Does Who You Are or Appear to Be Matter?: Understanding Identity-Based Harassment in Social VR Through the Lens of (Mis)Perceived Identity Revelation

KELSEA SCHULENBERG, Clemson University, USA
GUO FREEMAN, Clemson University, USA
LINGYUAN LI, Clemson University, USA
BART PIET KNIJNENBURG, Clemson University, USA

The popularity of social virtual reality (VR) platforms such as VRChat has led to growing concerns about new and more severe forms of online harassment targeting one's identity characteristics (i.e., *identity-based harassment*). Social VR users with marginalized identities (e.g., women, LGBTQIA+ individuals, and racial/ethnic minorities) have been reported as particularly vulnerable to such harassment. This is mainly because social VR can make one's offline identity known to others (i.e., what we term *identity revelation* in this work) through a unique combination of avatar design, voice use, and immersive full- or partial-body tracking. To address these safety concerns, there is an urgent need to unpack the complex dynamics surrounding how one's offline identity is (mis)perceived by others, and how these (mis)perceptions may affect identity-based harassment in social VR. This study thus utilizes a large-scale survey with 223 social VR users across six continents/regions of the world with varying social VR experiences and identities to investigate (1) the relationship between identity-based harassment in social VR and (mis)perceptions of selective identity revelation practices, (2) how embodying one's identity in social VR might actually be less risky than once thought, and (3) how who you are *does* still matter when it comes to identity-based harassment in social VR. It also highlights the need to better account for understudied aspects of identity-based harassment in social VR and to better educate social VR users on the interplay between harassment and *(mis)perceived* identity revelation in these spaces.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**;

Additional Key Words and Phrases: identity characteristics; identity revelation; online presentation; social virtual reality; online harassment

ACM Reference Format:
Kelsea Schulenberg, Guo Freeman, Lingyuan Li, and Bart Piet Knijnenburg. 2024. Does Who You Are or Appear to Be Matter?: Understanding Identity-Based Harassment in Social VR Through the Lens of (Mis)Perceived Identity Revelation. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 384 (November 2024), 39 pages. https://doi.org/10.1145/3686923

1 INTRODUCTION

Social virtual reality (VR) platforms are spaces where multiple users can interact with one another, typically through VR head-mounted displays, and with immersive 360-degree virtual content in 3D virtual spaces [56]. Examples include VRChat, Meta's Horizon Worlds, and RecRoom. Such

Authors' addresses: Kelsea Schulenberg, Clemson University, South Carolina, USA, 29634, kelseas@g.clemson.edu; Guo Freeman, Clemson University, Clemson, South Carolina, USA, 29634, guof@clemson.edu; Lingyuan Li, Clemson University, South Carolina, USA, 29634, lingyu2@clemson.edu; Bart Piet Knijnenburg, Clemson University, South Carolina, USA, 29634, bartk@clemson.edu.

2573-0142/2024/11-ART384
https://doi.org/10.1145/3686923

platforms are often praised for dramatically transforming online social interaction through a unique combination of features designed to enhance embodied and immersive social experiences. These include full- or partial-body tracked avatars (i.e., one's avatar body actions correspond to one's physical body actions in real-time); predominate voice communication; body language and gestures; and simulated immersive activities [27, 56, 96]. This unique combination allows users to craft novel approaches to **identity practices**, such as helping transgender individuals explore and experiment with their gender identities in more embodied, intimate, and personal ways [23, 27, 28, 31].

Unfortunately, engaging with one's identity in social VR can become a double-edged sword because it can make users more visible targets for **identity-based harassment**, i.e., harassment focused on one's identity characteristics. Prior works have indeed shown that social VR users who present certain identity characteristics through their avatar design, voice, or nonverbal cues frequently report experiencing such physicalized harassment as embodied sexual assault [7, 8, 30, 66, 68, 90, 92, 93]. These experiences appear to disproportionately target social VR users whose *offline identities* are considered marginalized within and outside of social VR (e.g., women, LGBTQIA+, racial/ethnic minorities, people with disabilities) [7, 8, 30, 86, 90, 92, 93, 110, 111]. In particular, this targeting seems to happen regardless of how marginalized individuals actually present themselves in social VR, as even misperceptions of their identities can be used against them. For example, despite their actual gender identity as women, transgender women have reported being misperceived and subsequently harassed in social VR for sounding "like a man" due to voice pitch expectations associated with their sex assigned at birth (male) [86].

As such, there is an urgent need to unpack the complex dynamics surrounding how one's offline identity comes to be (mis)perceived by others (i.e., what we term **identity revelation**), and how these (mis)perceptions may affect identity-based harassment in social VR. This understanding is crucial to creating safer social VR spaces for all, especially protecting people with marginalized identities. In this work, we use **identity revelation** as an umbrella categorization term to explore two phenomena: (1) situations where individuals present or disclose their own offline identity characteristics while in social VR to be correctly *perceived* by others, and (2) situations where an individual's offline identity characteristics are *misperceived* by others in social VR regardless of the user's actual presentation. While we acknowledge that there are additional nuances to the different ways one could communicate one's offline identity while in social VR (disclosing v. presenting v. gender assumptions v. being outed), our study's focus is not to parse out these intricacies. Rather, we aim to understand how identity-based harassment in social VR is affected differently by having one's offline identity characteristics in social VR *correctly perceived* vs. *misperceived* by others overall. In this sense, this lens of "(mis)perceived identity revelation" allows us to investigate both phenomena comprehensively.

Therefore, we conducted an anonymous online survey with 223 social VR users, who represent a geographically and culturally diverse sample from six countries/regions of the world and with varied genders, sexualities, and ethnicities. We specifically explore the following research questions:

RQ1: How do people present or disclose their various identity characteristics in social VR?

RQ1a: How frequently do people present or disclose their various identity characteristics in social VR, and to whom?

RQ1b: Which methods do they use to present or disclose their various identity characteristics in social VR?

RQ2: How are people's identity characteristics misperceived by others in social VR?

RQ2a: What identity characteristics do users report as being frequently misperceived by others in social VR?

RQ2b: What inform these misperceptions about other social VR users' identity characteristics?

RQ3: How are the identity revelation types and identity characteristics associated with harassment in social VR?

RQ3a: Which presented/disclosed or misperceived identity characteristics are associated with more frequent harassment in social VR?

RQ3b: When their identity characteristics are presented/disclosed or misperceived, which user groups may encounter more frequent harassment in social VR than others?

In answering these questions, we make three contributions to the fast-growing HCI and CSCW literature on understanding and mitigating harassment and harm in novel online social spaces such as social VR. First, by quantitatively investigating identity-based harassment in social VR using a non-U.S. centric sample, we bridge and extend upon the insights provided by prior, mostly qualitative works that often focus on U.S.-based populations [7, 8, 30, 65, 66, 90]. In doing so, we shed light on the global phenomenon of identity-based harassment targeting women, LGBTQIA+, and racial/ethnic minority user groups in social VR. Second, the large-scale and comprehensive nature of our study enables us to provide several previously un- or underexplored understandings, including: (1) unpacking the relationship between identity-based harassment in social VR and (mis)perceptions of selective identity revelation practices; (2) new evidence of how embodying one's offline identity in social VR may be less risky than previously thought by researchers and users alike; and (3) despite this more positive picture, different social VR user groups indeed experience significantly different frequencies of harassment for certain identity characteristics, namely by user gender identity, sexuality, and race/ethnicity. Third, we propose two primary implications for investigating and designing safer social VR spaces: re-orientating social VR research and development towards previously underexplored aspects of identity-based harassment and its connection to (mis)perceived identity revelation , and re-approaching how the interplay between the two should be communicated to social VR users.

2 RELATED WORKS

2.1 Existing Works on Understanding and Mitigating Harassment Risks in Social VR

In recent years, there has been a steady increase in HCI and CSCW research on how social VR is a novel, embodied online social space for people to engage with positive and supportive social interactions [2, 3, 23, 24, 27, 28, 31, 47, 54, 99, 108–111]. Yet, prior research has also made considerable efforts to investigate how these new spaces lead to more severe and physicalized forms of online harassment [7, 8, 18, 30, 66, 68, 86, 90, 92, 93, 111] and how to mitigate said harassment in these complicated virtual environments [22, 64, 80, 86–88]. Foundational works from Blackwell et al. [7, 8] and Freeman et al. [30] have warned that social VR's focus on embodiment, sense of presence, body tracking, and synchronous voice conversation can enable behaviors akin to physical assault. These behaviors, such as unwanted grabbing and groping, parallel offline harassment experiences (e.g., virtual "rape" in the Metaverse [86, 92, 93]). This, in turn, makes harassment in social VR particularly disruptive and disturbing to victims and bystanders alike [7, 8, 30].

As such, researchers have begun to closely examine existing social VR platforms and community structures for keeping users safe. For example, Sabri et al. [80] investigated how human moderators in social VR use platform-provided moderation tools to assess harassment in real time while facing various social, emotional, and technical challenges. Others have explored more novel approaches to social VR harassment mitigation, including human-AI collaborative efforts for moderation [87], AI moderation bots to keep children safe [22], and consent mechanics for combating interpersonal harm [88]. Unfortunately, these works collectively highlight how existing mitigation efforts are often complicated by a lack of a consensus on what types of behaviors constitute harassment given social VR users' diverse personal and cultural experiences [7, 8, 30]. Despite these conceptual

challenges, a common understanding in the offline world is that harassment is closely related to *unwelcome conduct targeting one or more characteristics of an individual's identity* [14]. These include race/ethnicity, gender, religion, sexuality, national origin, age, disability, and genetic information, amongst others [14].

In this sense, the first step towards understanding harassment targeting identity characteristics (i.e., identity-based harassment) in social VR, though, is to unpack further how social VR users' offline identities are (mis)perceived by other users (i.e., identity revelation). This is because one's online self-presentation acts as an interface for interactions [32], conveying personal values and identity perceptions [25, 26, 63, 107]. Other users then interpret these, influencing their interaction decisions [32]. As a result, the choices an individual makes when engaging with their identity fundamentally affect how they experience online activities and interact with others [9, 19, 85, 107]. This is especially relevant in immersive avatar-based online social environments (e.g., *Second Life* [9]) where the distinction between one's online avatar and offline self is blurred [9, 107]. This distinction becomes even more blurred in social VR due to its unique technical features to foster an enhanced sense of *embodiment*, i.e., how we can experience a virtual body representation as our own body within a virtual environment [91]. The following section further explores how applying this identity revelation lens helps advance our understandings of harassment in social VR.

2.2 The Importance of Exploring Harassment in Social VR Through the Lens of (Mis)Perceived Identity Revelation

2.2.1 Understandings of Identity-Based Harassment in Social VR are Rooted in Assumptions of Identity Revelation Consistent to One's Offline Identity. Collectively, prior works have identified an important assumption about the connection between identity revelation and identity-based harassment in social VR: users become targeted for harassment because they tend to present or disclose themselves in ways that are *consistent with their offline identity characteristics*. First, prior research on social VR has shown that users are often motivated and willing to present their identity characteristics *consistently* in social VR to resemble their offline identities [27, 28, 31]; and that most people feel comfortable disclosing their emotions, personal experiences, and demographic information to others in social VR [54, 97].

Working from these first insights, existing research has secondly highlighted several identity characteristics that are often targeted for identity-based harassment when presented or disclosed in social VR: age, or how old a user is or is perceived to be [3, 7, 8, 18, 30, 54, 66, 68, 97]; gender identity [6–8, 30, 54, 66, 68, 86, 97]; sex assigned at birth, which is often discussed in conjunction with gender identity (e.g., cisgender, transgender, and non-cisgender) [6–8, 28, 54, 66, 68, 86, 97]; sexuality [6–8, 30, 54, 66, 68, 97]; race/ethnicity [6–8, 30, 54, 66, 68, 97]; and disability [110, 111]. Yet, the majority of these works are either qualitative explorations with small samples (e.g., [6–8, 30, 54]) or quantitative explorations that focus on a limited number of identity characteristics (e.g., [66, 68]). Larger quantitative studies exploring both known *and* additional identity characteristics are needed, thus motivating our **RQ1**.

Additionally, a critically understudied consideration is how a harasser's *misperceptions* about someone's offline identity may also affect identity-based harassment in social VR, regardless of a user's actual presentation or disclosure practices. One notable example of this phenomenon is within Schulenberg et al.'s [86] work, where two women participants - one cisgender, one transgender - mentioned being harassed because the harasser misperceived them as being men *despite both participants presenting consistently with their offline identities as women while in social VR* [86]. This type of identity-based harassment is thought to be driven by a misperceived mismatch between one's voice ("too low") and the harassers' preconceived notions of what a person with their gender identity (woman) "should" sound like [86]. Transgender and non-cisgender individuals seem to

face such harassment in social VR more frequently, perhaps as a result of a hegemonic culture that preferences cis-gendered norms [23, 28, 30]. Yet, still little is known about how harassment in social VR arises from misperceptions about one's offline identity or how often this occurs in comparison to harassment associated with presenting or disclosing said identity. Therefore, we are motivated to explore **RQ2** and **RQ3**.

2.2.2 Understandings of Identity-Based Harassment in Social VR Need to Acknowledge How Selective Identity Revelation Is (Mis)Perceived. While the majority of research provides insights on how identity-based harassment is associated with identity revelation practices consistent with one's offline identity, there is a critical need to unpack how social VR users' selective identity revelation is (mis)perceived based on audience types, identity characteristics, and methods of identity revelation.

Grounded in Goffman's metaphor of *theatrical performance* [34], a large body of CSCW and HCI research has highlighted the *selective* nature of identity presentation online, such as how online users may carefully "craft" their identity characteristics [21, 45, 71] based on their "imagined audience" [48, 49]. This is especially evident in avatar-based systems (e.g., online games and virtual worlds) where users can, for example, select or customize avatars to construct new self-identities that may not align with their offline selves [19]. Similarly, limited prior works have shown that social VR users often distinguish between audience types when determining who is safe to interact with as a harassment mitigation strategy [86, 88]. For example, women social VR users may selectively hide their gender identity from unknown others to avoid sexual harassment targeting women users [86]. Yet, few, if any, social VR studies have made explicit attempts to explore how selective identity revelation practices might be comparatively different depending on their audience type. Furthermore, little is known about how such practices might relate to identity-based harassment when considering that even selective practices can be misperceived by others. Therefore, we are motivated to investigate **RQ1a** and **RQ2a**.

Prior works on identity practices in social VR have also collectively described three main methods for offline-consistent identity revelation: customizing avatar appearance to match offline identity characteristics (e.g., gender, race, age, skin tone, facial features, and disability features); using how one's voice sounds when using voice chat; and directly telling other users information about themselves [2, 3, 27, 31, 54, 110]. However, many of these qualitative studies tend to focus on the revelation of one particular characteristic of a person's identity at a time, such as one's gender identity [86] or disability [110, 111], or approach identity revelation generally without considering particular identity characteristics [54]. These approaches, while informative, leave room for quantitative studies like ours to explore how social VR users may use various methods to engage in selective identity revelation based on multiple types of identity characteristics, leading to **RQ1b** and **RQ2b**.

2.2.3 Understandings of Identity-Based Harassment in Social VR Need to Particularly Focus on Harassment Targeting Marginalized Identity Characteristics. Existing research has warned that individuals who present or disclose certain marginalized offline identity characteristics (e.g., women, LGBTQIA+, and racial/ethnic minorities) in social VR through avatar design, voice, or nonverbal cues could face an even greater harassment risk [7, 8, 30, 66, 68, 90]. Therefore, more HCI and CSCW research has begun to prioritize in-depth, focused explorations of how certain marginalized user groups experience and manage harassment in social VR, such as the experiences of women [86], people with disabilities (PWD) [110], and teenagers [18]. This focus on specific user groups serves two purposes.

First, it enables deeper explorations of the direct link between concealing a marginalized offline identity in social VR and harassment specifically targeting that identity characteristic. For example, Zhang et al.'s in-depth study of PWD in social VR found that individuals using avatars with disability

signifiers (i.e., avatar features that indicate the user's disability in the offline world) often experience harassment that specifically targets their disability, such as ableist language and disability mocking [111]. Likewise, physicalized and sexual harassment targeting women social VR users has been well documented, thus prompting women users to employ preventative mitigation strategies such as concealing their gender identity to prevent potential harassment [66, 68, 86, 92, 93]. Unfortunately, such strategies simultaneously serve to protect individual women users and re-enforce women's marginalization within social VR by reducing their visibility and ability to benefit from avatar-self connection [86].

