To Reveal or Conceal:
Privacy and Marginalization in Avatars

Mattea Sim
Indiana University
matsim@iu.edu

Basia Radka
University of Washington
basia@cs.washington.edu

Emi Yoshikawa
University of Washington
emiyosh@cs.washington.edu

Franziska Roesner
University of Washington
franzi@cs.washington.edu

Kurt Hugenberg
Indiana University
khugenb@iu.edu

Tadayoshi Kohno
University of Washington
yoshi@cs.washington.edu

Abstract

The present and future transition of lives and activities into virtual worlds — worlds in which people interact using avatars — creates novel privacy challenges and opportunities. Avatars present an opportunity for people to control the way they are represented to other users and the information shared or implied by that representation. Importantly, users with marginalized identities may have a unique set of concerns when choosing what information about themselves (and their identities) to conceal or expose in an avatar. We present a theoretical basis, supported by two empirical studies, to understand how marginalization impacts the ways in which people create avatars and perceive others' avatars: what information do people choose to reveal or conceal, and how do others react to these choices? In Study 1, participants from historically marginalized backgrounds felt more concerned about being devalued based on their identities in virtual worlds, which related to a lower desire to reveal their identities in an avatar, compared to non-marginalized participants. However, in Study 2 participants were often uncomfortable with others changing visible characteristics in an avatar, weighing concerns about others' anonymity with possible threats to their own safety and security online. Our findings demonstrate asymmetries in what information people prefer the self vs. others to reveal in their online representations: participants want privacy for themselves but to feel informed about others. Although avatars allow people to choose what information to reveal about themselves, people from marginalized backgrounds may still face backlash for concealing components of their identities to avoid harm.

Keywords

avatars, identity, marginalization, privacy, psychology

1 Introduction

A classic question in privacy is: what information about oneself does one conceal and what information does one expose, and to whom? This question is certainly applicable to the context of avatars. People have occupied avatars, or graphical representations of users, for decades. Whereas the earliest avatars were video game characters and other entities largely distinct from the user, more recent advances have seen avatars move closer to the users themselves, with countless customization options to better resemble the user [15], and mixed and virtual reality (i.e., VR) devices with sensors that can scan a user's face to create a hyper-realistic avatar [43], even mirroring users' real movements and expressions.

As avatar verisimilitude advances, and avatar usage becomes more ubiquitous, it is imperative to ask: what information about a user should (or should not) be revealed through their avatar? Should users be able to conceal and keep private parts of their visible identity? Should users be able to falsify their visible identity by adopting new realistic identities in an avatar? What are the implications to others if a user keeps part of their identity private? Avatars can afford users a great deal of control over their online privacy, with flexible options for representing social identities that are not available to most in the real world. For instance, an avatar can display a racial identity entirely different from its user's real racial identity. However, such anonymity necessarily affords opportunities for both benefits and threats — in the previous example, either identity protection or appropriation.[1] Conversely, hyper-realistic avatars often afford less privacy and anonymity, automatically revealing their user's "visible" social identities, a strategy which also affords both benefits (to interaction partners who want to feel informed) and threats (for protected identities being "outed").

We explore this tension from the perspective of marginalization. Users from marginalized backgrounds (i.e., groups of people who are systemically excluded from mainstream social, economic, and/or cultural life) may have unique privacy and safety concerns in the context of avatar representations. Indeed, in the real world, people with marginalized identities often face choices to protect or reveal components of their identities to navigate potentially stigmatizing and marginalizing experiences [11, 14, 26, 41, 56]. Unfortunately, virtual worlds are not free from these experiences. Many users from marginalized backgrounds face continued harassment and abuse on avatar platforms or social VR (i.e., virtual spaces where multiple users interact) [2, 10, 29, 46, 47, 59]. As such, marginalized users may navigate privacy in virtual contexts by choosing whether they want to *reveal* or *conceal* their marginalized identities in an avatar. In the current work, we explore how experiences with and concerns about marginalization impact decisions about privacy via avatars, and the implications of such choices for other users. Importantly, we conceptualize experiences of marginalization as including both prior

---

[1]In this paper, we use the term "anonymous" to refer to a user who intentionally conceals parts of their identities and/or adopts new identities in an avatar. An alternate term, though of less colloquial use, is "pseudonymous" [40].

experiences with marginalization *and* concerns about experiencing marginalization in the future, because the awareness that one's social group is marginalized (regardless of personal experiences) can trigger the same concerns and self-protective strategies [52, 53].

We studied U.S. participants who self-identified as either marginalized for one or more of their visible identities or *not* marginalized. Across two online survey studies, we compared marginalized versus non-marginalized participants' reported preferences on what to reveal or conceal in their own avatars (Study 1) and their perceptions of how *others* conceal identities in avatars (Study 2). We explore how marginalization concerns impact people's privacy-related intentions and expectations for their own and others' avatars. This project is a collaboration between two social psychologists (M.S. and K.H.) and four computer security and privacy researchers (B.R., E.Y., F.R., and T.K.). We leverage psychological theories of marginalization and identity to provide a basis by which to understand how people navigate novel choices about identity protection online. We seek to answer the following research questions:

(1) **RQ1:** How do people with marginalized identities, compared to those without marginalized identities, navigate *revealing* or *concealing* visible identities in an avatar?
(2) **RQ2:** Are there asymmetries between how people want to conceal or reveal parts of *their* marginalized identities through avatars and how people react to the ways *others* conceal or reveal marginalized identities through avatars?

The present work is important for several reasons. First, to foreshadow our findings, people with marginalized identities are, on average, more likely than non-marginalized people to want to conceal their identities in an avatar. However, people often feel uncomfortable with others concealing or adopting identities in their avatar. This asymmetry between wanting privacy for the self versus wanting to feel informed about others via avatars highlights important tensions that must be considered when designing avatar-based platforms. We set up these tensions in the following threat modeling section, discussing asymmetrical security and privacy needs for the self versus others in virtual worlds. Second, whereas much of the work on avatars and identity expression has investigated how a particular group navigates avatar creation and interaction, our work adopts a broader perspective to understand how marginalization can surface unique privacy and protection needs for avatars across a range of identities. This wide-scale survey complements approaches in prior work, identifying broad relationships between identity-related concerns and privacy-related concerns online.

2 Threat Model

A classic dilemma in computer security and privacy is that anonymity can come with both costs and benefits. Some users may choose anonymity (concealing identities or adopting new identities) to *avoid* security threats, whereas other users may choose anonymity to enact security threats. Users may also be unaware of others' motives for anonymity, which can surface its own security and privacy concerns. We thus highlight several security and privacy practices and threats from the perspective of both a user creating an avatar and a user interacting with other avatars. Although the same people represent both users, we expect that people may have asymmetrical privacy needs for the self versus others.

2.1 Users Creating Avatars

Users may choose to conceal identities or adopt new identities in an avatar as a means to protect the self from security threats. One way to maintain privacy in an avatar is to adopt unrealistic features, such as purple eyes, or non-human characters, such as animals or monsters. However, assuming a system allows avatars to be human-like, and assuming some users will want to conceal their identities from other users and the system itself, the resulting system must allow a user to create a realistic avatar that looks like a person but that is different from themselves. In this context, users may "customize" real social groups in their avatar to maintain privacy. For instance, people may adopt different identities to avoid harassment and abuse. This may be especially prominent amongst marginalized users concealing parts of their identities that are regularly targeted with such harassment, such as women disguising their gender to avoid excessive attention and harassment from other users. Users may also seek to avoid stalking or exploitation.

