

Springboard, Roadblock or “Crutch”? : How Transgender Users Leverage Voice Changers for Gender Presentation in Social Virtual Reality

Kassie C. Povinelli*

University of Wisconsin-Madison

Yuhang Zhao†

University of Wisconsin-Madison

ABSTRACT

Social virtual reality (VR) serves as a vital platform for transgender individuals to explore their identities through avatars and foster personal connections within online communities. However, it presents a challenge: the disconnect between avatar embodiment and voice representation, often leading to misgendering and harassment. Prior research acknowledges this issue but overlooks the potential solution of voice changers. We interviewed 13 transgender and gender-nonconforming users of social VR platforms, focusing on their experiences with and without voice changers. We found that using a voice changer not only reduces voice-related harassment, but also allows them to experience gender euphoria through both hearing their modified voice and the reactions of others to their modified voice, motivating them to pursue voice training and medication to achieve desired voices. Furthermore, we identified the technical barriers to current voice changer technology and potential improvements to alleviate the problems that transgender and gender-nonconforming users face.

Keywords: Social virtual reality, interview study, voice communication, transgender, gender-nonconforming

1 INTRODUCTION

Social Virtual Reality (VR) platforms provide an immersive digital space for individuals to interact with each other, where a combination of technologies such as head-mounted displays, real-time voice chat, full-body tracking, kinesthetic interactions, and traversable 3D worlds provide a Face-to-Face (FtF)-like interaction experience [55]. Increased adoption during the COVID-19 pandemic increases the importance of studying these interactions. In social VR, LGBTQ+ users find solace and empowerment through embodiment and community-building. They transcend physical limitations, forming connections with those who share similar experiences from across the world [4, 41].

While avatars in social VR serve as both representations of users and their interfaces for interacting with their environments and others, *voice* is the primary channel that can expose a user's identity in real life [22, 58]. For many transgender and gender-nonconforming (TGNC) users whose voices may not align with their gender identity, even though they could customize avatars to present their preferred gender identity, the voice-avatar mismatch can make them targets of harassment within social VR spaces [20, 22]. Furthermore, *voice dysphoria*—an emotional and psychological distress stemming from the incongruence between one's gender identity and voice—compounds the difficulties they face.

Voice training is a common solution to issues of voice-based misgendering and voice dysphoria, but the considerable daily exercises and outward practice required can pose a formidable obstacle in overcoming voice dysphoria for numerous TGNC individuals striving to

reconcile their voices with their gender identities. Voice training involves both active efforts to exercise vocal muscles and passive use of an altered but oftentimes nonpassing intermediary voice in social communications [49, 50]. While highly beneficial to TGNC individuals in the long-term, the overwhelming complexity, social risks of outwardly practicing (e.g., identity disclosure and harassment caused by speaking with intermediary voice), and time-commitment required by voice training prevents users from fully experiencing gender embodiment in social VR in the short-term [22].

As the challenges of voice representation become increasingly salient in social VR, the emergence of advanced voice changer technology offers a compelling solution. Voice changers hold particular significance for TGNC users who grapple with voice-avatar mismatches [22]. The ability to modify their voices convincingly offers a pathway to better align their virtual identities with their preferred gender identities, reducing the risk of harassment and diminishing the emotional toll of voice dysphoria. Recent advancements in AI-based real-time voice generation have transformed the landscape [59, 60, 65]. As a result, the artificiality that once marred voice changers has been greatly reduced, allowing for more authentic and immersive experiences with voice in social VR.

Despite the increasing prevalence of voice changer usage, the transformative potential of these technologies in supporting TGNC users and addressing issues such as voice dysphoria and voice-based harassment remains largely unexplored. There exists a notable research gap in understanding how voice changers are employed within the context of social VR, their impact on TGNC users' social experiences, and the technical and usability challenges faced by TGNC users when adopting these technologies.

To address this research gap, we conducted a comprehensive interview study involving 13 participants from diverse backgrounds within the TGNC community in social VR. Our findings offer valuable insights into the ways in which these users have harnessed voice changers, shedding light on the multifaceted impact of this technology on their virtual and real-world experiences. We found that the use of voice changers emerged as a transformative tool for many of our participants. It played a pivotal role in reducing instances of discrimination, bias, and hostility related to their identities in social VR environments. This newfound freedom allowed them to expand their boundaries and engage more comfortably in virtual social interactions. Beyond mitigating external challenges, voice changers also proved instrumental in alleviating internal struggles. Participants reported a significant reduction in feelings of voice dysphoria. The ability to hear a voice that resonated with their gender identity brought a profound sense of relief and congruence, bolstering their overall well-being.

While voice changers and voice training are the two typical methods used by TGNC users, our research revealed the tension and interaction between these two methods. Voice changers offered TGNC users the opportunity to “pass” in social VR, enabling them to be treated and perceived as their preferred gender identities. This experience, akin to a digital “test run,” provided a tangible demonstration of what passing could feel like, motivating them to explore voice training and other transition-related processes in the real world. Even basic (non-AI) voice changers acted as valuable “training wheels” for voice training, offering users a stepping stone towards

*e-mail: kassie.povinelli@wisc.edu

†e-mail: yuhang.zhao@cs.wisc.edu

greater voice authenticity through voice enhancements, rather than full conversions. Additionally, the modified voice that users created oftentimes served as a goal for voice training. However, we also found that voice changers can become a roadblock to voice training, with users highlighting several concerns. These include apprehensions about the perceptible change in their voice to others if they were to discontinue using voice changers, the feeling that AI voice changers create such drastic voice alterations that their own efforts seem less impactful, and worries about developing an overreliance on these tools as a “crutch,” which hinder consistent voice training.

Beyond the social impact, we also identified the usability issues of current voice changing technologies for TGNC users, regarding the technology complexity, capabilities, and affordability.

Our research provides five contributions. First, this is the first research that studies the impact of voice changers on TGNC users’ experiences in social VR. Second, we identify the goals of TGNC users when using voice changers and the factors influencing their decisions of when to use them. Third, we discuss the implications of passing as cisgender on gender exploration in social VR. Fourth, we derive design guidelines for effective and user-friendly voice changer applications. Fifth, we develop the idea of voice changers in social VR as a potential “springboard” for voice training, framing it as the next step in incorporating technology into voice training and therapy regimens.

2 RELATED WORKS

2.1 Marginalization of TGNC Users Online

Marginalization refers to the systematic social exclusion, discrimination, and disempowerment of individuals or groups based on characteristics such as gender identity, sexual orientation, disability, race, or ethnicity [26]. In the context of transgender and gender-nonconforming (TGNC) individuals, marginalization manifests in various forms, both online and offline, leading to detrimental consequences. The experiences of TGNC individuals in online spaces has been a subject of growing concern and investigation. Marginalization, discrimination, and harassment against TGNC users have been widely documented, shedding light on the challenges they face in digital environments [15, 37]. The marginalization of TGNC individuals both online and offline is associated with increased psychological distress, and increased coverage in media and legislation further increases feelings of stigmatization [9, 35].

Community building TGNC users, in response, create online communities for advocacy, safety, and visibility. Haimson et al. explored how trans users created their own website, “Trans Time,” to share their trans experiences without fear of the stigmatization present on other social media platforms, while focusing on aspects such as safety, privacy, and content warnings [32]. Jackson et al. explored how black transgender and queer women, using the “#Girls-LikeUs” network on Twitter, engage in activism to counter narratives surrounding them online and in the media [36].

Identity disclosure TGNC users control the pieces of their identity that they display on social media. They control who they come out to, how much of their identity they disclose, and how they construct their online identities and streams to fit with their transgender identities [31]. Haimson et al. explored how trans users use social media to come out on various forms of social media, such as blogs and Facebook pages, and the short and long-term impacts of disclosing identity on social media in comparison to real-life coming-out experiences [30]. Buss et al. investigated how transgender users engage in identity management on social media across platforms, finding that transgender users not only control their identity presentations through choosing what to post, but also curate their experiences by deciding what they see from others and who they interact with [11].

2.2 Experiences of Vulnerable Users in Social VR

Underrepresented users in social VR and virtual worlds frequently encounter discrimination, harassment, and exclusion based on their gender identity, sexual orientation, disability, race, or ethnicity. They may face offensive stereotypes, microaggressions, and outright hostility from other users, contributing to feelings of alienation and frustration [23, 51]. Additionally, the lack of diverse representation in virtual environments can reinforce real-world disparities, perpetuating harmful biases and prejudices [58, 70]. In response to these challenges, underrepresented users often form supportive communities, develop coping measures to counter stigmatization, and create new avenues for identity presentation within social VR and virtual worlds [4, 22, 41, 44, 69].