Second, this focus sheds light on how harassment in social VR becomes compounded by intersectional marginalized identities. This is important, as a significant amount of CSCW and HCI research has attended to how individuals with intersectional identities (e.g., black trans women or LGBTQIA+ youth) may face unique and more severe forms of online harassment [12, 36–40, 42, 62]. Deldari et al. especially highlighted how teenagers in social VR may be harassed for both their age *and* their disability (e.g., a stuttering teenager being harassed in voice chat) [18]. This intersectional targeting adds another layer of marginalization for teenagers, as they are already subjected to unfair power dynamics between themselves and adult users [52, 53] that can lead to severe harms such as virtual grooming [18].

Inspired to build upon the aforementioned qualitative studies, which typically focus on only one user group and identity characteristic at a time, our study seeks to understand how identity-based harassment in social VR differs across various user groups and identity characteristics by exploring **RQ3b**. While our study does not specifically compare how social VR users with or without intersectional identities face different frequencies of harassment, we provide valuable insights on how the frequency of harassment per identity characteristic might differ between marginalized and non-marginalized user groups, including by gender identity, sexuality, and race/ethnicity. In doing so, we provide the foundation for future work to further unpack intersectionality and identity-based harassment in social VR by first establishing a difference between marginalization and non-marginalization.

3 METHODS

3.1 Survey Design

To address our research questions, we designed an anonymous online survey on how one reveals their offline identity characteristics to others in social VR, how such identity revelations can be (mis)perceived, and how these (mis)perceptions connect to identity-based harassment in social VR. As a critical step, we first selected the offline identity characteristics to be included.

We acknowledge that the characteristics that make up the entirety of one's identity are complex and numerous, making it challenging for researchers to balance feasibility and thoroughness when conducting identity-focused research [94]. As this study is a foundational quantitative exploration, we utilized a deductive process to select the nine identity characteristics in our survey. This deductive selection process was grounded in three types of sources (see Table 3 in Appendix): (1) prior qualitative works and technical reports on (mis)perceived identity revelation and harassment in social VR as detailed in 2.2; (2) prior works on general online identity practices and identity-based harassment; and (3) our own observations of (mis)perceived identity revelation and harassment in social VR as both researchers and social VR users.

As a result, we include *Age, Disability, Gender Identity, Race, Sex Assigned at Birth*, and *Sexuality* because they have all been documented by prior works as identity characteristics that may lead to harassment online and specifically in social VR [7, 8, 30, 54, 66, 68, 86, 97, 110, 111]. It is important to note that this study treats *Gender Identity* and *Sex Assigned at Birth* as separate characteristics

because factors associated with societal expectations for different sexes (e.g., females have higher-pitched voices than males) can be targeted for harassment regardless of one's *actual* gender identity, e.g., a cis-gendered woman *and* a transgender woman being targeted for having too "low" of voices [86]. Therefore, an empirical exploration of how both *Gender Identity* and *Sex Assigned at Birth* may separately be linked to identity-based harassment provides a more thorough and accurate understanding of each. *Country of Origin* has not been widely explored in social VR research, yet has been identified by prior works on online identity-based harassment as a stigmatized identity characteristic [14, 20]. As our study intends to extend upon the U.S.-centric body of work on harassment in social VR through an international sample, including this characteristic is well-reasoned and beneficial to shaping foundational knowledge in the field.

Likewise, *Sub-community Membership* can and does lead to harassment in the online and offline world [58, 75, 78], yet has received little attention in prior works on social VR. During our own experiences speaking with social VR users as researchers and actual users, we have found that people who identify as members of a social VR sub-community, particularly the Furry community (i.e., individuals presenting online as variations of human-like animals [75]), feel targeted for harassment because of that identity. We thus include *Sub-community Membership* to gain empirical insight on whether this general characteristic is linked to identity-based harassment in social VR regardless of which specific sub-community one belongs to.

Finally, our discussions with other social VR users have highlighted that some users bothered or made fun of other users for the kinds of devices (i.e., equipment) they used to engage in social VR. Social VR users appear to be able to identify what device someone is using based on how someone's avatar does or does not move. Harassers then apparently use that information to make device-specific insults. For example, Quest users are called "bandwagoners." Research on identity practices in immersive avatar-based online systems also acknowledges that a device's technological affordances shape how one presents oneself and interacts with others [107], further justifying the inclusion of *Device Use* in our study.

Positionality Statement. In light of the use of our own experiences with social VR as mentioned above, it is necessary to share the context of our identity positionality and how our identities and cultural backgrounds may have influenced our choices regarding research and identity characteristic foci [4, 83]. All four authors identify as straight and cisgender. This can be a limitation as our study does seek to understand how user groups with marginalized identities (e.g., LGBTQIA+ individuals, women, racial/ethnic minorities) may experience identity-based harassment differently than non-marginalized user groups. However, three out of the four authors identify as women, including two Asian women who are considered ethnic minorities within the context of social VR's hegemonic White, U.S.-centric culture. Additionally, all but one woman author's country of origin is outside of the U.S. Three out of the four authors have extensive experience as users and researchers of social VR. As individuals whose intersectional identity characteristics are differently marginalized within and outside of social VR, our own experiences thus help us be aware of the unique considerations and harassment risks involved in (mis)perceived identity revelation in social VR.

3.2 Survey Structure

Grounded in the identity characteristics identified above, our survey consisted of five primary sections, three pertinent to this study's focus and described below. A full description of the survey design can be found in Appendix A.3. To orient participants to the study's context, participants were provided with a brief description of social VR that combined prior works' definitions (e.g., [7, 8, 30, 31, 86, 87]) at the very beginning of this survey (see A.2 in Appendix for exact wording). Screening questions were then used to remove non-social VR users from further participation.

3.2.1 Offline Identity Demographics & Social VR Use. Participants were then asked to share eight pieces of demographic information related to their offline identities: age, gender identity, sex assigned at birth, sexuality, race/ethnicity, disability, country of origin (i.e., where you were born), and country of residence (i.e., where you live). This information helped orient participants' thoughts towards their offline selves for the rest of the survey and facilitated subsequent data analysis, particularly in exploring identity-based harassment differences among user groups (**RQ3b**). Doing so allowed us to incorporate intersectionality by examining experiences of harassment among marginalized users—especially those already identified as more susceptible to harassment in social VR [7, 8, 30, 66, 68, 90]—in contrast to those of non-marginalized users. Five questions about participants' social VR platform usage and devices were also included.

3.2.2 (Mis)Perceived Identity Revelation & Identity-Based Harassment Subsections. The main body of the survey was designed to answer our research questions in two ways. First, it aimed to explore and compare between **multiple types of offline identity characteristics**. As such, nine subsections were created, as detailed in Section 3.1: 1. Age; 2. Country of Origin; 3. Disability; 4. Device Used; 5. Gender Identity; 6. Sub-community Membership; 7. Race; 8. Sex Assigned at Birth; and 9. Sexuality (see Table 3 in Appendix for characteristic descriptions). Each subsection contained the same types of (mis)perceived identity revelation and harassment measures (see 3.3 Measurements) and followed the same question flow pattern demonstrated in Figure 1. This allowed for comparisons across identity characteristics for all measures to address our research questions.

Second, the survey differentiated two different types of what we are categorizing as **identity revelation**: presenting or disclosing offline identity in social VR (i.e., having one's offline identity correctly perceived by others) and having misperceptions made about one's offline identity in social VR (i.e., having one's offline identity be incorrectly perceived by others). As such, each subsection would begin with asking participants if they had ever disclosed their **offline [identity characteristic]** in social VR, e.g., *"Have I ever disclosed my **offline gender identity** in social VR?"* for the *Gender Identity* subsection (leftmost blue square in Figure 1). Based upon their response, participants were then either directed to questions on their disclosure practices and related harassment for that characteristic (detailed in 3.3 Measurements) or moved straight to questions on "incorrect assumptions," i.e. misperceptions (see Figure 1).

Each subsection ended after participants completed the "incorrect assumptions" branch (see Figure 1). The order in which the nine subsections were presented to each participant was randomized evenly. The survey related to this study ended after all nine identity characteristic subsections were completed.

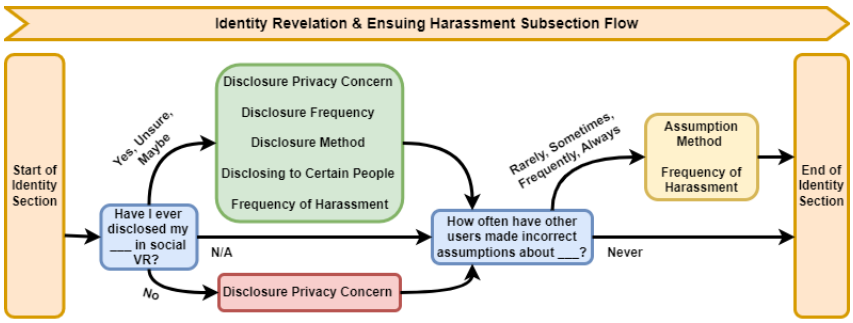

Fig. 1. Overall flow of each subsection regarding (mis)perceived identity revelation & ensuing harassment in the survey

3.3 Measurements

Our survey contained four types of measurements. All measures were developed specifically for this study and were informed by findings from prior qualitative research and technical reports concerning identity-based harassment in social VR [7, 8, 28, 30, 66, 67].

3.3.1 Frequency Measures. For this study, we developed a single-item measure that utilizes a 5-point Likert frequency scale (1 = Never - 0% of times..., 5 = Always - 100% of times...) to assess how frequently participants experience specific scenarios in social VR. This type of frequency scale was applied across nine survey subsections to evaluate four aspects: *Frequency of Identity Revelation* for (1) presenting or disclosing and (2) misperceptions; *Frequency of Harassment* for (3) presenting or disclosing and (4) misperceptions (see Figure 1). Explanations and operationalization of this frequency scale can be found in Appendix A.5.

Frequency of Identity Revelation. For each subsection, we used the above-mentioned single-item 5-point Likert frequency scale to ask participants how often they disclose (present) that identity characteristic to others in social VR (see Table 3 in Appendix). We used the same scale type to ask how often participants experience other users making incorrect assumptions (misperceptions) about the same identity characteristic.

Frequency of Harassment. We also used the above-mentioned single-item 5-point Likert frequency scale to ask participants how often other social VR users have harassed them because of their correctly perceived identity characteristics vs. misperceptions about their identity characteristics.

3.3.2 Methods of Identity Revelation. To understand the methods through which each identity characteristic is presented or disclosed and (mis)perceived in social VR, we used two sets of six single-item 5-point Likert agreement scales (1 = Strongly Disagree, 5 = Strongly Agree). These scales asked participants to indicate their level of agreement with using each method to (1) present or disclose that identity characteristic (*"I often disclose [identity characteristic] in social VR by ..."*) and for (2) having that identity characteristic misperceived (*"If I had to guess, I think other social VR users made incorrect assumptions about [identity characteristic] by..."*). Statements per method type and identity revelation type can be seen in Table 4 in Appendix A.4.

For this measure, we identified six methods based on a combination of prior qualitative works on self-presentation and harassment in social VR, general online identity presentation studies, our experiences as researchers and users of social VR, and documented social VR features. These methods include: 1. ***Avatar Looks***, which is the appearance of one's avatar; 2. ***Sound of Voice***, which is how one's voice sounds when using voice chat [7, 8, 30, 66, 68, 86, 90]; 3. ***Texting Style***, which is how one expresses themselves when text chatting other users (inspired by works on online harassment such as [61]); 4. ***User Bio***, which is a user's in-world profile visible to other users; 5. ***Username***, which is one's in-world username (inspired by works on identity presentation in social media such as [17]); and 6. ***Telling Others***, which is directly telling others one's offline identity or telling something to others that leads to misperceptions about one's offline identity [54, 97]. See Table 4 in Appendix A.4 for survey wording.

3.3.3 Audiences for Presenting or Disclosing Identity Characteristics. To understand to whom participants present or disclose their offline identity characteristics in social VR, we used two sets of five single-item 5-point Likert agreement scales (1 = Strongly Disagree, 5 = Strongly Agree). These scales assessed (1) **level of comfort** and (2) **frequency of disclosing (presenting)** that identity characteristic to various audience types. Participants were instructed to consider only their social VR-exclusive relationships (i.e., relationships not formed offline): *"When answering, keep in mind relationships you have that exist only in social VR (e.g., close friends you have in social VR but do not know offline)."* This meant that results from this measure could be specifically related back to one's

identity revelation practices in social VR, rather than those that take place outside of social VR or as a result of a pre-existing offline world relationship (e.g., family members).

For this measure, we identified five audience types based on a combination of prior qualitative works on social VR concerning harassment and mitigation [7, 8, 30, 86–88], self-presentation and self-disclosure [54, 97], social support [47], and collaborative work [24]. These include: 1. ***Strangers***; 2. ***Acquaintances or Casual Friends***; 3. ***Close Friends***; 4. ***Sexual or Romantic Partners***, i.e., a sexually or romantically intimate relationship with another user only in social VR, such as within Erotic Role Play (ERP) communities as described in [18, 86, 97]; and **5.*Work Colleagues,*** such as people who work with each other on projects or within VR collaborative workplace settings without any contact outside of social VR as described in [24, 88]. To account for the possibility of a participant not having one of these relationships in social VR, an *"N/A (Not Applicable)"* option was also offered for each item.

3.4 Recruitment and Data Collection

The survey was developed using Qualtrics, an industry-standard professional survey platform [59], and was distributed via an integrated link in Prolific in April 2023. Prolific is an international professional online data collection website commonly used by researchers for conducting anonymous online surveys and is known for providing reliable and high-quality data [70, 73]. We deliberately decided to recruit an international sample because prior exploratory work on identity-based harassment in social VR has often focused more on U.S.-centric samples despite social VR's international use [7, 8, 30, 66, 67]. Before distribution, the survey was pilot tested and reviewed by the authors. A pilot test was conducted with 20 Prolific users with social VR experience before full survey deployment. These participants were compensated $6 for completing the survey. Respondents were required to be proficient in English and at least 18 years of age at the time of the survey. Participants were also required to take a short pre-survey to identify if they were social VR users or had experiences in social VR. They were not eligible to take the subsequent main survey if they had not.

3.5 Sample

In total, 225 respondents were approved and compensated in Prolific. Of those 225, 2 participants were dropped from data analysis post-approval due to failed attention checks, leaving 223 final participants for this study. On average, respondents spent 28 minutes completing the survey. Respondents who completed the survey and passed the attention check criteria were compensated with USD $6.00 through Prolific. Table 1 summarizes the demographic information of the 223 valid responses for this study, which represents a geographically and culturally diverse sample with varied genders, sexuality, age groups, ethnicity, and locations.

Our 223 participants also varied in their social VR use. The majority (47.54%) of our participants had been using social VR for less than six months at the time of survey completion, with 27.8% having used it for six months - 1 year, 19.28% for 1 - 3 years, 4.04% for 3 - 6 years, and 1.35% for more than six years. Additionally, 71.3% of participants used social VR less than 5 hours per week on average at the time of survey completion, while 19.38% used it 6 - 9 hours per week, 6.28% used it for 10 - 19 hours per week, 2.69% used it for 20 - 29 hours per week, and 0.45% used it for 30 - 39 hours per week. In terms of device usage, 40.36% of participants at one point in time have used a desktop or PC without added equipment (e.g., SteamVR Suite) to engage in social VR, followed by the use of Oculus Quest 2 (39.1%), Oculus Quest 1 (30.49%), mobile device without added equipment (21.52%), HTC Vive (17.49%), Windows Mixed Reality headset (17.04%), Steam VR Tracking Base Stations (10.31%), and Valve Index (7.62%).