Unique threats may surface when people use avatar anonymity for more nefarious purposes. For instance, non-marginalized users could adopt marginalized identities in an avatar to engage in cultural appropriation or to spread counter-productive messages under the guise of belonging to a certain group. An example is a white user adopting a non-white avatar to claim that racism is not a problem. Some users may also adopt different identities to engage in predatory behavior, such as grooming.

2.2 Users Interacting with Others' Avatars

In avatar-based interactions, users may want to feel informed about others' identities to ensure their own security and privacy. For instance, a user may feel greater trust with someone who shares a social group identity, perhaps impacting the extent to which they reveal their own identity or identity-related experiences. Related to the threats outlined above, users may also want to be informed about others' identities to identify cultural appropriation or predatory behavior. In this way, avatar anonymity may make it more difficult for users to discern others' motives and intentions, and thus to feel safe. Users may have to consider whether others are concealing or adopting identities, and if yes, whether they are doing so for fun, for protection, or for harm. A user's own security and privacy practices with avatars may depend, in part, on these appraisals of others.

However, revealing identity information can also surface a variety of threats. Users who seek to protect their identities (e.g., to avoid marginalization) may be unintentionally "outed." For instance, someone exploring their gender identity in an avatar revealed as appearing to be a different gender in real life (regardless of how they actually identify) may experience abuse and/or discomfort. As another example, women, who are often the numerical minority in social VR environments, may be subject to excessive attention and harassment when forced to reveal their identities [30].

3 Background and Related Work

3.1 Theories of Identity and Marginalization

Social identities likely play an important role in virtual representations, which may mirror psychological experiences in the real

world. The social groups that we belong to (such as race, gender, age, social class, etc.) are part of what makes us who we are; these *social identities* contribute to our sense of self [28, 55]. Because social identities are often important to our self-concept, the knowledge that one's identities may be marginalized can be threatening to the self — an experience termed *identity threat* [53]. Identity threat can lead to a host of negative responses such as anxiety, vigilance, impaired working memory, and impaired task performance [13, 28, 45, 51]. People from historically marginalized backgrounds facing identity threat often employ various coping methods to protect the self from this threat, such as selectively valuing attributes depending on whether their group is positively or negatively stereotyped for that attribute [6] or distancing from the marginalized group [56].

People sometimes try to avoid devaluation or harm by *concealing* their marginalized identity [11, 14]. In the real world, concealing social group membership is easier for some than for others. Classic theories of stigma made distinctions between so-called "conceal-able" identities, such as sexual orientation, and "visible" identities, such as race [39, 41]. Importantly, no social identity falls strictly under one category. For instance, although sexual orientation was traditionally considered a concealable identity, many individuals experience this identity as visible; concealability is based on both the identity and individuals' subjective experiences [19]. Choosing not to reveal a concealable identity can be a self-protective strategy to avoid stereotyping, discrimination, and other harms [11, 41, 56]. For instance, in regions of the U.S. where non-religious identity (e.g., being atheist) is seen as more stigmatized, non-religious individuals are less likely to publicly identify as atheist [26].

However, people with "visible" marginalized identities do not typically have the same option to conceal their identity in the real world. Individuals with visible marginalized identities sometimes utilize other (often more temporary) forms of concealment, such as removing cues that might signal group membership on a resume [14]. However, avatars allow people to create representations of themselves that make traditionally visible identities into concealable ones, surfacing novel choices about how to participate online while also protecting oneself from marginalizing experiences. Motivating these studies, we hypothesize that people with marginalized identities may be more likely than non-marginalized people to want to conceal their identities via avatars. Conversely, perhaps non-marginalized people also want to conceal aspects of their identities, but for different reasons (e.g., for fun rather than for privacy and safety). Further, while marginalized people still face harassment and abuse for revealing their identities, they may also face interpersonal consequences for concealing identities if this anonymity surfaces safety concerns for others.

3.2 Identity Representation in Avatars

A robust literature is dedicated to investigating how users represent themselves through avatars. In fact, avatar creation and interaction involves a variety of psychological motives. For instance, users choose avatars to align with their actual and ideal selves [61], to adapt to the social context or for virtual exploration [23], and to fulfill psychological needs [8]. Because avatars can reveal a great deal of information about their users, researchers have also explored how users navigate tradeoffs between privacy and self-disclosure

via avatars [24, 54, 58]. Revealing personal information through avatars is seen as a means to connect with others, and can foster perceptions of authenticity in others as well [54]. When navigating these tradeoffs, users consider a variety of contextual factors like who they are interacting with, where they are, and what type of personal information to share [54, 58]. For instance, a person might share different information with close friends at an intimate party vs. with colleagues at a business meeting. In our studies, we explore avatar representations assuming users have one avatar to use across a variety of contexts, although we acknowledge motivations can differ across contexts, a point to which we return in the discussion.

Although the perspectives of people from marginalized backgrounds have not always been forefront to this research, the field of human-computer interaction broadly has begun to critically consider both conducting research with marginalized populations and designing inclusive interfaces [22, 49]. As such, a growing body of work centers users from marginalized groups and their unique experiences with avatars. For instance, some non-cisgender users report struggling to accurately represent themselves on platforms that support only binary and stereotypical gender options. However, some non-cisgender users have also used avatars to affirm their gender identities [9]. Users with disabilities similarly report a lack of inclusive customization options, and also adopt a range of self-representational strategies via avatars [25, 38, 60]. Some evidence within this body of work suggests users may also conceal identities or adopt higher-status identities in avatars to avoid harassment. For instance, women have reported disguising their gender identity to avoid excessive attention and harassment [30, 46]. Further, Black users who were numerically underrepresented in an online space created more white-prototypical avatars [20]. Users with disabilities also reported creating avatars without visible disabilities to avoid mistreatment [25].

Whereas concealing identities may be a self-protective strategy for some, other users may adopt identities in an avatar for more nefarious purposes. Predators have pretended to be younger and/or a different gender online to abuse children [34] (although in this case, not via avatars). Avatars can also introduce new forms of digital blackface, wherein Black avatars, created by and adopted by non-Black people, can reinforce harmful stereotypes and commodify Black bodies [33]. Indeed, some white users appropriate non-white characters in games as a form of "identity tourism" [36]. Thus, we also explore how participants react to non-marginalized users adopting marginalized identities to understand how other users' avatars affect safety concerns online.

Notably, much of the past research on identity and avatar representation used interviews or observational studies to deeply explore one group's experience within a specific framework or context. In the current work, we seek to build upon this prior knowledge by exploring how users from a variety of backgrounds might experience identity-related concerns, impacting both privacy intentions for the self and safety concerns about others. In other words, we theorize about the commonalities between groups by which concerns about marginalization might impact needs for privacy and protections in avatars, while acknowledging that different groups also have unique experiences with, and require unique design solutions for, inequality online.

4 Methodology

Across two studies, we used a combination of previously validated scales and novel scales designed to answer our research questions. Both surveys were piloted by the authors and external colleagues.

4.1 Ethics

Study materials were submitted to the university's Human Subjects Review Board (IRB), who deemed this research exempt because it poses no more than minimal risk to participants and meets a variety of other requirements. We strictly followed ethical procedures and guidelines in our studies. All participants voluntarily chose to take part in our studies. Participants were kept anonymous. Participants could skip questions they were uncomfortable answering and could leave the survey at any time.

