# "The World is Designed for Fluent People": Challenges, Benefits, and Opportunities of Videoconferencing Technologies for People Who Stutter

SHAOMEI WU, Almpower.org, USA

This work studies the experiences of people who stutter (PWS) with videoconferencing (VC) and VC technologies. Our interview study with 14 adults who stutter uncovers extra challenges introduced by current VC platforms to people who stutter. While some of the challenges are a direct result of the characteristics of stutter (e.g. people/systems mistaking pauses as end of turn), a bigger yet less visible challenge comes with the significant amount of emotional and cognitive efforts required to manage one's speech and identity over VC, in which people's existing communication strategies - such as body language and eye contact - are under-supported and their biggest discomfort - such as seeing oneself stutter - are exacerbated by preset features like self view. Overall, our work sheds light on the structural barriers and the opportunities for PWS to engage and enjoy virtual communications via VC technologies.

CCS Concepts: • Computer systems organization  $\rightarrow$  Embedded systems; Redundancy; Robotics; • Networks  $\rightarrow$  Network reliability.

Additional Key Words and Phrases: datasets, neural networks, gaze detection, text tagging

#### **ACM Reference Format:**

 

#### 1 INTRODUCTION

Stuttering is a complex neurodeveopmental condition that affects approximately 1% of the world's population [32]. Traditionally considered as a speech *disorder* characterized by atypical speech behaviors such as sound repetitions, prolongations, and speech blocks [32], recent research on stuttering has underscored its emotional and cognitive impact to people who stutter [3, 40]. Stuttering comes with substantial social penalties, including negative listener reactions, bullying and teasing, social harm and rejection, and stereotypes of being less intelligent, less capable, less attractive, less socially competent, and more anxious than fluent speakers [4, 8, 10, 12, 14, 17, 40, 45]. As a result, people who stutter (PWS) often develop strong emotional and cognitive reactions towards stuttering, including feelings of fear, guilt, shame, helplessness, social anxiety, self-stigma, as well as avoidance of certain sounds, words, situations, people, and relationships. Despite the success of a few notable people who stutter - such as President Biden, research shows that people who stutter are structurally disadvantaged and have an overall reduced quality of life: stuttering is associated with less satisfying personal relationships, higher risk of mental health problems, lower educational attainment, underemployment, and 20% - 35% reduced earnings compared to people who do not stutter [10, 20, 40].

Rooted in ableism and stigmatization, the marginalization of people who stutter is often exacerbated by and materialized through communication technology. For example, as many people who stutter find phone calls more difficult than in-person conversations, phone interviews and phone conversations at work create barriers to employment for people who stutter [20, 24]. As we enter a new era in which videoconferencing becomes the dominant and normalized mode for interpersonal and professional communications, it is crucial to understand its impact on people who stutter. Despite its widespread adoption, videoconferencing comes with unique challenges, such as the reduction of non-verbal

Author's address: Shaomei Wu, shaomei@gmail.com, AImpower.org, Mountain View, California, USA.

2022. Manuscript submitted to ACM

cues [2, 31], turn-taking confusion, connectivity/technical difficulties, and generally "Zoom Fatigue" [2]. While most of these challenges are also experienced by the general population, a recent study showed a disproportional effect of such challenges on people with aphasia that made it even more challenging for them to stay connected during the pandemic [31]. We thus hypothesize that people who stutter also face greater challenges with videoconferencing, and such challenges impact not only communication efficiency but also the social-emotional wellbeing of people who stutter. In this work, we explore the experience of people who stutter with videoconferencing technologies through interviews with adults who stutter. The interviews and data analysis were conducted to understand the challenges, benefits, and strategies for people who stutter during video conferences, in comparison to in-person meetings.

This paper provides a unique contribution to HCI and accessibility research by presenting to our knowledge the first formal study of the lived experiences of people who stutter with video conferences and videoconferencing technologies. Distinct from previous work on videoconferencing challenges for the general population [2, 18, 28], our study emphasizes the emotional and cognitive impact of such challenges beyond observable outcomes such as communication accuracy and efficiency. Rather it examines how videoconferencing technology interacts with core elements of stuttering experiences such as stigma, avoidance, and acceptance. In this way, our study uncovers challenges that are unique to PWS - such as signaling the intention to speak by speech itself, audience confusion between stuttering and connectivity issues, as well as known videoconferencing challenges that have a disproportional effect on PWS - such as the lack of connection with audience and the mental stress with seeing self on camera. While most of our participants reported spending extra effort to participate in video conferences due to such challenges, they also identified various benefits of videoconferencing for PWS, including the increased connectivity within the stuttering community, and general public empathy towards communication disruptions. Taken together, our work sheds light on the structural barriers and the opportunities for PWS to effectively communicate and emotionally connect with others via videoconferencing.

# 2 BACKGROUND

### 2.1 Stuttering

 Stuttering is a genetic, neurodevelopmental condition that impacts people who stutter in behavioral, emotional, and cognitive aspects [3]. Illustrated by the "stuttering iceberg" analogy, the observable behaviors associated with stuttering, such as speech disruptions and facial tension, are only the "tip of the iceberg", and most of the affective reactions to dysfluencies - including fear, guilt, shame, and helplessness - lie below the surface [38]. These negative affective reactions often lead to an increased level of social anxiety and self-stigma, and over time develop into cognitive reactions such as "avoidance strategies" to cope with the fear with stuttering and pass as fluent [3].

The hidden nature of these emotional and cognitive challenges also leads to limited visibility and public awareness of struggles and needs of the stuttering community. Despite clearly documented social and occupational disadvantages associated with stuttering, people who stutter are rarely offered or ask for reasonable accommodations, such as extra speaking time to account for unpredictable blocks [10]. As a result, people who stutter often need to go through interview processes with disabling barriers or get passed over on career opportunities that involve verbal communications [7]. This work aims to contribute to the public knowledge on the experience of people who stutter - and the barriers they face - in professional and social communications mediated by videoconferencing technologies.

Our research is also inspired by the recent breakthrough in stuttering research and therapy that emphasizes the subjective experience of stuttering rather than the perspectives and observations of the listeners [11, 40]. This epistemic shift led the field to understand that the biggest struggle with stuttering moments is not the dysfluencies but the feeling

of "being stuck" and "losing control" [33, 40], and people who stutter find it most satisfying when their speech is spontaneous, regardless of how fluent it is [11]. Combined with the theoretic framework from the social model of disability, these insights empowered the stuttering community to push back on the listener-oriented, fluency-focused notion of stuttering as an impaired, undesirable form of speech, and advocate for the right to stutter in pubic life [10]. Following a similar approach, we collect and study first-person accounts and reflections from people who stutter, of their experience with video conferences and videoconferencing technologies, to unpack the emotional and cognitive impact imposed by these technologies on people who stutter, beyond what is typically observed by third-parties, such as other meeting participants, employers, and communication researchers. Our findings highlight the significant hidden costs people who stutter have to pay in order to effectively participate in video conferences, calling for structural changes in videoconferencing technologies as well as the meeting culture to create a more inclusive and empowering digital communication environment for all.

### 2.2 Assistive Technologies for stuttering Speech

Despite the prevalence of stuttering [32], little attention and research efforts haven been devoted to understanding and improving the experiences of people with speech diversity with technologies. A search of the keywords "stutter" or "stammer" over the proceedings of CHI conferences in the past 10 years ('12 - '22) only returned 6 papers, and among them, only one was tangibly related to stuttering [46].

Although stuttering has been neglected by technological research, technology has played an active role in the marginalization of people who stutter. For example, as automatic speech recognition (ASR) systems become more and more prevalent through smart speakers and automated phone menus, people with stuttering speech suffer a degraded performance with such systems comparing to people who are fluent [30] and can be denied access to opportunities and services as a result.

While there have been some efforts improving the experiences of ASR systems for people with atypical speech, most these improvements were evaluated over static speech datasets using standardized metrics (e.g. word error rate, WER), without formal evaluations with users who stutter in real world scenarios [29, 30]. As a result, little is known about the user experience of these adapted systems. This disconnection between subjective user experience and ASR system evaluation is particularly alarming in our context, given the recent movement by the stuttering community to take agency in one's speaking experiences [19, 40].

On the other hand, current technical products designed for people who stutter tend to focus on "fixing" or "masking" stuttering speech rather than embracing it. They either ask a user to practice traditional fluency shaping techniques over a software (e.g. Stamurai<sup>1</sup>, BeneTalk<sup>2</sup>), or try to induce a temporal, non-stuttering voice of their users (e.g. SpeechEasy<sup>3</sup>, Whispp<sup>4</sup>). These products profit from the fear of stuttering and reinforce the notion that stuttering is unacceptable - something that the stuttering community are trying to push against - and have not been widely adopted by the community, perhaps unsurprisingly.