Variables	Total (N=223)	Percentage (%)
Gender		
Cis Man	146	65.47
Cis Woman	69	30.94
Trans Woman	1	0.45
Non-Binary/Third Gender/Agender	6	2.69
Not Provided	1	0.45
Age		
18 - 24	116	52.02
25 - 34	78	34.98
35 - 44	17	7.62
45 - 54	6	2.69
55 - 64	4	1.79
65 - 74	2	0.90
Sexuality		
Asexual	3	1.35
Bisexual	10	4.48
Gay	7	3.14
Lesbian	8	3.59
Pansexual	11	4.93
Straight	178	79.82
Not Provided	6	2.69
Disability		
Yes, identified as having a disability	16	7.17
No, not identified as having a disability	205	91.93
Not Provided	2	0.90

Variables	Total (N=223)	Percentage (%)
Race/Ethnicity		
Asian	6	2.69
Black/African American or African	50	22.42
Hispanic or Latino/a	28	12.56
White	129	57.85
Mixed	7	3.14
Not Provided	3	1.34
Country of Origin by Region		
Africa (except South Africa)	3	1.34
Asia	2	0.90
Australia	2	0.90
Central America (e.g., Mexico)	36	16.14
Europe	117	52.47
North America	11	4.93
South Africa	44	19.73
South America	8	3.59
Country of Residence by Region		
Australia	2	0.90
Central America (e.g., Mexico)	37	16.59
Europe	121	54.26
North America	12	5.38
South Africa	43	19.28
South America	8	3.59

Table 1. Demographic information of 223 valid respondents

Regarding social VR platforms, 84.3% of participants have used VRChat at one point in time, and 62.78% of participants (N = 140) named it their favorite platform. 47.09% of participants used Roblox with a VR headset and 17.49% (N = 39) identified it as their favorite platform. Rec Room was the third most used (19.28%) and the sixth most cited favorite platform (3.14%, N = 7). Other social VR platforms that participants have used are: Bigscreen (14.35%), Horizon Worlds (13.45%), Mozilla Hubs (12.56%), AltspaceVR (9.87%), Spatial (5.83%), and Neos VR (4.48%).

4 RESULTS

To address our RQs, we report results of the four measurements used for data analysis in this study: (1) *frequency of (mis)perceived identity revelation in social VR*; (2) *methods of (mis)perceived identity revelation in social VR*; (3) *audiences for presenting or disclosing identity characteristics in social VR*; and (4) *how (mis)perceived identity revelation is associated with the frequency of harassment in social VR*. Our survey design utilized a form of repeated measures that produced missing data (i.e., not all participants indicated presenting or disclosing or encountering misperceptions about all identity characteristic types). Therefore, our analysis leveraged linear mixed effects (LME) modeling because it is particularly appropriate for repeated measures on the same participant and retains robustness with missing data [1]. All LMEs conducted for this study first began by running a model for the dependent variable (e.g., frequency of harassment) with a random intercept to account for repeated measures. The identity characteristics are an independent variable in each of our statistical models (see Table 3 in Appendix for more details about each characteristic).

4.1 Frequency of (Mis)Perceived Identity Revelation in Social VR

To understand how frequently our participants present or disclose (**RQ1a**) and encounter misperceiptions about their identity characteristics in social VR (**RQ2a**), we added the identity revelation type, identity characteristics, and their interaction to our frequency of identity revelation model.

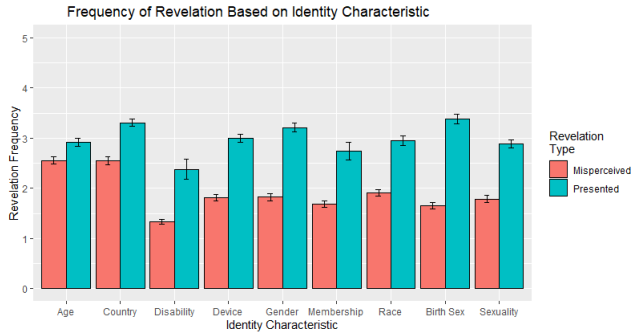

Fig. 2. Frequency of identity revelation based on (presented or disclosed vs. misperceived) and identity characteristic type. Error bars denote 95% confidence interval.

Both the main effects of identity revelation type ($p < .0001$) and identity characteristic type ($p < .0001$) and their interaction effect ($p < .0001$) significantly improved the linear model for frequency of identity revelation (see Table 5 in Appendix). Our results show that the frequency of participants' identity revelation in social VR is **significantly** impacted by (1) whether their identity is being presented or disclosed vs. misperceived, and (2) the type of identity characteristic, and (3) that the relative frequency of presentation or disclosure vs. misperception differs **significantly** per identity characteristic (see Table 5).

Addressing RQ1a, our results demonstrate that participants present or disclose some of their offline identity characteristics, like their *Sex Assigned at Birth* (*M = 3.38, SD = 1.16*) and *Country of Origin*) (*M = 3.3, SD = 0.96*), significantly more than other characteristics, such as whether they have an offline *Disability* (*M = 2.38, SD = 0.81*) (most noticeably within Figure 2). Furthermore, **addressing RQ2a**, participants overall rarely encounter other users misperceiving most of their identity characteristics in social VR. The only exceptions are *Age* (*M = 2.55, SD = 1.11*) and *Country of Origin* (*M = 2.55, SD = 1.15*), which appear to hover between "Rarely - About 30% of times [I've been in social VR]" (y-axis = 2.0) and "Sometimes - About 50% of times" (y-axis = 3.0) (see Figure 2).

Additionally, participants, on average, present or disclose their own identity characteristics significantly more often than they encounter others' misperceptions about their identity characteristics. This significant difference holds true—but is different in size—across all identity characteristics. Specifically, Figure 2 shows that this difference is particularly pronounced for *Gender Identity* and *Sex Assigned at Birth*, and is noticeably large for *Disability, Device, Sexuality*, and *Race*. Table 6 in the Appendix also provides a comprehensive list of frequency of (mis)perceived identity revelation in social VR for each identity characteristic.

4.2 Methods of Identity Revelation in Social VR

Next, we focus on which methods our participants frequently use to present or disclose specific identity characteristics (**RQ1b**) and ways through which others often misperceive their identity characteristics (**RQ2b**) in social VR (see Table 4 in Appendix for details on method types). To address these RQs, we added the method type, identity characteristics, and identity revelation type to our frequency of identity revelation model. All three significantly improved the model: Method, $p < .0001$; Identity Characteristics, $p < .0001$; and Identity Revelation type, $p < .0001$. We then added three two-way interactions between the three variables, all of which significantly improved the model, $p < .0001$. Finally, we added one three-way interaction between the three variables, which also significantly improved the model, $p < .0001$ (see Table 7 in Appendix). Our results show that

the frequency of of identity revelation is **significantly** impacted by (a) the method they use to present or disclose or the method other users use to misperceive their identity characteristics; (b) the specific identity characteristic; (c) the type of identity revelation; and (d) the combination of method, identity revelation type, and identity characteristic (see Table 7).

Addressing RQ1b, our results show that participants overall present or disclose their identity characteristics most frequently by *directly telling other social VR users* about their identity (*M = 3.69, SD = 1.15*). This is especially evident for offline identity characteristics that might be less easily translated into a virtual space (e.g., *Disability* status) or are less noticeable through other methods (e.g., what kind of *Device* is being used). Additionally, Figure 3 (see Table 8 in Appendix for more details) shows that people present or disclose their *Gender Identity* and their *Sex Assigned at Birth* most frequently through the way their *voice sounds when they use voice chat.* This result could potentially reflect the predominance of voice use in social VR. It may also be because voice pitch can be a more readily perceivable identity cues for cisgender individuals, which make up the majority of participants. Alternatively, participants reported presenting or disclosing least frequently through information contained in their *username* for all identity characteristics (*M = 2.05, SD = 1.27*), possibly because there are many ways to create a username with little to no connection to one's offline identity, e.g., randomly generated usernames.

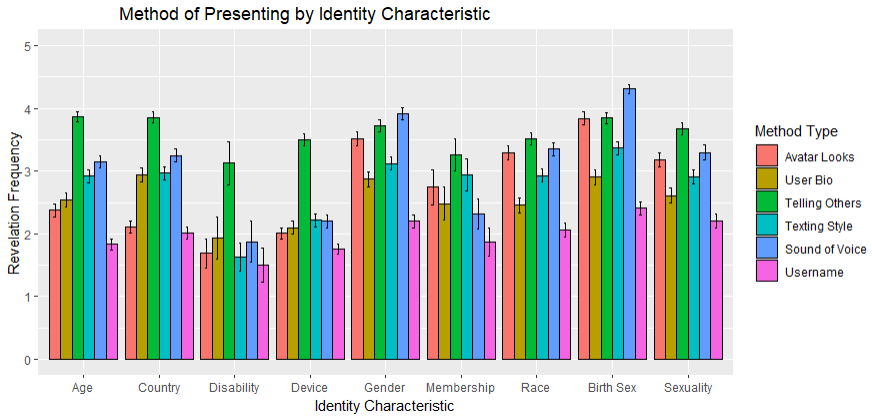

Fig. 3. Frequency of identity characteristics presented by methods of presenting or disclosing. Error bars denote 95% confidence interval.

Addressing RQ2b, Figure 4 (see Table 9 in Appendix for additional details) highlights that the methods used most to misperceive specific identity characteristics are varied. However, participants do believe that other users misperceive their identity most often overall because of the way that their *voice sounds when they use voice chat* (*M = 2.94, SD = 1.27*).

Specifically, four highlights emerge in our results. First, misperceptions of participants' *Gender Identity*, *Sex Assigned at Birth*, and *Sub-Community Membership* happen most often due to the way their *avatar looks*. Second, misperceptions of their *Sexuality* happen most often due to the way they *type or express themselves over text chat.* Third, misperceptions of their *Age*, *Country of Origin*, and *Race/Ethnicity* happen most often due to how their *voice sounds when they use voice chat.* Finally, misperceptions of their *Disability* status and what *Device* they use happen most often due to *unrelated information one shared with others* (see Table 9 for details). Alternatively, participants believe that *information in their user bios* is the overall least used method for having misperceptions about their identity (*M = 2.32, SD = 1.27*). This could possibly result from various reasons, such as

not having a bio, having little to no information in a bio, or having identity-consistent information in one's bio that would be difficult to misperceive.

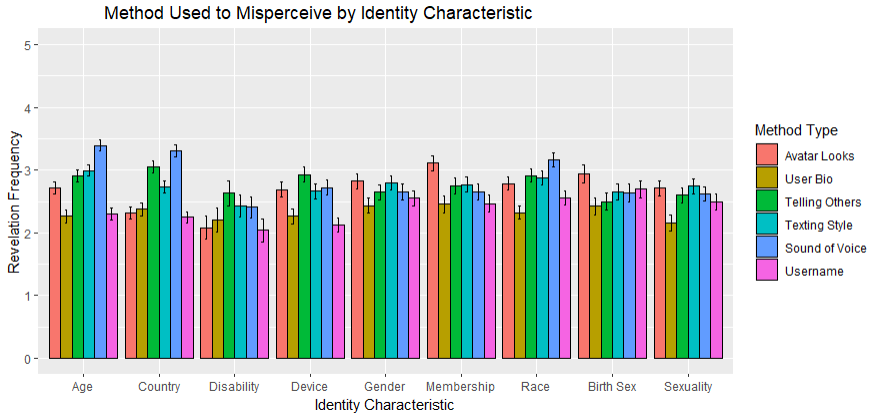

Fig. 4. Frequency of identity characteristics misperceived by others by methods of incorrectly assuming or misperceiving . Error bars denote 95% confidence interval.

4.3 Audiences for Presenting or Disclosing Identity Characteristics in Social VR

RQ1a also explores whom people choose to present or disclose their specific identity characteristics to in social VR. This exploration included two dimensions: comfort and frequency. We first ran a model for ***comfort*** with a random intercept. Then we added audience type ($p < .0001$), identity characteristics ($p < .0001$), and their interaction ($p < .001$) to the model, each of which significantly improved the linear model (see Table 10 in Appendix). We ran a separate model for ***frequency*** with a random intercept, adding audience type ($p < .0001$), identity characteristics ($p < .0001$), and their interaction ($p < .0001$) to the model, each of which significantly improved the linear model (see Table 10 and Table 11). Our results show that users' comfort with and frequency of presenting or disclosing their identities is **significantly** impacted by (1) who they are interacting with, and (2) the type of identity characteristic, and that (3) the relative comfort/frequency per audience type differs **significantly** per identity characteristic.

For ***comfort***, participants feel most comfortable overall presenting or disclosing identity characteristics to *close friends* ($M = 4.44, SD = 0.92$). In contrast, they feel the least comfortable overall presenting or disclosing to *strangers* ($M = 3.5, SD = 1.42$) (see Figure 5 and Table 12 in Appendix). Furthermore, Table 12 reveals that participants feel relatively comfortable with most audience types, except in the case of *strangers* for *Disability* and *work colleagues* for *Sub-community Membership*. On the latter, *work colleagues* being the least comfortable to present to for *Sub-community Membership* ($M = 2.75, SD = 1.29$) could reflect how such information can be sensitive and stigmatized even in casual or close relationship settings, much less professional environments. Additionally, *sexual or romantic partners* appears to be similarly comfortable as *close friends* for *Age*, *Country of Origin*, *Gender Identity*, *Race/Ethnicity*, *Sex Assigned at Birth*, and *Sexuality* (see Figure 5).

For ***frequency***, participants overall present or disclose their identities most frequently to *sexual or romantic partners* ($M = 4.04, SD = 1.29$). Similarly to comfort levels, *strangers* ($M = 4.04, SD = 1.29$) are least frequently presented or disclosed to (see Figure 6 and Table 13 in Appendix), with the exception of *work colleagues* for *Sub-community Membership* ($M = 2.33, SD = 1.34$).

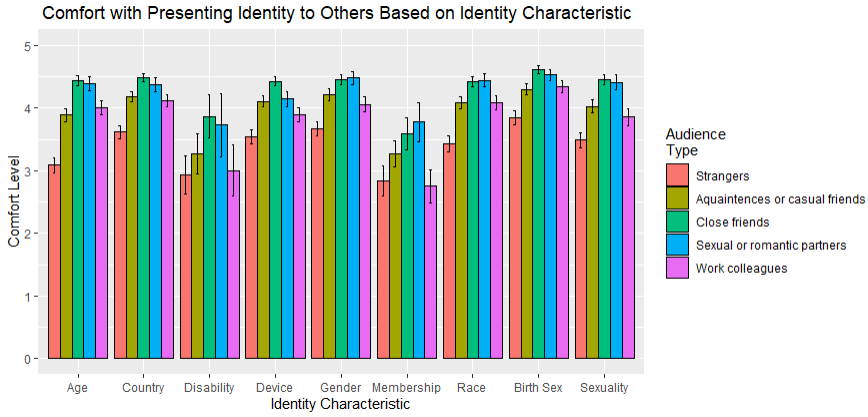

Fig. 5. Participants' comfort level with presenting or disclosing various identity characteristics to different audience types. Error bars denote 95% confidence interval.

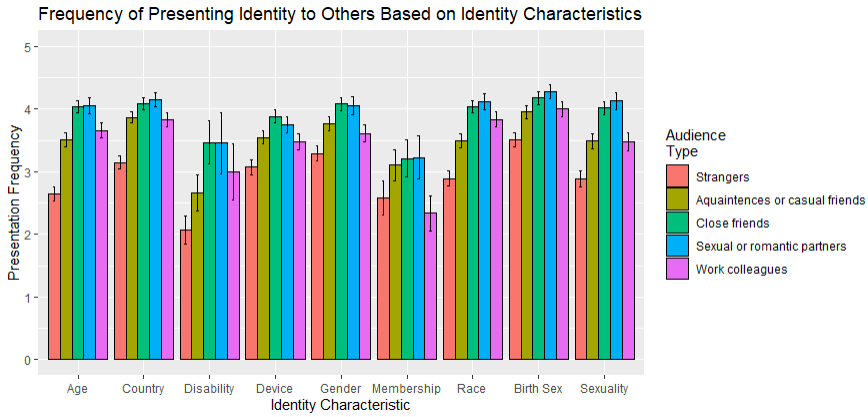

Fig. 6. Frequency of presenting or disclosing to each audience type based on identity characteristic. Error bars denote 95% confidence interval.

4.4 (Mis)Perceived Identity Revelation and Frequency of Harassment in Social VR

Built upon RQ1 and 2, RQ3 explores how (mis)perceived identity revelation is associated with the frequency of harassment in social VR (**RQ3a**), especially when considering various user groups (**RQ3b**). Separate models were run for the overall frequency of harassment and frequency of harassment by user group, i.e., based on a participant's offline ***gender identity***, ***sexuality***, and ***race/ethnicity***. User groups based on ***age***, ***country of origin***, and ***disability*** were also examined to compare the frequency of harassment but are not detailed further in this paper as LME results came back insignificant for these user groups' main and interaction effects.

