature [...] only works if that is what the audience is prepared to accept that is the practice. It only works if people are paying attention to it and interacting with it so that it’s a natural flow” (P3)

The participant discussed how even using the chat feature would be ineffective if the host did not interact with and acknowledge the user inputs through methods such as reading the message out-loud (P3). As a result they believed tools could be helpful but would not be a solution on their own. Another participant cautioned against a loss of authenticity:

"don't let AI do everything [...] let it be with someone, collaboration [because otherwise] the authenticity of a lot of things won't be there anymore" (P9)

This resistance towards the use of AI is in-line with other studies that find hesitancy regarding the use of this technology [78].

5.2.8 Post-based vs. Live: Learning How to Properly Use Social Audio Platforms. Participants also mentioned expectations they had from the platform (rather than from hosts). Discussing a mismatch between their expectation and the reality of using social-audio platforms, participants compared these platforms with other social media they used. Participants stated having times when they "have got a half hour free" (P13) and want to go "check out Clubhouse just like I wanted to go check out Facebook [or] Instagram, the same idea. I wanna be able to just go in there. [and] see what conversations are happening." (P13). They however expressed that this type of use does not seem to be possible, acknowledging that "I can't expect the moderator to be sitting there all day long for when [I] might suddenly want to jump in and find a good conversation. But that's kind of what we do with Facebook and Instagram" (P13). To help shift this mindset and improve understanding of how to use the platform, they suggested that "maybe Clubhouse should think about [having] a room that everyone can attend to kind of learn how to use Clubhouse better" (P13). While this participant was aware that Clubhouse had rooms discussing new features, they believed it would also be useful to teach users how to navigate the platform in a way that they "could get the most benefit from" (P13) it.

5.3 Study 2 Discussion

In this study, we showcased the diversity of preferences and expectations among users of social-audio platforms. This high degree of conflicting expectations could partly be attributed to the unclear ownership of social-audio events. In platforms such as Twitch or YouTube, videos/streams are seen as *belonging* to the streamer, with other users on the platforms being *viewers* [46]. While Twitch streamers might interact with comments on the stream, the comments are often not the main or only content of the stream. In other platforms such as X, Instagram, and Facebook each user owns their own content. While users can interact with other users' content, there is often a clear separation between who each piece of content belongs to. On social audio events, however, this distinction is not as clear. As attendees directly contribute to the content of social-audio events, they seem to have a sense of ownership and stronger opinions on how events should be conducted. Hosts, on the other hand, having created the event, perform actions inline with a sense of owning the event. This conflict in views often results in actions being taken by hosts that cause dissatisfaction in the attendees.

Another reason for the observed conflicting views on how events should be conducted could be that the goal of social audio events are not clear to all stakeholders, or that the same goals are not shared by them all. Participants could be coming to social-audio events to have real-time discussions about a topic they are interested in. For these users the topic, and the ability to have a discussion might be the driving factor. On the other hand, hosts might be looking to increase their following on the platform, and become influencers. Wanting a larger following or to become an influencer is not a concept unique to social-audio platforms. Influencers on other social media platforms such as Instagram have been studied in prior work [52, 79]. The interplay between how content is created in social-audio and the want to become an influencer, however, could introduce new challenges. While there are group dynamics and interactions between influencers and users on other platforms as well (e.g., comments on Instagram posts or Twitch streams), these dynamics don't directly change the content itself. For instance, a host might feel like they would need to talk a lot and dominate the conversation in order to attract more followers which would reduce the

amount of discussion within the event. Or they might feel like being argumentative would make them more entertaining and increase their appeal as an influencer, resulting in more speakers or hosts purposefully breaking rules or creating conflict. Similarly, hosts might refrain from enforcing some rules or being strict with rule breakers to avoid losing followers, which would again affect the flow of the event (i.e., the content) and lead to dissatisfaction. The issue of hosts' perception on these platforms, which we observed within our findings could also be partly related to this issue. Hosts demonstrated a heightened awareness of users' opinions and were particularly concerned about the perception of appearing "incompetent". This desire to project competence and authority could be reflected in their facilitation practices as well.

Some of the dissatisfaction expressed by users could stem from the host's lack of facilitation, conversation, and emotional intelligence skills. For example, participants shared that they do not feel heard, highlighting the need for facilitators to develop active listening skills. Active listening has been shown to be at the heart of emotional intelligence because it allows people to remain open to new perspectives, ideas, and experiences [83]. In our study, participants reported responses from hosts that were just aimed at stopping the disagreement by saying "Your idea also works" without showing genuine interest in the proposed idea or asking follow-up questions that could enrich the conversation. This could contribute to participants feeling dismissed.

Finally, our study highlighted several key differences between in-person facilitation and social audio-based facilitation. One notable difference is agenda sharing. Not sharing an agenda was found to be a common practice in social-audio events, even after they began. This contrasts with facilitation literature, which stresses the importance of starting any meeting with an agenda. Additionally, facilitation literature suggests that the group should have the opportunity to assess whether the agenda suits their needs and propose adjustments [7]. While this approach may lead some members to leave if the agenda doesn't meet their expectations, it is preferable to having dissatisfied participants remain in the discussion. Social audio platforms could improve their interfaces by allowing hosts to post an event agenda and highlight the topics being discussed at any given moment. This would help participants assess whether the room fits their needs and understand the ongoing discussion when they join.

In this study, we examined current practices of hosts on social audio platforms and the challenges they face. We found that some of the practices employed by hosts contribute to conflict and dissatisfaction. Additionally, we observed that certain views and practices, while not aligned with facilitation literature, are widely shared among hosts of these real-time audio events. This led us to question what might be influencing these views and behaviors. One possible influence could be guides or training materials that promote these practices. As a result, in our next study, we investigated the training materials available to social audio users.

6 Study 3: Analyzing a Corpus of Host Training Sessions

In our first study, we used a survey to explore the aspects of facilitation that social audio users found most challenging and where they observed the most issues. In our second study, we conducted co-design interviews to gain a deeper understanding of current practices and to explore reactions to potential technological solutions. Our findings revealed that hosts are a key source of the problems, making it important to investigate the reasons behind this. Prior work has shown that on-boarding conversations are used by new moderators [101]. In this study, we investigate what is being taught to hosts on social-audio platforms by analyzing 16 existing training sessions on Clubhouse. We examine whether the available training addresses the challenges identified in our previous studies or reinforces some of the conflicts we documented. We acknowledge that training sessions are not the only method hosts can use to learn how to facilitate; other methods may include understanding the community's values through participation and learning by doing.

6.1 Method

6.1.1 Training Material Selection. To find audio discussions focused on training hosts we searched for “moderation” on Clubhouse and selected a total of 16 events for analysis. The rationale behind searching for “moderation” instead of “facilitation” is that the role of guiding the conversation is advertised and recognized as being done by “moderators” on the application [74]. Additionally, the events held by official Clubhouse employees that teach methods for having better discussions are called “Lessons in Moderation”. Moreover, speakers in these sessions refer to the discussed techniques as “moderation methods”. Even though moderation keywords were used for these sessions, the content focuses on the discussion of facilitation techniques and practices.

Through our search on Clubhouse we found 16 events. The analyzed events were made up of eight that are held by official Clubhouse personnel and eight by other users on the platform. Where possible, events not associated with Clubhouse were selected such that they were held by different clubs. Our goal was to diversify the clubs that organize the events to ensure the content would be as heterogeneous as possible.