Users with Disabilities Individuals with disabilities navigating social VR and virtual worlds often encounter accessibility issues that limit their full participation. These platforms may not provide adequate support for various disabilities, such as mobility impairments or sensory challenges. There is a body of HCI research focused on designing VR interfaces that meet the needs of people with disabilities (PWD) [38, 46, 71]. Furthermore, users with disabilities oftentimes want ways to represent their disability identities in social VR. Zhang et al. explored how PWD represent themselves in social VR using their avatars, both through how they disclose their disabilities and how they use avatars to communicate with others using techniques such as VR-ASL (a version of American Sign Language designed for the limited hand gestures of VR) [69]. Further work identified the unique forms of verbal, physical, and environmental harassment that PWD face when disclosing their identities through the design of their avatars, including others mimicking their disabilities, insults, and others grabbing disability signifiers such as their wheelchairs and walking canes [70].

Women in Social VR Women in social VR and virtual worlds often face gender-based discrimination and harassment, reflecting broader societal issues. They may encounter unwelcome advances, objectification, or even verbal abuse from other users, and suffer from a lack of avatars representing their body types [19, 44, 47]. These negative experiences can deter women from fully enjoying and participating in virtual communities and perpetuates gender disparities in these digital spaces. Schulenberg et al. found that social VR environments pose unique threats to women, since their embodied avatars in the 3D space become targets of physical harassment and the predominant use of voice communication makes hiding their identities difficult [58].

LGBTQ+ Users LGBTQ+ users oftentimes lack sufficient offline social networks, and turn to online communities for information, support, and identity affirmation and formation [8, 13, 17]. These online communities, however, lack aspects of FtF interactions, making them unsatisfactory when compared to offline support and community [16, 53]. Li et al. explored how the immersive and embodied experiences made possible through social VR empowers LGBTQ+ users by simulating offline support [41]. Acena and Freeman found that this embodied multi-modal interaction in social VR not only felt more supportive, but also translated to real life benefits [4].

2.3 Avatars for Identity Representation in Social VR

A burgeoning body of research examines users’ utilization of avatars to depict themselves in social VR environments, as well as the consequences of avatar embodiment on their identities and interpersonal connections [21, 68]. Avatars, through hand tracking and partial and full-body tracking, enable communication alternatives to voice, such as VR sign language [69], built-in avatar gesture recognition to communicate basic emotions and responses through avatar faces (a feature built into many VRChat avatars [29]), and using advanced avatar and world shaders such as virtual keyboards and chat boxes or virtual pens [39, 66]. For marginalized users, such as transgender

people and cisgender women, this enables them to communicate while retaining privacy around their voices, which can otherwise make them targets of harassment [43]. In particular, transgender individuals utilize avatars within social VR platforms to synchronize their virtual representations with their preferred gender identities, providing a means for more intimate exploration and expression of their gender identities by bridging the gap to their physical bodies, while also empowering some to transition in real life [22].

2.4 Voice Training for TGNC Individuals

Voice training plays a pivotal role in the lives of many TGNC individuals, serving as a key aspect of gender presentation and identity affirmation. It involves altering multiple vocal features, including pitch, resonance, intonation, and articulation. The process of voice training involves both physiological and psychological adjustments, allowing TGNC individuals to increase their voice's congruence with their gender while also reducing the social stigma of having a gender-noncongruent voice, thus increasing quality of life [14, 33]. Physically altering one's voice may be achieved through individual exercises using online resources that compound videos, manuals, and regimens, such as r/transvoice's voice feminization and masculinization wikis, [2], with the help of a speech therapist and regular visits combined with daily vocal exercises [56, 57], or, in more complex cases, achieved through surgery on the vocal tract combined with intensive exercises afterward [48, 57].

Voice Training for Transfeminine Individuals Voice training for transfeminine individuals often encompasses more than just pitch elevation; it involves a comprehensive approach addressing physical, neurophysiological, and acoustical aspects of voice [40]. Morsomme and Remacle [45] explored the impact of ambulatory biofeedback in helping a transgender woman increase her voice pitch in daily activities, suggesting its potential to extend new vocal behaviors beyond clinical settings. Gelfer and Tice [27] conducted a study on the perceptual and acoustic outcomes of voice therapy for male-to-female transgender individuals, noting significant changes in how listeners perceived gender based on voice signals alone.

Voice Training for Transmasculine Individuals While voice feminization training is heavily covered academically, there isn't as much information on voice training for transmasculine individuals. This is because while feminizing hormone therapy for transfeminine individuals does not drastically change vocal characteristics, masculinizing hormone therapy for transgender men does [57]. Still, masculinizing hormone therapy on its own may not achieve some of the desired vocal results for transmasculine individuals, showing a need for vocal training alongside hormone therapy [34]. Furthermore, some transmasculine individuals may not have access to or may not wish to use masculinizing hormone therapy. In these cases, voice training or therapy are the only methods available to them, since surgical techniques focus exclusively on feminizing voice [57].

Modern Voice Training Tools and Techniques The evolution of voice training has seen the development of mobile applications and online tools, providing accessible resources for transgender individuals seeking voice feminization or masculinization. Ahmed et al. [6] critically analyzed mobile voice training apps, highlighting their tendency to reinforce binary gender norms and societal expectations. The community-based design of free and open-source software for transgender voice training, as described by Ahmed et al. [7], emphasizes the importance of user agency and collaborative development in creating effective voice training tools. Furthermore, Ahmed's study on trans competent interaction design [5] sheds light on the need for inclusive and adaptive technologies that cater to the diverse goals and experiences of the transgender community.

In addition to the vital but complex voice training, voice changers and social VR can potentially provide an easier and safer option

for TGNC users to experiment and achieve their desired voice. Our research thus investigates TGNC users' adoption of voice changers, as well as the interaction between voice training and voice changers in social VR environments.

3 METHOD

To understand how TGNC users use voice changer technology to curate their identity in social VR, we conducted an in-depth interview study with 13 TGNC users who used voice changers in social VR.

3.1 Participants

We recruited 13 participants who are transgender or gender-nonconforming, with ages ranging from 18 to 53 ($mean = 31.069$, $SD = 9.79$). Participants were primarily located in the United States, with two participants located in Brazil and one in Germany. Our participants cover diverse gender identities, including ten transgender female, two transmasculine nonbinary and one agender. All participants identified as transgender and five identified as gender-nonconforming. Table 1 shows participants' demographic information. Furthermore, five participants reported having disabilities, including four (P4, P8, P9, P11) identifying as neurodivergent (autism, ADHD) and one (P5) identifying as hard-of-hearing. One participant (P11) mentioned a stuttering disability that influenced her ability to speak with the voice changer. While using various types of voice changers (Table 2), all participants used VRChat as their primary social VR platform.

We designed our recruitment efforts to reach a diverse and representative sample of TGNC individuals who have engaged with social VR platforms and have experience using voice changers. We recruited participants through multiple channels, including Reddit (specifically the r/VRChat subreddit), Discord (VRChat and VRC Trans Academy Discord servers), and various LGBTQ+ organizations that facilitated online postings. To be eligible for participation, individuals had to be at least 18 years of age, identify as transgender or gender-nonconforming, possess past or present experience with social VR platforms, have experience using voice changers within Social VR environments, and be fluent in English. Before the formal study, we conducted a pilot study with a TGNC participant who had extensive technical knowledge and background in social VR development and voice changers. However, they did not use voice changer in social VR. We thus only used their data to refine the interview protocols and did not involve them in the final findings.

Table 1: Participant demographics, including Age, Race, Location (United States (US), Brazil (BR), or Germany (DE)), Gender (Female (F), Male (M), Agender (A), Nonbinary (NB)), whether the participant identifies as gender-nonconforming (GNC), and transition direction (Transfeminine (TF) or Transmasculine (TM))

ID	Age	Race	Loc.	Gender	GNC	Trans. Dir.
P1	53	Caucasian	US	F	No	TF
P2	28	Caucasian	US	A	Yes	N/A
P3	32	Hispanic	BR	F	No	TF
P4	20	Caucasian	US	NB	Yes	TM
P5	25	Caucasian	DE	F	No	TF
P6	22	Caucasian	US	F	No	TF
P7	18	Brazilian	BR	F	Yes	TF
P8	36	Caucasian	US	F	No	TF
P9	41	Caucasian	US	F	Yes	TF
P10	30	Caucasian	US	F	No	TF
P11	42	Caucasian	US	F	No	TF
P12	31	Mixed	US	NB	Yes	TM
P13	34	Caucasian	US	F	No	TF

3.2 Interviews

We conducted a single-session interview study that lasted for one to two hours on Zoom. The study included four sections, focusing on participants' demographic information, social VR experience, voice chat experience, and technical issues with voice changers, respectively. We detail all interview questions in Appendix A.