4.4.1 Overall Association between (Mis)perceived Identity Revelation and Frequency of Harassment in Social VR. To address **RQ3a**, we added identity revelation type, identity characteristics, and their interaction to our frequency of harassment model. Both the identity revelation type ($p = .0001$) and identity characteristics ($p < .0001$) significantly improved the linear model created for the frequency

of harassment, although the interaction between the two did not (p = 0.63). Our results show three overall patterns. First, there is a **significant** difference between how frequently social VR users are harassed for their presented or disclosed identities versus their misperceived identities (p < .0001). Second, there is a **significant** difference between how frequently social VR users are harassed for specific identity characteristics (p < .0001). Third, there is **no significant** interaction between identity characteristics and identity revelation type (p = .63) (see Table 14 in Appendix). Our results also show that social VR users, regardless of identity characteristic or identity revelation type, appear to experience harassment with relatively **low** frequency, with averages ranging between "Never - 0% of times I've been in social VR" (y-axis = 1.0), "Rarely - About 30% of times" (y-axis = 2.0), and "Sometimes - About 50% of times" (y-axis = 3.0) as shown in Figure 7.

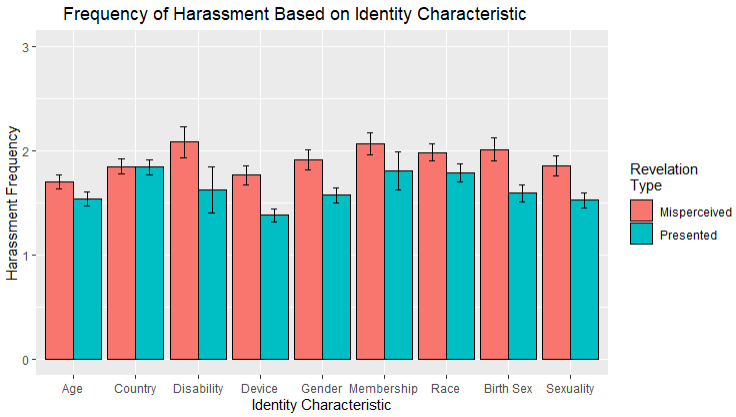

Fig. 7. Frequency of harassment based on identity revelation type and identity characteristics. Error bars denote 95% confidence interval. The axis was adjusted to 3.0 to increase visibility.

Therefore, one important highlight in our results is that social VR users seem to be harassed **significantly more frequently** overall for others' misperceptions about their offline identity characteristics (*M = 1.88, SD = 0.97*) than for the identity characteristics that they intentionally present or disclose to others (*M = 1.61, SD = 0.92*). However, the nature of this misperception does not necessarily mean that social VR users *do not* present or disclose their identities consistently when these misperceptions occur. Rather, there might be multiple plausible situations. For example, a social VR user may be presenting or disclosing consistently yet is harassed due to misperceptions of their identity characteristics (e.g., a cis or trans woman with a female-looking avatar who is harassed for being a man). Or, a user may not be presenting or disclosing consistently and is harassed for an identity characteristic inconsistent with their offline identity but consistent with their online presentation (e.g., a man with a female-looking avatar who is harassed for being a woman). Additionally, a user may not be presenting or disclosing consistently and is harassed for identity characteristics that are inconsistent with both their offline and online identity (e.g., a woman with a genderless avatar who is harassed for being a man).

Furthermore, regardless of whether an identity characteristic was presented or misperceived, participants reported being harassed most frequently for *Sub-community Membership*, followed by *Disability* and *Race/Ethnicity* (see Table 15 in Appendix). In other words, social VR users seem to be harassed most frequently for being a member of a sub-community (e.g., the Furry community), regardless of whether they actually belong to such a sub-community.

4.4.2 Association between (Mis)Perceived Identity Revelation and Frequency of Harassment in Social VR Across Different User Groups. **RQ3b** explores which user groups may encounter harassment more frequently than others in social VR. Given this, we used LME analysis to investigate the frequency of harassment in social VR across six user groups representing participant offline identity. However, only the following three show significant results: (1) ***gender identity***, (2) ***sexuality***, and (3) ***race/ethnicity***. In contrast, *age*, *country of origin*, and *disability* were analyzed but generated insignificant results for main and interaction effects for each, which are thus not discussed further in this section.

Additionally, participants who did not specify their gender (N = 1), sexuality (N = 6), or race/ethnicity (N = 3) (see Table 1) were labeled as *"Not Provided"* and excluded from analysis to prevent any skewing of results. Preliminary LME analyses by user group showed insignificant interaction effects between user groups and identity revelation type, as well as for the three-way interaction between user groups, identity revelation type, and identity characteristics. As such, analysis in this section utilized **combined frequency of harassment** resulting from both presented or disclosed and misperceived identity characteristics.

(1) Frequency of identity-based harassment in social VR across participant offline *gender identity*.

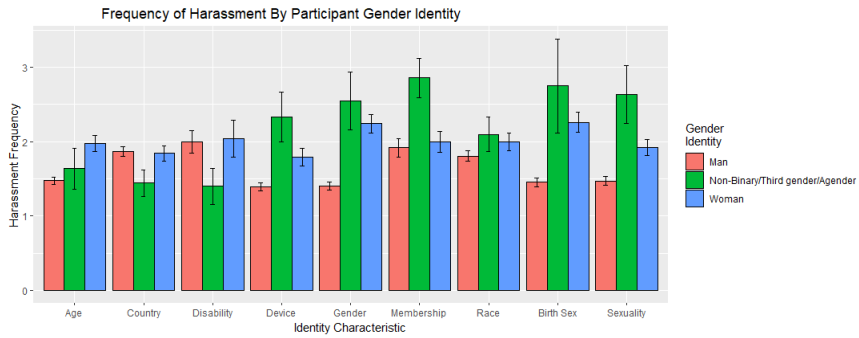

Fig. 8. Frequency of harassment based on identity characteristic type and participant gender identity. Error bars denote 95% confidence interval.

To examine the impact of a social VR user's gender identity on the frequency of identity-based harassment, we categorized gender into women (includes cisgender and transgender women participants), men (includes cisgender men, no transgender men in sample), and non-binary/third gender/agender. Using these categories, we added participant offline gender identity (i.e., women, men, and non-binary/...), identity characteristics, and their interaction to our frequency of harassment model. Gender identity ($p = .0001$), identity characteristics ($p < .0001$), and their interaction ($p < .0001$) all significantly improved the linear model created for frequency of harassment (see Table 16 in Appendix). Our results show that there is a **significant** difference in reported frequency of harassment overall between genders. There is also a **significant** difference in how frequently people of different genders are harassed for specific identity characteristics. Table 17 in the Appendix provides the average frequency of identity-based harassment for each identity characteristic by participants' self-reported offline gender identity.

Post-hoc tests on the model shown in Table 17 was also performed, which uncovered four significant patterns regarding the frequency of identity-based harassment in social VR across gender groups. Figure 8 provides a visualization of these significant results.

First, compared to men, women reported that they were **significantly** more harassed regardless of identity characteristic type in social VR (β = -0.33, SE = 0.1, t(-3.4), p = .002). Compared to men, non-binary/third gender/agender individuals reported that they were **almost significantly** more harassed in social VR regardless of identity characteristic type, though significance was not fully met (β = -0.58, SE = 0.27, t(-2.3), p = .068). **No significant difference** was found between women and non-binary/third gender/agender individuals concerning the overall reported frequency of harassment in social VR (p = .372).

Second, compared to men, women reported that they were **significantly** more harassed for the *Age* (β = -0.45, SE = 0.12, t(-3.74), p = .001), *Device Used* (β = -0.36, SE = 0.13, t(-2.9), p = .012), *Gender Identity* (β = -0.82, SE = 0.13, t(-6.52), p < 0.0001), *Sex Assigned at Birth* (β = -0.8, SE = 0.13, t(-6.4), p < .0001), and *Sexuality* (β = -0.36, SE = 0.13, t(-2.84), p = .009).

Third, compared to men, nonbinary/third gender/agender individuals reported that they were **significantly** more harassed for *Gender Identity* (β = -1.31, SE = 0.32, t(-4.03), p = .0001), *Sub-community Membership* (β = -1.03, SE = 0.36, t(-2.85), p = .014), *Sex Assigned at Birth* (β = -1.3, SE = 0.45, t(-3.12), p = .004), and *Sexuality* (β = -1.14, SE = 0.33, t(-3.5), p = .002).

Finally, compared to women, non-binary/third gender/agender individuals reported that they were **significantly** more harassed for *Sub-community Membership* (β = -1.02, SE = 0.37, t(-2.73), p = .014) and *Sexuality* (β = -0.77, SE = 0.37, t(-2.32), p = .021).

(2) Frequency of identity-based harassment in social VR across participant offline *sexuality*.

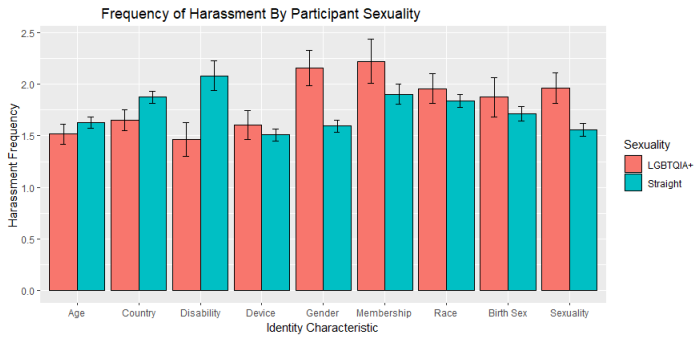

Fig. 9. Frequency of harassment based on identity characteristic type and participant sexuality. Error bars denote 95% confidence interval.

We also aimed to understand how a social VR user's offline sexuality may affect the frequency of identity-based harassment. However, we must acknowledge that individuals identifying as straight (N = 178) comprise most of our sample (almost 80%, see Table 1). While this creates statistical challenges, we believe there is still value in exploring the frequency of reported harassment for each sexuality type, keeping in mind that any differences seen in frequency are based on small sample numbers. Table 18 in the Appendix demonstrates the frequency of harassment by the unstratified version of participant offline sexuality (i.e., asexual, bisexual, gay, lesbian, pansexual, and straight) to explicate the frequency of harassment by various sexuality user groups.

In addition to these descriptive statistics, we re-stratified our participants into two broader categories of participant offline sexuality: straight (i.e., participants who identify as straight) and LGBTQIA+ (i.e., participants who identify as asexual, bisexual, gay, lesbian, or pansexual). Prior work on harassment online overall and in social VR specifically has often acknowledged that differences in harassment occurrences and severity are most stark when comparing LGBTQIA+

individuals with straight individuals (see Table 3 for sources). We thus believe that further analysis based on these two broader categories is both statistically sound to balance out the disparity in our sample and ethically sound, given prior works' observations.

Therefore, we added sexuality (i.e., straight and LGBTQIA+), identity characteristics, and their interaction into our frequency of harassment model. Our results show an **insignificant** main effect of sexuality (p = .217), but identity characteristics (p < .0001) and its interaction with sexuality (p < .0001) did **significantly** improve the model (Table 19 in Appendix). Essentially, while harassment frequency did not differ overall between straight and LGBTQIA+ participants, **significant** differences existed for specific identity characteristics (see Table 19). More specifically, post hoc tests show that, compared to straight individuals, LGBTQIA+ individuals were **significantly** more harassed for *Gender Identity* (β = -0.54, SE = 0.15, t(-3.57), p = .0004), *Sub-community Membership* (β = -0.44, SE = 0.18, t(-2.36), p = .018), and *Sexuality* (β = -0.4, SE = 0.15, t(-2.6), p = .0096). Figure 9 provides a visualization of these significant results.

(3) Frequency of identity-based harassment in social VR across participant offline *race/ethnicity*.

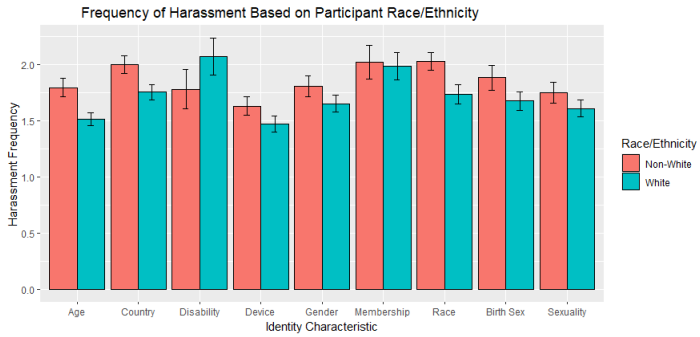

Fig. 10. Frequency of harassment based on identity characteristic type and participant race/ethnicity stratified into White and Non-White. Error bars denote 95% confidence interval. The axis was adjusted to 3.0 to increase visibility.

Lastly, we focused on unpacking how a social VR user's offline race/ethnicity may affect the frequency of identity-based harassment. Similarly to participant sexuality, we acknowledge that individuals identifying as White (N = 129) make up the majority of our sample at almost 58% (see Table 1). Additionally, some racial/ethnic identities, including Asian and Mixed, comprised less than 5% of our sample. Again, while imbalance can create statistical challenges, we provide descriptive statistics (see Table 20 in Appendix) for the unstratified user group for race/ethnicity (i.e., Asian, Black/African American or African, Hispanic or Latino/a, White, and Mixed) to highlight the frequency of reported harassment across various race/ethnicity user groups.

Indeed, prior work on harassment online and in social VR has revealed the differences in the occurrence and severity of harassment targeting non-White individuals compared to the White majority (see Table 3 in Appendix for sources). Therefore, similar to how we analyzed harassment by sexuality user groups, we also re-stratified our participants into two categories of participant offline race/ethnicity – White (i.e., participants who identify as White) and non-White (i.e., Asian, Black/African American or African, Hispanic or Latino/a, and Mixed).

We then added race/ethnicity (i.e., White and non-White), identity characteristics, and their interaction to our model for frequency of harassment. Race/ethnicity (p = .022), identity characteristics (p < .0001), and their interaction (p = .016) all **significantly** improved the model (see Table

21 in Appendix). Our results show that there is a **significant** difference in reported frequency of harassment overall between White and non-White individuals. There is also a **significant** difference in how frequently White and non-White individuals report being harassed for specific identity characteristics (see Table 21).

Further, post-hoc tests on the model shown in Table 21 were performed, which revealed three significant patterns. Figure 10 provides a visualization of these significant results. First, compared to White individuals, non-White individuals were **significantly** more harassed regardless of identity characteristic type (β = -0.21, SE = 0.09, t(-2.3), p = .022). Second, non-White individuals were **significantly** more harassed for *Age* (β = -0.28, SE = 0.12, t(-2.39), p = .017), *Device Used* (β = -0.27, SE = 0.12, t(-2.26), p = .024), and *Race/Ethnicity* (β = -0.36, SE = 0.12, t(-2.96), p = .003). Additionally, compared to White individuals, non-White individuals were **almost significantly** more harassed for *Country of Origin* (β = -0.21, SE = 0.11, t(-1.81), p = .071) and *Sex Assigned at Birth* (β = -0.23, SE = 0.12, t(-1.88), p = .06), though significance was not fully met for these factors. Finally, compared to non-White individuals, White individuals were **significantly** more harassed for *Disability* (β = -0.41, SE = 0.2, t(-2.05), p = .041).

5 DISCUSSION

In answering our research questions, Table 2 summarizes our key significant findings. Grounded in these results, in this section we present new understandings of identity-based harassment in social VR and focus on how users of different genders, sexualities, and races/ethnicities experience harassment frequency significantly differently (**5.1**). We then discuss promising directions for researching and designing for safer social VR spaces to mitigate identity-based harassment in the future (**5.2**).

5.1 New Understandings of Identity-Based Harassment in Social VR Through the Lens of (Mis)Perceived Identity Revelation

Using *(mis)perceived identity revelation* as a lens, our work sheds light on several previously under- and unexplored aspects of identity-based harassment in social VR. First, by breaking down how and to whom one's specific offline identity characteristics are (mis)perceived in social VR (**RQ1**), we offer new insights into the interplay of selective identity revelation and identity-based harassment in social VR (**5.1.1**). Second, by unpacking the unique impacts of misperceptions about one's identity characteristics on identity-based harassment (**RQ2 and 3a**), we reveal how embodying one's offline identity in social VR may be less risky than once thought (**5.1.2**). Finally, by explicating the frequency of harassment per identity characteristic by user group (**RQ3b**), our study provides critically needed quantitative evidence to validate a key assumption about identity-based harassment in social VR – who you are or appear to be ***does*** matter when it comes to identity-based harassment in social VR (**5.1.3**).