Overall, we see a tremendous gap in the technical research and investigation in the space of stuttering, and hope this work contribute to the literature by uncovering the needs and opportunities for technologies to empower the stuttering community, in a communication environment that is increasingly shaped and defined by digital technologies.

<sup>&</sup>lt;sup>1</sup>https://stamurai.com/

<sup>&</sup>lt;sup>2</sup>https://www.benetalk.com

<sup>3</sup>https://speecheasy.com

<sup>4</sup>https://whispp.com/

# 2.3 Videoconferencing and Its Challenges for Communication

Since its first introduction of this concept [1], videoconferencing and videoconferencing technologies have played a more and more significant role in interpersonal communications. Despite several well-known issues with videoconferencing such as the challenge with maintaining eye contact [37, 42] and surveillance effect of the camera feed [6, 16] - videoconferencing was widely adopted during the Covid-19 pandemic, and has since then become an increasingly dominant channel for professional and personal communications [21]. Recent work [2, 28] have examined the phenomenon and causes of "Zoom Fatigue", a term used to described the exhaustion as a result of excessive videoconferencing. [2, 18, 28] outline the major challenges with videoconferencing and contemporary videoconferencing technologies, including: nonverbal overload, close-up gaze, extended screen time, reduced mobility, and self-view.

Some recent work studied the use of videoconferencing technologies for communities with atypical communication patterns. For example, [31] investigated the experience and challenges of people with Zoom videoconferencing platform, uncovering additional challenges for this demographic introduced by the lack of non-verbal communication channels. [26, 36, 41] examined the accessibility of videoconferencing for d/Deaf and hard of hearing (DHH) people, highlighting unique challenges such as difficulty in lip reading, the design of identifying active speaker based on voice, and the additional delay between the signer and the ASL interpreter.

Our work contributes to this area of research by presenting the experiences of videoconferencing by people who stutter - a population that have traditionally faced systematic challenges in in-person communications. We hope our insights will uncover unique challenges for the stuttering community and inform the design and development of a more inclusive video-based communication environment for all.

# 3 METHOD

 We conducted semi-structured interviews with adults who stutter from the US and UK to learn about their experience of videoconferencing.

The participants were recruited directly by the author(s), through speech therapy groups, stuttering community events, and social contacts. We continued the study recruitment process until consistent high-level themes emerged. The interviews took place between February 2022 and August 2022 over Zoom videoconferencing system, and were recorded with explicit consent from the participants for later transcription. Understanding the multiple forms of suppression at play during professional and public communications, we prioritized the recruitment and inclusion of participants with multiply marginalized identities besides stuttering, such as, women, first-generation immigrants (P1, P4, P5, P6, P10), and people of color (P1, P3, P4, P5, P6, P8, P10). Cognizant of the potential challenges and discomfort for people who stutter to speak over Zoom, we made conscious efforts to create an accommodating environment for participants to speak freely and comfortably. For example, for multilingual participants, we conducted the interview in their preferred language (e.g. P4 in Mandarin Chinese) rather than in English. The transcription was later translated by one of the authors for analysis and reporting purposes. While respecting the participant's time, we always left ample time after the scheduled interview slot and let the participants know that there was no time pressure for them to speak fast or concisely. Additionally, all but two (P1, P10) interviews were conducted by a research team member who stutters and had moments of stuttering during the interviews. As a result, the length of the interviews varies, lasting between 45 mins to 1.5 hours. The participants were not compensated.

The interviews were structured with the following four components.

217

224

225

234 235

259

260

Self-identified ID Gender Country Occupation Stuttering Characteristics P1 F US Covert, filler words, blocks, word substitution, loss of eye contact Software engineer Covert, blocks, word substitute, raising heart rate, sweat P2 M UK Professional athlete F UK P3 Ophthalmologist resident Covert & mild, avoid sounds, reorder words P4 M US Postdoc researcher Blocks, word substitute, struggle more with starting vowels P5 F US SLP grad student Blocks, some repetition P6 M US Product designer Mild to moderate, blocks, backtracking, word substitution P7 F US Preschool teacher Blocks, facial tension P8 M US Medical student Filler words, pauses, tongue clicks F P9 UK University staff Covert, blocks, some repetition, word substitution F P10 US PhD student Blocks, repetition, some prolongation, facial tension P11 M US UX researcher Covert, speak slowly, word substitution P12 F US Geospatial Analyst Covert, mild, blocks on names P13 M US Program manager Blocks, word substitution, look away when stuttering P14 F UK Stutter openly with little struggle SLP

Table 2. VC Experience and Technology Use

ID	VC Frequency	Platforms
		(most to least used)
P1	Several times a day for work; weekly with family & friends	Zoom, Google Meet, MS Teams, WeChat
P2	4-5 times a week for work; daily with family & friends	Zoom, Microsoft Teams, Facetime
P3	Several times a week for work & school	Zoom, MS Teams, Google Hangout
P4	Several times a week for work; weekly with community	Zoom
P5	All classes for master program; weekly with community	Zoom, Google Hangout
P6	Several times a day for work	Zoom, Slack calls, MS Teams
P7	Daily for 3 months for work, weekly for therapy & church	Zoom, Google Meet
P8	1-3 times a day for work, several times weekly with community	Zoom, Google Meet, MS Teams
P9	Several times a day for work, weekly for therapy	Zoom, WebEx
P10	Several times a day for work & school	Zoom, weekly for church, FaceTime, WebEx
P11	Several times a day for work	Zoom
P12	Daily for work	MS Teams, WebEx
P13	Several times a day for work	Zoom
P14	Several times a week for work	Zoom

(1) Personal background and characteristics of one's stuttering. Stuttering is not monolithic, and our participants described their stuttering in terms of the speech and non-speech challenges with stuttering, impact of stuttering, coping behaviors, and current attitude towards stuttering. This helped us assess the representativeness of our participants to the stuttering community, and understand the challenges with videoconferencing in relation to the speech/behavior, affective, and cognitive aspects of stuttering. The key data collected from this part of the interview is reported on Table 1 <sup>5</sup>.

<sup>&</sup>lt;sup>5</sup>Several participants described their stuttering as "covert", a term used in speech therapy and stuttering research to characterize the type of stuttering with little or no disfluencies that can be effectively passed as fluent speech to the listener [?]. It is equivalent to "interiorized stammering" in the United Kingdom, and we use these two terms interchangeably.

(2) Use of video conferencing technologies. We asked about the frequency of videoconferencing, top use cases for

Shaomei Wu

technologies used (e.g. Zoom, Google Meet, Microsoft Teams, Skype, Facetime, etc).

(3) Experience of videoconferencing. We asked about participants general experience of videoconferencing in comparison to in-person meetings: whether they find one more challenging than the other, and in what situations. We also asked the participants to reflect on the top challenges and benefits they experienced with videoconferencing, as well as the role of stuttering in those experiences. We also inquired about the strategies participants have developed to manage their videoconferencing experiences.

videoconferencing (e.g. school, work, community, friends & family), as well as the types of videoconferencing

(4) Future of videoconferencing. We brainstormed with the participants for technical or non-technical ways to make videoconferencing experiences easier and more comfortable for them and/or for the stuttering community in general.

The interviews were transcribed and analyzed using inductive qualitative methods drawn from grounded theory [?]. The authors reviewed the interviews and performed thematic analysis for each interview. The authors then discussed the themes identified in individual interviews and used an affinity diagram to organize them into different areas of the findings we discuss below.

### 4 FINDINGS

Overall, all participants had used videoconferencing in professional and/or personal settings (see Table 2 for summary) and reported various degrees of satisfaction with their video conferencing experiences. While most of the participants agreed that video calls are easier than phone calls with no video, the preference for video conferencing versus in-person meetings varied, depending on the meeting context as well as the amount of experience participants had with video conferencing and the video conference platforms. For some, video conferencing is more suitable - or most similar to in-person experiences - for small groups or 1-on-1 conversations (P3, P4, P7, P9); and for others, it is easier to have larger meetings or public presentations over video conferencing platforms than in-person (P2, P6, P11, P13). While several participants (P3, P4, P10, P12) cited the lower expectation for active participation as a benefit of video conferencing, some (P1, P13) felt more comfortable when they served an active role with control over a meeting's agenda, content, and norms.

As the trend with videoconferencing persists, our participants' perception of it evolved as well. For those who only started serious videoconferencing since the pandemic (P2, P3, P4, P5, P7, P9, P14), they found their experience improved over time.