The first section focused on participants' demographic information. First, we asked about basic demographic information, including their age, ethnicity/racial background, gender identities (transgender identity, gender-nonconforming identity), and an open identity question. We then asked them about their real-life experiences with their voice, what voice changer they use or previously used in social VR, and how often they used the voice changer. Additionally, we asked participants if they encountered any social barriers in their real life because of their gender identity. For participants who had experiences with voice changers in social VR in the past, but no longer used them, we modified these questions to represent both their past and present experiences.

In the second section, we asked questions about participants' virtual world experiences, focusing on their experiences in social VR. We asked questions about their general social VR experiences, including what they usually did in social VR and the demographics of the people they hung out with. We asked participants how they presented themselves in social VR and for screenshots of their avatars. We asked various questions about participant's avatars, including questions on gender presentation and effectiveness of gender presentation if they indicated that the avatar represented their gender, and if they noticed differences in how other users treated them depending on the avatar they used if they indicated using multiple differing avatars. Participants also discussed both the benefits and barriers of using social VR as a TGNC user (depending on their identity).

The third section focused on using voice chat in social VR, and was split into two subsections. In the first subsection, we asked participants about experiences when using their *unmodified voice* (that is, without a voice changer), including who they used their unmodified voice with, why they used their unmodified voice in certain scenarios, and the impacts of using unmodified voice. For participants who avoided using unmodified voice, we asked them about the strategies they used to avoid communicating with their unmodified voice. In the second subsection, we asked participants about their voice changer use, including how they discovered and started using voice changers, how voice changers influenced their social VR experiences, and who they decided to turn the voice changer off with (if at any time they did). We also asked participants to send samples of their *modified voice* (that is, with a voice changer) or use their modified voice in the Zoom interview, and discussed the properties of the modified voice and how they tuned the voice changer to achieve a desired voice.

In the final section, we asked questions covering any technical aspects of the voice changer that were not mentioned in the previous section, including how the voice changer worked and the benefits and limitations of the voice changer software. We also asked how they wanted to improve the voice changer software. We finished the interview with a brainstorming session, where the interviewee and participant discussed possible social VR-based interfaces for voice changers, including implementation into the menu and avatar selection interfaces. All participants were compensated at a rate of \$25/hour upon the completion of the study.

3.3 Data Recording and Analysis

Upon obtaining informed consent from all participants, we recorded all 13 participants' interviews using Zoom's recording feature. These recorded interviews were then automatically transcribed by Zoom's transcription service. To ensure the accuracy and fidelity of the transcriptions, one researcher manually reviewed the transcripts while listening to the corresponding audio. Participants were given

the flexibility to use or turn off their cameras during the interviews.

When analyzing participant transcripts, we employed thematic analysis [10], a standard qualitative research method that aids in identifying recurring patterns and themes within the collected data. As opposed to quantitative statistical analysis, this method enables us to ground our analysis directly in participants' complex experience, and capture concepts, opinions, and the rationales behind them via bottom-up approaches, by highlighting and summarizing key information from the rich textual data with low-level codes and categorizing the codes into high-level themes to generate recurring patterns or findings.

Specifically, we selected two representative samples from the set of interviews. Two researchers coded the samples independently using open-coding and discussed and reconciled their codes, developing an initial codebook. Next, one researcher, using the codebook, coded the rest of the transcripts, and added new codes as needed upon the agreement between two researchers. We then categorized all the codes into high-level themes and subthemes using axial coding and affinity diagram. After the initial themes were identified, researchers cross-referenced the original data, the codebook, and the themes, to make final adjustments, ensuring that all codes fell in the correct themes. The initial codebook contained over 600 codes, which the researchers categorized into 11 themes. We then removed non-novel themes that had already been explored in previous works, and included 6 prominent and novel themes with 23 subthemes. Tables A.1 and A.2 in Appendix shows all themes, subthemes, and codes from our analysis. We elaborate all themes as our findings in Section 4.

3.4 Positionality Statement

We recognize that our personal identities and cultural backgrounds inevitably intersect with our roles as researchers and may influence various aspects of our work. It is crucial to acknowledge these intersections to ensure transparency and a comprehensive understanding of the study's context and potential biases.

The primary researcher for this study is a transgender woman. Her lived experience as a member of the transgender community played a pivotal role in shaping the research. As the sole interviewer, she brought her firsthand understanding of transgender and gender-nonconforming experiences to the forefront. This informed her approach to interview design, the development of questioning routes, and the nuanced interpretation of interview data.

The second researcher on our team is a cisgender woman with a background in social VR research. Her expertise in the field contributed to maintaining the quality and rigor of our research materials and procedures. She provided valuable insights into the broader landscape of social VR, ensuring that our study aligns with existing research and effectively addresses pertinent issues.

4 FINDINGS

4.1 Voice Changer Use Demographics

Our participants utilized various voice changer software options, including *Voicemod* [65], a tool that offers both AI and basic voice-changing capabilities, *Clownfish* [61], a more basic voice-changer that directly alters the audio on a frequency-level, *MorphVOX* [59, 60], a voice changer that uses neural networks for voice modification, *SuperVoiceChanger* [62], a basic, free piece of voice changing software, and *W-okada* [67], an open-source AI-based voice changer. Notably, Voicemod proved popular with six participants using it: P4 and P11 chose Voicemod Pro, P1, P2, and P10 opted for the non-Pro version, and P9 favored the Voicemod Persona AI voice feature specifically. Clownfish emerged as another common choice, selected by five participants (P2, P3, P5, P6, P8). P12 used MorphVox Pro, while two participants (P8, P13) used MorphVox Junior, and P7 selected SuperVoiceChanger. This diversity highlights the wide array of voice changer tools participants actively employed.

Table 2: Participant Experiences with Social VR and Voice Changers

ID	Voice Changer Experience	SocVR Frequency	Voice Changer
P1	Current	4-7 hours, 4 days/week	Voicemod
P2	Current	3-6 hours, 5 days/week	Voicemod, Clownfish
P3	Past	3-8 hours daily	Clownfish
P4	Current	evenings, 2-3 days/week	Voicemod Pro
P5	Current	weekly	Clownfish
P6	Current	6-12 hours/week	Clownfish
P7	Current	2-3 hours daily	SuperVoiceChanger
P8	Past	1-3 hours daily	Clownfish, MorphVOX Junior
P9	Current	variable	Voicemod Persona
P10	Current	weekly	Voicemod
P11	Current	limited	Voicemod Pro
P12	Past	2-8 hours/week	MorphVOX Pro
P13	Past	4-9 hours/week	MorphVOX Junior

In terms of the frequency of voice changer use, a notable trend emerged, with the majority of participants indicating near-constant utilization. Specifically, eleven participants (all except P2 and P10) reported using voice changers almost all the time, underlining the integral role of voice changers in their social VR interactions. Furthermore, participants demonstrated differences in terms of whom they were willing to use their unmodified voices with in social VR. Eight participants (P2, P3, P4, P5, P6, P7, P8, P11) expressed a willingness to use their unmodified voices with some or all of their close friends. On the other hand, five participants (P1, P9, P10, P12, P13) never turned off their voice changer even when with close friends. These findings highlight the complex interplay between voice changer usage, privacy, and interpersonal dynamics within the social VR landscape.

Interestingly, P10 used voice changers all the time in non-VR virtual social worlds, including roleplay environments like *Grand Theft Auto Online* [25] role-play servers and *Conan Exiles* [24], and other voice chat reliant games, but used voice changers minimally in social VR, estimating use at only 10% of the time. She regarded VRChat, her main social VR application, as a space for voice training, and thus tried to minimize her use of voice changers in the environment.

4.2 Motivations of Using Modified Voice

Participants approached the utilization and customization of voice changers with a range of distinct objectives in mind. While a common thread among all participants was the desire to authentically express their gender identity, many articulated specific motivations for harnessing this technology, including (1) adopting voice changers as a less arduous alternative to traditional voice training for achieving gender-affirming vocal presentations, (2) employing voice changers as a protective measure against potential instances of voice-related harassment, (3) aligning their virtual voice with their avatar's gender presentation, and (4) achieving an increased sense of personal comfort and confidence during social interactions in social VR. We elaborate on these motivations below.