5.1.1 New Insights on the Interplay of Selective Identity Revelation and Identity-Based Harassment in Social VR. Existing works on harassment in social VR often focus on how identity revelation that is consistent with offline identity affects identity-based harassment in social VR, with avatar design as the primary means for targeting others for such harassment [7, 8, 30]. Our work expands beyond these understandings by explicating both *selective* identity revelation (i.e., by audience type, identity characteristic, and method of revealing) and *misperceptions* along with such practices. In doing so, our results highlight that social VR users report more frequently engaging in more *selective* identity revelation rather than always *consistent*, especially in comparison to results found in prior qualitative works on social VR harassment [27, 54, 97]. This result is particularly important as it demonstrates social VR users' self-agency to control what and how they reveal their identity

RQ1: How do people present or disclose their various identity characteristics in social VR?	**Key Significant Findings**
RQ1a: How frequently do people present or disclose their various identity characteristics in social VR?	• Social VR users present or disclose their ***Sex Assigned at Birth*** most frequently and present or disclose their ***Disability*** information least frequently.
... and to whom?	• Social VR users overall feel most **comfortable** presenting or disclosing identity characteristics to *close friends* and least comfortable presenting or disclosing to *strangers*. • Social VR users overall present or disclose their identity characteristics most **frequently** to *sexual or romantic partners* and least frequently to *strangers*.
RQ1b: Which methods do they use to present or disclose their various identity characteristics in social VR?	• The most common method for presenting or disclosing *Age*, *Country of Origin*, *Device Used*, *Race/Ethnicity*, *Sexuality*, *Sub-community Membership*, and *Disability* to others in social VR is **directly telling other users** that information; • The most common method for presenting or disclosing offline *Gender Identity* and *Sex Assigned at Birth* to others in social VR is through the way their **voice sounds** when they use voice chat; • Information contained in social VR users' **username** is the least frequently used method for presenting or disclosing all identity characteristics.
RQ2: How are people's identity characteristics misperceived by others in social VR?	**Key Significant Findings**
RQ2a: What identity characteristics do users report as being frequently misperceived by others in social VR?	• Social VR users feel that others **most often** misperceive their *Age* and *Country of Origin*, and **least often** misperceive whether they have a *Disability* and their *Sex Assigned at Birth*.
RQ2b: What informs these misperceptions about other social VR users' identity characteristics?	• The **sound of a social VR user's voice** is the most frequently used method that causes misperceptions of identity characteristics. In contrast, **information in a user's bio** is the least frequently used method.
RQ3: How are the presented or misperceived identity characteristics associated with harassment in social VR?	**Key Significant Findings**
RQ3a: Which presented/disclosed or misperceived identity characteristics are associated with more frequent harassment in social VR?	• Social VR users are harassed more frequently for identity characteristics that are **misperceived** by others rather than for those they actually present or disclose. • Concerning presented or disclosed identity characteristics, social VR users are harassed most frequently for ***Country of Origin***. • Concerning misperceived identity characteristics, social VR users are harassed most frequently for ***Disability.*** • Regardless of correctly perceived or misperceived identity characteristics, social VR users are harassed most frequently for ***Sub-community Membership***.
RQ3b: When their identity characteristics are presented/disclosed or misperceived, which user groups may encounter more frequent harassment in social VR than others?	• **Women** are harassed more frequently than men for *Age*, *Device Used*, *Gender Identity*, *Sex Assigned at Birth*, and *Sexuality*; • **Nonbinary/third gender/agender** individuals are harassed more frequently than men for *Gender Identity*, *Sub-community Membership*, *Sex Assigned at Birth*, and *Sexuality*, and more frequently than women for *Sub-community Membership* and *Sexuality*. • **LGBTQIA+** individuals are harassed more frequently than straight individuals for *Gender Identity*, *Sub-community Membership*, and *Sexuality*. • **Non-White** individuals are harassed more frequently than White individuals for *Age*, *Device Used*, and *Race/Ethnicity*, and almost significantly more for *Country of Origin* and *Sex Assigned at Birth*. • However, **White** individuals are harassed more frequently than **non-White** individuals for *Disability*.

Table 2. Summary of Key Significant Findings

characteristics to others (e.g., as a voluntary, intentional, and direct disclosure rather than indirect signaling through avatar design and body movement), which we explain below.

First, our study adds nuance to existing understandings of identity-based harassment in social VR by highlighting how social VR users *selectively* reveal identity characteristics based on their "imagined audience" [21, 45, 48, 49, 71]. Indeed, our study found that the interaction between identity characteristic type and audience type is significant. This means that *what* is being revealed to *whom* significantly affects the level of comfort and frequency of identity revelation that is consistent with one's offline identity. This critical insight may better inform how future social VR harassment mitigation strategies should be designed to account for this selectivity, such as consent mechanics

to protect social VR users from harassment based on their specific relationship to other users as shown in [88].

Second, compared to prior works that mainly highlight gender, sexuality, and race as the main factors for identity-based harassment in social VR [7, 8, 30, 66, 68], our results show a surprising observation: regardless of being correctly perceived or misperceived by other users, having a *sub-community membership* is the most frequently targeted identity characteristic for identity-based harassment in social VR compared to gender identity, sexuality, or race/ethnicity (see 4.4.1, Figure 14, and Table 15 in Appendix). Indeed, the sub-community of Furries has become a prevalent user group in social VR spaces, which often consists of individuals of already-identified vulnerable populations (i.e., LGBTQIA+ individuals [75]). However, the study of sub-communities in social VR as a particularly vulnerable population is still critically underexplored.

In fact, prior works on online harassment targeting the sub-community membership of Furries have found that they are often reluctant to disclose their Furry identity to others, partially due to imagined and realized fears of encountering scorn, disdain, or threats to their jobs and relationships because of their identity [5, 75, 78] (i.e., criteria for identity-based harassment [14]). Our results simultaneously diverge from and reflect these patterns. On the one hand, our social VR user participants reported that sexual or romantic partners are the audience type they feel most comfortable presenting or disclosing to and most frequently present their sub-community membership to (see Figures 10 and 11). This seems to suggest that social VR users may not fear losing this type of relationship or being harassed by this type of audience for revealing their Furry identity. On the other hand, our participants reported feeling least comfortable presenting or disclosing their sub-community identity, and doing so least frequently, to work colleagues (see Figures 10 and 11). This suggests that social VR users may indeed fear losing their jobs or being harassed by one's coworkers once they reveal their sub-community membership. Taken together, our study further highlights the vulnerability and sensitivity of sub-community membership identity in social VR and the harassment risks stemming from revealing this identity characteristic to specific audiences.

5.1.2 Presenting or Disclosing One's Offline Identity in Social VR May be Less Risky Than Once Thought. Given how harassment in social VR is widely reported and described in prior qualitative research such as [7, 8, 30], popular media sources such as [93], and technical reports such as [66, 68], it is understandable why harassment may be viewed as almost inevitable in social VR spaces when embodying one's own identity. Our study paints a more positive and safer picture of social VR. First, regardless of presenting or disclosing or encountering others' misperceptions of their identity characteristics, social VR users reported ***relatively low frequencies of harassment overall*** as explained in 4.4.1 (see also Figure 14). Second, while our participants reported presenting or disclosing their identities significantly more frequently overall than encountering others' misperceptions of their identity, they were ***harassed significantly more frequently overall for others' misperceptions about them than for what they actually present*** (see 4.4.1, Figure 14).

These results thus demonstrate that choosing to present and fully embody one's identity does not always lead to identity-based harassment in social VR. Indeed, many studies have pointed out how embodying one's identity in social VR may benefit a user's life by creating unique opportunities for exploring and experimenting with new self-identities [23, 27, 28, 31], enhancing visibility [23], and seeking more natural and realistic social support [47]. However, the very features for embodying one's identity in social VR may also lead to harassment that feels more realistic compared to other online social platforms, such as embodied physicalized harassment [47]. Therefore, past works have argued that some social VR users may be forced to decide between fully embodying their identities in social VR for those novel opportunities and protecting themselves from harassment targeting their identities. Yet, as our results have shown, there is some hope that the various social

and emotional benefits of *consistent* identity presentation in social VR may still outweigh the risk of potential identity-based harassment.

*5.1.3 Who You Are or Appear to Be **Does** Matter When It Comes to Identity-Based Harassment in Social VR.* While our results paint a seemingly more positive picture of identity-based harassment in social VR, critically reflecting upon this positivity is also important. Above all, many of the 223 participants in our study identify as cis men (65.47%), straight (79.82%), and White (57.85%). However, our sample is more internationally representative than typical US-centered social VR studies. Therefore, it is essential to explicate how our participants who are often considered as marginalized in social VR (i.e., women, non-White individuals, and LGBTQIA+ individuals [7, 8, 30, 66, 68]) specifically report harassment frequency. Indeed, our quantitative results in 4.4.2, 4.4.3, and 4.4.4 further confirm and support the aforementioned existing qualitative studies and technical reports by showing ***significant differences* in how different social VR user groups experience frequency of harassment for certain identity characteristics**, regardless of whether they present their identities or encounter misperceptions about their identities.

First, compared to men, women reported that they were significantly more harassed regardless of the identity characteristic type they presented or were misperceived. Non-binary/third-gender/agender individuals reported being more significantly harassed than men overall as well (see 4.4.2 and Figure 8). This is not surprising, as the disproportionate frequency of harassment against women in online environments such as social media, live streaming, and competitive gaming is exceedingly well-documented (e.g., [10, 13, 29, 50, 51, 69, 81, 89, 100–102]). Qualitative works and technical reports on harassment in social VR have drawn similar conclusions about the disproportionate harassment women face in these spaces, particularly sexual harassment [7, 8, 30, 86]. Our study thus provides critical evidence that these insights on women's marginalization in social VR hold true within a larger, international sample with statistically significant results.

These gendered differences also uncover previously understudied patterns. For example, women were harassed significantly more frequently for *age* and *device use*, two identity characteristics under or unexplored in prior works on social VR. Additionally, non-binary/third gender/agender participants were harassed for *sub-community membership* and *sexuality* significantly more frequently than both men and women, regardless of whether this perception was accurate (i.e., presented) or inaccurate (i.e., misperceived) to their actual sub-community membership and sexuality identities. This result thus further highlights how non-cisgender (including transgender, non-binary and genderfluid [28]) individuals often encounter online harassment more frequently compared to individuals who identify explicitly as within the binary [41, 44, 76, 82].

Second, as with gender, individuals identifying as LGBTQIA+ have also been well-documented as being targets for harassment in online environments, particularly within environments that place heavy emphasis on heteronormative values and norms [11, 33, 74, 77, 79]. Here again, prior qualitative and technical report works on harassment in social VR echo these sentiments [7, 8, 30, 66]. Our results further provide critical validating evidence that social VR users who identify as LGBTQIA+ (specifically gay, bisexual, asexual, lesbian, or pansexual) do indeed report being harassed significantly more frequently than straight individuals, but only for the presented or misperceived identity characteristics of *gender identity*, *sub-community membership*, and *sexuality*. It is interesting to note that *sub-community membership* is again more significantly associated with individuals who fall outside of heteronormative standards, perhaps because LGBTQIA+ individuals often actually are or are perceived to be members of sub-communities such as Furries [75].

Finally, prior works have shown racial minorities to be particularly vulnerable to harassment in social VR [23, 27, 30]. Yet, much of these works have utilized U.S.-based samples, making it difficult to determine how social VR users on a more international level might experience race-based

harassment. Using a sample of 223 individuals from 32 countries across six continents/regions of the world (see Table 1), our study thus simultaneously extends prior works by demonstrating that non-White individuals (e.g., Asian, Hispanic or Latino/a, Black/African American or African, and Mixed) report experiencing harassment *significantly more frequently* than White individuals overall. This is especially impactful given the international nature of our study.

5.2 Implications for Investigating and Designing for Safer Social VR Spaces

In light of our results on (mis)perceived identity revelations and their relation to identity-based harassment, we propose two primary implications for investigating and designing safer social VR spaces: (1) re-orienting research and development efforts to place more attention on previously underexplored aspects of identity-based harassment in social VR, and (2) better communicating how the interplay between (mis)perceived identity revelation and identity-based harassment to social VR users.

5.2.1 Re-Orienting Research and Development on Harassment in Social VR Spaces. Our work first demonstrates the value of re-orienting existing research and development on identity-based harassment in social VR spaces towards several previously underexplored factors. Above all, greater attention should be paid to how a user's voice is far more impactful on revealing one's identity than previously thought in existing work on social VR. In particular, our findings highlight that how one's voice sounds is significantly associated with misperceptions of one's identity characteristics, which may lead to more frequent harassment. Therefore, it becomes apparent that protecting users' voices from others should be prioritized to promote safer social VR spaces. Additionally, by taking a more *selective* lens and by focusing on more identity characteristics beyond the typical trifecta (i.e., gender, sexuality, and race), it becomes clear that social VR researchers and developers alike must further investigate challenges and risks for sub-communities (e.g., Furries). More specifically, while our study clearly shows how this identity characteristic is particularly targeted, more investigation is needed to understand how their harassment experiences and protection needs might be unique compared to other populations and variations of identity-based harassment.

5.2.2 Educating Social VR Users on the Interplay between (Mis)Perceived Identity Revelation and Identity-Based Harassment. Many social VR platforms seek to establish community norms surrounding safety and appropriate conduct in social VR through the publication of community guidelines on their official websites, e.g., VRChat's Community Guidelines page [103] and Horizon Worlds' Safety page [57]. Yet, these guidelines offer relatively little education on how one can or should approach protecting oneself from identity-based harassment specifically. Through the investigation of *selective* identity revelation, our results have uncovered various patterns concerning to whom social VR users are comfortable presenting or disclosing, how they present, and the implications these selective processes have on presenting or disclosing versus misperceiving, all of which should be communicated to social VR users directly. All of these patterns also inherently require users to leverage self-agency to gain profound knowledge about how (mis)perceived identity revelation can serve as a double-edged sword in social VR (e.g., if the benefits of presenting or disclosing one's offline identity outweigh the risks of identity-based harassment). On this matter, our results provide a hopeful answer: a social VR user is more likely to face harassment for an identity characteristic that is *not even true to their identity*, rather than for an identity characteristic that is true to them and intentionally presented. Educating users about this interplay between identity-revelation and identity-based harassment in social VR - including methods for presenting or disclosing, audiences one might present to, and harassment risks for certain marginalized user groups - should be included within social VR platforms' community guidelines and safety material. In doing so, social VR users can both be fully aware of the potential risks of identity-based harassment and still have the chance

to make informed decisions about how they wish to control their identity presentation in social VR selectively.

5.3 Limitations and Future Work

Despite our thorough recruitment on a well-verified research platform from 32 different countries across six continents/regions around the world, a primary limitation of this study is the skew towards White, straight men. While future work must keep seeking to recruit more diverse social VR users, the use of LME and stratified categorization of gender, sexuality, and race still allowed us to provide valuable insights into how marginalized user groups did experience harassment significantly more frequently than their non-marginalized counterparts in various ways. Additionally, other important identity characteristics and user groups that were not included in our study should be further explored regarding identity-based harassment in social VR, including but not limited to religion, socioeconomic class, and stigmatized physical characteristics. This study also did not compare children vs. adults. As minors are developmentally different than adults when dealing with online risks [72, 104], they may also suffer from more traumatic experiences and severe psychological and emotional damages when they become targeted for online harassment [16, 105]. Nevertheless, future work could replicate the essence of our research design to focus on children in social VR.

Furthermore, our significant results concerning *Sub-community Membership*'s impacts on identity-based harassment calls for future works to explicate how different sub-community memberships may lead to various degrees of identity-based harassment in social VR. Our study also shows the value of further investigating how misperceptions can affect identity-based harassment overall. This foundation thus justifies future work to compare more nuanced types of identity revelation. Finally, while this study specifically sought to establish a baseline quantitative understanding of various identity characteristics and user groups, our future work plans to focus on *how people with complex intersectional identities may encounter unique identity-based harassment* in social VR to further unpack structural oppression and unbalanced power dynamics in these new online social spaces.