All selected events were recorded by the hosts of the session themselves and made publicly available on Clubhouse. Participants in the event were informed that the event was being recorded as it was happening. No recordings were made by the research team. The recorded sessions are available on the application, as well as through the web version of Clubhouse for public access. The content of these sessions is not searchable through the web as the transcripts of the audio sessions are not publicly available online. However, we were not able to get consent from participants after the fact, because the recordings do not provide a full list of event participants. This study was approved by our institution’s IRB.

In total, the 16 rooms are 18.3 hours long. The rooms have an average of nine speakers (maximum = 16, minimum = 2) and an average of 1,890 attendees (maximum = 5,525, minimum = 11). All rooms were held between November 10, 2021 and May 8, 2023. Events that were held by Clubhouse personnel (N=8) usually had a 20-30 minute teaching portion, with the remainder of the hour being dedicated to question and answers. Sessions held by other users on the platform (N=8) had different formats, with some being a discussion between a pre-selected panel of speakers, while others invited participants to contribute or ask questions. While some platforms have mandatory training sessions for their moderators [53], as hosts on Clubhouse are not employees, participation in these training sessions is not mandatory.

In addition to Clubhouse, we searched for training material on X-Spaces and Discord Audio. However, we were not able to find pre-recorded audio events aimed at training hosts on these platforms. We were able to find two online text-based documents published by X-Spaces⁸ and Discord⁹ that discussed moderation in their voice-based spaces. However these guides focused on the features of these spaces, or practices such as finding evidence for reports of harms, which have already been investigated [56]. As a result, these guides were not included in our findings.

6.1.2 Analysis. As we are interested in aspects of teachings that are specific to social audio, we performed inductive coding [109] on the transcripts of recordings. Our goal was to identify teachings that are unique to social audio, or teachings that conflict with the existing facilitation literature. This will help us identify issues the facilitation and moderation literature need to address to help hosts of voice-based spaces better perform their tasks. These findings were grouped into six codes which will be discussed in Section 6.2.

⁸<https://help.x.com/en/using-x/spaces-hosting>

⁹<https://discord.com/community/how-to-moderate-voice-channels-discord>

6.2 Findings

Through our analysis we found strategies for various aspects of facilitation discussed including time management, engaging participants, and specific features of the application to assist hosts. In this section, however, we will focus on tips that address novel requirements of social audio spaces. The training session that discussed each point is referenced using the notation of (*R* + session #). We do not include number of mentions for each issue because the eight sessions that were held by Clubhouse personnel often recited the same instructions. As a result, these numbers could be misleading as to the importance of each issue based on whether they were brought up in the official or non-official sessions. The presented issues are ordered based on the level of specificity of the task they address, going from higher level issues to lower level ones.

6.2.1 *Owning the Sessions vs. Facilitating a Discussion.* Some of the investigated training sessions explicitly state that the facilitator owns the space:

“you really are the captain of your ship. What you want in your room is what you want” (**R4**)

While these sessions emphasize that it is important to “try to make space for people [...] so that they can share their input in their voice, in your spaces” (**R4**), they still view the space as being owned by the host. This view could encourage a number of behaviors from the hosts. For instance, hosts are encouraged to remove users instantly if they believe the user’s behavior disrupts their session (**R1**). Additionally, this view allows them to believe it is justified to have specific participation practices that might not allow everyone to take part in the discussion:

“I’ve been criticized a lot by people in other cultures [...] who say that my [session management] technique discriminates against them. And it does. My conducting my sessions in English discriminates against non-English speaking people as well [...] There are requirements to participate in certain kinds of sessions and if you don’t have what’s needed for that session, my attitude is that’s fine. There are plenty of other sessions you can go to” (**R10**)

While some instructors expressed that it is important to address participant needs when possible (e.g., “really thinking about what does this person need and being intentional about asking how can I support them and what do they need from me? And if it’s in my locus of control to give, then I’m more than happy to do that” (**R5**)), another outcome of this sense of ownership of events is that some events are designed to address the needs of the host rather than the needs of participants:

“I’m always thinking about, ok, what’s happening in my business? How can I utilize Clubhouse to help grow it?” (**R16**)

Within these business-centered sessions, participants are then viewed as potential customers rather than discussants: “not everybody who’s gonna come into your room [in] Clubhouse is going to be the person that you’re really looking for” (**R16**). Another behavior that potentially results from this view of social audio events is the participation of hosts in the content rather than managing the process. In line with the facilitation literature, training sessions on Clubhouse (**R1**) also make a distinction between the process of running the room and the content being shared. Stating that “you need to have a solid process in place in order to ensure that your content shines through” (**R1**). However, these classes tend to leave the level of participation in the content up to the individual host, with statements such as “some moderators just like to be the conductor of the orchestra, they direct the flow of conversation but they don’t necessarily inject their opinion or their thoughts heavily [...] Other moderators are more likely to be equal participants in a conversation” (**R1**). Instructors do not argue against either approach:

“the common feeling by most people is that a moderator is supposed to always be neutral [...] that is not authentic to that way of thinking [as] we don’t remove our identities just because we’re moderating something. Our identities don’t disappear. Our lived experiences don’t disappear. And trying to pretend like that’s not a factor, or someone saying something that’s potentially hurtful to us personally [...] is not impacting us, in my experience, generally tends to backfire [...] I’m a moderator in this, but I am also a participant by simple virtue of the fact that I’m in this conversation and being a part of the discussion that is happening in here” (R6)

This finding closely relates to the conflicting views observed in our previous study. It’s also worth noting that this sense of ownership of the session is often not present in the facilitation literature [7]. This literature often views the facilitator as the person who helps the group reach its goal. In this view, the facilitator does not own the space but instead tries “to help make groups perform more effectively by soliciting the leadership skills and potential of all members” [7]. Studies also suggest that the facilitator “must manage the group, the activities and the learning, but not dominate the group” [64]. Thus it is important to note this difference in views of the host role and how that might influence the group dynamics in social audio settings. Previous studies of computer-mediated communication have suggested that power relations are reinforced in computer-mediated communications [104]. This rings true with our findings which show how the host has full control of whether or how attendees are allowed to participate.

6.2.2 Strategic Selection of Speakers & Handling Users with Malicious Intent. Having experienced cases in which people “will enter your rooms for the sole purpose of disruption” (R2), one of the methods suggested and used by the instructors in these training sessions was to use “the general go-to [...] to check bios and to see if their social media’s connected and [...] they seem to post things that are normal” (R2). Hosts were also encouraged to pay attention to whether the users who wanted to come up on stage and speak had “a picture or a bio, a name that [...] seems reasonably real or don’t obviously seem unreal” (R2). While this is an example of proactive behavior by hosts, it requires hosts to perform a large number of tasks that could take away from their ability to run the session efficiently.

The reason for this challenge could be attributed to a number of differences between social-audio and traditional computer-mediated communications. Firstly, in most sessions that are not held on social audio (e.g., online or in-person classes, conferences, or meetings), participants are often made up of a limited pre-selected number of representatives, as well as others who might reasonably be assumed to be from a specific population (e.g., students from a particular university participating in that university’s town hall). Secondly, participants’ awareness of the potential long-term relationships with others in the non-social-audio session could lead them to have more collaborative prosocial behaviors as not behaving in such a way could have long-term consequences. These assumptions, however, are not true in social audio platforms. As a result, hosts are required to ensure participants are selected appropriately.