Less Effort for Gender Identity Presentation Six participants (P3, P5, P6, P9, P12, P13) acknowledged that voice changers offered them a more manageable alternative to the demanding processes of medical transition and rigorous voice training to achieve their desired vocal presentation of their gender identities. While having limited support in current social VR platforms, and thus not a full solution to voice difficulties, voice changers offered our participants an accessible voice matching their identities. For transgender women and transfeminine nonbinary individuals (all participants other than

P2, P4, and P12), feminizing hormone therapy (MtF HRT) does not change their voice, and achieving a feminine voice through voice training can be extremely difficult, as training requires constant vocal exercises and safe spaces where they can practice using their trained voices. For the transmasculine participants (P4, P12), Testosterone-based HRT provided a route to an affirming voice, but was difficult to acquire as a result of medical gatekeeping and financial costs. When P4's voice started dropping because of testosterone, he no longer felt a need for the voice changer.

Avoiding Harassment Six participants (P4, P5, P8, P9, P11, P12) openly discussed the harassment they endured due to their unmodified voices in social VR. Three participants (P2, P4, P8) explicitly mentioned that even if they used an avatar aligning to their gender identity, others would often determine their gender based on their voice. The adoption of a voice changer offered an avenue to evade both the stigmatization directed at transgender users in social VR and, in the case of participants transitioning to a more masculine identity (P4, P12), the harassment typically directed at women.

Matching Voice to Avatar Eleven participants (except for P1 and P11) mentioned avatar-voice mismatch as a concern when presenting their gender identities and interacting with others in social VR. Using a voice changer to match their voice with the avatar reduced harassment from others, alleviated dysphoria, and made them more comfortable in their online identities. P8, who also uses her avatar for streaming from VRChat, mentioned how her unmodified voice doesn't sound right coming from her avatar, leading to voice-based gender dysphoria: *"I really do feel like my voice, my actual voice does not match my avatar. It's really dysphoric. When I think about it, I would much rather have a much more feminine voice. And it really, it really makes it hard to look at my VODs when I'm streaming because I'm hearing my voice coming from an anime girl and it's just like, yeah, why can't you just be more feminine?"*

Confidence in Voice Three participants (P3, P7, P9) mentioned that using a voice changer increased their confidence when speaking with others, while five (P4, P6, P8, P9, P11) participants stated that they would not interact with strangers in social VR without a voice changer. P9 highlights this increase in confidence: *"[T]hey're not hearing me, they're hearing a more idealized voice they're hearing, it will change how I feel about myself. Because it's not me they're hearing. And in my eyes, that's fine. And that's what enables me to engage like I do, because they're not hearing me."*

4.3 Social Affirmation of Identity via Passing

Ten participants (except for P2, P6, P10) with binary-related identities (that is, they identified with or in proximity to a binary identity) mentioned that voice changers enabled them to have experiences with *passing*—the social feedback of being treated as cisgender in their gender identity. These experiences of passing differ from just being around individuals that gendered them appropriately, in that even the unconscious biases separating transgender and cisgender identities disappeared.

Social feedback from others can play a significant role in affirming the gender identities of transgender individuals. Utilizing a passing voice changer cultivates an environment in which transgender users consistently receive affirmation of their gender identity, as others perceive and treat them in alignment with that identity. For several participants (P3, P5, P7, P8, P11, P13), the experiences of being perceived as cisgender and hearing themselves with a passing voice served to reinforce their gender identities. Additionally, for three participants (P1, P12, P13), the social feedback they received from the use of voice changers contributed to their process of self-discovery and bolstered their confidence in their new gender identities.

4.4 Experiences of Sexism and Preferential Treatment

Four participants (P4, P7, P12, P13) mentioned experiences with sexism and preferential gender treatment in social VR as a result of passing as cisgender. For the transmasculine participants (P4, P12), using the voice changer swapped their experiences from preferential treatment and sexism to being treated as “one of the guys.” For transfeminine participants (P7, P13), using the voice changer created situations where they were targets of sexism or preferentially treated by male VRChat users.

Transmasculine Experiences P4 and P12 mentioned that when not using the voice changer, they became targets of sexism and harassment based on their perceived genders. Other users perceived them as women, even when they were using male avatars. P12 described experiences of harassment when using their unmodified voice in social VR: *“I was immediately like, gendered in that phrase, it was like someone goes, someone insulted... They were like, they referred to me as bitch.”*

P4 highlighted that using a voice changer makes sexist harassment far less frequent, allowing him to be taken seriously by male users: *“I no longer experience instances like that kid coming up to me and asking me why I was a woman in that [male] avatar. People make less sort of comments about me that are in that nature. If someone’s gonna make fun of me, it’s because I did something ridiculous in VR as a joke. And not because of how I sound or my identity.”*

Both P4 and P12 had experiences of inclusion in male-dominated social circles, where utilizing voice changers allowed them to navigate and interact within these spaces. P4 articulated this phenomenon, stating: *“[P]eople are more comfortable with including me or like having myself sort of insert my own take to the joke more like, I find that the public worlds guys will primarily want to hang out with other guys and will be a bit more rude to women and so when I’m using the voice mod I get included in that guys group.”*

Both P4 and P12 recognized a loss in preferential treatment, with P12 experiencing harassment from males because his modified voice, while passing as male, sounded younger than his age. Both participants viewed this loss in preferential treatment positively. As P12 highlighted, preferential treatment based on being perceived as a woman is a source of disgust: *“I actually feel disgusted when I feel like I’m getting special treatment from cis guys, just because they perceive me to be a female or feminine.”*

Transfeminine Experiences For transfeminine participants, using voice changers often led to the barriers of being perceived as cisgender women in social VR environments. While this enabled them to explore a gender presentation aligned with their gender identities, it also exposed them to sexist behavior and harassment. P13 shared insights into how presenting as cisgender woman with a voice changer brought about both positive and negative experiences: *“It will also reveal to you what the negatives of being your preferred gender are very quickly. You know, people think you’re a girl and they find you attractive. Guys of all stripes will try to interact with you very poorly, and in a wide range of ways.”*

4.5 Voice Changers and Voice Training

Participants held divergent perspectives regarding the relationship between voice changers and voice training. Some participants perceived voice changers as impediments to their transition, viewing them as a roadblock. Others thought voice changers offered useful support, but could cause overreliance, acting as a “crutch.” Conversely, many participants regarded voice changers as valuable aids in their voice training journey. For certain individuals, the altered voice served as a benchmark for their voice training goals, while for others, particularly those using simpler, non-AI voice-changing technologies, the voice changer acted as a form of vocal “training wheels.”

Voice Changers as a Roadblock P3, P9, and P10 felt that voice changers posed certain roadblocks to voice training. P9 split voice training into two categories: passive voice training, where a transgender individual practices their voice by using it regularly in conversation, and active voice training, which includes vocal exercises and listening to a recording of one’s voice through headphones. The use of voice changer can largely discourage a transgender user’s passive voice training. P9 stressed that as the “almost ideal” voice from the voice changer becomes one’s identity in social VR, and passive voice training necessitates the use of the partially trained but intermediary voice, overcoming the fear of others hearing a different, “non-ideal” voice in passive voice training became a significant challenge for those reliant on voice changers: *“[P]eople have gotten used to your real voice. And they’re now hearing your altered voice like your intentionally changed sound. So it’s a lot like dropping the Voicemod and using a different voice.”*

Voice Changers as a “Crutch” P10, while using voice changers heavily in non-VR social online environments, made an effort to minimize her usage of voice changers in social VR. Treating social VR as an opportunity to practice using her new voice, P10 felt that voice changers sometimes acted as a barrier, but other times provided a necessary crutch—a support that can cause overreliance: *“The better term [to describe the voice changer] might be a crutch. Because I can always run back to this type of scenario... It lowers my motivation. If I keep it [on] it would lower my motivation if I rely on it too much. In certain scenarios, if that makes sense. [It lowers] my motivation to learn and, and work on my voice. The fact that it’s there is always good for if I’m too tired to even try.”*

P5 highlighted how, while useful, her overreliance on an AI voice changer prevented her from voice training in social VR. While the voice changer allowed her to access an affirming voice in social VR, she acknowledged that voice training would be necessary to affirm her identity outside the virtual world. *“I’ve known I would have to do [voice training] since the very beginning, but I just didn’t because I thought you know, have my computer adjusted for me. Which you know, as you get older, life exists outside of computer screen, and suddenly you can’t do that anymore.”* However, P5 also mentioned that this issue did not pertain to basic voice changing software like Clownfish, which could only change pitch and thus required voice training efforts alongside them.