6 CONCLUSION

In recognition of how one's identity can be linked to harassment, our study takes a vital step towards understanding harassment in social VR on a broad scale, especially regarding identity-based harassment. Our study is particularly impactful by taking to heart the notion that harassment is culturally contextual. In doing so, we intentionally recruited an international, non-U.S. centric sample from a wide range of countries, which significantly complements existing works on harassment in social VR that often focus on U.S.- based populations [7, 8, 30, 65, 66, 90]. Our study thus uncovers how various factors surrounding (mis)perceived identity revelations are significantly impacted by the nature of the revelation (i.e., whether one is presenting or disclosing or being misperceived by others) and the type of identity characteristic being revealed. We also provide critically needed evidence of previously unexplored targets for harassment (i.e., sub-community membership) and highlight the continued marginalization and targeting of women, LGBTQIA+, and racial/ethnic minority user groups for identity-based harassment in social VR as a global phenomenon. Therefore, in hopes of designing safer social VR spaces, our study calls for social VR researchers and developers alike to redirect their attentions toward previously underexplored aspects of identity revelation and identity-based harassment in social VR and advocates for the explicit education of users on leveraging self-agency to fully understand how (mis)perceived identity revelation can serve as a double-edged sword in social VR.

ACKNOWLEDGMENTS

We thank our participants and the anonymous reviewers. We also especially thank Christopher Flathmann for his highly valuable help with our data analysis. This work was supported by the National Science Foundation under award 2112878.

REFERENCES

[1] Paritosh Bahirat, Yangyang He, Abhilash Menon, and Bart Knijnenburg. 2018. A data-driven approach to developing IoT privacy-setting interfaces. In *23rd International Conference on Intelligent User Interfaces*. 165–176.

[2] Steven Baker, Ryan M Kelly, Jenny Waycott, Romina Carrasco, Thuong Hoang, Frances Batchelor, Elizabeth Ozanne, Briony Dow, Jeni Warburton, and Frank Vetere. 2019. Interrogating social virtual reality as a communication medium for older adults. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–24.

[3] Steven Baker, Jenny Waycott, Romina Carrasco, Ryan M Kelly, Anthony John Jones, Jack Lilley, Briony Dow, Frances Batchelor, Thuong Hoang, and Frank Vetere. 2021. Avatar-Mediated Communication in Social VR: An In-depth Exploration of Older Adult Interaction in an Emerging Communication Platform. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[4] Shaowen Bardzell and Jeffrey Bardzell. 2011. Towards a feminist HCI methodology: social science, feminism, and HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 675–684.

[5] Penny Bernstein, Nicholas Paolone, Justin Higner, Kathleen Gerbasi, Samuel Conway, Adam Privitera, and Laura Scaletta. 2008. Furries from A to Z (Anthropomorphism to Zoomorphism). *Society & Animals* 16, 3 (2008), 197–222.

[6] Lindsay Blackwell, Jill Dimond, Sarita Schoenebeck, and Cliff Lampe. 2017. Classification and its consequences for online harassment: Design insights from heartmob. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–19.

[7] Lindsay Blackwell, Nicole Ellison, Natasha Elliott-Deflo, and Raz Schwartz. 2019. Harassment in social virtual reality: Challenges for platform governance. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–25.

[8] Lindsay Blackwell, Nicole Ellison, Natasha Elliott-Deflo, and Raz Schwartz. 2019. Harassment in social VR: Implications for design. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 854–855.

[9] Tom Boellstorff. 2015. *Coming of age in Second Life: An anthropologist explores the virtually human*. Princeton University Press.

[10] Jason T Bowey, Ansgar E Depping, and Regan L Mandryk. 2017. Don't talk dirty to me: How sexist beliefs affect experience in sexist games. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 1530–1543.

[11] Jeremy Brenner-Levoy. 2019. *Virtually Masculine: Queer Men's Experiences with Harassment in Online Video Games*. Ph. D. Dissertation. University of Cincinnati.

[12] Kalyani Chadha, Linda Steiner, Jessica Vitak, and Zahra Ashktorab. 2020. Women's Responses to Online Harassment. *International Journal of Communication (19328036)* 14 (2020).

[13] Despoina Chatzakou, Nicolas Kourtellis, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Athena Vakali. 2017. Hate is not binary: Studying abusive behavior of# gamergate on twitter. In *Proceedings of the 28th ACM conference on hypertext and social media*. 65–74.

[14] U.S. Equal Employment Opportunity Commission. 2021. *Harassment*. https://www.eeoc.gov/harassment

[15] Robyn M Cooper and Warren J Blumenfeld. 2012. Responses to cyberbullying: A descriptive analysis of the frequency of and impact on LGBT and allied youth. *Journal of LGBT Youth* 9, 2 (2012), 153–177.

[16] Jennifer E Copp, Elizabeth A Mumford, and Bruce G Taylor. 2021. Online sexual harassment and cyberbullying in a nationally representative sample of teens: Prevalence, predictors, and consequences. *Journal of adolescence* 93 (2021), 202–211.

[17] Shanley Corvite, Ben Zefeng Zhang, and Oliver L Haimson. 2022. Social Media's Role During Identity Changes Related to Major Life Events. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–22.

[18] Elmira Deldari, Diana Freed, Julio Poveda, and Yaxing Yao. 2023. An investigation of teenager experiences in social virtual reality from teenagers', parents', and bystanders' perspectives. In *Proceedings of the Nineteenth USENIX Conference on Usable Privacy and Security*. 1–17.

[19] Nicolas Ducheneaut, Ming-Hui Wen, Nicholas Yee, and Greg Wadley. 2009. *Body and Mind: A Study of Avatar Personalization in Three Virtual Worlds*. Association for Computing Machinery, New York, NY, USA, 1151–1160. https://doi.org/10.1145/1518701.1518877

[20] Maeve Duggan. 2017. Online harassment 2017. (2017).

[21] Shelly D. Farnham and Elizabeth F. Churchill. 2011. Faceted Identity, Faceted Lives: Social and Technical Issues with Being Yourself Online. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (Hangzhou, China) *(CSCW '11)*. Association for Computing Machinery, New York, NY, USA, 359–368. https:

//doi.org/10.1145/1958824.1958880

[22] Cristina Fiani, Robin Bretin, Mark McGill, and Mohamed Khamis. 2023. Big Buddy: Exploring Child Reactions and Parental Perceptions towards a Simulated Embodied Moderating System for Social Virtual Reality. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*. 1–13.

[23] Guo Freeman and Dane Acena. 2022. " Acting Out" Queer Identity: The Embodied Visibility in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–32.

[24] Guo Freeman, Dane Acena, Nathan J McNeese, and Kelsea Schulenberg. 2022. Working Together Apart through Embodiment: Engaging in Everyday Collaborative Activities in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–25.

[25] Guo Freeman, Jeffrey Bardzell, and Shaowen Bardzell. 2016. Revisiting computer-mediated intimacy: In-game marriage and dyadic gameplay in Audition. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 4325–4336.

[26] Guo Freeman, Jeffrey Bardzell, Shaowen Bardzell, and Susan C. Herring. 2015. Simulating Marriage: Gender Roles and Emerging Intimacy in an Online Game. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1191–1200. https://doi.org/10.1145/2675133.2675192

[27] Guo Freeman and Divine Maloney. 2021. Body, avatar, and me: The presentation and perception of self in social virtual reality. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–27.

[28] Guo Freeman, Divine Maloney, Dane Acena, and Catherine Barwulor. 2022. (Re) discovering the Physical Body Online: Strategies and Challenges to Approach Non-Cisgender Identity in Social Virtual Reality. In *CHI Conference on Human Factors in Computing Systems*. 1–15.

[29] Guo Freeman and Donghee Yvette Wohn. 2020. Streaming your Identity: Navigating the Presentation of Gender and Sexuality through Live Streaming. *Computer Supported Cooperative Work (CSCW)* 29, 6 (2020), 795–825.

[30] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Dane Acena. 2022. Disturbing the Peace: Experiencing and Mitigating Emerging Harassment in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (2022), 1–30.

[31] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Alexandra Adkins. 2020. My body, my avatar: How people perceive their avatars in social virtual reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[32] Markus Friedl. 2002. *Online game interactivity theory with cdrom*. Charles River Media, Inc.

[33] Laura E Gillin and Margaret L Signorella. 2023. Attitudes Toward Sexual Orientation and Gender Identity in Online Multiplayer Gaming Spaces. *Psychological reports* (2023), 00332941231153798.

[34] Erving Goffman et al. 1978. *The presentation of self in everyday life*. Harmondsworth London.

[35] Nitesh Goyal, Leslie Park, and Lucy Vasserman. 2022. "You have to prove the threat is real": Understanding the needs of Female Journalists and Activists to Document and Report Online Harassment. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[36] Kishonna Gray. 2014. *Race, gender, and deviance in Xbox live: Theoretical perspectives from the virtual margins*. Routledge.

[37] Kishonna L Gray. 2012. Intersecting oppressions and online communities: Examining the experiences of women of color in Xbox Live. *Information, Communication & Society* 15, 3 (2012), 411–428.

[38] Kishonna L Gray. 2018. Gaming out online: Black lesbian identity development and community building in Xbox Live. *Journal of lesbian studies* 22, 3 (2018), 282–296.

[39] Jon 'Catalyst' Grey. 2011. BurnYourBra discusses the difficulties of being a female gamer. https://www.eventhubs.com/news/2011/apr/15/dmgburnyourbra-discusses-difficulties-being-female-gamer/

[40] Lucy Hackworth. 2018. Limitations of "just gender": The need for an intersectional reframing of online harassment discourse and research. *Mediating misogyny: Gender, technology, and harassment* (2018), 51–70.

[41] Oliver L Haimson. 2016. Transgender experiences with online harassment. In *Social Computing Symposium*.

[42] Tyler Hatchel, Cagil Torgal, America J El Sheikh, Luz E Robinson, Alberto Valido, and Dorothy L Espelage. 2021. LGBTQ youth and digital media: online risks. In *Child and Adolescent Online Risk Exposure*. Elsevier, 303–325.

[43] Jane Im, Sarita Schoenebeck, Marilyn Iriarte, Gabriel Grill, Daricia Wilkinson, Amna Batool, Rahaf Alharbi, Audrey Funwie, Tergel Gankhuu, Eric Gilbert, et al. 2022. Women's Perspectives on Harm and Justice after Online Harassment. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–23.

[44] Shagun Jhaver, Sucheta Ghoshal, Amy Bruckman, and Eric Gilbert. 2018. Online harassment and content moderation: The case of blocklists. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 2 (2018), 1–33.

[45] Sanjay Kairam, Mike Brzozowski, David Huffaker, and Ed Chi. 2012. Talking in circles: selective sharing in google+. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1065–1074.

[46] Amanda Lenhart, Michele Ybarra, Kathryn Zickuhr, and Myeshia Price-Feeney. 2016. *Online harassment, digital abuse, and cyberstalking in America.* Data and Society Research Institute.

[47] Lingyuan Li, Guo Freeman, Kelsea Schulenberg, and Dane Acena. 2023. " 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*. 1–16.

[48] Eden Litt. 2012. Knock, knock. Who's there? The imagined audience. *Journal of broadcasting & electronic media* 56, 3 (2012), 330–345.

[49] Eden Litt and Eszter Hargittai. 2016. The imagined audience on social network sites. *Social Media+ Society* 2, 1 (2016), 2056305116633482.

[50] Daniel Madden and Casper Harteveld. 2021. "Constant Pressure of Having to Perform": Exploring Player Health Concerns in Esports. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[51] Daniel Madden, Yuxuan Liu, Haowei Yu, Mustafa Feyyaz Sonbudak, Giovanni M Troiano, and Casper Harteveld. 2021. "Why Are You Playing Games? You Are a Girl!": Exploring Gender Biases in Esports. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.

[52] Divine Maloney, Guo Freeman, and Andrew Robb. 2020. It is complicated: Interacting with children in social virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 343–347.

[53] Divine Maloney, Guo Freeman, and Andrew Robb. 2020. A Virtual Space for All: Exploring Children's Experience in Social Virtual Reality. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 472–483.

[54] Divine Maloney, Samaneh Zamanifard, and Guo Freeman. 2020. Anonymity vs. familiarity: Self-disclosure and privacy in social virtual reality. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*. 1–9.

[55] Andy Marra. 2013. Out online: the experiences of LGBT youth on the internet. *Gay, Lesbian, and Straight Education Network* (2013).

[56] Joshua McVeigh-Schultz, Anya Kolesnichenko, and Katherine Isbister. 2019. Shaping pro-social interaction in VR: an emerging design framework. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[57] Meta. 2022. Safety and privacy in Meta Horizon Worlds. https://www.meta.com/help/quest/articles/horizon/safety-and-privacy-in-horizon-worlds/index-safety-and-privacy/

[58] Steven E Mock, Courtney N Plante, Stephen Reysen, and Kathleen C Gerbasi. 2013. Deeper leisure involvement as a coping resource in a stigmatized leisure context. *Leisure/Loisir* 37, 2 (2013), 111–126.

[59] Andras Molnar. 2019. SMARTRIQS: A simple method allowing real-time respondent interaction in Qualtrics surveys. *Journal of Behavioral and Experimental Finance* 22 (2019), 161–169.

[60] Tyler Musgrave, Alia Cummings, and Sarita Schoenebeck. 2022. Experiences of Harm, Healing, and Joy among Black Women and Femmes on Social Media. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[61] Lisa Nakamura. 2013. *Cybertypes: Race, ethnicity, and identity on the Internet.* Routledge.

[62] Lisa Nakamura. 2014. Gender and race online. *Society and the Internet. How networks of information and communication are changing our lives. Oxford* (2014), 81–95.

[63] Kristine L Nowak and Christian Rauh. 2005. The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. *Journal of Computer-Mediated Communication* 11, 1 (2005), 153–178.

[64] Joseph O'Hagan, Florian Mathis, and Mark McGill. 2023. User Reviews as a Reporting Mechanism for Emergent Issues Within Social VR Communities. In *Proceedings of the 22nd International Conference on Mobile and Ubiquitous Multimedia*. 236–243.

[65] J Outlaw and B Duckles. 2017. Why Woman Don't Like Social Virtual Reality.

[66] Jessica Outlaw and Beth Duckles. 2017. Why Women Don't Like Social Virtual Reality: A Study of Safety, Usability, and Self-expression in Social VR. https://www.extendedmind.io/why-women-dont-like-social-virtual-reality

[67] Jessica Outlaw and Beth Duckles. 2017. Why Women Don't Like Social Virtual Reality: A Study of Safety, Usability, and Self-expression in Social VR. *The Extended Mind* (2017).

[68] Jessica Outlaw and Beth Duckles. 2018. Virtual Harassment: The Social Experience of 600+ Regular Virtual Reality (VR) Users. https://virtualrealitypop.com/virtual-harassment-the-social-experience-of-600-regular-virtual-reality-vr-users-23b1b4ef884e

[69] Benjamin Paaßen, Thekla Morgenroth, and Michelle Stratemeyer. 2017. What is a true gamer? The male gamer stereotype and the marginalization of women in video game culture. *Sex Roles* 76, 7 (2017), 421–435.

[70] Stefan Palan and Christian Schitter. 2018. Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17 (2018), 22–27.

[71] Elliot T Panek, Yioryos Nardis, and Sara Konrath. 2013. Mirror or Megaphone?: How relationships between narcissism and social networking site use differ on Facebook and Twitter. *Computers in Human Behavior* 29, 5 (2013), 2004–2012.