6.2.3 Increasing Participation in Spaces. While social audio platforms differ from other social media platforms by offering live discussions instead of post-based interactions, they remain social spaces. As such, several sessions discussed strategies for attracting more participants to events hosted on these platforms. One suggested strategy was for hosts to expand their network within the platform by participating in other sessions and collaborating with different hosts:

“if you want people to come to your room and get involved [...] participate in other clubs or other conversations beforehand [...] that allows] you to meet people who maybe

have similar interests and find potential collaborators for your conversations, it also helps you get probably a few followers if you actively participate" (R1)

Another method for reaching more users that was suggested in the training sessions was to have more people on the "stage" (i.e., having them in the subset of users who have the ability to speak), in other words:

"more people on a stage is generally better [... because of] how the product works. So when someone comes up on stage to be a speaker, [the application] sends out notifications to their followers that they're speaking in a room. And so that automatically lets more people know that there's this conversation going on" (R1)

While hearing different perspectives is mentioned as the other benefit of having more people on the stage, that seemed to be a secondary reason for this choice based on how the idea was presented. This should serve as a reminder that while having a good discussion on these platforms might be desired, they are still social media spaces in which users hope to gain a following.

"having a diverse group of moderators who are maybe active or have friends or connections or network in different kind of pockets of the Clubhouse social graph can be really helpful" (R2)

Holding regularly scheduled rooms in order to create a community (R16), as well as teasing upcoming content within the session (R16) were some of the strategies that were suggested to help keep participants from leaving soon after they join a room. Finally, Clubhouse employees did suggest that they are "working on ways to improve distribution and reach" (R1) when it came to finding participants for different rooms.

Strategies aimed at boosting group participation, such as those discussed here, align more with marketing-oriented tasks rather than traditional facilitation practices. Facilitation literature often centers around understanding and guiding a predetermined audience, while the emphasis in this context is on methods to increase the overall number of participants [7]. In contrast, social audio hosts are compelled to integrate effective marketing and outreach strategies into their responsibilities. However, these strategies are not new to social media spaces as large numbers of users on each platform are aiming to increase their following and become "influencers" [52].

6.2.4 Managing New Arrivals Through "Room Resets". Another characteristic of social audio is that participants can join a session at any point during the conversation. This is in contrast with most other sessions in which participants are expected to be present from the beginning (e.g., work meetings or classes). Due to this difference, hosts need to have a way to inform users of what is happening in the space. This is often done through "room resets" when the host lays out the goal of the session, as well as what has transpired up to that point in the discussion. Hosts are also encouraged to "customize [room resets] for how far into the room you are, [and] how many people have come in to the room" (R10). The type of content in the reset can also change based on the goal of the room:

"if you're having a conversation, that is a natural laid back conversation and you're not promoting anything. I only do room resets when people come into the room and the room reset is letting the people in the room know what we're talking about, where we are, and where we're going" (R10)

Resets in business related sessions however, are encouraged to include calls to action. As most sessions assume participation from the beginning, the facilitation literature often includes steps to ensure participants are aware of the goal and context of the meeting as part of the introduction [7]. It is also encouraged to ensure that these goals are clearly stated and very specific [7]. Resets, however, are not a common practice in the facilitation literature.

6.2.5 Addressing the Need to Step Away During a Session. Conversations in social audio spaces occasionally run for prolonged durations. Additionally, since hosting social audio sessions is not a paid position and not what most individuals do as their profession, hosts might be in locations or situations that might require them to attend to other tasks as well:

“You never know what else you have going on, you may need to step away for a second and to be able to trust someone to give over that responsibility to [is helpful]” (R4)

As a result, while the training sessions recognize the possibility of running sessions alone, they state that “the best practice is to have at least two moderators to kind of carry [the] load” (R1). To ensure a successful collaboration, they suggest (a) being on the same page about the goal and expectations of the room, (b) being clear about who will play what role in the room, and (c) selecting co-moderators you trust because there have been “instances where hostile takeovers can happen” (R1). Issue between co-hosts was an issue that was also discussed in our previous study, however, the specific issue of hi-jacking events was not previously brought up.. To address room hijacks, the Clubhouse training sessions revealed that “we had kind of quietly shipped a little update where you couldn’t actually remove the room creator” (R3).

To coordinate among hosts, instructors suggested using the private chat features on the platform (R1) or other applications (R16) to communicate with their co-hosts and resolve any issues in the space or between the hosts. Hosts are also encouraged to “have a follow up” (R16) after the meeting to review any issues that might have occurred.

6.2.6 Acknowledging the Importance of Listeners. While the majority of teaching focused on users who participate in the content, one training session discussed the importance of listeners in spaces and how without them there would be no conversation:

“we sometimes forget that without the folks listening to you, you won’t have a room. I think that we sometimes forget that and don’t address the people that are in the room listening. We kind of focus on everyone that’s on stage and not acknowledge that the folks that are listening, no matter what part of the room they’re in [...] make a big part of the room as well, because they are participating in the chat sometimes and their mere presence brings an experience to the room” (R13)

Even though the majority of work on online participation has focused on active participants, prior work [19, 80] has acknowledged the importance of “listeners” as they make up the majority of users on the internet [19]. Compared to post-based platforms, the importance of listeners is even higher in social audio platforms as these spaces are built around the idea of having the community come to listen to specific discussions.

6.3 Study 3 Discussion

Through this study we found further evidence for the mismatch between participant and host expectations of and goals for spaces. In Section 5.3, we discussed how wanting to be an influencer might be one reason behind this mismatch. This study provided further evidence for this desire, and showed that hosts are being taught tricks to help improve metrics such as their number of attendees and followers. These strategies closely tie to the concept of attention economy [22]. Influencers on all social media platforms are looking to capture the attention of users. In most post-based platforms, however, seeing a post or tweet by an influencer is providing the attention they are seeking. On other platforms such as YouTube, Twitch, or the social audio platforms being discussed here, merely seeing the existence of a video or event is not enough. Influencers need to ensure that users would click through to view the content within. This additional step might be one of the reasons why numerous techniques are being taught to hosts to improve their statistics.

Another issue we had previously touched on was the mismatch between the practices of social audio hosts and the facilitation literature. While presenting the teachings in existing training sessions, we also pointed out how a number of the lessons were in contrast with the existing literature. One major difference that could greatly affect the quality of events was the participation of hosts in the content as well as the process. This active participation in the content of sessions by hosts connects to concepts of neutrality and multi-partiality in the facilitation literature [11, 89]. Some work in the facilitation literature acknowledges a need for the facilitator to have expertise in the topic of discussion but asserts that facilitators need to be neutral in relation to the issue [106]. Neutrality or impartiality encourages not favoring anyone. Multi-partiality, on the other hand, advocates for favoring everyone equally [5]. In both views, however, the facilitator is not encouraged to side with specific participants or treat participants differently unlike what is experienced by users of Clubhouse.