Voice Changers for Setting Voice Training Goals Five participants (P5, P7, P9, P11, P13) regarded their modified voices by the voice changer as either a voice training goal or an ideal vocal representation they aimed to achieve. By establishing a target voice, particularly in situations where the voice changer retained elements of their natural voices, it served as a tangible model for attainable vocal results through training. P5, for instance, expressed: *“[The modified voice] almost gave me a goal to strive for. And it gives me the feeling of what it’s like to actually pass on the internet. And that is invaluable.”*

Voice Changer as “Training Wheels” for Voice Training Six participants (P1, P3, P5, P6, P8, P13) trained their original voice alongside the voice changer, with social VR as a training ground. Of these participants, all were using voice changers that preserved aspects of their original voice, rather than synthesizing an entirely new one. P1 employed a basic voice changer to solely adjust the pitch of her voice. Gradually, she reduced the pitch modification while simultaneously refining her voice through training. Over time, she succeeded in lowering the pitch modification to a point where she felt comfortable turning off the voice changer entirely, relying solely on her trained voice. This approach circumvented the challenges associated with using an unmodified voice, while also mitigating the abrupt and noticeable changes in one’s vocal presentation to others mentioned by P9. For P5, the limitations of non-AI voice changers like Clownfish, which could only change pitch, opened up

new avenues for voice training, since training other aspects of voice, such as resonance, could be done alongside the voice changer, which would offer a necessary support when focusing on specific areas of voice training: “[Y]ou have to practice it whilst listening to how it’s changing your voice in real time. It’s almost like enhancement, rather than a converter, if that makes sense.”

4.6 Limitations of Voice Changers

Several participants encountered various technical challenges while configuring and using their voice changers. These difficulties encompassed complex and convoluted settings that impeded a clear understanding of the voice changer’s functionality. Many participants also grappled with voice changer artifacts that not only caused annoyance to others but also risked exposing their true identities. Additionally, intelligibility issues, which occurred when others could not understand them when the voice changer was in effect, either because of glitchiness, low bitrate, or other voice artifacts, further complicated their experiences. Those who experimented with AI-based voice changers reported performance problems stemming from system strain, further exacerbating the technical hurdles they faced.

Difficulties with Settings Several participants encountered significant challenges during the configuration of their voice-changing software. Specifically, three participants (P6, P8, P11) expressed frustration with the complexity of voice changer settings. These interfaces were often made of multiple interacting components controlled by sliders, with little connection to real-life vocal characteristics like resonance or weight, as shown in the interfaces of SuperVoiceChanger (Figure A.1b), VoiceMod (Figure A.1a), and Clownfish (Figure A.2b). P3 and P8 detailed their experiences of discovering functional voice changer settings through trial-and-error, while P3 and P4 noted their reliance on assistance from friends in this process. Moreover, three participants (P3, P7, P12) voiced concerns regarding the available voice presets. Examples of these pre-loaded presets include MorphVox’s “Man”, “Woman”, and “Tiny Folks” voice presets (Figure A.2a), and the pre-loaded anime-style presets of W-okada (Figure A.2c). For example, P12 highlighted the lack of humanlike voice changers in his search for presets, “[A] lot of options just sounded like caricatures of a voice or something like that. Like an animated or like inhuman comed or some, like very animated version of an actual human voice.”

AI Voice Changers Four participants (P3, P5, P9, P11) raised concerns regarding the compatibility of AI-based voice changers, such as Voicemod Persona, with VRChat. They reported instances where both VRChat and the AI voice changer struggled to efficiently allocate system resources, resulting in undesirable consequences such as audio glitches and robotic-sounding voices. These performance disruptions manifested as various issues, ranging from users detecting the use of a voice changer to complete voice changer malfunctions. Notably, these performance challenges persisted even for participants like P9, who possessed a high-performance system equipped with a 12th generation Intel processor, underscoring the severity of the problem.

Besides off-the-shelf voice changer software, P11 used an open-source AI voice changer, named W-okada [67]. While having extensive programming background, P11 still found the voice changer overly-complex. W-okada required her to follow a YouTube tutorial, since the instructions for setup were in Japanese. Furthermore, she had to collect voice samples for training the AI, deal with command-line interfaces, and train each AI voice, which took over 100 hours. The results of this extensive work, however, were immediately appealing to the researcher—when the participant entered the interview, the researcher was unaware that the participant was using a voice changer. In addition, the relatively easy task of blending multiple AI voices together using the voice changer allowed P11 to create a

unique and humanlike voice. W-okada, unlike other AI voice changers, runs on the GPU. With an RTX 3060 12GB, the participant was able to both run VRChat and the voice changer at the same time without straining her system’s resources.

Paywalls Many of the most powerful voice modification tools, such as VoiceMod Pro and MorphVox Pro, are locked behind subscription services and paywalls. For transgender users, many of whom may not be able to afford expensive subscription services, the lack of powerful free options forces them to settle for less equipped voice changers. Five participants (P1, P2, P7, P11, P13) mentioned that paywalls or subscription services stopped them from trying fully-featured versions of voice changers. P13 highlighted how the companies behind these voice changers understand the market, and lock gender-swapping features behind paywalls: “They didn’t offer any good gender changing options in the free version, because of course they didn’t. They knew what people wanted to use it for. And they knew how they could get people to pay.”

Clocking While most participants reached relatively natural-sounding voices with their voice changers, five (P3, P5, P6, P8, P9) referenced “clockable” voice changers, that is, voice changers that others could tell were in use, as an issue. P5 and P9 mentioned accidentally using their unmodified voices when the voice changer was not enabled as a factor which could out them to others, with P9 mentioning a clear attitude shift when others heard her unmodified voice. For P3, the imperfect nature of the voice changer pushed her to pursue voice training, since she was jealous of others who did not have to worry about the voice changer giving away their identities: “I would say that I pass on my style because whenever I use voice changers I would get addressed correctly. Until some people got really smart and started figuring out it was a voice changer. And that’s one of the main reasons I wanted to stop too, because I started to get a little bit more jealous of people that were able to mark their voice using these voice mechanics, these voice skills [voice training] that I mentioned before.”

5 DISCUSSION

We present the first study to examine voice changers as a solution to TGNC users’ issues with voice-based dysphoria, misgendering, and harassment in social VR. We identified four reasons why TGNC individuals use voice changers in social VR (Section 4.2). These reasons include (1) as an alternative to difficult voice training, (2) as a protective measure against harassment, (3) to align their voice with their gender presentation, and (4) to achieve a sense of personal comfort and confidence when speaking with others. We also found that voice changer use affirmed TGNC users’ identities by enabling them to pass, and, for some, played a role in their exploration of their gender identities (Section 4.3). A notable impact of passing as cis was that TGNC users experienced swapped experiences of gender biases and sexism (Section 4.4).

We found that TGNC individuals had mixed views of how voice changers interacted with voice training. Some users felt that voice changers acted as a roadblock, making voice training more difficult to pursue. Others felt that the voice changers provided necessary support when they couldn’t or didn’t want to voice train, but could cause overreliance, acting as a “crutch.” Still, many users recognized voice changers as a springboard for voice training, providing a supporting role during the voice training process, both as voice “training wheels” by allowing individuals to safely and incrementally alter their voice, and as a voice training goal (Section 4.5).

Besides social impact, TGNC voice changer users noted that voice changing software had various technical problems and limitations. Settings were vague and confusing, AI voice changers were either too resource intensive—or in the case of W-Okada, impossible to use without programming knowledge—and regularly locked behind paywalls and subscriptions (Section 4.6).

In this section, we discuss the unique possibilities offered by voice changers as a supplement to avatars for gender exploration in social VR, design implications for creating effective and user-friendly voice changers that meet TGNC users' needs, and considerations for leveraging voice changers to support voice training.

5.1 Gender Exploration via Voice Changers in Social VR

Previous work evaluated avatars as a tool for gender exploration and affirmation, but noted that the main form of communication, voice chat, posed risks of outing TGNC users to others [22]. Our findings regarding participants' cis-passing experiences illustrate how the combination of social VR avatars and voice changers can effectively simulate gender presentations that might otherwise require costly and labor-intensive medical transitions and voice training in real life. This lower-effort and accessible solution to changing others' perception of one's gender poses unique benefits for transgender identity exploration. While previous studies have explored the impact of avatar embodiment effects on identity and gender exploration [21, 54], these studies are limited to an individual's perceptions of one's body, and do not explore the impacts of being treated as the avatar's gender in the social VR environment.

Multiple participants said that using a voice changer strengthened their decision to transition, since they enjoyed the social aspects of presenting as the gender they aligned their voice and avatar with. These virtual experiences differ from typical first-experiences of attempting social transition, where hurdles such as difficulty with getting gendered appropriately by others, being treated as something separate from the goal gender (e.g., getting treated as a transgender woman rather than just as a woman), and the stigma of being transgender hinder gender exploration. External factors, such as social stigma and non-acceptance are major drivers in cessation of transition [64]. By providing an idea of what an end-game for transitioning could be like, virtual experiences of passing can be profoundly affirming and supportive, providing individuals early in transition with the opportunity to build confidence in their decision to transition. Social VR, when used with a voice changer, offers a unique platform for transgender identity exploration that complements traditional real-world experiences, ultimately contributing to a more informed and empowered decision-making process for individuals on their gender journey.