[72] Jessica Pater and Elizabeth Mynatt. 2017. Defining digital self-harm. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1501–1513.
[73] Eyal Peer, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer. 2022. Data quality of platforms and panels for online behavioral research. *Behavior Research Methods* 54, 4 (2022), 1643–1662.
[74] Amanda Phillips and Bonnie Ruberg. 2018. Not Gay as in Happy: Queer Resistance and Video Games (Introduction). *Game Studies* 18, 2018 (2018), 1–2.
[75] Courtney N Plante, Sharon Roberts, Stephen Reysen, and Kathleen Gerbasi. 2014. Interaction of socio-structural characteristics predicts identity concealment and self-esteem in stigmatized minority group members. *Current Psychology* 33 (2014), 3–19.
[76] Anastasia Powell, Adrian J Scott, and Nicola Henry. 2020. Digital harassment and abuse: Experiences of sexuality and gender minority adults. *European Journal of Criminology* 17, 2 (2020), 199–223.
[77] Alexis Pulos. 2013. Confronting heteronormativity in online games: A critical discourse analysis of LGBTQ sexuality in World of Warcraft. *Games and Culture* 8, 2 (2013), 77–97.
[78] Sharon E Roberts, Courtney Plante, Kathleen Gerbasi, and Stephen Reysen. 2015. Clinical interaction with anthropomorphic phenomenon: Notes for health professionals about interacting with clients who possess this unusual identity. *Health & Social Work* 40, 2 (2015), e42–e50.
[79] Bonnie Ruberg and Adrienne Shaw. 2017. *Queer game studies*. U of Minnesota Press.
[80] Nazanin Sabri, Bella Chen, Annabelle Teoh, Steven P Dow, Kristen Vaccaro, and Mai Elsherief. 2023. Challenges of Moderating Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–20.
[81] Nithya Sambasivan, Amna Batool, Nova Ahmed, Tara Matthews, Kurt Thomas, Laura Sanely Gaytán-Lugo, David Nemer, Elie Bursztein, Elizabeth Churchill, and Sunny Consolvo. 2019. " They Don't Leave Us Alone Anywhere We Go" Gender and Digital Abuse in South Asia. In *proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
[82] Morgan Klaus Scheuerman, Stacy M Branham, and Foad Hamidi. 2018. Safe spaces and safe places: Unpacking technology-mediated experiences of safety and harm with transgender people. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–27.
[83] Ari Schlesinger, W Keith Edwards, and Rebecca E Grinter. 2017. Intersectional HCI: Engaging identity through gender, race, and class. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 5412–5427.
[84] Sarita Schoenebeck, Carol F Scott, Emma Grace Hurley, Tammy Chang, and Ellen Selkie. 2021. Youth trust in social media companies and expectations of justice: Accountability and repair after online harassment. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–18.
[85] Ralph Schroeder. 2012. *The social life of avatars: Presence and interaction in shared virtual environments*. Springer Science & Business Media.
[86] Kelsea Schulenberg, Guo Freeman, Lingyuan Li, and Catherine Barwulor. 2023. "Creepy Towards My Avatar Body, Creepy Towards My Body": How Women Experience and Manage Harassment Risks in Social Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 236 (oct 2023), 29 pages. https://doi.org/10.1145/3610027
[87] Kelsea Schulenberg, Lingyuan Li, Guo Freeman, Samaneh Zamanifard, and Nathan J McNeese. 2023. Towards Leveraging AI-based Moderation to Address Emergent Harassment in Social Virtual Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
[88] Kelsea Schulenberg, Lingyuan Li, Caitlin Lancaster, Douglas Zytko, and Guo Freeman. 2023. "We Don't Want a Bird Cage, We Want Guardrails": Understanding & Designing for Preventing Interpersonal Harm in Social VR through the Lens of Consent. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW2, Article 323 (oct 2023), 30 pages. https://doi.org/10.1145/3610172
[89] Orit Shaer, Lauren Westendorf, Nicholas A Knouf, and Claudia Pederson. 2017. Understanding gaming perceptions and experiences in a women's college community. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 1544–1557.
[90] Ketaki Shriram and Raz Schwartz. 2017. All are welcome: Using VR ethnography to explore harassment behavior in immersive social virtual reality. In *2017 IEEE Virtual Reality (VR)*. IEEE, 225–226.
[91] Mel Slater, Daniel Pérez Marcos, Henrik Ehrsson, and Maria V Sanchez-Vives. 2009. Inducing illusory ownership of a virtual body. *Frontiers in neuroscience* (2009), 29.
[92] Weilun Soon. 2022. A researcher's avatar was sexually assaulted on a metaverse platform owned by Meta. https://www.businessinsider.com/researcher-claims-her-avatar-was-raped-on-metas-metaverse-platform-2022-5
[93] Hannah Sparks. 2021. Woman claims she was virtually 'groped' in Meta's VR metaverse. https://nypost.com/2021/12/17/woman-claims-she-was-virtually-groped-in-meta-vr-metaverse/
[94] Jan E Stets and Peter J Burke. 2000. Identity theory and social identity theory. *Social psychology quarterly* (2000), 224–237.

[95] Sharifa Sultana, Mitrasree Deb, Ananya Bhattacharjee, Shaid Hasan, SM Raihanul Alam, Trishna Chakraborty, Prianka Roy, Samira Fairuz Ahmed, Aparna Moitra, M Ashraful Amin, et al. 2021. 'Unmochon': A Tool to Combat Online Sexual Harassment over Facebook Messenger. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.

[96] Philipp Sykownik, Linda Graf, Christoph Zils, and Maic Masuch. 2021. The Most Social Platform Ever? A Survey about Activities amp; Motives of Social VR Users. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. 546–554. https://doi.org/10.1109/VR50410.2021.00079

[97] Philipp Sykownik, Divine Maloney, Guo Freeman, and Maic Masuch. 2022. Something personal from the Metaverse: Goals, topics, and contextual factors of self-disclosure in commercial social VR. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[98] Jirassaya Uttarapong, Jie Cai, and Donghee Yvette Wohn. 2021. Harassment experiences of women and LGBTQ live streamers and how they handled negativity. In *ACM international conference on interactive media experiences*. 7–19.

[99] Vincent van Brakel, Miguel Barreda-Ángeles, and Tilo Hartmann. 2023. Feelings of presence and perceived social support in social virtual reality platforms. *Computers in Human Behavior* 139 (2023), 107523.

[100] Aditya Vashistha, Abhinav Garg, Richard Anderson, and Agha Ali Raza. 2019. Threats, abuses, flirting, and blackmail: Gender inequity in social media voice forums. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[101] Jessica Vitak, Kalyani Chadha, Linda Steiner, and Zahra Ashktorab. 2017. Identifying women's experiences with and strategies for mitigating negative effects of online harassment. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1231–1245.

[102] Vogels, Emily A. 2021. *The State of Online Harassment*. Technical Report. Pew Research Center, Washington, D.C. https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/

[103] VRChat. 2022. Community guidelines. https://hello.vrchat.com/community-guidelines#:~:text=VRChat%20Moderation%20will%20take%20action,boards%2C%20and%20other%20online%20portals.

[104] Michel Walrave, Ini Vanwesenbeeck, and Wannes Heirman. 2012. Connecting and protecting? Comparing predictors of self-disclosure and privacy settings use between adolescents and adults. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 6, 1 (2012).

[105] Pamela Wisniewski, Heng Xu, Mary Beth Rosson, Daniel F Perkins, and John M Carroll. 2016. Dear diary: Teens reflect on their weekly online risk experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3919–3930.

[106] Janis Wolak, Kimberly J Mitchell, and David Finkelhor. 2006. Online Victimization of Youth: Five Years Later. (2006).

[107] Nick Yee, Jeremy N Bailenson, Mark Urbanek, Francis Chang, and Dan Merget. 2007. The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior* 10, 1 (2007), 115–121.

[108] Samaneh Zamanifard and Guo Freeman. 2019. " 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*. ACM, 438–442.

[109] Samaneh Zamanifard and Guo Freeman. 2023. A Surprise Birthday Party in VR: Leveraging Social Virtual Reality to Maintain Existing Close Ties over Distance. In *Information for a Better World: Normality, Virtuality, Physicality, Inclusivity: 18th International Conference, iConference 2023, Virtual Event, March 13–17, 2023, Proceedings, Part II*. Springer, 268–285.

[110] Kexin Zhang, Elmira Deldari, Zhicong Lu, Yaxing Yao, and Yuhang Zhao. 2022. "It's Just Part of Me:" Understanding Avatar Diversity and Self-presentation of People with Disabilities in Social Virtual Reality. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–16.

[111] Kexin Zhang, Elmira Deldari, Yaxing Yao, and Yuhang Zhao. 2023. A Diary Study in Social Virtual Reality: Impact of Avatars with Disability Signifiers on the Social Experiences of People with Disabilities. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*. 1–17.

A ADDITIONAL INFORMATION ABOUT SURVEY DESIGN

A.1 Summary of Identity Characteristics Included in the Survey

Identity Characteristics	Shorthand	Description	Sources
1. Age	Age	One's age in the offline world, (i.e., how old you are).	[3, 7, 8, 14, 15, 18, 20, 30, 46, 54, 66, 68, 84, 97, 106]
2. Country of Origin	Country	One's birth country in the offline world (i.e., where you were born).	[14, 20]
3. Disability	Disability	Whether one identifies as a person with a disability (PWD) offline.	[14, 20, 110, 111]
4. Device Use	Device	The type of device one uses when engaging in social VR (e.g., Oculus Quest 2, HTC Vive).	[14, 107]
5. Gender Identity	Gender	One's gender identity in the offline world (e.g., woman, man, non-binary, agender, genderfluid).	[6–8, 14, 15, 20, 30, 35, 43, 46, 54, 55, 60, 66, 68, 86, 95, 97, 98, 101]
6. Sub-Community Membership	Membership	One's association with a specific community within social VR (e.g., Furry communities).	[58, 75, 78]
7. Race/Ethnicity	Race	One's race/ethnicity in the offline world (e.g., Asian, Black/African..., Hispanic or...).	[6–8, 14, 20, 30, 54, 60, 66, 68, 97]
8. Sex Assigned at Birth	Birth Sex	One's biological sex at birth in the offline world (i.e., female, male, or intersex).	[6–8, 14, 15, 20, 28, 30, 46, 54, 55, 66, 68, 86, 97, 98]
9. Sexuality	Sexuality	One's sexual identity in the offline world (e.g., asexual, gay, lesbian, bisexual, panasexual, straight).	[6–8, 14, 15, 20, 30, 46, 54, 55, 66, 68, 97, 98]

Table 3. Identity characteristics included in the survey. The column "Sources" indicates prior works that have identified a certain identity characteristic as pertinent to identity-based harassment online generally and/or in social VR, which motivated us to include this specific identity characteristic in this research.

A.2 Definition of Social VR Provided to Participants in Survey

"Social virtual reality (VR) can be defined as online platforms where multiple users can interact with one another through VR head-mounted displays in 3D virtual spaces. Such platforms include VRChat, AltspaceVR, Meta's Horizon Worlds, RecRoom, etc. These platforms allow social VR users to enjoy offline-like social activities (e.g., walking in public spaces, playing a game, watching a movie, participating in a concert, and having a party) in a highly realistic and immersive simulated 3D virtual environment through the predominant use of real-time voice chat, partial or full-body track avatars, and more customized avatar design."

A.3 Detailed Overview of Survey Design

To address our research questions, we designed an anonymous online survey consisting of five primary sections: (1) consent form and filtering questions; (2) demographic and social VR use questions; (3) general trust and self-disclosure questions; (4) identity revelation and ensuing harassment questions (see Figure 1); and (5) coping strategy questions. Note that sections (3) and (5) are part of a broader project and are out of the scope of this paper. Therefore, below, we mainly explain Sections (1), (2), and (4) of the survey design.

In **Section 1**, upon consent agreement, participants were presented with screening questions, such as if they had ever used social virtual reality (VR) before, how long they had been engaging in social VR, and how many hours per week they use social VR. Participants who answered "No", "I do not engage in social VR," or "I do not use social VR" to the respective screening questions would not be allowed to proceed to the rest of the survey.

Section 2 contained eight demographic questions, including age, gender identity, sex assigned at birth, sexual identity, race/ethnicity, disability, country of origin (i.e., where you were born), and country of residence (i.e., where you live). As shown in Section 2, prior work has highlighted that individuals who present or who are perceived as women, non-White populations, and LGBTQIA+

individuals may face a higher risk of harassment in social VR [7, 8, 30, 66, 68, 90]. As such, answers to

Section 4 is the main focus of the survey. This section contained nine subsections representing nine identity characteristics that we have identified based on findings from prior qualitative research on self-presentation [54, 97] and harassment in social VR [7, 8, 28, 30] and technical reports on social VR harassment [65, 66], including 1. Age; 2. Country of Origin; 3. Disability; 4. Device Used; 5. Gender Identity; 6. Sub-community Membership; 7. Race; 8. Sex Assigned at Birth; and 9. Sexuality. The order in which these nine subsections were presented to each participant was randomized evenly.

Figure 1 provides an overview of the survey flow for each subsection in Section 4. Each identity characteristic subsection started with a question asking if they had ever presented or disclosed that identity characteristic to others in social VR. People who answered “Yes,” “Unsure,” or “Maybe” were then presented with questions about their disclosure practices (frequency, methods, and audience), level of concern for their privacy in the context of disclosure, and frequency of harassment after they present or disclose this characteristic to others in social VR. People who answered “No” were only asked about their level of concern in the context of potential disclosure.

Regardless of their initial branch path, all participants were also asked how often other users had made *misperceptions* about that identity characteristic. Participants who answered "Never" to this question were not asked further questions within the subsection. In contrast, all other answers directed participants to follow-up questions (i.e., assumption-related questions and frequency of harassment after such assumptions). The framing of "incorrect assumptions" thus was used in the survey to better delineate between presentation or disclosure and perception/assumptions questions that might otherwise have looked relatively similar or confusing to participants.

It is important to note that putting all presentation/disclosure-related questions before incorrect assumptions helps define the latter’s meaning through its relation to the former. In other words, if ‘disclosure’ questions refer to times when other users got to know the participant’s accurate offline identity, then ‘incorrect assumptions’ must refer to times when other users’ thoughts about the participant’s offline identity were not accurate. Indeed, incorrect assumptions can be made due to several different scenarios. For instance, if another social VR user thinks that our participant is presenting or disclosing in social VR consistently with their offline identity but they are not, that would be a type of incorrect assumption. As our study is the first of its kind to our knowledge to explore misperceptions systematically, we chose to particularly focus on establishing foundational knowledge about whether this type of identity revelation affects identity-based harassment in social VR rather than unpacking the intricacies of how the ‘incorrect assumption’ (i.e., misperception) happens in the first place.

A.4 Summary of Methods of Identity Revelation Included in the Survey

Method Type	Presentation Description	Misperception Description
Avatar Looks	*...through the way my avatar looks.*	*... because of the way my avatar looks.*
Sound of Voice	*... through the way my voice sounds when I use voice chat to talk to other users.*	*... because of the way my voice sounds when I use voice chat.*
Texting Style	*... through the way I type or express myself when I text chat other users.*	*... because of the way I type or express myself when I text chat.*
User Bio	*... by putting this information in my social VR bio that other users can see when they interact with me.*	*... because of the information I put in my social VR bio.*
Username	*... by putting this information in my social VR username.*	*... because of my username.*
Telling Others	*... by directly telling another social VR user this information, either verbally or through text chat.*	*... because of information I told them, either verbally or through text chat.*

Table 4. Methods of revealing identity characteristics presented in the survey.

A.5 Operationalizing the Frequency Scale in This Survey

To set clear parameters for how frequency should be determined when asking about the occurrence of events or phenomena in social VR, all frequency measures given to participants in this study were accompanied by the following explanation: *"For the purposes of this study, frequency basically refers to how often something happens every time you enter social VR for a session: Never - 0% of times I've been in social VR.; Rarely - About 30% of times.; Sometimes - About 50% of times.; Frequently - About 70% of times.; Always - 100% of times I've been in social VR."*

This operationalization of frequency was utilized for four reasons. First, *"every time you enter social VR for a session"* provides participants with a contextual frame of reference that is readily understandable and specific to the immersive qualities of social VR environments (*"enter"*). Second, *"how often something happens every time you enter"* provides participants with a manageable temporal framework to work from rather than asking them to think of all of their time in social VR in its entirety. Third, it provides both qualitative (Never, Rarely, etc.) and quantitative (0%, 30%, etc.) explanations to aid in participant clarity and allow researchers to use more than one explanation framing depending on need. Finally, the single-item nature of the measure prompts participants to isolate their evaluations of frequency to one type of occurrence at a time. In doing so, this measure aims to reduce issues of comparative frequency (i.e., rating an occurrence by how much more or less frequently it happens in comparison to a different occurrence) that can occur when being asked about more than one occurrence at a time and/or about different occurrences that happen simultaneously.

B STATISTIC RESULTS FOR 4.1 FREQUENCY OF MIS(PERCEIVED) IDENTITY REVELATION IN SOCIAL VR

Model	$\Delta\chi^2$	Δdf	***p-value***
Frequency of Revelation (1-pid)			
+ Identity Revelation Type	820.83	4	<.0001
+ Identity Characteristics	260.99	12	<.0001
+ Identity Revelation Type: Identity Characteristics	116.10	20	<.0001

Table 5. Linear model for effects of identity revelation type and identity characteristic type on frequency of revelation. Each model is built upon and compared to the one listed above.