Another issue that was brought up both in existing practices and teachings was the importance of co-hosting. Co-hosting or co-moderation is not unique to social audio platforms. Moderators on Twitch, for instance, often work in teams [15]. The mix between content and process creates complications here as well. In most other platforms the creator (e.g., Twitch streamer) is not considered part of the moderation team while the content creation (e.g., streaming) is taking place. Similarly, moderators are not thought to be producing content while engaged in their role as moderators. On social audio events on the other hand, the majority, if not all hosts contribute to the content. This duality could create complications on top of those faced by moderation teams on other platforms. One way to address this issue would be to truly separate content from process, much like the facilitation literature suggests. Different groups of users would then be engaged in different roles. One potential challenge standing in the way of this separation is the fact that hosts currently often engage in hosting and facilitation duties for the popularity and status it could bring. However, this issue is also not unique to social-audio. Streamers on Twitch who might also have similar aims, have been able to find methods of moderator selection from their community that both support the sense of community and ensure the roles are filled [121].

Training sessions also brought up how co-hosts should be communicating with one another, with some suggesting the use of other platforms for this communication. Currently social audio applications do not offer specific features to assist co-host collaborations. Real-time collaboration between hosts has also been studied on streaming [15, 16] and social virtual reality [90] platforms. These studies have also found that moderators in these spaces often rely on external communication platforms (e.g., Discord) to satisfy their needs [15, 90]. This reliance could indicate a need for the development of within-platform collaboration tools.

7 Overall Reflections

7.1 Design Recommendations

7.1.1 The Design of Multiple Interfaces. Our findings highlighted the differing expectations of attendees and hosts. To address these varying needs, one potential solution could be designing separate interfaces for each group—one for hosts and another for listeners/speakers—providing unique information tailored to each group. Currently, social audio platforms offer similar types of information to both groups, which may not fully meet their distinct requirements.

The mismatch between the level of participation users wanted and the participation they were allowed to have was one issue. The interface dedicated to attendees (i.e., speakers and listeners) could include such information through the use of tags. Examples of event tag types could include the style of the conversation (e.g., casual chat, Q&A, panel discussion, networking, interview, etc.) and the expected level of participation (e.g., listener only, interactive, etc.). Using tags or labels has

been leveraged by other social media platforms [51, 81]. Reddit submissions and Github issues, for instance, employ labels to provide more information about the content of the item without the need for the user to open the post or issue [51, 81]. A similar mechanism could be particularly useful in social audio settings, especially since some contextual information may be unavailable to users who join an event late. In non-ephemeral settings, content remains accessible for users to explore, and tags primarily serve as a summarization tool to expedite navigation. However, in ephemeral social media platforms like real-time audio spaces, information can quickly become unavailable. For example, if a host explains the level of participation or agenda at the start of an event without documenting it, this information will not be accessible to users who join midway unless the host repeats it

Given the host's need to manage participation and encourage contribution, the interface provided for hosts could provide different types of information in line with these needs. Speech contribution visualization is one example of the type of information hosts might find useful which has been explored in in-person [10] and online video-conferencing platforms [66]. For instance, the Conversation Clock designed to visualize turn-taking can be transformed into a radial timeline, where each segment represents a speaker's contribution [10]. Participation histograms that dynamically update to show speaking time could also be used to influence group participation through shared displays [28]. Word clouds [116] could also be used to highlight frequently used terms by speakers, and assist the host in summarizing the discussion. Additionally, such adaptations can enhance the host's understanding of conversation dynamics in social-audio spaces.

7.1.2 Signaling mechanisms: Through our co-design interviews, we also identified instances where users wanted to share information with the host, such as indicating whether they did or did not wish to participate in the verbal discussion. Additionally, it could be useful for users to easily express their level of satisfaction with the discussion. Thus, incorporating signaling mechanisms may add value. Social translucence and signaling theory, concepts previously explored in the HCI literature, provide a foundation for designing such features [30, 60]. Information signaling could be explicit or implicit. Social audio platforms have already implemented some signals such as emoji reactions and hand raising to request speaking privileges which are explicit. Other signals, however, could be developed to speed up communication of needs, help hosts in their decision-making process, and lower the stakes for participation. As some users might feel intimidated to speak or be in an environment they don't feel comfortable speaking in (e.g., a noisy space), such methods could potentially encourage more users to participate. Some design recommendations include status indicators to let users set their status (e.g., "Listening intently", "Have a question", "Agreeing") to give hosts a better sense of audience engagement.

The need for mechanisms that could allow users to have lower-stakes participation also connects with prior work on legitimate peripheral participation [32, 43, 45, 58, 70, 85]. Legitimate peripheral participation "concerns the process by which newcomers become part of a community of practice" [70]. This concept has been studied in the context of learning [123] and participation in online communities [32] with suggestions for tools to support passive participation. New users typically begin their participation in online spaces through activities that do not require them to expose themselves [85]. Social-audio platforms, however, introduce two challenges: (1) limited availability of participation methods, and (2) how exposed users might feel speaking and revealing their voice. Future work could investigate text-to-voice and voice-modification technologies in the context of voice-based events [59, 112]. Other platforms have explored ways of reducing the weight of contributions to encourage newcomers. Halfaker et. al. [45] for instance designed a system to elicit lightweight contributions from Wikipedia readers. Tools to assist the transition from lurking [75] to contributing could be designed in voice-based spaces as well.

7.1.3 Managing Participation. Participants in this study expressed that they often felt excluded from opportunities to participate in events. This was attributed either to the presence of too many individuals in the space or to the hosts' speaker-selection strategies (e.g., only allowing their friends to speak). Previous research has suggested developing mechanisms to divide virtual participants into smaller groups to facilitate meaningful communication and participation [46, 57]. While these studies discuss different strategies, noting the importance of taking existing relationships, histories, and interests into account [46], our findings also point out the importance of maintaining diversity and differing points of view when creating the smaller groups. Another challenge when it comes to such groupings of users in social-audio events is the potential need for a host to be present within each subgroup. This need is due to the fact that different content would now be produced within each sub-group rather than each group interacting with shared content (such as a shared stream/video). One solution for this need could be designing a way for one person to be able to monitor multiple sessions in parallel. One example of such a tool is "Groupnamics", a visual classroom support tool that helps manage parallel group interactions more effectively [93]. Through the analysis of vocal activity in different rooms, the system is able to identify spaces where intervention might be required [93].

7.1.4 Selecting Speakers Properly. We observed a number of issues when it came to speaker selection in social-audio events. One issue was the difficulty of ensuring the person allowed to speak would not cause harms. Hosts were recommended to take a number of precautions, such as evaluating each user's profile and affiliated social media links prior to granting them speaking privileges. However, this approval process would take time and attention away from managing the session, and could be taxing for the host. Additionally, the information provided on users profiles are generally written by the users themselves and not authenticated, making them easily manipulable.

Reputation systems and badges could be one way of reducing the amount of harm users selected to speak might cause without requiring the host to conduct a deep-dive on their profile. Reputation systems have been explored on other social media platforms such as Reddit [81]. Other than preventing harms, badges could also be developed to indicate viewpoints, participation type, and other factors that might influence the decision making of hosts when selecting speakers. A tool to summarize the contributions made by speakers could be another way of helping hosts select speakers who might offer a new or different perspective [35, 118, 124]. Recommendation of speakers according to specific characteristics could also be a tool that hosts might find useful. As diversity of expertise, opinions, and speakers was important to users, hosts might find value in being recommended speakers that match one or more of these criteria.

Another issue with regards to speaker selection that was brought up, was attendees feeling like hosts were biased and unfair when it came to deciding who was allowed to speak (e.g., hosts only allowing their friends to speak). As a result, hosts could provide transparency in how they select speakers, possibly through guidelines that explain the criteria for selecting speakers. This transparency could build trust and understanding among participants [61].