4.2 Study 1 Procedure

Study 1 was designed to answer RQ1: how do people with marginalized identities, compared to those without marginalized identities, navigate the choice to reveal or conceal their identities in an avatar? Full instructions and measures for both studies are available in Appendix A and Appendix B. After consenting to participate, participants read definitions of social identities and marginalization before identifying whether they had experienced or believed they might experience marginalization based on one or more of their social identities that are visible to other people, with binary-choice options of "Yes" or "No." We include both prior and anticipated experiences with marginalization because both can elicit identity threat and subsequent self-protective strategies [52, 53]. These questions allow us to understand both whether people have experienced marginalization themselves and whether they belong to groups who have historically experienced such marginalization.[2] On the same page, we asked participants who had answered "Yes" to indicate the social category or categories for which they have experienced marginalization from a provided list. See Table 1 for the percent of marginalized participants (i.e., those who selected "Yes") who are marginalized for each identity. Participants could check as many options as were applicable and could choose to provide their own free response. This question and the response options allowed participants to clarify their own experiences of visibility with their identities [19]. These questions also categorized participants into what we refer to as self-identified marginalized or non-marginalized

---

[2]We chose to define marginalization in terms of people's psychological experiences (prior or anticipated), rather than sorting participants ourselves based on their identification into specific historically marginalized groups [35]. This allows us to include a wider range of participants and experiences in our studies, and explore how the psychological experience of marginalization impacts motivations online — and indeed, these psychological experiences are impactful for identity threat-related outcomes [35]. However, people who do not belong to historically marginalized groups could choose to identify themselves as "marginalized," for instance if white participants believe they are discriminated against for their race. We found that a small subset of participants (13 in Study 1, 10 in Study 2) identified themselves as marginalized for their race, and also identified themselves as monoracial white. However, we are cautious to over-interpret these data, as we acknowledge that some groups (e.g., people who are Middle Eastern, North African, or Hispanic American) have historically needed to identify themselves as "white" or "other" on racial demographic U.S. Census questions, yet may perceive themselves (and be perceived by others) as people of color (other groups, such as Jewish people, may also identify themselves as both white and marginalized for their race). We include participants in analyses regardless of how they identify on demographic questions, but acknowledge the possibility and limitation that a small number of participants may perceive themselves as marginalized despite not belonging to historically marginalized groups.

Table 1: Percent of self-identified marginalized participants in Studies 1 and 2 who experience or believe they may experience marginalization for each identity.

	Study 1	Study 2
Race	43.7%	45.5%
Ethnicity	31.0%	24.8%
Sex	44.2%	46.5%
Gender identity	14.2%	17.8%
Age	17.3%	21.8%
Disability	6.6%	9.9%
Body weight, shape, or size	34.5%	38.6%
Height	17.3%	9.9%
Pregnancy status	0.0%	3.0%
Religion	15.7%	13.9%
Sexual orientation	16.2%	17.8%
Social class	21.3%	24.8%
Another identity(ies) not listed	0.0%	4.0%

Note. Percentages are calculated from the number of participants who selected "yes" to the self-identified marginalization question.

groups, allowing us to understand how experiences and concerns about marginalization broadly impact avatar representation, with the ability to conduct more granular group-level analyses as well.

Participants read that we were interested in how people might represent themselves through avatars if they had to use a single avatar across several contexts. Although users in some online communities normatively represent themselves through fantastical or otherwise unrealistic avatars, our studies were intended to explore how people might create avatars to be used in a variety of increasingly virtual contexts, where more humanlike or realistic avatars may become the norm. Participants were also told their avatar would represent themselves to other users they interact with online, specifying a more socially interactive context (e.g., work, school, socializing).

We first sought to understand participants' identity-related concerns in the virtual world. Participants responded to a validated 6-item identity threat scale adapted from social psychology research [17] assessing the degree to which people worry about being devalued based on their identities in the virtual world. For these questions, participants were asked to imagine their avatar was identical to their actual appearance.

Participants then reported the degree to which they wanted their avatar to match their real appearance across several items. Participants responded to 6 items measuring their overall desire to reveal their identities in an avatar, with items adapted from previous research [20, 23]. Next, participants rated the extent to which they wanted the appearance of their avatar to perfectly match their actual self along 12 specific characteristics on a slider scale from 1 (Not at all like my actual self) to 100 (Exactly like my actual self). We used slider scales (common in psychological research) because we believed a series of 12 continuous slider scales may be quicker and easier for participants than 12 Likert scales [57]. We did not observe any statistical implications of this choice, consistent with prior research showing comparable means across Likert and slider

scales [21]. Further, the 12 specific characteristics were not meant as an exhaustive list and were supplemented by two open-ended questions allowing participants to explain more about how they want their avatar to look 1) different from themselves and 2) similar to themselves, and why.

Finally, participants responded to questions assessing their prior experience with avatars and concerns about how accurately they could represent themselves in an avatar.[3] Participants reported demographic information and completed the study. Participants took an average of 8.43 minutes (SD = 3.52) to complete the survey.

4.3 Study 2 Procedure

Study 2 was designed to answer RQ2 in tandem with Study 1: are there (a)symmetries between how people want to conceal or reveal their own identities in an avatar and how people react to the ways *others'* conceal or reveal identities in an avatar? Here again, participants were asked to imagine users had a single avatar to use across a variety of contexts. Participants imagined they were getting to know someone in a virtual world where they were both represented by avatars.

We manipulated on a within-subjects basis whether this person was *concealing* a marginalized identity or *adopting* a marginalized identity in their avatar, because we believed this distinction may affect perceivers' comfort and because both are likely to occur [33, 34]. The viewing order of the two conditions was randomized across participants. Both conditions began by describing that the person you've been getting to know online has an avatar with identities that are different from their identities in the real world. In the conceal condition, participants read that their real identity is typically marginalized in society, but their avatar has an identity that is different from their marginalized identity. In the adopt condition, participants read that their real identity is not marginalized, but their avatar has an identity that is typically marginalized in society. Participants were not told which specific identities they were concealing or adopting.

In each condition, participants rated how uncomfortable they would be if this person's avatar looked different from their actual self across 12 characteristics on a slider scale from 0 (Not at all uncomfortable) to 100 (Very uncomfortable). This question assessed the same characteristics as in Study 1 to explore (a)symmetries in the characteristics people want to reveal and the characteristics people prefer others to reveal. Participants also responded to two open-ended questions considering the characteristics they were most 1) uncomfortable and 2) comfortable with others changing in an avatar, and discussed what they think someone's motives would be for changing these characteristics. Participants responded to all questions in a given condition before moving on to the next condition.

Finally, participants reported how many avatars they had created in the past and self-identified as marginalized or not using the same question from Study 1. Participants reported demographic information and completed the study. Participants took an average of 8.92 minutes (SD = 3.91) to complete the survey.

---

[3]Consistent with past research, marginalized participants (compared to non-marginalized participants) were more worried that avatar customization tools would not allow them to create representative avatars, and reported more negative past experiences with avatar customization tools, $ps < .030$, $ds > 0.26$.

4.4 Participants

We conducted an a priori power analysis using G*Power to determine how many participants were needed to detect a moderate effect size in both Studies 1 and 2. This analysis determined that 200 participants were needed to detect an effect size of d=0.40 at 80% power in an independent-samples t-test. We thus aimed to collect at least 100 participants who identified themselves as marginalized (across any combination of identities, not within a specific identity) and at least 100 participants who identified themselves as non-marginalized. See Table 2 for participant demographics in Studies 1 and 2.