Video calls were a nightmare for me, at least at the beginning. But now I am getting used to this nightmare. (P4)

At this point, I think I am about 50:50 [Zoom: in-person]. Just the amount of time we used it, just the practice we have had [since the pandemic] (P9)

However, the participants with the most remote work and video conferencing experience found themselves much more accustomed to speaking in video calls, sometimes at the cost of lower comfort level with in-person interactions:

With work, on VC, I stutter much less because I got used to. I do stutter, but much less. In person, just an aspect that we haven't had too much in person calls or conversations recently. Because I work on software, I have been working remotely for the past few years, I lost some muscle memory of that kind of interactions.

[...] I haven't had too much opportunities to interact with people in person recently.[...] Having small talks with people in person, as an introvert, that was difficult. (P6)

In terms of me leading a meeting, or facilitating something, events like if I'm in the hot seat, talking on a panel, or something like that, I, at this point - that I never would have said this before the pandemic - but I would actually rather do it virtual.[...] I actually don't have a lot of experience facilitating, or panel, in person, because a lot of those opportunities came to me during the pandemic. The idea of doing a live TED Talk freaks me out, but I've just done a half hour presentation over the computer, and I loved it! (P13)

In the rest of this section, we will describe the benefits of videoconferencing as identified by our participants, the challenges they experienced, and the strategies they have developed to alleviate these challenges. All the benefits and challenges of videoconferencing were discussed in comparison to their in-person equivalent, unless otherwise specified. Although the majority of our participants draw their insights heavily from virtual meetings at work or school - a setting where most of their video calls took place, many of the findings are generally applicable to other context such as social and community activities.

#### 4.1 Benefits

 Although the usage and context for videoconferencing varies, all of our participants saw some benefits of videoconferencing, throughout the Covid pandemic and extending into the future. While all the benefits noted by our participants are applicable to people who do not stutter, some of these benefits are particularly appreciated in the context of stuttering, which we will highlight in the subsections below.

4.1.1 General Videoconferencing Benefits. Similar to everyone else who has gone through the pandemic and experienced the shift towards online communications, our participants appreciated the utility of videoconferencing in maintaining social connections and empowering distributed collaboration across geographical boundaries. As expressed by P9, video conferencing "is a vital way to keep in touch with everybody, especially during lockdown". And P11 mentioned, "I am a fan of the COVID shift, I think it has opened the talent pool to much broader geographies... It has democratized the access to talent."

Like the general public, our participants noticed the trend in remote and hybrid work, accepting the fact that videoconferencing would be an increasingly important channel for human communications: "Zoom, and virtual, and hybrid working, is never going away. There is still collaboration across coasts, across countries, that is not gonna go away" (P13). In any case, videoconferencing was preferred by our participants over audio-only phone calls - a situation noted as most challenging by many people who stutter [24], including our participants. Our participants chose to embrace videoconferencing - together with its challenges - in the post-pandemic era, and even considered it an opportunity to redefine meeting dynamics and norms. For example, P13 has been hosting training on virtual meeting etiquette and best practices at his workplace, leading the effort to build a more inclusive, humanized communication environment over videoconferencing:

I want everyone to use the same, or similar practices, when it comes to communicating through the computer. Because if we all use the same, best practices, everyone is better off, not only the people you presenting to. [...] If I model behavior, it's like, "I like he did that!", that might trickle to the next meeting. (P13)

4.1.2 More Control on One's Environment. People who stutter often have more speech difficulties in new and unfamiliar environments [15]. Videoconferencing makes it possible for people who stutter to participate in a wide range of speaking situations while staying in a consistent, familiar environment that is customized to their preferences and comfort. Many of our participants appreciated this aspect of videoconferencing and noted its positive effect on their speech and emotional wellbeing.

Several of our participants spent intentional efforts setting up the physical environment they video conference in, as a way to achieve psychological comfort and eliminate potential stress brought by stuttering:

I feel comfortable, I'm in my house, I'm in my chair, I'm very comfortable with my setup here. I have two screens. (P13)

I certainly try to create a working environment that feels very positive to me, so choosing a blouse I like, a shirt, or doing my hair how I like it, or choosing artwork in my office that makes me feel comfortable. It's a way I can shape my environment to make me feel more comfortable, and also feel like I have a personality that I am presenting. So that you are not focusing on my stutter. (P12)

I can manage my energy a little bit better on VC, because you are in your own environment. For people who stutter, going to a bar is very challenging, the office can have a similar effect.[...] you just have more control on VC than in person environment. (P6)

The physical distance with other videoconferencing participants also created a psychological "safety distance" that protected our participants. As expressed by P5, "on Zoom, you see the person, but that person is still some distance. I feel it's still safer for me. It's less intimidating to meet person on Zoom that in person."

Besides the control over their physical environment, videoconferencing technology also allowed the participants to customize their communication environment. By adjusting the position, layout, and size of the display of their audience and conversation partners, several of our participated indicated gaining the sense of "in control", especially when speaking to authority figures and in bigger groups - settings that are generally more challenging for people who stutter [15]. As P6 explained: "the other thing with VC is that it can be one or with a thousand people, you can change how many people you view on the screen. VC makes it easy for you to speak to a larger group. [...] I will be more anxious in person for larger group meetings, maybe even holding the mic is a different dynamic, versus on Zoom, it doesn't feel too much different to me, one person or a thousand people".

4.1.3 More Control on One's Self Representation. Beyond the control over their own environments, our participants also believed that video conferencing technologies have provided them more control over how they were perceived by others as well

In additional to speech disfluencies, people who stutter often struggle with stuttering associated secondary behaviors, such as facial tension, flushing, sweating, and sudden head and body movements [3]. While most of these behaviors are involuntary, they can trigger negative social reactions due to the stigma towards stuttering [3]. Several of our participants called out the benefit of videoconferencing as to hide their secondary behaviors more easily when they stuttered. For example, P1 perspires more when she stutters, and often worries that others would see or smell her sweat at in-person meetings. Videconferencing largely eliminated this worry for her, so now she had one less thing to worry about during meetings. Similarly, P10 noticed, "I sometimes get really shaky if I'm presenting in person. By if I'm presenting on Zoom, I am already sitting down so it's okay. And nobody can see me even if I'm shaking".

Manuscript submitted to ACM

"The World is Designed for Fluent People": Challenges, Benefits, and Opportunities of Videoconferencing Technologies for People Who Stutter

Videoconferencing also allows our participate to utilize existing or new strategies to manage their speech, without drawing unnecessary attention to their struggles. For example, both P1 and P12 mentioned the strategy of rehearsing what they wanted to say beforehand with both microphone and video turned off. P12 gave the example of rehearsing her self-introduction in a video call:

When I know I have to introduce myself, but I am a little nervous because it's a new group, or in front of leadership, and so I want to impress them. So I won't have my audio on, I won't have my video on, I will just like quietly say to myself and I'll practice my breath work with it. [...] I try to do that, so that I can add the skill set I have to try to mitigate any disfluencies, [...] just to make myself more confident. So when that [self-introduction] does happen, I can say, okay, I think this is what I need to do to get through this. I will probably do that, like, 30 seconds before, or even before the meeting starts, so I feel a little more confident. (P12)

Other speech and identity management strategies that worked particularly well over videoconferencing include squeezing a stressball (P1), tapping the feet (P1, P6), turning off the camera when struggling with facial tension (P6, P12), blaming the Internet connection for speech blocks and long pauses (P2, P3, P11). and using fluency-inducing technologies (P4). Our participants were acutely aware of the potential biases and discrimination against their stuttering, as well as other marginalized identities of theirs, and appreciated the possibility to blend in with the majority through their self-curated virtual presence:

In my mind, it's easier to control how I am being perceived when all you can see is a square. [...] I'm 5'3, I am a small Caucasian woman, there is nothing remarkable about me. [...] I'm not tall, and I'm not a man. A lot of my colleagues are men, my clients are men, [...] The leadership are typically male, so I want to look like tall when I try to introduce myself. That - in my mind - creates a little bit more of an even playing field. If the only thing you can see about someone is their background and they are out there, I think that kind of makes me feel more comfortable. It's a very controlled setting. (P12)

P4 used DAF Pro, an smartphone App that enables the user to hear themselves speaking with a 60ms delay, for virtual job interviews, and managed to have perfect fluency during the interviews. However, this strategy did not come without a cost. After starting on the job and experiencing the effect of delayed auditory feedback (DAF) worn out - a common phenomenon for DAF users, P4 struggled with not only speech difficulties, but also the feeling of embarrassment and guilt at the workplace:

I felt a little bit guilty that I spoke very fluently during the interviews by using the DAF app, but showing stutter afterwards. I don't want my supervisor to think that I cheated. I wanted to be an honest person, but during the interview I did want to cover up the fact that I stutter, so I didn't disclose during the interview, and I was also a bit worried that my stuttering would impact whether they give me the offer. I am a bit embarrassed now. (P4)