7.2 Improving Facilitation Practices Through Training

Analyzing our results and noting the significant dissatisfaction with hosts' performances, we can speculate that hosts may either be unfamiliar with best facilitation practices or unable to implement these best practices effectively. One potential direction for future work is to assist hosts in learning and practicing facilitation skills. This support could be provided through recommended or mandatory courses, as well as through practice sessions where hosts can experiment with different strategies. Given the potential difficulty in scheduling live practice sessions, an alternative

approach could involve using AI agents (e.g., leveraging large language models) with which hosts can interact to provide feedback on facilitation utterances in a rehearsal or real-time context. A similar method has been successfully implemented for practicing conflict resolution strategies in prior research [102].

7.3 Conclusions

In this study, we investigated the facilitation challenges faced by social audio platform users. We initiated our research by conducting a survey to rank various facilitation tasks according to their implementation difficulty for hosts. Following this, we carried out co-design interviews to gain deeper insights into current practices and identify existing tensions within the community. Lastly, we analyzed existing training sessions to assess how current teachings may contribute to the observed challenges. Throughout each study, we included diverse user groups, such as listeners, speakers, and hosts, to ensure a comprehensive understanding of the current landscape.

The findings reveal that many users experience dissatisfaction due to a lack of meaningful engagement, self-promotional behavior by hosts, and inadequate tools for managing discussions and conflicts. These issues underscore the need for better facilitation training and tools that can support hosts in creating more inclusive and productive environments. This entails developing features to enable diverse forms of participation, improving speaker management, and providing robust conflict resolution tools. Overall, our study contributes to a deeper understanding of the hosts' role in social audio platforms and offers practical insights for platform designers to foster more rewarding conversations and heightened user satisfaction.

References

- [1] Henrik Horn Andersen, Iben Nelson, and Kare Ronex. 2020. *Virtual facilitation: Create more engagement and impact*. John Wiley & Sons, www.wiley.com.
- [2] Terry Anderson. 2004. Teaching in an online learning context. *Theory and practice of online learning* 273 (2004), 273–294.
- [3] Angwin, Julia and Grassege, Hannes. 2017. Facebook's Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children. <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms>. Online; accessed 26 November 2023.
- [4] Andrew Arshat and Daniel Etcovitch. 2018. The human cost of online content moderation.
- [5] Vicki Assegued. 2018. I'm Not Neutral about Neutrality: I'm Partial to Multi-Partiality. *Am. J. Mediation* 11 (2018), 1.
- [6] Association of University Centers on Disabilities. 2023. HOW TO FACILITATE GROUPS. https://www.aucd.org/docs/urc/Leadership_Institute/2013%20UCEDD%20LI/USB/Thursday/Breakouts/Conflict/facilitator%20handbook%20best.pdf. Online; accessed 18 May 2023.
- [7] Brian Auvine, Mary Extrom, and Betsy Densmore. 1977. *A manual for group facilitators*. Center for Conflict Resolution Madison, WI, Route 1 Box 156, Rutledge, MO 63563.
- [8] Belle Wong, J.D. 2023. Top Social Media Statistics And Trends Of 2023. <https://www.forbes.com/advisor/business/social-media-statistics/>. Online; accessed 26 November 2023.
- [9] Zane L Berge. 1995. Facilitating computer conferencing: Recommendations from the field. *Educational technology* 35, 1 (1995), 22–30.
- [10] Tony Bergstrom and Karrie Karahalios. 2007. Conversation Clock: Visualizing audio patterns in co-located groups. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. IEEE, 78–78.
- [11] Ida Boesten. 2021. *Neutrality: a cornerstone of facilitation practice?* SLU, Dept. of Urban and Rural Development, Uppsala, Sweden.
- [12] Charles F Bond and Linda J Titus. 1983. Social facilitation: A meta-analysis of 241 studies. *Psychological bulletin* 94, 2 (1983), 265.
- [13] Robert P Bostrom, Robert Anson, and Vikki K Clawson. 1993. Group facilitation and group support systems. *Group support systems: New perspectives* 8 (1993), 146–168.
- [14] Jie Cai and Donghee Yvette Wohn. 2019. What are Effective Strategies of Handling Harassment on Twitch? Users' Perspectives. In *Companion Publication of the 2019 Conference on Computer Supported Cooperative Work and Social Computing (Austin, TX, USA) (CSCW '19 Companion)*. Association for Computing Machinery, New York, NY, USA, 166–170. doi:10.1145/3311957.3359478

- [15] Jie Cai and Donghee Yvette Wohn. 2022. Coordination and Collaboration: How Do Volunteer Moderators Work as a Team in Live Streaming Communities?. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 300, 14 pages. doi:10.1145/3491102.3517628
- [16] Jie Cai and Donghee Yvette Wohn. 2023. Understanding Moderators' Conflict and Conflict Management Strategies with Streamers in Live Streaming Communities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 669, 12 pages. doi:10.1145/3544548.3580982
- [17] Sandra F Camacho. 2002. Addressing conflict rooted in diversity: The role of the facilitator. *Social Work with Groups* 24, 3-4 (2002), 135–152.
- [18] John M Carroll. 2003. *Making use: scenario-based design of human-computer interactions*. MIT press.
- [19] Kate Crawford. 2011. Listening, not Lurking: The Neglected Form of Participation. *Cultures of Participation* 1 (March 2011), 63–77. <https://www.microsoft.com/en-us/research/publication/listening-not-lurking-the-neglected-form-of-participation/>
- [20] Kathryn D Cross. 1996. An analysis of the concept facilitation. *Nurse Education Today* 16, 5 (1996), 350–355.
- [21] Ruth Cross and Louise Warwick-Booth. 2016. Using storyboards in participatory research. *Nurse researcher* 23, 3 (2016), 8–12.
- [22] Thomas H Davenport and John C Beck. 2001. The attention economy. *Ubiquity* 2001, May (2001), 1–es.
- [23] Scott Davidoff, Min Kyung Lee, Anind K Dey, and John Zimmerman. 2007. Rapidly exploring application design through speed dating. In *UbiComp 2007: Ubiquitous Computing: 9th International Conference, UbiComp 2007, Innsbruck, Austria, September 16-19, 2007. Proceedings 9*. Springer, 429–446.
- [24] Mariëlle den Hengst and Mark Adkins. 2005. The demand rate of facilitation functions. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*. IEEE, 43c–43c.
- [25] Mariëlle den Hengst and Mark Adkins. 2007. Which collaboration patterns are most challenging: A global survey of facilitators. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. 17–17. doi:10.1109/HICSS.2007.614
- [26] Thiago Dias Oliva, Dennys Marcelo Antoniali, and Alessandra Gomes. 2021. Fighting hate speech, silencing drag queens? artificial intelligence in content moderation and risks to LGBTQ voices online. *Sexuality & Culture* 25 (2021), 700–732.
- [27] Ángel Díaz and Laura Hecht-Felella. 2021. Double standards in social media content moderation. *Brennan Center for Justice at New York University School of Law*. <https://www.brennancenter.org/our-work/research-reports/double-standards-socialmedia-content-moderation> 1 (2021), 1–40.
- [28] Joan Morris DiMicco, Anna Pandolfo, and Walter Bender. 2004. Influencing group participation with a shared display. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) (CSCW '04). Association for Computing Machinery, New York, NY, USA, 614–623. doi:10.1145/1031607.1031713
- [29] Bryan Dosono and Bryan Semaan. 2019. Moderation Practices as Emotional Labor in Sustaining Online Communities: The Case of AAPI Identity Work on Reddit. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3290605.3300372
- [30] Thomas Erickson and Wendy A. Kellogg. 2000. Social translucence: an approach to designing systems that support social processes. *ACM Trans. Comput.-Hum. Interact.* 7, 1 (mar 2000), 59–83. doi:10.1145/344949.345004
- [31] Jennifer Fereday and Eimear Muir-Cochrane. 2006. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods* 5, 1 (2006), 80–92.
- [32] Casey Fiesler, Shannon Morrison, R. Benjamin Shapiro, and Amy S. Bruckman. 2017. Growing Their Own: Legitimate Peripheral Participation for Computational Learning in an Online Fandom Community. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, Oregon, USA) (CSCW '17). Association for Computing Machinery, New York, NY, USA, 1375–1386. doi:10.1145/2998181.2998210
- [33] Andrea Forte, Vanesa Larco, and Amy Bruckman. 2009. Decentralization in Wikipedia governance. *Journal of Management Information Systems* 26, 1 (2009), 49–72.
- [34] Nancy Fried Foster, Nora Dimmock, and Alison Bersani. 2008. Participatory design of websites with web design workshops. *Code4Lib Journal* 2 (2008).
- [35] Nikhil Garg, Benoit Favre, Korbinian Reidhammer, and Dilek Hakkani Tür. 2009. Clusterrank: a graph based method for meeting summarization. (2009).
- [36] Sarah A. Gilbert. 2020. "I run the world's largest historical outreach project and it's on a cesspool of a website." Moderating a Public Scholarship Site on Reddit: A Case Study of r/AskHistorians. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 19 (May 2020), 27 pages. doi:10.1145/3392822