5.2 Design Implications for Voice Changer Usability

Our study identified several key challenges around the use of voice changers, including the scarcity of available voice presets, the intricate nature of voice changer setup, the performance instability arising from system resource constraints, and the affordability of subscription-based models for accessing essential features. These challenges highlight areas where improvements can be made to enhance the usability of voice changers for TGNC users.

Designing better foundations Voice changing software presets often revolve around generic character archetypes, featuring options like anime-inspired female personas or stereotypical dorky male voices. Unfortunately, these presets fall short of meeting the unique needs of transgender users who are in search of voice modifications that truly reflect their individual identities. Two of the voice changers used by participants avoided this issue in different ways, but were imperfect in other ways. The neural network voice transformer found in MorphVox allows transgender users to maintain certain elements of their original voices while achieving desired modifications, but the sliders for further customization are complex and difficult to understand. W-Okada, which enables users to blend multiple AI voices to craft a personalized and unique voice, offers transgender users a practical means of establishing realistic voice goals and fostering a stronger sense of identity affirmation, but is extremely difficult to use given the command-line interfaces

required for voice model generation and merging and lack of documentation. While not a real-time solution, AVOCUS, by Byeon et al., presents an alternative voice alteration system that enables target voice searching with an understandable hashtag system and customization based on the user's own voice, preserving vocal features [12]. To support more effective voice modification, developers should consider integrating design foundations produced by prior research (e.g., AVOCUS) alongside the specific needs of target groups, such as TGNC users, to produce more tailored tools.

Keep it simple Our findings highlight the challenges encountered by numerous transgender users when attempting to create convincing human-like voices using voice changers equipped with an overwhelming array of options. Not to mention that our recruitment criteria limited our ability to capture the experiences of transgender individuals who faced difficulties in developing a functional voice and subsequently abandoned their efforts. Future voice changer design should simplify the voice adjustment process by limiting the number of parameters; abstracting features into understandable categories for transgender users that mimic aspects of voice training, such as voice pitch and resonance/weight; involving neural-network training on users' voices for best-fit approximations of gender swaps, such as in MorphVox; or providing AI presets and allowing users to merge them using a single weight value for each.

Addressing processor constraints AI voice changers, despite their capacity to generate highly realistic human voices, can exert significant pressure on the CPU. This strain on the CPU becomes problematic, particularly in social VR applications such as VRChat, where the CPU is responsible for managing numerous users with intricate avatars within the virtual environment. Our research findings indicate that GPU-based AI voice changers, exemplified by W-Okada, do not encounter this problem. Since individuals employing PCs for social VR typically possess gaming-grade GPUs, directing the development of future AI voice changers towards GPU compatibility should effectively address this concern. Still, developers should keep in mind the hardware available to users and the types of software that run alongside their voice changers. When the application is less strenuous, such as voicechat or videochat applications (e.g., Discord, Zoom), CPU voice changers may be the optimal choice, since they do not assume that users have access to powerful GPUs, but in more CPU-intensive applications (e.g., VRChat), developers should consider adding GPU-compatible modes.

A community-driven solution Developers lock gender-swapping settings in voice changers behind paywalls, and transgender users are regularly not in positions to afford expensive voice changer subscriptions. A community-driven solution, integrating voice model crowdsourcing networks like AI Hub, a Discord community for sharing voice models, alongside a piece of open-source software, like W-Okada, would satisfy the needs of these users, while also providing a platform for researchers to build on.

5.3 Voice Changers as a Support for Voice Training

Voice training is a vital approach for TGNC users to align their voice with their gender identities. Given the arduous process of voice training and difficulties with measuring intermediary success, TGNC individuals have developed communities, applications, and online video-based and text-based guides for voice training techniques. Historically, modern media technologies have been integral to sharing insights on transgender voice training, with "Melania Speaks!" providing the first full series on transgender voice training on VHS and DVD in the 1990s [1]. New methods range from voice training video series on YouTube [63], Discord communities for voice training such as "TransVoice" and "Scientifically Augmented Voice" [3], and guides on Reddit for attaining specific goals [2], to full-fledged mobile applications which automatically determine feminine, masculine, and androgynous qualities of their voices [6, 7]. In

our study, we identified social VR as a new avenue for voice training, with voice changers as a possible support, crutch, and roadblock.

As potential support, participants reported a range of experiences with voice changers in connection to their voice training goals. Some used the modified versions of their own voices from the voice changer as a step towards their vocal targets. Others found that using voice changers motivated them to start formal voice training. However, some chose to stop using voice changers and instead used social VR as a platform for practicing their voice. There were also those who combined the use of voice changers with voice training, combining alterations in their voice changer with alterations in their voice to avoid suspicion from others. Based on these findings, a future work could explore voice changing technology in combination with social VR as an important addition to modern transgender voice therapy.

Social VR as a Safe Environment for Passive Training While useful for exercising vocal muscles and ascertaining the feeling of a target voice, the exercises supported by modern voice training media, applications, and community events focus primarily on active voice training, where most if not all cognition is directed towards maintaining the desired voice. In contrast, passive voice training requires using the trained, intermediary voice in day-to-day communication, a task which normalizes practicing the new voice in parallel with performing other cognitively challenging tasks, which is a major barrier for voice training to many individuals and requires additional techniques, such as “voice-resetting” [18, 42]. Passive voice training is important since it regularizes consistently using a trained voice, and thus can overwrite vocal habits of the previous voice [52]. However, this would typically force TGNC individuals to come out to others in real life, since real-world interactions would be the primary forum for passively training. Being out as a TGNC person, especially with a non-passing voice, poses enormous risks, including social, economic, and physical risks [28]. While still not free from risks of harassment, social VR environments provide a virtual emulation of real-world interactions in which TGNC individuals can practice passive voice training without the risks of coming out in real life.

Voice Changers as a Safety Net Voice changers provide a way to address the concerns of others hearing an intermediary, non-passing voice during voice training in social VR by altering the voice others hear to be more masculine or feminine. This allows TGNC users to avoid harassment in public virtual worlds while also maintaining comfort in that the voice that others hear represents themselves. By separating voice presentation from practice, such as by having others hear the modified voice while the practicing individual hears their trained voice, TGNC individuals engaging in voice training can enjoy the best of both worlds.

An Iterative Design for Transgender Voice Therapy Combining the findings from our study alongside modern voice training techniques, we recommend a framework for voice training to be explored in future work, consisting of (1) using voice changers for discovering a participant’s target voice, (2) using voice changers to modify and improve the participant’s voice to the target voice, (3) comparing the current trained voice with the target voice to develop statistics to find approachable iterations for voice training in areas such as pitch and resonance, and (4) having participants practice their new voice in a social VR environment, all while providing live statistics and feedback and enhancing their voice with a voice changer to avoid harassment and increase participant comfort. This approach would (1) make voice training more approachable by presenting steps the participant could take in achieving their goal voice, instead of presenting the process as a single journey and (2) allow participants to manage their safety and comfort more effectively in passive voice training scenarios, making passive voice training more

accessible to participants who do not want to or cannot present their identities in real life.

5.4 Limitations and Future Work

This study has several limitations that warrant consideration. Our participants exclusively used VRChat, and as such, the generalizability of their experiences to other social VR platforms like RecRoom and Horizon Worlds remains uncertain. Future research should endeavor to explore a more diverse range of social VR platforms to determine if the findings hold true across different virtual environments. Moreover, the study featured a limited number of participants, which inherently restricts the scope of the study and potentially limits its applicability to a broader user base. To address this limitation, a future research endeavor could extend the study duration and recruitment efforts to include a larger and more diverse sample of users who employ voice changers for gender identity presentation.

Another noteworthy limitation is that our study did not delve into the experiences of TGNC users who abandoned the use of voice changers and the reasons behind their decisions. A subsequent research project could adopt a survey-based approach designed around our findings while widening the eligibility criteria (e.g., including participants who used or tried but abandoned voice changers). This approach would provide a more comprehensive understanding of how to design inclusive and easy-to-use voice changers that cater to the needs of TGNC users effectively.

An additional limitation of our study is the exclusive use of qualitative methodology. While this approach allowed us to cover nuanced personal and social phenomena as they occurred in our exploratory research, the lack of quantitative data prevents us from providing measurable, statistical insights into the use of voice changers by TGNC users in social VR. Future research could benefit from a mixed-methods approach, integrating quantitative data such as online surveys to create a more holistic understanding of the role and impact of voice changers for TGNC users in social VR.