4.1.4 Reduced Barrier to "Show Up". Research has shown that adults who stutter suffer from heightened social anxiety and are more likely to avoid social situations as a result [22]. The avoidance behaviors (e.g. not showing up, avoiding speaking, avoiding eye contact) would then elicit negative responses from others and reinforce existing social anxiety [34]. Modern speech therapy research and practice has accumulated evidence that by reducing avoidance behaviors, people who stutter can break from this vicious cycle to live and speak much more comfortably [9, 22]. Videoconferencing has reduced the barriers to join meetings. P8 noted "the ease of joining" as one of the biggest advantages of Zoom and

commended Zoom for "make it easiest possible for people to find and join meetings". The convenience of videoconferencing has also greatly reduced the physical labor for the participants to attend to social events and meetings that they would otherwise pass over:

I can definitely see that as a good tool to kind of help me show up to places. There were so many times [in the past] that I just didn't show up to things. [...] I feel that [Zoom] is an advantage cuz you don't have to get to your car and travel [...] I think for me, because you don't have to travel, you don't have to dress up, just dress up top of your body and put whatever pants on underneath, it just feels easier to say that I can come. (P7)

While stuttering is often a very isolating and heavily stigmatized experience, our participants were able to leverage videoconferencing platforms to connect with others who stutter, and build communities that are safe yet supportive. P5 shared that her journey of self-acceptance, started with a weekly global support group meeting over Google Hangout that she participated in without turning on her camera or microphone for several months, before her "coming out" as a person who stutters and eventually hosting the Google Hangout support group. For PWS who experience social and self stigma towards stuttering, the anonymity and passive participation afforded by videoconferencing technologies pave the way for them to take the first step of reaching out to the stuttering community. Inspired by her own journey, P5 started a Chinese-English bilingual stuttering support group over Zoom, with people joining from China, US, and the Europe. While almost all the attendants from China joined with a pseudonym and camera off at the beginning, P5 noticed that people gradually became more comfortable showing their faces and introducing themselves after a few calls. For P5, "Zoom play an important part in my education and in my healing journey", and this sentiment is shared by many other participants of our interview study. P1, P4, P9, and P12 started acceptance-based speech therapy and meeting other people who stutter since the pandemic, over videoconferencing platforms. P8 remarked the community connections that were facilitated through videoconferencing: "the world has become more and more connected. Now, on a regular basis, I talk to people across the world, because people are used to Zoom. I know people across the world who stutter, that is a blessing."

4.1.5 General Trend towards More Inclusive Meeting Behaviors. As COVID-19 disrupted lives and blurred work-life boundaries, research also showed that people have developed increased empathy towards others since the outbreak of the pandemic [?]. In the context of video conferencing, our participants noticed that people have become more patient and more understanding with speaking-related challenges, which alleviated some pressure for people who stutter to participate fluently. P7 felt "people are more used to things like pauses". And P3 noticed, "even fluent speakers have difficulties on Zoom, [...] having challenge of being heard is more understood now".

Most videoconferencing technologies, as well as the telework environment, have also provided more options to participate in meetings. All but two participants mentioned the "hand raising" function in Zoom, and found it effective at getting people's attention when it was understood and enforced as the norm. Chat is another feature that some participants appreciated. As P7 explained, "try to get into a conversation with your voice would be difficult; using the chat, it can be a good way to ease into that". Compared to in-person conversations, P2 felt it was helpful to have the option to type in the chat, as "in some way you can avoid having to speak, it does give you that option; whereas when you are face to face, [...] it's be a bit strange to text them, or email the message.". However, several participants also found the chat "ancillary"(P11), "unnoticable"(P3), and "distracting"(P4), and would not participate through chat unless someone was actively monitoring and addressing it (P1, P3, P10, P11). The reaction feature in Zoom was not as popular, although a

Manuscript submitted to ACM

469 470

471

472 473

474

475

476 477

478

480

481

482

483 484

485

486

487

488 489

490

491

492

494

495

496

497

498 499

500 501

502

503

504

508

509 510

511

512

513

514 515

516

517

518

"The World is Designed for Fluent People": Challenges, Benefits, and Opportunities of Videoconferencing Technologies for People Who Stutter

few participants (P8, P10, P13) did use it as a way to actively participate in virtual meetings. Besides the functionality of videoconferencing platforms, participants also made use of the asynchronous communication channels:

In distributed work, there are other ways for you to speak up - quote unquote "speak up". You can speak up in documents, you can speak up in posts. I think that's extremely valuable. I think it is definitely something I have leaned on more. [...] There are just more options to speak up in different ways. (P6)

Overall, our participants saw a cultural shift towards more inclusive meeting expectations and behaviors that empowered people who stutter to speak up:

10 years ago, it was perfectly acceptable to just have one person speak in the entire meeting; but now, if there is only one person speaking, I will definitely call it out. (P6)

# 4.2 Challenges

Despite benefits, videoconferencing and videoconferencing technologies also introduced additional challenges for people who stutter to engage and participate in meetings and conversations. While some of the challenges are a direct result of speech related communication difficulties (e.g. people/systems mistaking a pause as the end of turn), the bigger part of the challenge comes from the significant amount of emotional and cognitive effort required to manage one's speech and identity over voice-centric platforms like Zoom, in which our participants's existing communication strategies - such as body language and emotional connections with the audience - are under supported, whereas their struggle and embarrassment with stuttering were directly exposed by the zoom-in view of one's face and the preset feature of self-view. As P5 pointed out:

On Zoom, your voice is so much important for you to communicate than before. They can not see your body, your gestures, your words carry more meaning, you have to impress people with your words, for someone who stutters, that's a disadvantage. (P5)

Such challenges not only limited the participation of PWS during meetings, but also created emotional and cognitive burdens that reinforce some of the most negative impact of stuttering, preventing PWS from engaging and enjoying communications via videoconferencing technologies. P7 mentioned, "I have the same [social anxiety] in person, but it is heightened in VC." And P2 described his experience with video conferencing as "exhausting" and "not rewarding". While that exhaustion from VC meetings was shared among many people who do not stutter [2], P2 noted that the exhaustion was exacerbated for people who stutter as "I have to focus so much on trying to say the right thing at the right time and then also alongside managing the stammer".

4.2.1 Self-view. Numerous studies have shown seeing oneself in a mirror can induce self-evaluation and distress [2, 18, 43], and the effect is stronger for certain social groups such as women and Asian, as comparing to men and White, respectively [35]. Not surprisingly, the "self view" function - turned on by default on most video conferencing platforms - stands out as one of the top challenges with videoconferencing in our data. Almost all of our participants indicated some discomfort with the self-view, finding it "distracting" (P1, P7, P9, P19, P14) and anxiety-triggering. P7 related the self-view with traumatizing experience in the past when she had to watch her own recorded speech at speech therapy. P5 considered self-view as a "curse" in which "you have to face your worst fear as a person who stutters", noting that "before, when you talk to a person, you don't necessarily see your own face; seeing your own face puts the person who stutter in a direct confrontation with your stuttering".

 12 Shaomei Wu

Although the challenge with self view is not unique to people who stutter [2], our participants reported heightened mental stress when confronted with the self-view in video conferences. As P11 illustrated, "during COVID, people would share that, for the first time, they were very self-conscious in meetings, because they see themselves, and hear themselves, and I was like, 'oh my god, this is the first time that you have been that conscious about how you talk?' Because that's how I am all the time. But now I am like 2x, because I am concerned about how I am coming off but then I am also seeing how I coming off. And it was just like that much mental energy." As a result, seeing the self-view can make people who stutter struggle more with their speech: "definitely it is harder [with the self-view], I can feel the tension but now I can see it myself, from the other person's point of view. You are in the moment, you are also seeing it, that made it more difficult." (P7)

Participants also found the self-view distracting as it directed more attention to their own speech behaviors and less to their conversation partners: "seeing yourself is definitely a different experience, especially for someone for stuttering. I realized there is something I do with my face that I didn't know. [...] Sometimes, seeing myself, and what I am doing with my mouth, distracts myself." (P7)

To mitigate these challenge with the self-view, more than half of the participants had turned it off, at least for some meetings (P1, P4, P8, P9, P10, P12, P13, P14), or tried to avoid looking at it even when it is on (P4, P12). P13 explained his rationale behind turning off self view: "I turn off my self-view, every meeting of it, so I don't see myself. Cuz I hate to see myself, I hate to see myself talk, I don't need to see. For me, personally, because I really don't like seeing myself talk, it makes me more self-conscious, it makes me look at myself more. Then, generously speaking, when you turn off your self view, it's more like mimic a real life conversation, because in real life conversations, you are not looking at yourself all the time, unless you are talking to a mirror! You can actually be more engaged with the person you are talking to, if you hide yourself, or no longer caring about yourself, only caring about the person you are with." In fact, the impact of the self-view was so detrimental, that P13 deliberately avoided videconferencing platforms that do not let people hide their self views, and questioned the very idea of self-view in videoconferencing:

There are still some platforms, including Room, which is a product from Meta, where you can't turn off your self-view, so you have to watch yourself. So, any meeting I setup, I do Zoom, because, I can do that [turning off self-view]. I'd be curious to hear what the rationale is for that as a preset feature, cuz it doesn't make sense to me. (P13)

Nevertheless, some participants found a silver lining and started to accept the self-view as a therapeutic tool for them to self-monitor and desensatize stuttering. P7 explained her current approach with the self-view: "I am able to look at it as information: this is what I look like when I struggle; when I am tired, my eye contact is not great, I close my eyes, I block more. It's interesting to get more feedback, that I normally wouldn't get.[...] For me, it's helpful. How I am gonna use that information? I am blocking, I am holding back, if I am gonna change that, what I need to do. I use it [self-view] in an informative way.".