In Study 1, participants (N = 301) were U.S. undergraduates who completed the survey online in exchange for course credit. At one of our institutions, all students in lower-level psychology courses can optionally participate in research or complete other research assignments (e.g., reading research papers) to gain educational exposure to research and earn course credit. Thus, participants in Study 1 chose to take part in this study for course credit (but were not required to do so), a practice common in psychology and approved by the institution's IRB. None of the authors had an instructional role in these courses, and students self-selected from a list of available studies posted to an online research portal for university students. 197 participants identified as marginalized, and 104 participants identified as not marginalized. In Study 1, all participants were included in analyses because no responses indicated inattention.

In Study 2, participants (N = 218) were U.S. crowdworkers who chose to take part in the survey advertised online via Prolific. Each worker was paid $1.87 to participate in a 6-minute survey, with compensation set to match the highest minimum wage amongst the authors' respective cities. We conservatively excluded 24 participants from analyses who had at least one response indicating inattention (coded as "N/A" by two researchers), including participants whose responses indicated they did not appropriately differentiate between the conditions, resulting in 194 participants included in analyses. 101 participants identified as marginalized, 89 identified as not marginalized, and 4 participants did not respond to this question. We concluded that 89 participants identifying as non-marginalized was sufficiently close to our original goal of 100 participants in this group.

4.5 Qualitative Analysis

We adopted an inductive approach to thematic coding [27] to code participants' open-ended responses to six questions (two in Study 1, four in Study 2). Three principal investigators evaluated the codes and themes throughout the process and were involved in iterative refinement of codes and organization. The primary investigator first read through all responses to become familiar with the data, taking notes on each conceptually distinct response. Next, the investigator organized these responses into meaningful groups, and used this organization to identify broader themes by which to understand the responses. Two coders independently read all responses and recorded the frequency with which each theme emerged across participants, using this step to review and iterate on the themes in the context of the entire data set. A given participant's response could be classified into any number of themes. As such, we analyze

Table 2: Breakdown of participant demographics in Studies 1 and 2 by gender, age, and race.

Gender	Study 1	Study 2	Age	Study 1	Study 2	Race	Study 1	Study 2
Woman	63.1%	45.4%	18-24	97.7%	17.5%	White	74.1%	75.3%
Man	32.6%	49.0%	25-34	0.3%	32.1%	Black/African American	8.3%	14.9%
Non-binary	2.3%	2.1%	35-44		22.2%	Multiple identities selected	8.3%	7.2%
Multiple identities selected	1%	2.1%	45-54		17.5%	East Asian	7.6%	6.2%
Genderqueer	0.3%	1.5%	55-64		7.2%	South Asian	9.6%	2.6%
Agender	0.7%	0.5%	65+		2.1%	Identity not listed	4.7%	2.6%
Identity not listed/Self-described	1.0%					American Indian/Alaskan Native	1.3%	3.1%
						Native Hawaiian/Pacific Islander	1.7%	1.0%

Note. Percentages can sum to greater than 100 because some participants have multiple gender or racial identities. We also include the percentage of participants who chose multiple identities.

the frequency with which each theme emerged within marginalized and non-marginalized participant groups. The two coders met to discuss and resolve code disagreements. We calculated interrater reliability across studies (number of agreed cells/total number of cells): Study 1 IRR=.98; Study 2 IRR=.95. Finally, we summarized the nature of responses within each theme via brief definitions, reviewed and refined by all investigators, and resulting in our final codebook. The full codebook of themes and descriptions are available in the online supplemental materials: https://osf.io/5u6eg/.

Participants whose responses were unintelligible, or indicated they were not paying attention or did not understand the instructions, were categorized as "N/A". We excluded participants with responses coded as "N/A" from data analyses (Study 1=0, Study 2=24).

5   Study 1 Results

5.1   People with Marginalized Identities are Less Willing to Reveal their Identities in Avatars than Non-Marginalized People

We sought to answer RQ1, comparing how people with marginalized identities (vs. non-marginalized people) navigate the choice to reveal or conceal their identities in an avatar, and if this choice is associated with identity threat concerns. We created a composite identity threat scale from the 6 identity threat items (α = .90) and a composite willingness to reveal identities scale from the 6 reveal-willingness items (α = .78), and conducted independent samples t-tests comparing across marginalized (n = 197) and non-marginalized (n = 104) participants on each scale.

Marginalized participants reported feeling significantly more identity threat, or concerns about being devalued based on their identities in the virtual world (M = 2.49, SD = 0.97), than did non-marginalized participants (M = 1.97, SD = 0.84), $t(299)$ = 4.638, $p < .001$, 95% CI [0.32, 0.80], d = 0.56. Marginalized participants were also significantly less willing to reveal their identities through an avatar (M = 3.41, SD = 0.79) compared to non-marginalized participants (M = 3.64, SD = 0.78), $t(299)$ = -2.377, p = .018, 95% CI [-0.53, -0.05], d = -0.29. Identity threat and willingness to reveal identities were significantly correlated, both overall, r = -.375, $p < .001$, and amongst marginalized participants exclusively, r = -.409, $p < .001$. Put simply, the more participants (and especially marginalized participants) were concerned about being devalued in the virtual world

Figure 1: Study 1 participants' average desire for their avatar to match their actual self on each characteristic compared to Study 2 participants' average discomfort with other users' avatars looking different from their actual selves on each characteristic. Pregnancy status not shown because of low response rate to this question in Study 1.

based on their identities, the less they wanted to reveal their identities in an avatar.

5.2   Identity Can Affect Expression of Specific Identity-Signaling Characteristics

We explored identity representation and privacy more deeply through participants' intentions to reveal more specific identity-signaling characteristics. Figure 1 shows the degree to which participants wanted their avatar to perfectly match their actual selves across different characteristics. Participants on average wanted to more accurately match characteristics such as sex, race, and age, whereas they wanted relatively less accurate matching for body-related characteristics such as shape, size, and weight.

It is likely that the desire to reveal or conceal specific characteristics depends on the identities for which people have experienced or believe they may experience marginalization. Thus, we conducted a series of comparisons focusing on specific identities and characteristics that may signal these identities as outcome variables. We

chose three identities to investigate more closely: sex, race, and body weight/shape/size, because they were the most commonly selected identities for which our participants identified themselves as marginalized. Notably, people at the intersection of different identities (e.g., race and sex) likely have distinct experiences and responses, however, we were underpowered to test for intersectional effects. We report all comparisons conducted in this exploration.

Looking first at sex, we compared participants who identified themselves as marginalized for their sex (n = 87) to those who did not identify themselves as marginalized for their sex (n = 214).[4] These groups did not significantly differ in their desire to have their avatar's sex perfectly match their actual sex, $t(298) = 0.608$, $p = .544$, 95% CI [-0.17, 0.33], $d = 0.08$.