- [37] Tarleton Gillespie. 2018. *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press, New Haven, Connecticut.
- [38] Tarleton Gillespie. 2020. Content moderation, AI, and the question of scale. *Big Data & Society* 7, 2 (2020), 2053951720943234.
- [39] Robert Gorwa, Reuben Binns, and Christian Katzenbach. 2020. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society* 7, 1 (2020), 2053951719897945.
- [40] Anthony F Grasha. 1994. A matter of style: The teacher as expert, formal authority, personal model, facilitator, and delegator. *College teaching* 42, 4 (1994), 142–149.
- [41] Lyndsay Green. 1998. Playing Croquet with Flamingos: A Guide to Moderating Online Conferences. *Non-Journal N/A* (1998), 33.
- [42] James Grimmelman. 2015. The virtues of moderation. *Yale J L & Tech.* 17 (2015), 42.
- [43] Mark Guzdial and Allison Elliott Tew. 2006. Imagineering Inauthentic Legitimate Peripheral Participation: An Instructional Design Approach for Motivating Computing Education. In *Proceedings of the Second International Workshop on Computing Education Research* (Canterbury, United Kingdom) (ICER '06). Association for Computing Machinery, New York, NY, USA, 51–58. doi:10.1145/1151588.1151597
- [44] Oliver L Haimson, Daniel Delmonaco, Peipei Nie, and Andrea Wegner. 2021. Disproportionate removals and differing content moderation experiences for conservative, transgender, and black social media users: Marginalization and moderation gray areas. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–35.
- [45] Aaron Halfaker, Os Keyes, and Dario Taraborelli. 2013. Making Peripheral Participation Legitimate: Reader Engagement Experiments in Wikipedia. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) (CSCW '13). Association for Computing Machinery, New York, NY, USA, 849–860. doi:10.1145/2441776.2441872
- [46] William A. Hamilton, Oliver Garretson, and Andruud Kerne. 2014. Streaming on twitch: fostering participatory communities of play within live mixed media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI '14). Association for Computing Machinery, New York, NY, USA, 1315–1324. doi:10.1145/2556288.2557048
- [47] Gill Harvey, Alison Loftus-Hills, Jo Rycroft-Malone, Angie Titchen, Alison Kitson, Brendan McCormack, and Kate Seers. 2002. Getting evidence into practice: the role and function of facilitation. *Journal of advanced nursing* 37, 6 (2002), 577–588.
- [48] John Heron. 1999. *The complete facilitator's handbook*. Kogan Page Publishers, London, UK.
- [49] Ed Hootstein. 2002. Wearing four pairs of shoes: The roles of e-learning facilitators. In *E-learn: World conference on e-learning in corporate, government, healthcare, and higher education*. Association for the Advancement of Computing in Education (AACE), AACE, Montreal, Canada, 457–462.
- [50] International Association of Facilitators. 2024. Library of facilitation techniques. <https://www.sessionlab.com/library>. Online; accessed 1 January 2024.
- [51] Javier Luis Cánovas Izquierdo, Valerio Cosentino, Belén Rolandi, Alexandre Bergel, and Jordi Cabot. 2015. GiLA: GitHub label analyzer. In *2015 IEEE 22nd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*. IEEE, 479–483.
- [52] Loes Janssen, Alexander P Schouten, and Emmelyn AJ Croes. 2022. Influencer advertising on Instagram: product-influencer fit and number of followers affect advertising outcomes and influencer evaluations via credibility and identification. *International journal of advertising* 41, 1 (2022), 101–127.
- [53] Jason Koebler. 2024. How Facebook Trains Content Moderators. <https://www.vice.com/en/article/43z7gj/how-facebook-trains-content-moderators>. Online; accessed 3 January 2024.
- [54] Shagun Jhaver, Iris Birman, Eric Gilbert, and Amy Bruckman. 2019. Human-machine collaboration for content regulation: The case of reddit automoderator. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 5 (2019), 1–35.
- [55] Shagun Jhaver, Amy Bruckman, and Eric Gilbert. 2019. Does transparency in moderation really matter? User behavior after content removal explanations on reddit. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–27.
- [56] Jialun Aaron Jiang, Charles Kiene, Skyler Middler, Jed R Brubaker, and Casey Fiesler. 2019. Moderation challenges in voice-based online communities on discord. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–23.
- [57] Quentin Jones and Sheizaf Rafaeli. 2000. Time to split, virtually: 'Discourse architecture' and 'community building' create vibrant virtual publics. *Electronic Markets* 10, 4 (2000), 214–223.
- [58] David A. Joyner. 2020. Peripheral and Semi-Peripheral Community: A New Design Challenge for Learning at Scale. In *Proceedings of the Seventh ACM Conference on Learning @ Scale* (Virtual Event, USA) (L@S '20). Association for Computing Machinery, New York, NY, USA, 313–316. doi:10.1145/3386527.3406736