4.2.2 Turn Taking. While the connectivity issue and the lack of non-verbal cues over videoconferencing have created difficulties for taking turns and jumping into a conversation for everyone [2], such difficulties are multiplied for our participants. By default, many videoconferencing platforms rely on audible sound to detect and switch the current speaker, making the first sound/word crucial to signal one's turn. However, several of our participants found themselves struggle the most with initializing a sentence. With a limited channel for non-verbal communication strategies such as body language over videoconferencing, they would often be held back by that very first word:

I find things just like saying "hello", that's probably the hardest bit. [...] Then once you get into the flow of things, it becomes easier. [...] The hardest things for me is starting the sentence of a conversation. (P2)

As a stutterer, it is harder to get your first sound out. Body language is more visible in person. That's particularly challenging. I am trying to say something but before I can say something, somebody also already got the word out. (P6)

On the other hand, the limited perception of other meeting participants on videoconferencing platforms also made it harder for others to support people who stutter on turn taking. P14, a SLP who has been conducting group therapy sessions with PWS via Zoom since the pandemic, noted that:

With timing, there is sometimes a delay, and if you deal with that, in a group situation, and you're stammering, maybe it is hard to indicate that you want to speak. When you are in the same room, you can see more of the body language that you want to speak, versus when we are just in the screen, it's harder to read the room, read the group in the same way, get a sense that that person is wanting to talk now. I think we are more focused on just looking at the one person who is talking, and not having the peripheral vision of the whole group.... that's true for everyone, not just for people who stammer, but I wonder whether your stammer adds an extra layer of difficulty, that you can't get in. (P14)

Even after PWS successfully cut into the conversation, they could face higher risk of losing their turn involuntarily because of their stutter. Without sufficient non-verbal cues, P10 was concerned that people would assume she's finished when she was in fact having a long pause due to stutter. Some participants noticed that it could be challenging to differentiate a stuttering block with the loss of Internet connectivity in some situations, especially when the block is long and silent: "if the person is not able to get any sound out, they can't say to people, I am not finished yet. that's the loss of body on Zoom. because often you can tell somebody is still trying to speak based on what they are doing with their body." (P14)

To overcome the challenge with jumping into the conversation with speech, our participants leveraged videoconferencing platform features such as "hand raising" and "chat" extensively, and found the hand-raising function relatively effective at signaling their desire to have a turn. However, the participants also reported that the "hand raising" was sometimes ignored/unseen - especially when the speaker was in presentation mode. P10 sometimes raised her physical hand in front of the camera at the same time as she pressed the hand-raising button, in the hopes for attracting more attention to her willingness to participate. Some participants (P1, P10) also shared that they would not feel comfortable to be the first/only one to use the hand-raising function, if others in the meeting were getting their turns through voice.

Another strategy our participants adopted for easier turn taking is to proactively setup the structure and norms of participation during virtual meetings they hosted. For example, P1 tried always preparing an agenda for the larger meetings she hosted, with speaker name and timing assigned for each agenda item to ensure people on the agenda (including herself) have their turns. P14 would ask people who did not speak to type out their ideas in the chat, and reserved time to read and respond to all chat messages. P13 would give people a heads-up on who would go next during round-the-room introductions or updates. And P9 would proactively check in with people who did not get a turn or people who had expertise on the topic for things to add.

Our participants also developed various strategies to protect one's turn during stuttering moments. For instance, using, and getting everyone else to use, the hand-raising button was found helpful by a few participants, as it often created a break between speakers, giving PWS more time to unblock themselves before losing their turn. P14 would check with people when in doubt, asking whether they were finished, and stopped other people from cutting in before

getting an affirmative answer. P10 leaned on the non-verbal signals available on Zoom, keeping her camera on and making her facial expression more pronounced even though it felt "tiring" to have camera close-up for a prolonged period of time.

4.2.3 Emotional Connection with Others. With a strong association between stuttering and social anxiety, people who stutter are more sensitive to negative evaluations from others, and more likely to engage with safety behaviors such as loss of eye contact [22]. While the reduction of social cues during video calls has made everyone feel less connected to their conversation partners [2], the lack of emotional support from others could exacerbate the social anxiety experienced by PWS, causing further behavioral and emotional struggles.

Preconditioned by their experiences with social discrimination towards stuttering, our participants were more likely to associate the lack of responses from the audience with the loss of interest or patience with their speech:

I think when you have a stammer, people can be at times, not intentionally, but unintentionally a bit more impatient with me. [...] In the context of a zoom call, they have the same impatience stance toward you when you are talking, and then they clock off straight away. That you start stammering on the first word and then they go "here we go". I definitely think it is heightened for stammers. [...] Whereas in a face-to-face interaction, [..] people can't clock off as easily without being rude. Whereas in a zoom meeting you can kind of clock out and it's not necessarily, obviously rude. (P2)

Even when the audience gave responses, they could be interpreted more negatively, especially when non-verbal cues such as facial expression and eye contact were inaccessible. P11 shared the experience of giving a Zoom presentation to his CEO, who used phrases like "alright, let's move on" to communicate that he had comprehended the content of the presentation:

I think that's a unique challenge for people who stutter, because anytime we are talking, and someone indicates "let's move on", we will always interpret it as "oh they are commenting on my speech, rather than my content". I would say that I am self conscious about it, I spent a lot more time worrying about it. I stayed up last night writing out my slide, because I want to be super crisp, even though I don't do that for any other meetings, because I actually find my stutter is worse when I am reading off something, so there is added burden. (P11)

When asked about most satisfying meeting experiences in the past, several of the participants recalled group meetings or conversations where they clearly perceived the appreciation and attention by the listeners. For example, R4 called out one presentation he gave: "one girl in the project was listening very carefully, and I could tell from her eye contact the warmth, acknowledgement, and appreciation. I felt so much better immediately. I was basically fluent throughout the talk."

Our participants found eye contact crucial in building human connections: "Eye contact is everything. Good eye contact means much more than what you say. Eye contact is indicative of the enthusiasm of your conversation" (P8). Even though many of the participants have a tendency to look away when they stutter, they were aware of the power of eye contact and worked on improving eye contact with others. However, the design of most modern videoconferencing platforms has made effective eye contact almost impossible, especially in group settings:

It's hard for people to know who to look at on Zoom. In terms of eye contact, who do we keep eye contact with. Even if we all know who we want to keep eye contact with, do they know that? How can they tell, they probably can't. (P14)

730

731 732

733

734

735

736 737

738

739 740

741

742

743

744 745

746

747 748

749

750

751

752 753

754

755 756

757

758

759 760

761

762

763

764 765

766

767

768

769

770 771

772

773 774

775

776

777

778 779

780

As a result, our participants were often deprived from an important source of the positive social support that they sought out in-person from others in the videoconferencing setting:

I kind of grew to like the fact that, eye contact - when someone is looking at you - it's kind of like you have the floor, [...] you just had as much right as the person next to you to say what you have to say, the story you want to tell. [...] If I am in person, I can pick someone to say what I want to say. In zoom, I don't know what to look. I can look at someone's face, but I don't have the connection, the feedback that they are looking at me, even though that I know when I look at their face, I know they are listening, but I just can not get the feedback of the eye contact. (P7)

Videoconferencing also disabled some of the participants' existing strategies for social and emotional support in in-person meetings. For example, when attending in-person group meetings, P3 and P10 would chose to sit next to friendly, familiar people to feel more relaxed. Small talk and chitchat right before a meeting starts is another strategy that P10 and P14 took to serve a similar purpose. P11 have developed a personal "charisma" to compensate for his stutter, and found himself "less effective on VC" due to the limits videoconferencing put on communicating "body language, energy, and interpersonal chemistry":

I like to shake hands, I will give people a hug if I know them. If I am comfortable, I will talk with my hands, I will also be open, I will lean back on my chair, I will think out loud, try to model that this is a space where I hope you can be yourself. I can do that via VC, too, but I think it's much more noticeable when I am in person. (P11)

To compensate for the lost connections with others, our participants extensively utilize available channels to make a conscious effort in communicating their emotions and intentions. For example, several participants (P1, P8) deliberately lifted the position of their camera to the eye level so that they could mimic the in-person eye contact. Almost all of our participants indicated that they were making an effort for maintaining eye contact over video conferences, and some even turned off the self view to better direct their gaze to the conversation partners. Some participants tried to make their facial expression more salient by lightening up their faces (P1, P8), putting on make-up (P3), and positioning the camera for a close-up view (P8, P10). And some participants (P1, P8, P10, P13) leveraged the reaction feature on Zoom to both give and gather direct, positive emotional feedback. Despite potential speech challenges, some participants actively describe their body language and the intention behind it to avoid misinterpretations. For example, P13 would explicitly tell the speaker that he was looking away only to open up the document the speaker just mentioned.