Next, we compared participants identified as marginalized (n = 86) or not (n = 215) for their race. Participants marginalized for their race were descriptively (but not statistically significantly) less likely to want their avatar's race to perfectly match their actual race (M = 89.06, SD = 21.30) compared to those not marginalized for their race (M = 93.14, SD = 17.45), $t(299) = -1.718$, $p = .087$, 95% CI [-0.47, 0.03], $d = -0.22$. There were no significant differences across groups in the desire for their avatar to perfectly match their actual facial features or skin tone, $ps > .350$, $ds < 0.15$. However, participants who were marginalized for their race were significantly less likely to want their avatar to perfectly match their actual hairstyle (M = 73.07, SD = 29.03) compared to those not marginalized for their race (M = 79.76, SD = 25.13), $t(298) = -1.992$, $p = .047$, 95% CI [-0.51, -0.003], $d = -0.25$.

Finally, we compared across those identified as marginalized (n = 68) or not (n = 231) for their body. Participants marginalized for their body were significantly less likely to want their avatar to perfectly resemble their body shape, body size, *and* body weight, compared to participants not marginalized for their body $ps < .015$, $ds > 0.35$. Thus, people experiencing or concerned about body-related marginalization report a lower desire to accurately represent their body compared to those who are not marginalized for their body. These effects were consistent and moderate-to-large in magnitude, suggesting that privacy related to body attributes in avatars may be an important area for future work.

5.3 People Have Unique Motives for Identity Expression

We also explored themes that emerged in open-ended responses. We identified 16 themes describing how and why participants wanted their avatar to look different from themselves (i.e., concealing; see Figure 2), and 19 themes describing how and why participants wanted their avatar to look similar to themselves (i.e., revealing; see Figure 3).

Amongst the "concealing" themes, participants most frequently listed specific features they might conceal, often without elaborating on specific motives. In contrast, some participants had no desire

---

[4]We distinguish between sex assigned at birth and gender identity in this paper, and discuss both in ways that are consistent with participants' responses. The majority of participants identified themselves as marginalized for their sex assigned at birth, more so than for their gender identity. We thus use the term "sex" here and in other places to be consistent with the option selected by participants. Indeed, people can experience marginalization for both their sex and gender identity, and we allow participants to make distinctions appropriate for their own identities.

Figure 2: Study 1 concealing themes and percent of marginalized and non-marginalized participants whose responses were classified into each theme.
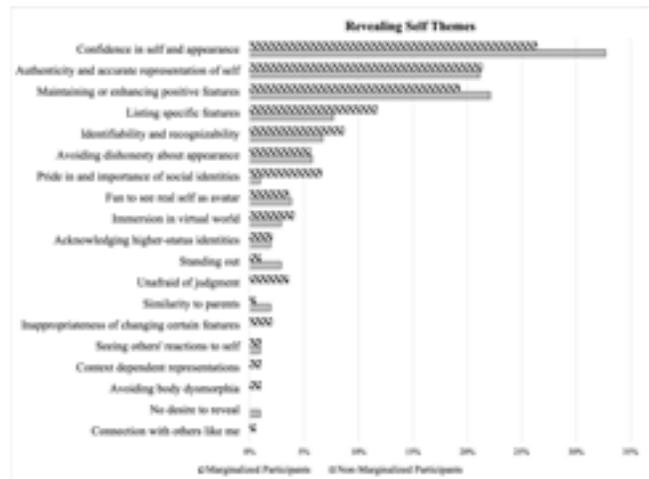

Figure 3: Study 1 revealing themes and percent of marginalized and non-marginalized participants whose responses were classified into each theme.

for privacy via identity concealment. 13% of non-marginalized participants and 8% of marginalized participants discussed having no desire to conceal their identities in their avatar. Relatedly, some participants discussed changing identities in an avatar to explore creative self-expressions and outlandish characters, rather than to maintain privacy.

> "I do not see why a virtually created avatar should be limited to the visual aspect that exist in real life. If I am constructing a virtual identity for myself, then I would want it to represent my personality (instead of just my physical disposition due to my genetics). Also, it might be fun to have other types of lifeforms such

as Vampires, Giants, Hobbits, etc. instead of just limiting our appearance to humans." *Non-Marginalized (Study 1)*

Participants also discussed a desire to conceal insecurities in their avatars. Two unique dimensions related to concealing insecurities were surfaced as their own themes: representing ideal selves and goals, and body image and body dysmorphia. In the former theme, avatars allow users to create an ideal version of themselves that highlights changes they are striving for in the real world. In the latter theme, many participants discussed body-related insecurities specifically, with some also mentioning the prevalence of thin or idealized avatar shapes.

"I might make my avatar look more idealized in terms of body shape and size because it is something I am sometimes insecure about. I also typically think that avatars will inevitably look more idealized." *Non-Marginalized (Study 1)*

Importantly, some marginalized participants also discussed concealing characteristics to avoid experiences of marginalization. Quotes below demonstrate how some participants choose to participate online while protecting themselves from marginalizing experiences.

"Because of the discrimination, micro aggressive behaviors, and racism I've experienced before, I would not want my avatar to look like me. Instead, I would give it features that are more accepted within today's society (European features)." *Marginalized for race and ethnicity (Study 1)*

"I would make my avatar look different from myself because I feel like thinner body shapes are appreciated and respected more which would make communicating on it easier." *Marginalized for sex, body, height, and sexual orientation (Study 1)*

Some participants also discussed concealing identities due to general privacy concerns.

"I just generally don't want people to know what I look like online. The reason I play video games or things of that nature is because it's supposed to be an escape from the real world. If I carry real-world issues or parts of me into a virtual world, they can be exploited there just like they are here." *Marginalized for race, ethnicity, sex, gender identity, and sexual orientation (Study 1)*

Amongst the "revealing" themes, participants discussed a variety of motivations or dimensions for which they did not want privacy protections in an avatar. Some participants discussed maintaining or enhancing their best features, or listed features they wanted to reveal without discussing underlying motivations.

A subset of revealing themes emerged relating to confidence, authenticity, and social identity. In some cases, the occurrence of these themes differed descriptively across participants' identities. For instance, some participants want to represent their real appearance in avatars because they are confident and self-assured, the most common theme discussed by 33% of non-marginalized participants and 26% of marginalized participants.

Marginalized participants in particular discussed confidence related to their identities. In this unique theme, participants want to reveal social identities they are proud of and that are important to their sense of self, as illustrated in the quotes below.

"Race, gender, facial features, etc. are parts of my identity I cannot and would not erase. I am proud of them." *Marginalized for race, ethnicity, sex, religion, and sexual orientation (Study 1)*

"I like my identity as a feminine woman and I would want my avatar to look like me because I feel intertwined with my soul and body and I think it expresses my personality." *Marginalized for sex (Study 1)*

Whereas some may choose to conceal marginalized identities to protect against devaluation and harm, others wanted to reveal these same identities as a point of pride. Marginalized users may feel tension between these competing motivations while customizing an avatar. Indeed, concealing identities that are important to our sense of self can trigger a host of negative emotions [1, 37], yet the privacy tradeoff of revealing identities could also be anxiety-provoking.

Participants also discussed avoiding appearing dishonest online, with some concerns that identity-related privacy in an avatar may be construed as lying or manipulation. A desire for authenticity emerged as another related theme. In one example, such themes were coupled with representing marginalized social identities as a means to connect with others like themselves online, while also recognizing complex social dynamics with representing more privileged intersectional identities.