Moreover, participants in our study recounted past experiences, which might have resulted in some details and difficulties related to voice changer setup being overlooked or inaccurately remembered. Future research could involve participants who are new to voice-changing technology and assign them the task of setting up a voice changer to gain more precise insights into the setup process and associated challenges.

Finally, participants primarily remembered salient and recognizable behaviors of others when recalling experiences of harassment, misgendering, or other negative encounters tied to using their unmodified voices in social VR. Future research efforts could adopt an alternative approach by capturing these experiences through participants recording their social VR sessions while using and not using voice changers, similar to those employed in Zhang et al. [70]. This method would provide a more comprehensive and detailed perspective on the challenges faced by users in real-time scenarios.

6 CONCLUSION

In this paper, we conducted interviews to delve into the experiences of TGNC users employing voice changers in social virtual reality (VR) environments. Our findings have shed light on several critical aspects, including the diverse goals that TGNC individuals aim to achieve with voice changers, the marked disparities in their experiences with and without these voice-modifying tools, the profound impact of “passing” as cisgender in social VR for those who may not have this opportunity in real-life settings, attitudes around the complex relationship between voice training and voice changers, and the limitations of current voice changers for TGNC users.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant No. IIS-2328182.

REFERENCES

- [1] Melanie speaks! on developing a female voice, 1996, 1993.
- [2] wiki - transvoice. <https://www.reddit.com/r/transvoice/wiki/index/>, 2021. Accessed: 2024-01-16.
- [3] U. D. account). Collecting the trans voice discord servers in one place. https://www.reddit.com/r/transvoice/comments/odw2z1/collecting_the_trans_voice_discord_servers_in_one/, 2021. Accessed: 2024-01-16.
- [4] D. Acena and G. Freeman. "in my safe space": Social support for lgbtq users in social virtual reality. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI EA '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411763.3451673
- [5] A. A. Ahmed. Trans Competent Interaction Design: A Qualitative Study on Voice, Identity, and Technology. *Interacting with Computers*, 30(1):53–71, 11 2017. doi: 10.1093/iwc/iwx018
- [6] A. A. Ahmed, L. Kim, and A. L. Hoffmann. 'this app can help you change your voice': Authenticity and authority in mobile applications for transgender voice training. *Convergence*, 28(5):1283–1302, 2022. doi: 10.1177/13548565221079459
- [7] A. A. Ahmed, B. Kok, C. Howard, and K. Still. Online community-based design of free and open source software for transgender voice training. *Proc. ACM Hum.-Comput. Interact.*, 4(CSCW3), jan 2021. doi: 10.1145/3434167
- [8] K. Bessière, A. Seay, and S. Kiesler. The ideal elf: Identity exploration in world of warcraft. *Cyberpsychology behavior : the impact of the Internet, multimedia and virtual reality on behavior and society*, 10:530–5, 09 2007. doi: 10.1089/cpb.2007.9994
- [9] W. O. Bockting, M. H. Miner, R. E. Swinburne Romine, A. Hamilton, and E. Coleman. Stigma, mental health, and resilience in an online sample of the us transgender population. *American Journal of Public Health*, 103(5):943–951, 2013. PMID: 23488522. doi: 10.2105/AJPH.2013.301241
- [10] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.
- [11] J. Buss, H. Le, and O. L. Haimson. Transgender identity management across social media platforms. *Media, Culture & Society*, 44(1):22–38, 2022. doi: 10.1177/01634437211027106
- [12] H. J. Byeon, S. Ha, and U. Oh. Avocus: A voice customization system for online personas. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI EA '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544549.3585892
- [13] M. Carrasco and A. Kerne. Queer visibility: Supporting lgbt+ selective visibility on social media. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–12. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173824
- [14] Cleveland Clinic. Transgender voice therapy. <https://my.clevelandclinic.org/health/treatments/24510-transgender-voice-therapy>, 2022. Accessed: 2024-01-16.
- [15] S. L. Craig, A. D. Eaton, L. B. McNroy, V. W. Y. Leung, and S. Krishnan. Can social media participation enhance lgbtq+ youth well-being? development of the social media benefits scale. *Social Media + Society*, 7(1):2056305121988931, 2021. doi: 10.1177/2056305121988931
- [16] Y. Cui, N. Yamashita, M. Liu, and Y.-C. Lee. "so close, yet so far": Exploring sexual-minority women's relationship-building via online dating in china. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3517624
- [17] J. N. Fish, L. B. McNroy, M. S. Pacey, N. D. Williams, S. Henderson, D. S. Levine, and R. N. Edsall. "i'm kinda stuck at home with unsupportive parents right now": Lgbtq youths' experiences with covid-19 and the importance of online support. *Journal of Adolescent Health*, 67(3):450–452, 2020. doi: 10.1016/j.jadohealth.2020.06.002
- [18] O. Flanagan. 5 best ways to practice your new voice in conversation - vocal feminization. <https://www.youtube.com/watch?v=V1nDaGZ6AKs>, 2023. Video clip: [Start 0:00, End 1:09]. Accessed: 2024-01-17.
- [19] J. Fox and W. Y. Tang. Sexism in online video games: The role of conformity to masculine norms and social dominance orientation. *Computers in Human Behavior*, 33:314–320, 2014. doi: 10.1016/j.chb.2013.07.014
- [20] G. Freeman and D. Acena. "acting out" queer identity: The embodied visibility in social virtual reality. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2), nov 2022. doi: 10.1145/3555153
- [21] G. Freeman and D. Maloney. Body, avatar, and me: The presentation and perception of self in social virtual reality. *Proc. ACM Hum.-Comput. Interact.*, 4(CSCW3), jan 2021. doi: 10.1145/3432938
- [22] G. Freeman, D. Maloney, D. Acena, and C. Barwulor. (re)discovering the physical body online: Strategies and challenges to approach non-cisgender identity in social virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3502082
- [23] G. Freeman, S. Zamanifard, D. Maloney, and D. Acena. Disturbing the peace: Experiencing and mitigating emerging harassment in social virtual reality. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW1):1–30, 2022.
- [24] Funcom. Conan exiles. <https://www.conanexiles.com/>, 2018. Accessed: 2024-01-17.
- [25] R. Games. Grand theft auto online. <https://www.rockstargames.com/gta-online>, 2013. Accessed: 2024-01-17.
- [26] E. Garrett, Willie. Marginalized populations. https://www.mnpsych.org/index.php?option=com_dailyplanetblog&view=entry&category=division%20news&id=71:marginalized-populations, 2017. Accessed: 2023-10-04.
- [27] M. P. Gelfer and R. M. Tice. Perceptual and acoustic outcomes of voice therapy for male-to-female transgender individuals immediately after therapy and 15 months later. *Journal of Voice*, 27(3):335–347, 2013. doi: 10.1016/j.jvoice.2012.07.009
- [28] GLAAD. Unsafe in america: New glad data shows unprecedented threats and attacks against lgbtq americans. <https://glaad.org/releases/unsafe-america-new-glaad-data-shows-unprecedented-threats-and-attacks-against-lgbtq>, 2023. Accessed: 2024-01-17.
- [29] hai-vr. Combogestureexpressions: Attach face expressions to hand gestures. <https://github.com/hai-vr/combo-gesture-expressions-av3>, 2024. Accessed: 2024-01-17.
- [30] O. L. Haimson. *The Social Complexities of Transgender Identity Disclosure on Social Media*. PhD thesis, 2018. Copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2023-03-03.
- [31] O. L. Haimson, J. R. Brubaker, L. Dombrowski, and G. R. Hayes. Disclosure, stress, and support during gender transition on facebook. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, p. 1176–1190. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2675133.2675152
- [32] O. L. Haimson, J. Buss, Z. Weinger, D. L. Starks, D. Gorrell, and B. S. Baron. Trans time: Safety, privacy, and content warnings on a transgender-specific social media site. *Proc. ACM Hum.-Comput. Interact.*, 4(CSCW2), oct 2020. doi: 10.1145/3415195
- [33] A. B. Hancock, J. Krissinger, and K. Owen. Voice perceptions and quality of life of transgender people. *Journal of Voice*, 25(5):553–558, 2011. doi: 10.1016/j.jvoice.2010.07.013
- [34] C. Hodges-Simeon, G. Grail, G. Albert, M. Groll, C. Stepp, J. Carré, and S. Arnocky. Testosterone therapy masculinizes speech and gender presentation in transgender men. *Scientific Reports*, 11, 02 2021. doi: 10.1038/s41598-021-82134-2
- [35] J. M. Hughto, D. Pletta, L. Gordon, S. Cahill, M. J. Mimiaga, and S. L. Reisner. Negative transgender-related media messages are associated with adverse mental health outcomes in a multistate study of transgender adults. *LGBT health*, 8(1):32–41, 2021.
- [36] S. J. Jackson, M. Bailey, and B. F. Welles. #girlslkeus: Trans advocacy and community building online. *New Media & Society*, 20(5):1868–1888, 2018. doi: 10.1177/1461444817709276
- [37] S. Jhaver, S. Ghoshal, A. Bruckman, and E. Gilbert. Online harassment