Our participants also leveraged their identity as PWS to better connect with others in virtual meetings. Most of the participants had proactively disclosed their stutter in high-stress video conferences (e.g. job interviews, presentations, orientations) to build the connection with their audience, and found that effective at reducing mental stress and bringing in audience's emotional support, even though it did not change how fluent they were. Some participants purposefully embraced the vulnerability that came with the identity as a person who stutters, as a way to empower others to all be more open and collaborative in virtual meetings:

As soon as I say something deeply personal, [...] about my stutter, [...] it's not top of mind for other people, and when they hear me saying that, they go, 'this guy is being open, okay, maybe I will be open, too!' I think the whole modeling behavior is huge over Zoom, and over virtual communications. Modeling behavior is massive, because all you see is this much of me, and you are constantly seeing me. I always want to model the behavior of being open and collaborative. (P13)

#### 5 DISCUSSION

# 5.1 The Hidden Cost of Videoconferencing for People Who Stutter

Despite the benefits identified by our participants, videoconferencing has introduced significant emotional and cognitive costs to people who stutter.

The constant close-up view of their facial features and speaking behaviors by self and others has contributed to heightened self-consciousness and more negative thoughts. Although the challenge with "Zoom gaze" is widespread [2, 18], people who stutter are more likely to pay disproportional attention to "negative" behaviors (e.g. stuttered words, facial tension) that reinforce existing self stigma and social anxiety [22]. The increased difficulty with turn taking over videoconferencing platforms posed structural barriers for people who stutter to have their voices heard and points across, deepening people's existing feeling of social isolation and reject, and preventing participants such as P6 and P12 from seeing themselves as leaders. The uncertainty with turn taking and audience reactions further contributes to the sense of loss of control, one of the defining characteristics of stuttering and a direct cause of many negative emotional and cognitive reactions when people stutter [40]. While the emotional connection with their conversation partners was highlighted by several of our participants as the hallmark of their most rewarding communication experience, our participants are systematically disadvantaged in seeking and sharing emotional support now, as their previous strategies - such as physical proximity, hugs, and good eye contact - were largely unsupported by today's videoconferencing technology.

To overcome these VC challenges, people who stutter had to adopt strategies that often require extra time, labor, and mental efforts, on top of the existing cognitive and emotional loads associated with stuttering. For example, our participants made an effort to put on make-up, well-position themselves in front of the camera, give others more verbal and non-verbal feedback, and over-prepare themselves with the content and agenda of meetings. As P11 put it, "it is an extraordinary mental effort, [..] my brain is always in the 5th gear, like 150mph, I can't get it to stop." Even the use of the hand-raising feature and a predetermined speaking order worked at the expense of spontaneity, one of the top predictors for how satisfying a speaking experience is to people who stutter, ahead of even fluency [11]. It is perhaps not surprising that our participants reported feeling videoconferencing particularly "exhausting", "draining", and "unrewarding", something that they - while still participating in - did "not look forward to".

Even the named benefits of videoconferencing could lead to questionable long-term outcomes for the stuttering community. For example, the convenience and comfort of a familiar, controlled videoconferencing environment could potentially disincentivize PWS from engaging in in-person meetings and social interactions. The ability to hide one's stuttering behaviors and identity via videoconferencing is also a double-edged sword. Although it does serve PWS with better impression and identity management at the moment, it could also hold people back from accepting their stutter and stuttering identity, reinforcing negative emotions associated with stuttering such as embarrassment, guilt, and fear [9]. Collectively, if people who stutter all manage to hide their stutter and stutterer identity during video calls, speech-related challenges would be even less understood and further marginalized by the mainstream society. While VC reduced the barriers for PWS to find and join the stuttering community, the bonding and commitment within the community might be weakened due to the difficulty in forming emotional connections via video conferences, making the community more fragmented and superficial.

To summarize, videoconferencing and videoconferencing technologies have substantially changed the dynamics and the structure of interpersonal communications, charging potentially profound emotional, cognitive, and social costs to people who stutter. The very design of the videoconferencing technologies that induced such costs (e.g. lack of

"The World is Designed for Fluent People": Challenges, Benefits, and Opportunities of Videoconferencing Technologies for People Who Stutter

non-verbal communication support), has also helped render these costs **invisible**, preventing public awareness on the structural barriers for PWS to participate and engage in the age of videoconferencing.

While the research on technology-mediated communications tend to evaluate the effectiveness of communication from the perspective of a 3rd party observer, the importance of communication participants' subjective experience is often overlooked. Similarly, mainstream assistive technologies were typically designed to enhance "productivity, efficiency, normalcy, and speed" [25], without sufficiently attending to the user's emotional needs [27]. We argue that the subjective experiences of marginalized users - such as people who stutter - should be respected and prioritized over external observations in the research and design of videoconferencing technologies, in order to create an inclusive and equitable communication environment for all. Foregrounding the lived experiences of people who stutter not only offers us insights of communication technologies that were inaccessible from an observer perspective, but also serves epistemic justice to the stuttering community which - similar to many other historically marginalized groups - had long been treated as epistemic subjects rather than the knowers.

## 5.2 The Future of Videoconferencing

 Discussing and re-envisioning videoconferencing tools with the stuttering community also creates space and opportunities for videoconferencing technologies and video-mediated communications that have not been explored before.

5.2.1 Enhance Non-verbal Communications. Our findings highlight the importance of non-verbal communications for people who stutter in offline and online meetings. While the current design of videoconferencing technologies has over-indexed on verbal channels, our participants saw many ways that non-verbal cues can be better leveraged by videoconferencing platforms. For example, the camera can detect and communicate meaningful body movements and facial expressions, such as leaning forward, clapping, and smiling. To reduce the mental stress from "Zoom gaze" while maintaining communicative facial expression and eye contact, VC systems could deploy filters/avatars that capture exaggerate these features in a context appropriate way (i.e. no cat filter for court appearance).

5.2.2 Support For Atypical Speech. Videoconferencing platforms also need to accommodate and empower more diverse speech patterns.

Several of our participants noted that Zoom's auto-captioning and translation function worked poorly for people who stutter. Although there have been some research efforts in improving the performance of speech recognition models for stuttering speech [??], more investment is required in this domain to close the performance gap between stuttering and fluent speeches, especially in a more dynamic setting like video conferences.

Similarly to how the hand-raising button signals one's intention to speak, videoconferencing platforms could also design a way for the user to express "I'm not done, please hold the floor" during a stuttering moment. The message could potentially be customized to provide more context about stuttering and educate others on respectful ways to support someone during the stuttering moments (e.g. stop repeatedly asking "are you still there?").

The idea of a "voice mask" that auto-filters disfluencies from stuttering speech was discussed heavily by our participants. While most participants saw the benefit of having the option to sound more fluent in certain situations, all but two participants also mentioned that they would not use it themselves, as technologies like this could potentially undermine their self-acceptance and the space for disfluent speech in our society.

5.2.3 Videoconferencing as Speech Therapy. Videoconferencing platforms could also serve as a therapeutic tool for people who stutter, offering them insights and support during stuttering moments and reinforcing positive mental images. For example, the platform could remind PWS to maintain eye contact during stuttering moment if that is something they are working on, or display tips like "keep moving forward" when a severe speech block is detected.

#### 5.3 Intersectionality

Our research also highlights the need for intersectionality [13] in understanding and designing technological experiences with marginalized communities. For instance, over half of our interview and focus groups participants are women, and surfaced the tension between "taking up space" with their speech disfluencies and the socialized "supportive" role in conversations for women [47]. A third of our interview participants are first-generation immigrants and non-native English speakers working in an English environment. They have experienced confusion by others of their stuttering with a lack of language proficiency. They also reported higher levels of pressure to perform and establish themselves in the workplace, which could cause additional stress in both virtual and in-person meetings on top of the challenges brought on by stuttering and language barriers. Many of our female participants and participants with ethnic and racial minorities also recalled being talked over at video conferences or being labeled as "quiet" by their colleagues - an experience well-documented for minority groups at work [23, 44] but extra difficult for our participants to push against when they are already burdened with existing speech and emotional challenges from stuttering.