"I would like to represent my sexuality and gender identity through my style in subtle ways to find people like me. I am gender nonconforming and would like to represent that through my avatar because it will help me to feel more authentic and less dysphoric. I would leave my skin tone and race the same because I do not face discrimination for my race as a white person and I feel that representing my race accurately helps others to better understand what social dynamics may be at play." *Marginalized for sex, gender identity, body, and sexual orientation (Study 1)*

As in the example above, many participants considered how others would react to their avatars, often coming to unique and nuanced conclusions about tradeoffs between privacy, authenticity, and marginalizing experiences. Whereas some highlighted authenticity and accuracy, others highlighted privacy and protection. Next, we explore the interpersonal privacy and security implications of concealing and adopting identities in an avatar.

6 Study 2 Results

In Study 2, we explore RQ2, or whether perceptions of others' avatars align with or diverge from intentions for one's own avatar.

6.1 Symmetries between What People Want to Reveal and What People Want Others to Reveal

We first explored participants' comfort level with other users' concealing or adopting new identity characteristics, descriptively comparing these reactions with the characteristics a different set of participants wanted to reveal via avatars in Study 1. We make only descriptive comparisons because Studies 1 and 2 were comprised of two separate samples of participants.

Figure 1 shows participants' comfort level with other users changing specific characteristics in an avatar, collapsed across the conceal and adopt conditions. At a broad level, there are clear *symmetries* between what people want to reveal and their comfort with what others reveal. Participants were most uncomfortable overall with other users changing their race, sex, or age (although comfort can depend on whether someone is concealing or adopting marginalized identities, as we explore more in the next section); in Study 1, these were the characteristics that participants on average most wanted to match their actual identities. Participants were also relatively more comfortable with other users changing their body shape, size, and weight, which were the characteristics participants in Study 1 wanted to match their actual identities the least.

At a broad level, some participants also discussed feeling comfortable with other users changing their identities in an avatar, often regardless of the motive. These participants did not voice security concerns in relation to others' avatars, often because they do not view avatars as accurate representations of users.

> "Again, unless I'm potentially involved with this person on an intimate level, what I think regarding their avatar and their actual appearance doesn't matter. And even if I have an intimate relationship with this person, whose characteristics may not match what they chose for an avatar, it doesn't matter a lot to me and I don't see a reason why I'd be uncomfortable or comfortable with any choices they may have made for a fictitious representation of themselves." *Marginalized for sex, age, and social class (Study 2)*

> "I would be completely comfortable because online is not reality so a person should be able to choose at their discretion who they want to be online." *Non-Marginalized (Study 2)*

6.2 Asymmetries Arise When Exploring Other Users' Identities and Participants' Own Identities

Distinguishing between other users adopting vs. concealing marginalized identities, and between participants' own marginalized vs. non-marginalized status, reveals more nuanced perceptions of avatars. First, we expected that participants' comfort may depend on whether other users were concealing or adopting marginalized identities. Indeed, participants were more uncomfortable overall with others adopting a marginalized identity, as opposed to concealing a marginalized identity, as indicated by a significant main effect of condition, $F(1,188) = 23.651$, $p < .001$, $n_p^2 = .112$. Other users adopting marginalized identities may raise more apparent privacy and security-related concerns.

We also compared between marginalized and non-marginalized participants. Here, one hypothesis is that because marginalized participants in Study 1 reported a lower willingness to reveal their identities in avatars (a tendency related to identity threat), perhaps a new sample of marginalized participants in Study 2 will be more comfortable with other marginalized users choosing to conceal their identities as well. There was a significant interaction between participants' marginalized status and the condition, $F(1,188) = 5.257$, $p = .023$, $n_p^2 = .027$. However, this interaction was driven by differences in the *adopting* condition. In particular, marginalized participants were descriptively more uncomfortable than non-marginalized participants with other users adopting marginalized identities in avatars, $t(188) = 1.761$, $p = .080$, 95% CI [-.03, 0.54], $d = 0.26$. However, this difference in comfort was *reduced* when others were concealing marginalized identities in avatars, $t(188) = 0.329$, $p = .742$, 95% CI [-.24, 0.33], $d = 0.05$. In other words, marginalized participants tended to feel more uncomfortable with others adopting, as opposed to concealing, marginalized identities via avatars. However, marginalized and non-marginalized participants felt similarly uncomfortable with others concealing marginalized identities online. See Figure 4 for participants' comfort level across identity, condition, and characteristic.

## 6.3 People Often Assume Nefarious Motives for Changing Identities in an Avatar

We explored open-ended responses discussing others' motives for concealing or adopting identities via avatars. We discovered nine unique themes in the context of characteristics that participants are *uncomfortable* with other users changing; see Figure 5. These themes were fairly consistent across both the conceal and adopt conditions, and thus are described as one class of "uncomfortable" themes.

Rather than or in addition to speculating about other users' motives, many participants expressed beliefs about certain characteristics (e.g., race, sex) they deemed entirely inappropriate to change in an avatar, regardless of whether the user was concealing or adopting a marginalized identity.

> "Digital avatars should share some important characteristics with their real self (e.g.; sex, race, skin tone, pregnancy status, etc.)." *Non-Marginalized (Study 2)*

When considering user's motives, participants most commonly believed other users would have malicious motives, such as stealing, manipulating, catfishing, or grooming, regardless of whether the user was concealing or adopting a marginalized identity. In other words, participants most commonly identified motives that would pose direct threats to their own or other's safety and security in avatar-based interactions. For instance, many participants discussed predatory motives for age and sex changes in avatars.

> "We already have predators on the internet lying about their age and sex to take advantage of young people, and having an avatar will only help further that lie. (Please note that when I talk about predators, I'm thinking more in the way of grown men pretending to
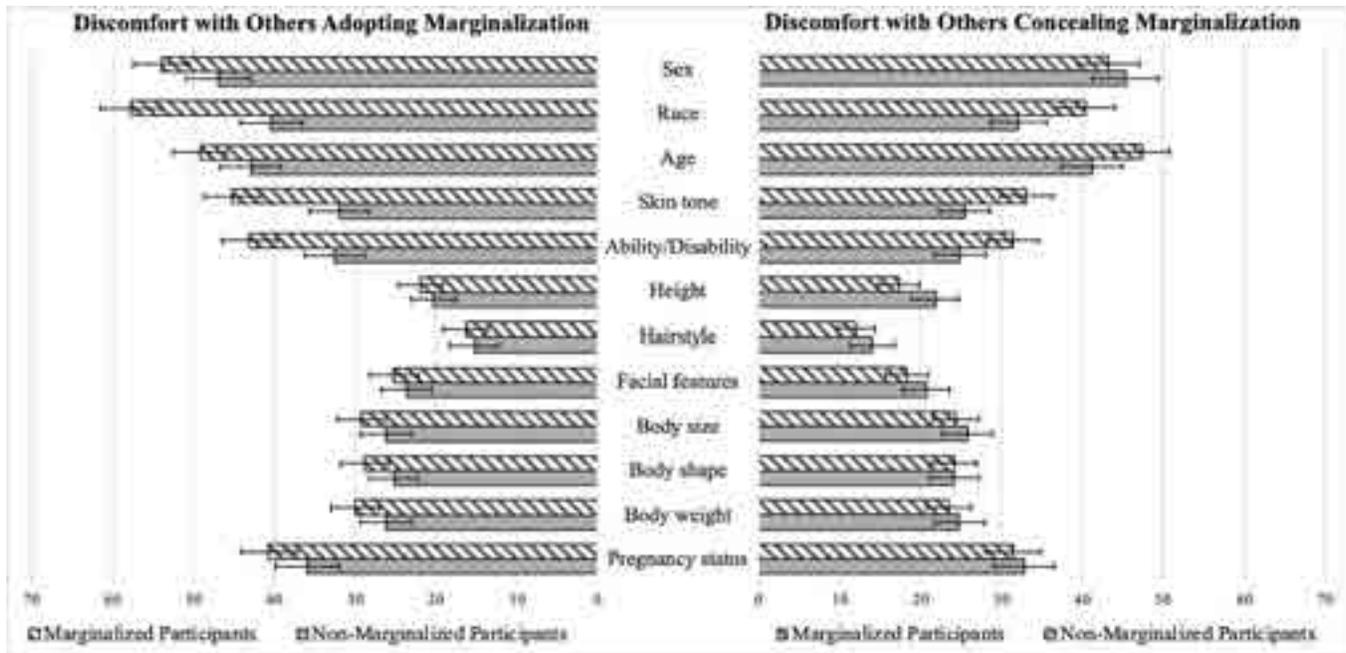
Figure 4: Participants' comfort level with other users changing characteristics in their avatar, by condition and self-identified marginalization status. Error bars are represented by standard error of the mean.