- [59] Wonjune Kang, Margaret A. Hughes, and Deb Roy. 2024. Anonymization of Voices in Spaces for Civic Dialogue: Measuring Impact on Empathy, Trust, and Feeling Heard. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 482 (Nov. 2024), 22 pages. doi:10.1145/3687021
- [60] Ray Karasek III and Phil Bryant. 2012. Signaling theory: Past, present, and future. *Academy of Strategic Management Journal* 11, 1 (2012), 91.
- [61] Norbert L. Kerr and R. Scott Tindale. 2004. Group performance and decision making. *Annu. Rev. Psychol.* 55, 1 (2004), 623–655.
- [62] Sara Kiesler, Robert Kraut, Paul Resnick, and Aniket Kittur. 2012. Regulating behavior in online communities. *Building successful online communities: Evidence-based social design* 1 (2012), 4–2.
- [63] Joel Kiskola, Thomas Olsson, Aleks H. Syrjämäki, Anna Rantasila, Mirja Ilves, Poika Isokoski, and Veikko Surakka. 2022. Online Survey on Novel Designs for Supporting Self-Reflection and Emotion Regulation in Online News Commenting. In *Proceedings of the 25th International Academic Mindtrek Conference* (Tampere, Finland) (*Academic Mindtrek '22*). Association for Computing Machinery, New York, NY, USA, 278–312. doi:10.1145/3569219.3569411
- [64] Mark Kitchen. 2012. Facilitating small groups: how to encourage student learning. *The clinical teacher* 9, 1 (2012), 3–8.
- [65] Alison Kitson, Gill Harvey, and Brendan McCormack. 1998. Enabling the implementation of evidence based practice: a conceptual framework. *BMJ Quality & Safety* 7, 3 (1998), 149–158.
- [66] Steffi Kohl, André Calero Valdez, and Kay Schröder. 2023. Using Speech Contribution Visualization to Improve Team Performance of Divergent Thinking Tasks. In *Proceedings of the 15th Conference on Creativity and Cognition* (Virtual Event, USA) (*C&C '23*). Association for Computing Machinery, New York, NY, USA, 319–324. doi:10.1145/3591196.3596824
- [67] Peter Kollock and Marc Smith. 1996. Managing the virtual commons: Cooperation and conflict in computer communities. *Pragmatics and Beyond new series* (1996), 109–128.
- [68] Steinar Kvale and Svend Brinkmann. 2009. *Interviews: Learning the craft of qualitative research interviewing*. sage.
- [69] Kyle Langvardt. 2017. Regulating online content moderation. *Geo. LJ* 106 (2017), 1353.
- [70] Jean Lave and Etienne Wenger. 1991. *Situated learning: Legitimate peripheral participation*. Cambridge university press, Cambridge.
- [71] Heidi Levitt, Mike Butler, and Travis Hill. 2006. What clients find helpful in psychotherapy: Developing principles for facilitating moment-to-moment change. *Journal of counseling Psychology* 53, 3 (2006), 314.
- [72] Hanlin Li, Brent Hecht, and Stevie Chancellor. 2022. All that's happening behind the scenes: Putting the spotlight on volunteer moderator labor in reddit. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 584–595.
- [73] R. Mason. 1991. Moderating educational computer conferencing. Retrieved on July 20, 2002.
- [74] MICHAEL STELZNER. 2024. How to Start a Clubhouse Room and Moderate Like a Pro. <https://www.socialmediaexaminer.com/how-to-start-a-clubhouse-room-and-moderate-like-a-pro/>. Online; accessed 2 January 2024.
- [75] Michael Muller. 2012. Lurking as Personal Trait or Situational Disposition: Lurking and Contributing in Enterprise Social Media. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA) (*CSCW '12*). Association for Computing Machinery, New York, NY, USA, 253–256. doi:10.1145/2145204.2145245
- [76] David R. Musicant, Yuqing Ren, James A. Johnson, and John Riedl. 2011. Mentoring in Wikipedia: a clash of cultures. In *Proceedings of the 7th international symposium on Wikis and Open Collaboration*. 173–182.
- [77] Sarah Myers West. 2018. Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society* 20, 11 (2018), 4366–4383.
- [78] Tom Nadarzynski, Oliver Miles, Aimee Cowie, and Damien Ridge. 2019. Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Digital health* 5 (2019), 2055207619871808.
- [79] Anika Nissen, Colin Conrad, and Aaron Newman. 2023. Are You Human? Investigating the Perceptions and Evaluations of Virtual Versus Human Instagram Influencers. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (*CHI '23*). Association for Computing Machinery, New York, NY, USA, Article 659, 14 pages. doi:10.1145/3544548.3580943
- [80] Blair Nonnecke and Jenny Preece. 2001. Why lurkers lurk. In *AMCIS 2001 Proceedings*. AMCIS, Boston, MA, USA, 1521–1530.
- [81] Alexandre Parmentier and Robin Cohen. 2019. Learning User Reputation on Reddit. In *IEEE/WIC/ACM International Conference on Web Intelligence* (Thessaloniki, Greece) (*WI '19*). Association for Computing Machinery, New York, NY, USA, 242–247. doi:10.1145/3350546.3352524
- [82] Don Paulsen. 2004. Leadership Essentials: Facilitation Skills for Improving Group Effectiveness. In *Proceedings of the 32nd Annual ACM SIGUCCS Conference on User Services* (Baltimore, MD, USA) (*SIGUCCS '04*). Association for Computing Machinery, New York, NY, USA, 153–160. doi:10.1145/1027802.1027840