### **6 LIMITATIONS AND FUTURE WORK**

This work has a few limitations. First, the fact that our participants were recruited from stuttering community events and group therapy program determined that all the participants had at least partially accepted their identity as people who stutter, and were comfortable enough to talk about their experience and vulnerability with stuttering. As a highly stigmatized identity, many adults who stutter do not self identify as a person who stutters or speak openly about their stuttering experiences [5], and their perspectives and needs with videoconferencing technologies might be different from the participants in our study. Second, although we did not set recruitment criteria on stuttering severity, all participants had relatively mild speech disfluencies, with a large percentage of the participants stuttering covertly. For future work, more efforts are needed to reach out to the PWS population with more speech challenges, as they might have different types of communication challenges within the videoconferencing context. Third, while our recruitment was targeted at adults with videoconferencing experiences, all our participants were relatively young (the oldest participant is in their 40s). The intersectional experiences with older adults who stutter would be an interesting topic for future research. Finally, although we tried to cover diverse demographics in our study, all participants were recruited from and resided in the US and UK, a geographical region with a relatively greater awareness and acceptance of stuttering [39]. Future work should explore and compare the videoconferencing experiences for PWS across countries/cultures.

### 7 CONCLUSION

Although typically considered a speech disorder, stuttering comes with strong emotional and cognitive reactions, and can have a profound effect on how people engage and experience conversations in-person and virtually. In this paper, we present one of the first studies that examines the experience of people who stutter with videoconferencing and videoconferencing technologies, based on interviews of 14 adults who stutter.

Our findings reveal that, despite benefits, videoconferencing presents extra challenges for people who stutter due to the reduction of non-verbal communication channels and the constant surveillance by self and others in the call.

Manuscript submitted to ACM

"The World is Designed for Fluent People": Challenges, Benefits, and Opportunities of Videoconferencing Technologies for People Who Stutter

Although these issues are also challenging for fluent people, they are greatly exacerbated when interacting with the characteristics and nature of stuttering. Confronted with the social stigma and negative assumptions associated with stuttering, people who stutter spend more time and effort curating their physical environment and virtual presence for video conferences, and have adopted different strategies to compensate the lost non-verbal and seek/provide emotional support in video calls. Besides being technically challenging, videoconferencing also become an mentally draining experience, creating significant - yet invisible - emotional, cognitive, and social barriers for people who stutter to engage and enjoy. Even though most of our participants were generally accepting of their speech and their identity as people who stutter, videoconferencing brings new struggles when their speech behavior violates the hardcoded assumptions made about human speech and communication behaviors (e.g. the length of a pause, the existence of blocks) by the technology. They are also more likely to be emotionally impacted by the lack of attention and feedback from the audience, and further develop social anxiety and negative self-image as a result.

We hope our findings and discussion shed light on the gap between current videoconferencing technologies and the needs of people who stutter, and inspire the research and development of more inclusive communication environment for all.

#### **8 ACKNOWLEDGEMENT**

First of all, we thank all our interview participants for their trust, openness and valuable insights. We also thank our volunteers, Ben Lickly and Yijing Lu, for conducting, analyzing, and coding some of the interviews. Last but not least, we thank Lindsay Reynolds for her early feedback and guidance on the research design and analysis.

#### **REFERENCES**

- [1] Joel S. Angiolillo, Harry E. Blanchard, and Edmond W. Israelski. 1993. Video Telephony. <u>AT&T Technical Journal</u> 72, 3 (1993), 7–20. https://doi.org/10.1002/j.1538-7305.1993.tb00536.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1993.tb00536.x
- [2] Jeremy N. Bailenson. 2021. Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue. <u>Technology, Mind, and Behavior</u> 2, 1 (feb 23 2021). https://tmb.apaopen.org/pub/nonverbal-overload.
- [3] O. Bloodstein, N.B. Ratner, and S.B. Brundage. 2021. <u>A Handbook on Stuttering, Seventh Edition</u>. Plural Publishing, Incorporated. https://books.google.com/books?id=Abw0EAAAQBAJ
- [4] John Van Borsel, Marie Brepoels, and Janne De Coene. 2011. Stuttering, attractiveness, and romantic relationships: The perception of adolescents and young adults. <u>Journal of Fluency Disorders</u> 36, 1 (2011), 41–50.
- [5] Michael P. Boyle and Rodney M. Gabel. 2020. "Openness and progress with communication and confidence have all gone hand in hand": Reflections on the experience of transitioning between concealment and openness among adults who stutter. <u>Journal of Fluency Disorders</u> 65 (2020), 105781. https://doi.org/10.1016/j.jfludis.2020.105781
- [6] Jed R. Brubaker, Gina Venolia, and John C. Tang. 2012. Focusing on Shared Experiences: Moving beyond the Camera in Video Communication. In <u>Proceedings of the Designing Interactive Systems Conference</u> (Newcastle Upon Tyne, United Kingdom) (<u>DIS '12</u>). Association for Computing Machinery, New York, NY, USA, 96–105. https://doi.org/10.1145/2317956.2317973
- [7] Clare Butler. 2014. Wanted straight talkers: stammering and aesthetic labour. Work Employment Society 28 (10 2014). https://doi.org/10.1177/0950017013501956
- [8] C. T. Byrd, M. McGill, Z. Gkalitsiou, and C. Cappellini. 2017. The effects of self-disclosure on male and female perceptions of individuals who stutter. American Journal of Speech-Language Pathology 26, 1 (2017), 69–80.
- [9] C. Cheasman, R. Everard, and S. Simpson. 2013. <u>Stammering Therapy from the Inside: New Perspectives on Working with Young People and Adults.</u>
   J & R Press. <a href="https://books.google.com/books?id=QZVdMwEACAAJ">https://books.google.com/books?id=QZVdMwEACAAJ</a>
- [10] Christopher Constantino, Patrick Campbell, and Sam Simpson. 2022. Stuttering and the social model. <u>Journal of Communication Disorders</u> 96 (2022), 106200. https://doi.org/10.1016/j.jcomdis.2022.106200
- [11] Christopher Dominick Constantino, Naomi Eichorn, Eugene H. Buder, J. Gayle Beck, and Walter H. Manning. 2020. The Speaker's Experience of Stuttering: Measuring Spontaneity. <u>Journal of Speech</u>, <u>Language</u>, and <u>Hearing Research</u> 63 (2020), 983-1001. Issue 4.
- [12] C. D. Constantino, W. H. Manning, and S. N. Nordstrom. 2017. Rethinking covert stuttering. Journal of Fluency Disorders 53 (2017), 26-40.
- [13] Kimberle Crenshaw. 1989. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. <u>The University of Chicago Legal Forum</u> 140 (1989), 139–167.

[14] S. Davis, P. Howell, and F. Cooke. 2002. Sociodynamic relationships between children who stutter and their non-stuttering classmates. <u>Journal of Child Psychology and Psychiatry</u>, and Allied Disciplines 43, 7 (2002), 939–947.