be teens for nefarious purpose, NOT trans people with avatars matching their correct gender presentation.)" *Marginalized for body and social class (Study 2)*

Notably, while many participants were concerned about the safety implications of adults pretending to be children online, others also felt uncomfortable with children adopting adult avatars, whereby they may unwittingly find *themselves* interacting with children online. In this way, some participants felt uncomfortable with characteristic changes, regardless of the directionality (e.g., an adult pretending to be a child or vice versa), but perhaps for different reasons.

Participants surfaced multiple complex motives within this malicious impersonations theme. Some weighed the motives and contexts for their (dis)comfort. For instance, whereas some viewed privacy as more appropriate in work contexts, the same behavior in social contexts was viewed with more skepticism.

> "It depends on the context in which the avatar was used for. If it's for work, then people might change their gender so they're evaluated for their work, typically females might change to male. But I'd be concerned if it was in a social context and someone was using it to get info under false pretenses. You tend to try to find common ground with people and if they lie about it, that means one thing. Otherwise if its just everyone has gender neutral avatars at work, that's okay." *Marginalized for sex and body (Study 2)*

Some participants, and especially marginalized participants, identified that other users may conceal identities to avoid marginalizing
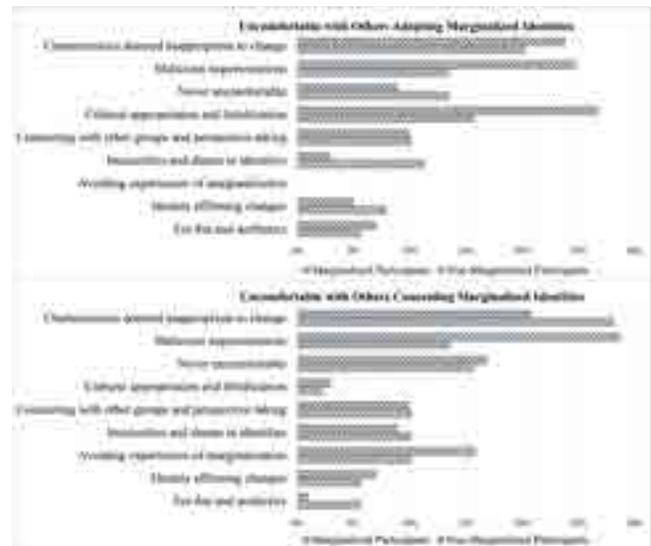
Figure 5: Percent of marginalized and non-marginalized participants whose responses were classified into each "uncomfortable" theme, by condition.

experiences. Even so, some participants expressed discomfort with this self-protective motive.

> "It does not matter if it is to make them seem like they are not part of a marginalized group. It still seems dishonest and manipulative." *Non-Marginalized (Study 2)*

Thus, users from marginalized backgrounds may still face backlash (perhaps even from other marginalized users) for protecting parts of their identities online, as others' anonymity surfaces concerns for one's own safety online.

We separated one unique theme related to malicious intentions into its own theme, emerging predominantly in the context of other users *adopting* marginalized identities: cultural appropriation and fetishization. Many participants, particularly marginalized participants, discussed the unique concerns emerging from this motive.

> "I'd be most uncomfortable with someone presenting themselves as a marginalized race or as a disabled person because it seems to scream stolen valor and appropriation — someone looking to put on a marginalized identity as a costume that denigrates those who actually have that identity." *Marginalized for gender identity, body, and sexual orientation (Study 2)*

6.4 People are Comfortable with Changes Deemed Less Foundational to Identity

We discovered ten unique themes in the context of characteristics participants are *comfortable* with other users changing (see Figure 6). Themes were fairly consistent across the conceal and adopt conditions, and thus are described as one class of "comfortable" themes.

Participants most commonly assumed others would change characteristics (often perceived "flexible" characteristics) for fun or experimentation. In the example below, one participant discussed privacy and what is reasonable and safe to conceal online in this theme.

> "I said I wouldn't mind if someone's avatar had a totally different height, weight, etc. than the person actually had, because, to me, that's not important for everyone on the net to know. I think that sometimes people might change these things just to make a fun image and that it's not usually meant to deceive." *Non-Marginalized (Study 2)*

Relatedly, many participants discussed specific characteristics, such as hairstyle, often deemed unimportant to interactions, more malleable in vivo, or less foundational to identity. Interestingly, many participants believed certain characteristics (e.g., hairstyle, body) were not associated with marginalization or identity, indicating such characteristics may not be viewed as consequential to privacy.

> "Many people have frequently-changing characteristics that may not always be reflected in their avatar. Incidentally, I don't really see how some of the things on this list could be considered "marginalized", such as hairstyle or height." *Non-Marginalized (Study 2)*

Notably, 11% of participants discussed feeling comfortable with other users concealing identities to avoid marginalization, comparable to the 13% of participants who discussed their discomfort with this theme. Although participants identified this self-protective motive for identity concealment online, some may weigh others' protection against possible safety concerns.
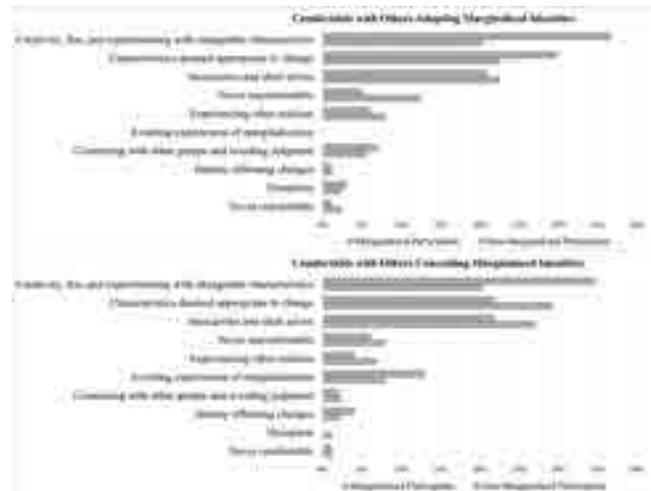

Figure 6: Percent of marginalized and non-marginalized participants whose responses were classified into each "comfortable" theme, by condition.

Broadly, this class of themes revealed that participants felt more comfortable with avatar changes that did not pose apparent threats to their own security and safety online.