- and content moderation: The case of blocklists. *ACM Trans. Comput.-Hum. Interact.*, 25(2), mar 2018. doi: 10.1145/3185593
- [38] T. F. Ji, B. Cochran, and Y. Zhao. Vrbubble: Enhancing peripheral awareness of avatars for people with visual impairments in social virtual reality. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3517428.3544821
 - [39] Kevtendo. Keyboard girl avatar — vrchat. <https://youtu.be/q0nA-1XDgDU?si=Zv43cZa2CBsyX0Nt>, 2021. Accessed: 2024-01-17.
 - [40] H.-T. Kim. Vocal feminization for transgender women: Current strategies and patient perspectives. *International Journal of General Medicine*, 13:43–52, 2020. doi: 10.2147/IJGM.S205102
 - [41] L. Li, G. Freeman, K. Schulenberg, and D. Acena. "we cried on each other's shoulders": How lgbtq+ individuals experience social support in social virtual reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3581530
 - [42] J. Majdalani. How to maintain a deeper voice in conversation. <https://www.youtube.com/watch?v=5n9ZR8ZrcAo>, 2022. Accessed: 2024-01-17.
 - [43] D. Maloney, G. Freeman, and D. Wohn. "talking without a voice": Understanding non-verbal communication in social virtual reality. *Proceedings of the ACM on Human-Computer Interaction*, 4:1–25, 10 2020. doi: 10.1145/3415246
 - [44] M. E. Morris, D. K. Rosner, P. S. Nurius, and H. M. Dolev. "i don't want to hide behind an avatar": Self-representation in social vr among women in midlife. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, DIS '23, p. 537–546. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3563657.3596129
 - [45] D. Morsomme and A. Remacle. Can ambulatory biofeedback help a transgender woman speak at a higher pitch? *Logopedics Phoniatrics Vocology*, 47(2):125–132, 2022. PMID: 33660586. doi: 10.1080/14015439.2021.1881158
 - [46] M. Mott, E. Cutrell, M. Gonzalez Franco, C. Holz, E. Ofek, R. Stoakley, and M. Ringel Morris. Accessible by design: An opportunity for virtual reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 451–454, Oct 2019. doi: 10.1109/ISMAR-Adjunct.2019.00122
 - [47] B. Nic Giolla Easpaig. An exploratory study of sexism in online gaming communities: mapping contested digital terrain. *Community Psychology in Global Perspective*, 4(2):119–135, 2018.
 - [48] I. Nolan, S. Morrison, O. Arowojolu, C. Crowe, J. Massie, R. Adler, S. Chalet, and D. Francis. The role of voice therapy and phonosurgery in transgender vocal feminization. *Journal of Craniofacial Surgery*, Publish Ahead of Print:1, 01 2019. doi: 10.1097/SCS.0000000000005132
 - [49] J. Oates, R. Adler, S. Hirsch, and J. Pickering. Evidence-based practice in voice training for trans women. *Voice and communication therapy for the transgender/gender diverse client: A comprehensive clinical guide*, pp. 87–103, 2019.
 - [50] J. Oates and G. Dacakis. Transgender voice and communication: Research evidence underpinning voice intervention for male-to-female transsexual women. *Perspectives on Voice and Voice Disorders*, 25(2):48–58, 2015. doi: 10.1044/vvd25.2.48
 - [51] O. C. Oguine and K. Badillo-Urquiola. You don't belong here: Ableist microaggressions on adolescents with disability (ies) and special needs in social virtual reality (vr). In *CHI'23: Workshop on Combating Toxicity, Harassment, and Abuse in Online Social Spaces, Hamburg, Germany*, 2023.
 - [52] Penn Medicine Becker ENT Allergy. Common challenges faced by transfeminine speakers during voice training. <https://www.beckerentandallergy.com/blog/common-challenges-faced-by-transfeminine-speakers-during-voice-training>, 2024. Accessed: 2024-01-16.
 - [53] S. A. Rains, S. R. Brunner, C. Akers, C. A. Pavlich, and S. Goktas. Computer-mediated communication (cmc) and social support: Testing the effects of using cmc on support outcomes. *Journal of Social and Personal Relationships*, 34(8):1186–1205, 2017. doi: 10.1177/0265407516670533
 - [54] Z. Reyes and J. Fisher. The impacts of virtual reality avatar creation and embodiment on transgender and genderqueer individuals in games: A grounded theory analysis of survey and interview data from transgender and genderqueer individuals about their experiences with avatar creation interfaces in virtual reality. In *Proceedings of the 17th International Conference on the Foundations of Digital Games*, FDG '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3555858.3555882
 - [55] S. Riches, S. Elghany, P. Garety, M. Rus-Calafell, and L. Valmaggia. Factors affecting sense of presence in a virtual reality social environment: A qualitative study. *Cyberpsychology, Behavior, and Social Networking*, 22(4):288–292, 2019. PMID: 30802148. doi: 10.1089/cyber.2018.0128
 - [56] Royal College of Speech & Language Therapists. Trans voice – clinical information for slts. <https://www.rcslt.org/speech-and-language-therapy/clinical-information/trans-voice/>, 2017. Accessed: 2024-01-16.
 - [57] S. Schneider and M. Courey. Transgender voice and communication - vocal health and considerations. <https://transcare.ucsf.edu/guidelines/vocal-health>, June 2016. Accessed: 2024-01-16.
 - [58] K. Schulenberg, G. Freeman, L. Li, and C. Barwulor. "creepy towards my avatar body, creepy towards my body": How women experience and manage harassment risks in social virtual reality. 2023.
 - [59] Screaming Bee Inc. Morphvox junior. <https://screamingbee.com/morphvox-free-voice-changer>, n.d. Accessed: 2023-10-04.
 - [60] Screaming Bee Inc. Morphvox pro. <https://screamingbee.com/morphvox-voice-changer>, n.d. Accessed: 2023-10-04.
 - [61] Shark Labs. Clownfish voice changer. <https://clownfish-translator.com/voicechanger/>, n.d. Accessed: 2023-10-04.
 - [62] SuperVoiceChanger Inc. Supervoic changer. <https://supervoicchanger.en.uptodown.com/windows>, n.d. Accessed: 2023-10-04.
 - [63] TransVoiceLessons. Transvoicelessons youtube channel. <https://www.youtube.com/@TransVoiceLessons>, 2024. Accessed: [insert date of access here]2024-01-16.
 - [64] J. L. Turban, S. S. Loo, A. N. Almazan, and A. S. Keuroghlian. Factors leading to "detransition" among transgender and gender diverse people in the united states: A mixed-methods analysis. *LGBT Health*, 8(4):273–280, 2021. PMID: 33794108. doi: 10.1089/lgbt.2020.0437
 - [65] VoiceMod. Voicemod. <https://www.voicemod.net/>, n.d. Accessed: 2023-10-04.
 - [66] VRChat. Simple pen system - udon example scene. <https://creators.vrchat.com/worlds/examples/udon-example-scene/simple-pen-system/>, 2023. Accessed: 2024-01-17.
 - [67] w okada. W-okada voice changer. <https://github.com/w-okada/voice-changer>, n.d. Accessed: 2023-10-04.
 - [68] S. Zamanifard and G. Freeman. "the togetherness that we crave": Experiencing social vr in long distance relationships. In *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*, CSCW '19, p. 438–442. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3311957.3359453
 - [69] K. Zhang, E. Deldari, Z. Lu, Y. Yao, and Y. Zhao. "it's just part of me": understanding avatar diversity and self-presentation of people with disabilities in social virtual reality. In *The 24th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, oct 2022. doi: 10.1145/3517428.3544829
 - [70] K. Zhang, E. Deldari, Y. Yao, and Y. Zhao. A diary study in social virtual reality: Impact of avatars with disability signifiers on the social experiences of people with disabilities. 08 2023. doi: 10.1145/3597638.3608388
 - [71] Y. Zhao, C. L. Bennett, H. Benko, E. Cutrell, C. Holz, M. R. Morris, and M. Sinclair. Enabling people with visual impairments to navigate virtual reality with a haptic and auditory cane simulation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, p. 1–14. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173690