- [83] Michelle E Pence and Andrea J Vickery. 2012. The roles of personality and trait emotional intelligence in the active-empathic listening process: Evidence from correlational and regression analyses. *International Journal of Listening* 26, 3 (2012), 159–174.
- [84] Pilling, David and Murgia, Madhumita. 2023. ‘You can’t unsee it’: the content moderators taking on Facebook. <https://www.ft.com/content/afeb56f2-9ba5-4103-890d-91291aea4caa>. Online; accessed 26 November 2023.
- [85] Jennifer Preece and Ben Shneiderman. 2009. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS transactions on human-computer interaction* 1, 1 (2009), 13–32.
- [86] RAVI IYER. 2022. Content Moderation is a Dead End. <https://psychoftech.substack.com/p/content-moderation-is-a-dead-end>. Online; accessed 26 November 2023.
- [87] Sarah T Roberts. 2014. *Behind the screen: The hidden digital labor of commercial content moderation*. University of Illinois at Urbana-Champaign.
- [88] Sarah T Roberts. 2019. *Behind the screen*. Yale University Press, New Haven, Connecticut.
- [89] Robbie Routenberg, Elizabeth Thompson, and Rhian Waterberg. 2013. When neutrality is not enough: Wrestling with the challenges of multipartiality. In *The Art of Effective Facilitation*. Routledge, New York, 173–197.
- [90] Nazanin Sabri, Bella Chen, Annabelle Teoh, Steven P. Dow, Kristen Vaccaro, and Mai Elsherief. 2023. Challenges of Moderating Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (, Hamburg, Germany,) (CHI ’23). Association for Computing Machinery, New York, NY, USA, Article 384, 20 pages. doi:10.1145/3544548.3581329
- [91] Elizabeth B.-N. Sanders, Eva Brandt, and Thomas Binder. 2010. A Framework for Organizing the Tools and Techniques of Participatory Design. In *Proceedings of the 11th Biennial Participatory Design Conference* (Sydney, Australia) (PDC ’10). Association for Computing Machinery, New York, NY, USA, 195–198. doi:10.1145/1900441.1900476
- [92] Elizabeth B-N Sanders and Pieter Jan Stappers. 2008. Co-creation and the new landscapes of design. *Co-design* 4, 1 (2008), 5–18.
- [93] Arissa J. Sato, Zefan Sramek, and Koji Yatani. 2023. Groupnatics: Designing an Interface for Overviewing and Managing Parallel Group Discussions in an Online Classroom. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (, Hamburg, Germany,) (CHI ’23). Association for Computing Machinery, New York, NY, USA, Article 701, 18 pages. doi:10.1145/3544548.3581322
- [94] Charlotte Schluger, Jonathan P. Chang, Cristian Danescu-Niculescu-Mizil, and Karen Levy. 2022. Proactive Moderation of Online Discussions: Existing Practices and the Potential for Algorithmic Support. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 370 (nov 2022), 27 pages. doi:10.1145/3555095
- [95] Charlotte Schluger, Jonathan P Chang, Cristian Danescu-Niculescu-Mizil, and Karen Levy. 2022. Proactive moderation of online discussions: Existing practices and the potential for algorithmic support. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–27.
- [96] Sandor P Schuman. 1996. The role of facilitation in collaborative groups. *Creating collaborative advantage*. Sage, Thousand Oaks, California, USA N/A (1996), 126–140.
- [97] Joseph Seering. 2020. Reconsidering Self-Moderation: The Role of Research in Supporting Community-Based Models for Online Content Moderation. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 107 (oct 2020), 28 pages. doi:10.1145/3415178
- [98] Joseph Seering and Sanjay R. Kairam. 2022. Who Moderates on Twitch and What Do They Do? Quantifying Practices in Community Moderation on Twitch. *Proc. ACM Hum.-Comput. Interact.* 7, GROUP, Article 18 (dec 2022), 18 pages. doi:10.1145/3567568
- [99] Joseph Seering, Geoff Kaufman, and Stevie Chancellor. 2022. Metaphors in moderation. *New Media & Society* 24, 3 (2022), 621–640.
- [100] Joseph Seering, Robert Kraut, and Laura Dabbish. 2017. Shaping pro and anti-social behavior on twitch through moderation and example-setting. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 111–125.
- [101] Joseph Seering, Tony Wang, Jina Yoon, and Geoff Kaufman. 2019. Moderator engagement and community development in the age of algorithms. *New media & society* 21, 7 (2019), 1417–1443.
- [102] Omar Shaikh, Valentino Emil Chai, Michele Gelfand, Diyi Yang, and Michael S. Bernstein. 2024. Rehearsal: Simulating Conflict to Teach Conflict Resolution. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI ’24). Association for Computing Machinery, New York, NY, USA, Article 920, 20 pages. doi:10.1145/3613904.3642159
- [103] Guy Simon. 2020. OpenWeb tests the impact of “nudges” in online discussions. *OpenWeb Blog* (2020).
- [104] Russell Spears and Martin Lea. 1994. Panacea or panopticon? The hidden power in computer-mediated communication. *Communication research* 21, 4 (1994), 427–459.
- [105] Kumar Bhargav Srinivasan, Cristian Danescu-Niculescu-Mizil, Lillian Lee, and Chenhao Tan. 2019. Content removal as a moderation strategy: Compliance and other outcomes in the changemyview community. *Proceedings of the ACM*

- on *Human-Computer Interaction* 3, CSCW (2019), 1–21.
- [106] Susanne Stollkleemann and Martin Welp. 2006. *Stakeholder dialogues in natural resources management: theory and practice*. Springer, Springer Berlin, Heidelberg.
 - [107] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science* 12, 2 (1988), 257–285.
 - [108] H. Tarmizi, G. de Vreede, and I. Zigurs. 2006. Identifying Challenges for Facilitation in Communities of Practice. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'06)*, Vol. 1. IEEE, Piscataway, NJ, 26a–26a. doi:10.1109/HICSS.2006.210
 - [109] David R Thomas. 2003. A general inductive approach for qualitative data analysis. *American Journal of Evaluation* 27, 2 (2003), 237–246.
 - [110] Kurt Thomas, Devdatta Akhawe, Michael Bailey, Dan Boneh, Elie Bursztein, Sunny Consolvo, Nicola Dell, Zakir Durumeric, Patrick Gage Kelley, Deepak Kumar, Damon McCoy, Sarah Meiklejohn, Thomas Ristenpart, and Gianluca Stringhini. 2021. SoK: Hate, Harassment, and the Changing Landscape of Online Abuse. In *2021 IEEE Symposium on Security and Privacy (SP)*. 247–267. doi:10.1109/SP40001.2021.00028
 - [111] Thomas Stackpole. 2022. Content Moderation Is Terrible by Design. <https://hbr.org/2022/11/content-moderation-is-terrible-by-design>. Online; accessed 26 November 2023.
 - [112] TOM SIMONITE. 2021. These Deepfake Voices Can Help Trans Gamers. <https://www.wired.com/story/deepfake-voices-help-trans-gamers/>. Online; accessed 22 January 2025.
 - [113] trainerbubble. 2008. Why conflict can be difficult to resolve. <https://www.trainerbubble.com/why-conflict-can-be-difficult-to/>. Online; accessed 18 May 2023.
 - [114] Kristen Vaccaro, Christian Sandvig, and Karrie Karahalios. 2020. "At the End of the Day Facebook Does What It Wants" How Users Experience Contesting Algorithmic Content Moderation. *Proceedings of the ACM on human-computer interaction* 4, CSCW2 (2020), 1–22.
 - [115] Kristen Vaccaro, Ziang Xiao, Kevin Hamilton, and Karrie Karahalios. 2021. Contestability for content moderation. *Proceedings of the ACM on human-computer interaction* 5, CSCW2 (2021), 1–28.
 - [116] Fernanda B Viegas, Martin Wattenberg, and Jonathan Feinberg. 2009. Participatory visualization with wordle. *IEEE transactions on visualization and computer graphics* 15, 6 (2009), 1137–1144.
 - [117] Fernanda B Viégas, Martin Wattenberg, and Matthew M McKeon. 2007. The hidden order of Wikipedia. In *Online Communities and Social Computing: Second International Conference, OCSC 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007. Proceedings 2*. Springer, 445–454.
 - [118] Lu Wang and Claire Cardie. 2013. Domain-Independent Abstract Generation for Focused Meeting Summarization. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Hinrich Schuetze, Pascale Fung, and Massimo Poesio (Eds.). Association for Computational Linguistics, Sofia, Bulgaria, 1395–1405. <https://aclanthology.org/P13-1137>
 - [119] Qiyun Wang. 2008. Student-facilitators' roles in moderating online discussions. *British Journal of Educational Technology* 39, 5 (2008), 859–874.
 - [120] Dorothy Wardale. 2013. Towards a model of effective group facilitation. *Leadership & Organization Development Journal* 34, 2 (2013), 112–129.
 - [121] Donghee Yvette Wohn. 2019. Volunteer Moderators in Twitch Micro Communities: How They Get Involved, the Roles They Play, and the Emotional Labor They Experience. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3290605.3300390
 - [122] X. 2024. About X Spaces. <https://help.x.com/en/using-x/spaces>. Online; accessed 11 July 2024.
 - [123] Yutaka Yamauchi. 2009. Power of Peripheral Designers: How Users Learn to Design. In *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology* (Philadelphia, Pennsylvania) (DESRIST '09). Association for Computing Machinery, New York, NY, USA, Article 13, 12 pages. doi:10.1145/1555619.1555637
 - [124] Ming Zhong, Da Yin, Tao Yu, Ahmad Zaidi, Mutethia Mutuma, Rahul Jha, Ahmed Hassan Awadallah, Asli Celikyilmaz, Yang Liu, Xipeng Qiu, et al. 2021. QMSum: A new benchmark for query-based multi-domain meeting summarization. *arXiv preprint arXiv:2104.05938* (2021).

Received January 2024; revised July 2024; accepted October 2024