- [15] Janine Diehl, Michael P. Robb, John G. Lewis, and Tika Ormond. 2019. Situational speaking anxiety in adults who stutter. Speech, Language and Hearing 22, 2 (2019), 100–110. https://doi.org/10.1080/2050571X.2018.1441782 arXiv:https://doi.org/10.1080/2050571X.2018.1441782
- [16] Chris Elsden, David Chatting, Michael Duggan, Andrew Carl Dwyer, and Pip Thornton. 2022. Zoom Obscura: Counterfunctional Design for Video-Conferencing. 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 143, 17 pages. https://doi.org/10.1145/3491102.3501973
  - [17] A. M. Ferguson, J. M. Roche, and H. S. Arnold. 2019. Social judgments of digitally manipulated stuttered speech: An evaluation of self-disclosure on cognition. <u>Journal of Speech</u>, <u>Language</u>, and <u>Hearing Research</u> 62, 11 (2019), 3986–4000. https://doi.org/10.1044/2019\_jslhr-s-19-0088
  - [18] Autumm Gaine. 2020. The Zoom Gaze. https://reallifemag.com/the-zoom-gaze/, note=Online; accessed 2022-08-16.
- 998 [19] Hope Gerlach-Houck and Christopher D. Constantino. 2022. Interrupting Ableism in Stuttering Therapy and Research: Practical Suggestions. Perspectives of the ASHA Special Interest Groups 7, 2 (2022), 357–374. https://doi.org/10.1044/2021\_PERSP-21-00109 arXiv:https://pubs.asha.org/doi/pdf/10.1044/2021\_PERSP 21 00109
- 1001 [20] Gerlach H, Totty E, Subramanian A, and Zebrowski P. 2018. Stuttering and Labor Market Outcomes in the United States. J Speech Lang Hear Res.
  1002 61, 7 (2018), 1649–1663. https://doi.org/10.1044/2018\_JSLHR-S-17-0353 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6195060/.
  - [21] RUTH IGIELNIK. 2022. As telework continues for many U.S. workers, no sign of widespread 'Zoom fatigue'. https://www.pewresearch.org/fact-tank/2022/05/04/as-telework-continues-for-many-u-s-workers-no-sign-of-widespread-zoom-fatigue/.
  - [22] Lisa Iverach and Ronald M. Rapee. 2014. Social anxiety disorder and stuttering: Current status and future directions. <u>Journal of Fluency Disorders</u> 40 (2014), 69–82. https://doi.org/10.1016/j.jfludis.2013.08.003 Anxiety and stuttering.
  - [23] Kelly F. Jackson. 2011. Derald Wing Sue, Microaggressions in Everyday Life: Race, Gender, and Sexual Orientation. <u>Social Service Review</u> 85, 3 (2011), 519–521. https://doi.org/10.1086/663007 arXiv:https://doi.org/10.1086/663007
  - [24] Sarah E James, Shelagh M Brumfitt, and Peter A Cudd. 1999. Communicating by telephone: Views of a group of people with stuttering impairment. Journal of Fluency Discorders 24, 4 (1999), 299–312.
- 1010 [25] Alison Kafer. 2019. Crip Kin, Manifesting. Catalyst: Feminism, Theory, Technoscience 5 (04 2019), 1–37. https://doi.org/10.28968/cftt.v5i1.29618
- 101 [26] Raja S. Kushalnagar and Christian Vogler. 2020. Teleconference Accessibility and Guidelines for Deaf and Hard of Hearing Users. In The 22nd
  1012 International ACM SIGACCESS Conference on Computers and Accessibility (Virtual Event, Greece) (ASSETS '20). Association for Computing
  1013 Machinery, New York, NY, USA, Article 9, 6 pages. https://doi.org/10.1145/3373625.3417299
  - [27] Franklin Mingzhe Li, Franchesca Spektor, Meng Xia, Mina Huh, Peter Cederberg, Yuqi Gong, Kristen Shinohara, and Patrick Carrington. 2022. "It Feels Like Taking a Gamble": Exploring Perceptions, Practices, and Challenges of Using Makeup and Cosmetics for People with Visual Impairments. 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 266, 15 pages. https://doi.org/10.1145/3491102.3517490
- [28] Geert Lovink. 2020. The anatomy of Zoom fatigue. https://www.eurozine.com/the-anatomy-of-zoom-fatigue/, note=Online; accessed 2022-08-16.
  - [29] Robert L. MacDonald, Pan-Pan Jiang, Julie Cattiau, Rus Heywood, Richard Cave, Katie Seaver, Marilyn A. Ladewig, Jimmy Tobin, Michael P. Brenner, Philip C. Nelson, Jordan R. Green, and Katrin Tomanek. 2021. Disordered Speech Data Collection: Lessons Learned at 1 Million Utterances from Project Euphonia. In Proc. Interspeech 2021. 4833–4837. https://doi.org/10.21437/Interspeech.2021-697
- 1021 [30] Vikramjit Mitra, Zifang Huang, Colin Lea, Lauren Tooley, Panayiotis Georgiou, Sachin Kajarekar, and Jefferey Bigham. 2021. Analysis and Tuning
  1022 of a Voice Assistant System for Dysfluent Speech. https://arxiv.org/pdf/2106.11759.pdf
- 1023 [31] Timothy Neate, Vasiliki Kladouchou, Stephanie Wilson, and Shehzmani Shams. 2022. "Just Not Together": The Experience of Videoconferencing for
  1024 People with Aphasia during the Covid-19 Pandemic. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (New
  1025 Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 606, 16 pages. https://doi.org/10.1145/3491102.
  1026 3502017
- 1026 [32] National Institute on Deafness and Other Communication Disorders. 2016. NIDCD Factsheet: Stuttering. https://www.nidcd.nih.gov/health/stuttering.

  Retrieved on April 19th, 2022..
- [33] Laura Plexico, Walter H. Manning, and Anthony DiLollo. 2005. A phenomenological understanding of successful stuttering management. <u>Journal</u>
   of Fluency Disorders 30, 1 (2005), 1–22. https://doi.org/10.1016/j.jfludis.2004.12.001
- 1030 [34] Ronald M. Rapee and Richard G. Heimberg. 1997. A cognitive-behavioral model of anxiety in social phobia. Behaviour Research and Therapy 35, 8
  1031 (1997), 741–756. https://doi.org/10.1016/S0005-7967(97)00022-3
  - [35] Rabindra Ratan, Dave B. Miller, and Jeremy N. Bailenson. 2022. Facial Appearance Dissatisfaction Explains Differences in Zoom Fatigue. Cyberpsychology, Behavior, and Social Networking 25, 2 (2022), 124–129. https://doi.org/10.1089/cyber.2021.0112 arXiv:https://doi.org/10.1089/cyber.2021.0112 PMID: 34842445.
- 1034
  [36] Jazz Rui Xia Ang, Ping Liu, Emma McDonnell, and Sarah Coppola. 2022. "In This Online Environment, We're Limited": Exploring Inclusive Video
  Conferencing Design for Signers. 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 609, 16 pages. https://doi.org/10.1145/3491102.3517488
- [37] Abigail Sellen, Bill Buxton, and John Arnott. 1992. Using Spatial Cues to Improve Videoconferencing. In Proceedings of the SIGCHI Conference on
  Human Factors in Computing Systems (Monterey, California, USA) (CHI '92). Association for Computing Machinery, New York, NY, USA, 651–652.
  https://doi.org/10.1145/142750.143070

991

992

993

994

995

996

1003

1004

1005

1006

1007

1008

1009

1014

1015

1016

1019

1020

1032

- [38] J.G. Sheehan. 1970. Stuttering: Research and Therapy. Harper & Row. https://books.google.com/books?id=0KhrAAAAMAAJ
- [39] Glen Tellis. 2002. Multicultural Aspects of Stuttering. Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse Populations 8 (07 2002), 8–11. https://doi.org/10.1044/cds8.2.8
- [40] Seth E. Tichenor and J. Scott Yaruss. 2019. Stuttering as Defined by Adults Who Stutter. <u>Journal of Speech, Language, and Hearing Research</u> 62 (2019), 4356–4369. https://doi.org/10.1044/2019\_JSLHR-19-00137
- [41] Christian Vogler, Paula Tucker, and Norman Williams. 2013. Mixed Local and Remote Participation in Teleconferences from a Deaf and Hard of Hearing Perspective. In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (Bellevue, Washington) (ASSETS '13). Association for Computing Machinery, New York, NY, USA, Article 30, 5 pages. https://doi.org/10.1145/2513383.2517035
- [42] Shirley Weitz. 1976. Visual Behavior: <i>Gaze and Mutual Gaze</i> Michael Argyle and Mark Cook. Cambridge University Press, New York, 1976. xii, 210 pp., illus. \$18.50. <a href="Science">Science</a> 194, 4260 (1976), 54–55. <a href="https://doi.org/10.1126/science.194.4260.54">https://doi.org/10.1126/science.194.4260.54</a> arXiv:https://www.science.org/doi/pdf/10.1126/science.194.4260.54
- [43] Robert A. Wicklund. 1975. Objective Self-Awareness11Much of the research reported in this paper as well as the writing of this paper were supported by NSF Grant GS-31890. Sharon S. Brehm, William J. Ickes, Michael F. Scheier, and Melvin L. Snyder are acknowledged for their suggestions and insightful criticisms. Advances in Experimental Social Psychology, Vol. 8. Academic Press, 233–275. https://doi.org/10.1016/S0065-2601(08)60252-X
- [44] Nicola Woods. 1989. Talking shop: Sex and status as determinants of floor apportionment in a work setting. <u>Women in their speech communities</u> (1989), 141–157.
- [45] Patricia M. Zebrowski. 2016. The High Cost of Stuttering. The ASHA Leader 21, 7 (2016), 6-7.
- [46] Xinlei Zhang, Takashi Miyaki, and Jun Rekimoto. 2021. JustSpeak: Automated, User-Configurable, Interactive Agents for Speech Tutoring. Proc. ACM Hum.-Comput. Interact. 5, EICS, Article 202 (may 2021), 24 pages. https://doi.org/10.1145/3459744
- [47] Don H Zimmerman and Candace West. 1975. Sex Roles, Interruptions and Silences in Conversation: In Thorne, B. & Henley, N.(Eds.), Language and sex: Difference and Dominance (pp. 105-129).