Most -> Least Freq. Revealed - Misperceived	***M***	***SD***	**Most -> Least Freq. Revealed - Presented**	***M***	***SD***	**Most -> Least Freq. Revealed - Overall**	***M***	***SD***
Age	2.55	1.11	Sex Assigned at Birth	3.38	1.16	**Country of Origin**	**2.88**	**1.33**
Country of Origin	2.55	1.15	Country of Origin	3.3	0.96	**Age**	**2.7**	**1.08**
Race/Ethnicity	1.91	0.95	Gender Identity	3.21	1.08	**Gender Identity**	**2.89**	**1.26**
Gender Identity	1.83	1.05	Device Used	2.99	1.01	**Sex Assigned at Birth**	**2.36**	**1.35**
Device Used	1.82	1.0	Race/Ethnicity	2.95	1.11	**Race/Ethnicity**	**2.32**	**1.13**
Sexuality	1.78	1.03	Age	2.92	0.99	**Device Used**	**2.3**	**0.06**
Sub-community Membership	1.69	0.99	Sexuality	2.88	0.99	**Sexuality**	**2.21**	**1.15**
Sex Assigned at Birth	1.65	0.96	Sub-community Membership	2.74	0.96	**Sub-community Membership**	**1.82**	**1.04**
Disability	1.33	0.69	Disability	2.38	0.81	**Disability**	**1.4**	**0.75**

Table 6. Average frequency of revelation by identity characteristic for **Misperceived** (left column clusters), **Presented** (middle column clusters), and Misperceived & Presented combined i.e. **Overall** (right column clusters).

C STATISTIC RESULTS FOR 4.2 METHODS OF (MIS)PERCEIVED IDENTITY REVELATION IN SOCIAL VR

Model	$\Delta\chi^2$	Δdf	*p-value*
Frequency of Revelation (1-pid)			
+ Method	1102.08	8	<.0001
+ Identity Characteristics	468.75	16	<.0001
+ Identity Revelation Type	157.49	17	<.0001
+ Method: Identity Revelation Type	304.21	22	<.0001
+ Method: Identity Characteristics	351.7	62	<.0001
+ Identity Revelation Type: Identity Characteristics	214.78	70	<.0001
+ Identity Revelation Type: Identity Characteristics: Method	172.25	110	<.0001

Table 7. Linear model for effects of identity revelation type, identity characteristics, and method of presenting or disclosing/misperceiving on the frequency of revelation. Each model is built upon and compared to the one listed above.

Identity Characteristic	Most Common	*M*	*SD*	Least Common	*M*	*SD*
Age	Telling Others	3.87	1.03	Username	1.84	1.13
Country of Origin	Telling Others	3.85	1.15	Username	2.01	1.29
Disability	Telling Others	3.13	1.36	Username	1.5	1.1
Device Used	Telling Others	3.5	1.2	Username	1.76	1.08
Gender Identity	Sound of Voice	3.92	1.92	Username	2.2	1.32
Sub-community Membership	Telling Others	3.3	1.44	Username	1.87	1.26
Race/Ethnicity	Telling Others	3.51	1.17	Username	2.06	1.31
Sex Assigned at Birth	Sound of Voice	4.31	0.94	Username	2.41	1.34
Sexuality	Telling Others	3.68	1.13	Username	2.2	1.34
Overall for presenting or disclosing	**Telling Others**	**3.69**	**1.15**	**Username**	**2.05**	**1.27**

Table 8. The most and least common methods used by participants to **present** their identity characteristics to others in social VR.

Identity Characteristic	Most Common	*M*	*SD*	Least Common	*M*	*SD*
Age	Sound of Voice	3.4	1.19	Bio Info	2.27	1.35
Country of Origin	Sound of Voice	3.3	1.21	Username	2.25	1.23
Disability	Telling Others	2.63	1.38	Username	2.04	1.3
Device Used	Telling Others	2.93	1.25	Username	2.12	1.13
Gender Identity	Avatar Looks	2.82	1.27	Bio Info	2.45	1.23
Sub-community Membership	Avatar Looks	3.12	1.19	Bio Info	2.46	1.27
Race/Ethnicity	Sound of Voice	3.17	1.28	Bio Info	2.32	1.19
Sex Assigned at Birth	Avatar Looks	2.93	1.39	Bio Info	2.42	1.27
Sexuality	Text Chat	2.75	1.28	Bio Info	2.16	1.26
Overall for Misperceiving	**Sound of Voice**	**2.94**	**1.29**	**Bio Info**	**2.32**	**1.27**

Table 9. The most and least common methods our participants believe other social VR users used to **misperceive** their identity characteristics incorrectly.

D STATISTIC RESULTS FOR 4.3 AUDIENCES FOR PRESENTING OR DISCLOSING IDENTITY CHARACTERISTICS IN SOCIAL VR

Model	$\Delta\chi^2$	Δdf	*p-value*
Comfort with presenting or disclosing to Others (1-pid)			
+ Audience Type	656.2	7	<.0001
+ Identity Characteristics	206.76	15	<.0001
+ Audience Type: Identity Characteristics	59.33	47	<.001

Table 10. Linear model for effects of audience type and identity characteristics on the comfort level with presenting or disclosing to others. Each model is built upon and compared to the one listed above.

Model	$\Delta\chi^2$	Δdf	*p-value*
Frequency of presenting or disclosing to Others (1-pid)			
+ Audience Type	571.133	7	<.0001
+ Identity Characteristics	191.67	15	<.0001
+ Audience Type: Identity Characteristics	66.13	47	<.0001

Table 11. Linear model for effects of audience type and identity characteristics on the level of frequency presenting or disclosing to others. Each model is built upon and compared to the one listed above.

Identity Characteristic	Most Comfortable	*M*	*SD*	Least Comfortable	*M*	*SD*
Age	Close Friends	4.43	0.95	Strangers	3.09	1.46
Country of Origin	Close Friends	4.48	0.84	Strangers	3.62	1.33
Disability	Close Friends	3.87	1.36	Strangers	2.93	1.32
Device Used	Close Friends	4.42	0.88	Strangers	3.54	1.39
Gender Identity	Sexual or Romantic Partners	4.48	1.05	Strangers	3.68	1.37
Sub-community Membership	Sexual or Romantic Partners	3.77	1.45	Work Colleagues	2.75	1.29
Race/Ethnicity	Sexual or Romantic Partners	4.44	0.98	Strangers	3.43	1.49
Sex Assigned at Birth	Close Friends	4.61	0.74	Strangers	3.85	1.36
Sexuality	Close Friends	4.46	0.92	Strangers	3.49	1.43
Overall for Comfort	**Close Friends**	**4.44**	**0.92**	**Strangers**	**3.5**	**1.42**

Table 12. Types of audience participants feel the most and least comfortable **presenting or disclosing** their identity characteristics to.

Identity Characteristic	Most Frequently	*M*	*SD*	Least Frequently	*M*	*SD*
Age	Sexual or Romantic Partners	4.05	1.29	Strangers	2.64	1.45
Country of Origin	Sexual or Romantic Partners	4.14	1.2	Strangers	4.15	1.4
Disability	Close Friends	3.45	1.63	Strangers	2.07	0.88
Device Used	Close Friends	3.88	1.26	Strangers	3.07	1.43
Gender Identity	Close Friends	4.08	1.4	Strangers	3.29	1.47
Sub-community Membership	Sexual or Romantic Partners	3.23	1.63	Work Colleagues	2.33	1.34
Race/Ethnicity	Sexual or Romantic Partners	4.11	1.21	Strangers	2.89	1.46
Sex Assigned at Birth	Sexual or Romantic Partners	4.27	1.13	Strangers	3.5	1.41
Sexuality	Sexual or Romantic Partners	4.13	1.23	Strangers	2.89	1.48
Overall for Frequency	**Sexual or Romantic Partners**	**4.04**	**1.29**	**Strangers**	**3.04**	**1.46**

Table 13. Types of audience our participants **present or disclose** their identities to most and least frequently.

E STATISTIC RESULTS FOR 4.4.1 OVERALL ASSOCIATION BETWEEN (MIS)PERCEIVED IDENTITY REVELATION AND FREQUENCY OF HARASSMENT IN SOCIAL VR

Model	$\Delta\chi^2$	Δdf	***p-value***
Frequency of Harassment (1-pid)			
+ Identity Revelation Type	14.47	4	.0001
+ Identity Characteristics	69.12	12	<.0001
+ Identity Revelation Type: Identity Characteristics	6.12	20	.63

Table 14. Linear model for effects of identity revelation type and identity characteristic type on frequency of harassment. Each model is built upon and compared to the one listed above.

Most -> Least Freq. Harassed - Misperceived	***M***	***SD***	**Most -> Least Freq. Harassed - Presented**	***M***	***SD***	**Most -> Least Freq. Harassed - Overall**	***M***	***SD***
Disability	2.08	1.02	Country of Origin	1.84	0.99	**Sub-Community Membership**	**2.0**	**1.03**
Sub-Community Membership	2.07	1.04	Sub-Community Membership	1.81	1.01	**Disability**	**1.97**	**1.0**
Sex Assigned at Birth	2.01	1.06	Race/Ethnicity	1.79	1.0	**Race/Ethnicity**	**1.88**	**0.97**
Race/Ethnicity	1.98	0.93	Disability	1.63	0.89	**Country of Origin**	**1.85**	**0.96**
Gender Identity	1.92	0.98	Sex Assigned at Birth	1.59	0.98	**Sex Assigned at Birth**	**1.75**	**1.03**
Sexuality	1.85	0.95	Gender Identity	1.57	0.91	**Gender Identity**	**1.71**	**0.95**
Country of Origin	1.85	0.93	Age	1.54	0.82	**Sexuality**	**1.67**	**0.91**
Device Used	1.77	0.94	Sexuality	1.53	0.85	**Age**	**1.62**	**0.87**
Age	1.7	0.91	Device Used	1.38	0.76	**Device Used**	**1.54**	**0.86**

Table 15. Average frequency of harassment by identity characteristic for **Misperceived** (left column clusters), **Presented** (middle column clusters), and Misperceived & Presented combined i.e. **Overall** (right column clusters).

F STATISTIC RESULTS FOR 4.4.2 ASSOCIATION BETWEEN (MIS)PERCEIVED IDENTITY REVELATION AND FREQUENCY OF HARASSMENT IN SOCIAL VR ACROSS DIFFERENT USER GROUPS: BY GENDER IDENTITY

Model	$\Delta\chi^2$	Δdf	***p-value***
Frequency of Harassment (1-pid)			
+ Gender Identity	18.25	2	.0001
+ Identity Characteristics	73.66	8	<.0001
+ Gender Identity: Identity Characteristics	142.08	16	<.0001

Table 16. Linear model for effects of participant gender identity and identity characteristic type on frequency of harassment. Each model is built upon and compared to the one listed above.

Identity Characteristic	Most Freq. Harassed	*M*	*SD*	Middle	*M*	*SD*	Least Freq. Harassed	*M*	*SD*
Age	Women	1.98	1.07	Non-Binary/...	1.64	0.92	Men	1.48	0.73
Country of Origin	Men	1.87	0.95	Women	1.84	1.01	Non-Binary/...	1.44	0.53
Disability	Women	2.04	1.2	Men	2.0	0.89	Non-Binary/...	1.4	0.55
Device Used	Non-Binary/...	2.33	0.82	Women	1.8	1.08	Men	1.39	0.69
Gender Identity	Non-Binary/...	2.55	1.2	Women	2.24	1.09	Men	1.4	0.68
Sub-community Membership	Non-Binary/...	2.86	0.69	Women	2.0	0.91	Men	1.92	1.09
Race/Ethnicity	Non-Binary/...	2.1	0.74	Women	2.0	1.09	Men	1.81	0.9
Sex Assigned at Birth	Non-Binary/...	2.75	1.26	Women	2.26	1.25	Men	1.46	0.75
Sexuality	Non-Binary/...	2.64	1.29	Women	1.92	0.98	Men	1.47	0.75
Overall Freq. Harassment by Gender Identity	**Non-Binary/...**	**2.18**	**1.05**	**Women**	**2.0**	**1.08**	**Men**	**1.6**	**0.84**

Table 17. Average frequency of identity-based harassment (**presented & misperceived combined**) for each identity characteristic by participants' self-reported gender identity. *Note: "Nonbinary/..." in this table represents participants who self-reported as Non-Binary or Third gender or Agender.*

G STATISTIC RESULTS FOR 4.4.2 ASSOCIATION BETWEEN (MIS)PERCEIVED IDENTITY REVELATION AND FREQUENCY OF HARASSMENT IN SOCIAL VR ACROSS DIFFERENT USER GROUPS: BY SEXUALITY

Identity Characteristic	Most Freq. Harassed	*M*	*SD*	Least Freq. Harassed	*M*	*SD*
Age	Asexual	2.33	1.15	Pansexual	1.22	0.43
Country of Origin	Bisexual	2.0	0.79	Lesbian	1.33	0.49
Disability	Straight	2.08	1.02	Lesbian	1.0	0.0
Device Used	Gay	2.13	1.36	Bisexual	1.22	0.44
Gender Identity	Asexual	2.8	1.1	Straight	1.59	0.84
Sub-community Membership	Asexual	3.0	0.0	Gay	1.4	0.89
Race/Ethnicity	Gay	2.71	1.5	Bisexual	1.57	0.65
Sex Assigned at Birth	Asexual	2.33	1.15	Lesbian	1.56	1.01
Sexuality	Asexual	2.5	0.58	Straight	1.56	0.82
Overall Freq. Harassment by Sexuality	**Asexual**	**2.31**	**0.86**	**Bisexual**	**1.67**	**0.89**

Table 18. Average frequency of harassment (**presented & misperceived combined**) for each identity characteristic by participants' self-reported Sexuality.

Model	$\Delta\chi^2$	Δdf	*p-value*
Frequency of Harassment (1-pid)			
+ Sexuality	1.52	1	.217
+ Identity Characteristics	71.36	8	<.0001
+ Sexuality Stratified: Identity Characteristics	47.14.8	8	<.0001

Table 19. Linear model for effects of participant sexuality stratified and identity characteristic type on the frequency of identity-based harassment. Each model is built upon and compared to the one listed above.

H STATISTIC RESULTS FOR 4.4.2 ASSOCIATION BETWEEN (MIS)PERCEIVED IDENTITY REVELATION AND FREQUENCY OF HARASSMENT IN SOCIAL VR ACROSS DIFFERENT USER GROUPS: BY RACE

Identity Characteristic	**Most Freq. Harassed**	***M***	***SD***	**Least Freq. Harassed**	***M***	***SD***
Age	Mixed	2.36	1.12	Asian	1.44	0.73
Country of Origin	Hispanic & Latino/a	2.32	0.91	White	1.76	0.97
Disability	White	2.07	1.07	*Mixed	1.67	0.58
Device Used	Black/...	1.69	0.97	Mixed	1.45	0.69
Gender Identity	Mixed	2.2	1.23	Asian	1.43	0.53
Sub-community Membership	Mixed	2.75	0.96	Black/...	1.63	0.69
Race/Ethnicity	Asian	2.33	1.0	White	1.74	0.97
Sex Assigned at Birth	Hispanic & Latino/a	2.07	1.21	White	1.68	1.0
Sexuality	Mixed	2.5	1.2	Asian	1.0	0.0
Overall Freq. Harassment by Race/Ethnicity	**Mixed**	**2.05**	**1.0**	**White**	**1.67**	**0.94**

Table 20. Average frequency of harassment (**presented & misperceived combined**) for each identity characteristic by participants' self-reported Race/Ethnicity. *"Black/..." represents participants who self-reported their race/ethnicity as Black/African American or African. *No participants who self-reported as Asian indicated having had Disability presented or misperceived. Thus, Asian was excluded.*

Model	$\Delta\chi^2$	Δdf	***p-value***
Frequency of Harassment (1-pid)			
+ Race/Ethnicity	5.28	1	.022
+ Identity Characteristics	73.93	8	<.0001
+ Race/Ethnicity:Identity Characteristics	18.8	8	.016

Table 21. Linear model for effects of stratified participant race/ethnicity and identity characteristic type on frequency of harassment. Each model is built upon and compared to the one listed above.

Received July 2023; revised January 2024; accepted March 2024