7 Discussion

What people reveal or conceal in online spaces is a key question for security and privacy in computing, and the preferences of marginalized individuals are particularly important in this space. Across two studies, we investigated how people navigate privacy for themselves via avatars (Study 1) and how people perceive *others'* privacy decisions via avatars (Study 2). We find that people with marginalized identities (compared to people without) are less willing to reveal their identities in avatars, in part to avoid experiences of marginalization online. However, both marginalized and non-marginalized people were often uncomfortable with other users changing identities in an avatar, even when recognizing they may do so to avoid marginalization. In fact, people frequently assumed others would have malicious motives for both concealing and adopting marginalized identities in an avatar, identifying threats with implications for their own security and safety online. This revealed a tension between stronger desires for one's own privacy juxtaposed against stronger desires for clear information about others.

We also observed important differences between marginalized and non-marginalized participants. In the context of avatar customization, marginalized people more often discussed protecting against experiences of marginalization, whereas non-marginalized people more frequently discussed feeling confident and self-assured and having no desire to conceal their identities. In the context of avatar interactions, marginalized participants more often discussed other users' malicious intentions for choosing anonymity in avatars. Marginalized and non-marginalized people may experience unique privacy-related concerns in virtual worlds, revealing a great deal of nuance in the choice to conceal and reveal identities across groups. Indeed, avatars can afford or revoke an individual's privacy online,

with complex implications for both those who create avatars and those who interact with others' avatars. We find that exploring this question through the lens of marginalization surfaces important considerations for designing safe avatar-based worlds.

This work builds upon a recent wave of research investigating how users from marginalized backgrounds navigate identity representation in virtual worlds. Researchers are increasingly considering identity-related experiences in social VR [9, 25, 30, 44, 60]. We leverage social psychological theories to understand how marginalization broadly impacts concerns about privacy and protections in avatars, which may be useful in guiding continuing research with users from marginalized backgrounds. Although different populations will have their own unique experiences with these platforms, theories of identity can provide insight into deep similarities across groups for informing secure computing design.

Applying psychological theories to avatars also facilitates future interdisciplinary research. Broadly, psychological theories make distinctions between different dimensions of stigma (e.g., concealability, perceived controllability), which may influence both how people navigate avatar customization and how people react to other users' avatars. In this work, we focused on people with "visible" marginalized identities because virtual worlds provide a novel context in which these identities may be revealed or concealed. Whereas concealing one's racial identity is often difficult in vivo, this becomes possible in online spaces. Yet many open questions remain to integrate secure computing with theories of stigma. For instance, do people with typically "concealable" identities have unique design needs for avatars? The answer is likely yes. As one example, people with disabilities report a need for customization options to reveal typically invisible aspects of their identities in safe virtual spaces [25]. Exploring theories of identity may uncover novel questions and design solutions in virtual worlds.

7.1   Limitations and Future Directions

The current work has several limitations and open questions that may motivate future research. First, we adopted a broad approach to explore how concerns about marginalization affect users' privacy-related choices and motivations online. Preliminary findings from Study 1 suggest these broader effects may apply to specific groups as well — for instance, people marginalized for their bodies were less willing to reveal their body shape, body size, and body weight in an avatar, as compared to people who are not marginalized for their bodies. Future work should more deeply explore these groups' interactions with avatar-based worlds, especially as there is a lack of research centering anti-fat stigma in computing (see [25]). Although our work does not deeply explore specific group's unique experiences with avatars, many researchers are investigating just these questions [9, 20, 25, 38, 46, 60]. Our theory-driven approach fills an important gap in the literature, although it will be integral to continue to investigate the unique experiences of specific groups and uncover design solutions appropriate to different groups' needs.

Second, our studies used self-report methods, which are common in psychological research, provide internal validity, and can be insightful tools to assess motivations and psychological states. Whereas interviews or observational methods can dive deeply into a specific context, our goal was to understand the overall landscape

and identify broad relationships between identity concerns and privacy and safety concerns online. A variety of methodological approaches will help us develop a more comprehensive understanding of this space.

Relatedly, Study 2 employed a within-subjects design in which participants imagined two hypothetical scenarios. Imagining two hypothetical scenarios may be less immersive or more confusing than a between-subjects design. However, we note that the vast majority of participants were able to distinguish between these scenarios, evidenced by robust and significant differences in comfort across the two conditions and distinct themes emerging across conditions (e.g., cultural appropriation versus avoiding experiences of marginalization). In some cases, it was clear that participants did not fully understand the hypothetical scenarios, and these participants were excluded from analyses (see subsection 4.4, subsection 4.5).

Third, we made descriptive comparisons between participants in Studies 1 and 2, but these samples differed in their demographics (undergraduates versus crowdworkers, respectively). These sample differences could have affected the observed asymmetries between privacy and transparency, though this disconnect is still meaningful and impactful, even if between different groups of people. To explore one implication further, participants' prior experience with avatars could vary across studies. We did not control for prior experience in our studies because we do not believe expertise is necessary for such deep psychological processes to manifest in the novel context of virtual worlds — and indeed, such processes could play out for both high and low experience users alike. With these limitations in mind, we find that participants appear to have similar degrees of experience across studies. On average, participants reported between a little and a moderate amount of experience creating avatars in Study 1 (M = 2.58, SD = 1.18), and having created between a few and a moderate amount of avatars in Study 2 (M = 2.33, SD = 0.80). The observed asymmetries are important and reflect real dynamics — the tension between wanting privacy for the self but avoiding threats from others — that could extend across age and context. Centrally, there were important lessons *within* each study that stand alone, even without comparing across studies.

Fourth, we acknowledge limitations arising from specific survey questions. The self-identifying marginalization question intentionally included both prior and anticipated experiences of marginalization. An awareness that one's social group experiences marginalization is enough to trigger concerns about experiencing marginalization [52, 53]. The following question asked participants to indicate the identities for which they had experienced marginalization. This question could have been more precisely worded, though we believe participants understood due its direct connection to the preceding question. Further, in Study 2, asking participants to rate how *uncomfortable* they felt may have biased more negative responses (though subsequent open-ended questions invited more nuanced reflection).

Finally, our studies assumed users would have one avatar to use across multiple contexts. However, self-representation will likely vary across contexts. Whereas some platforms employ fixed user identities, others permit changes. People also vary in the extent to which they experience marginalization across contexts. We anticipate that our findings will be especially relevant in more socially interactive contexts. Indeed, evaluation concerns and safety

concerns are likely to occur in situations of intergroup contact [4]. Virtual worlds may expose users to people they would not normally interact with, surfacing such concerns. Participants' responses lend support to the importance of context. For instance, participants discussed authenticity, avoiding dishonesty, and concerns about being judged or left out, all of which suggest interpersonal and/or intergroup motivations. In Study 2, others' avatars were often discussed in terms of their impact on relationships. In contrast, several participants discussed adopting more unrealistic avatars in gaming contexts, or matching the artistic style of the game. Thus, if a virtual world has an unrealistic avatar aesthetic, is predominately gaming-focused, or is non-social, unique motivations may emerge. Future work should continue to uncover the role of context on avatar representations and perceptions.

7.2 Lessons for Design

Researchers have suggested a variety of security and privacy enhancing design solutions in social VR and other avatar-based platforms, such as hardware that protects users' biometric information [12], voice modulators to anonymize users [31], and non-identifiable avatar options to conceal identities [31], accompanying other developments aimed at mitigating harassment (with mixed results), like personal space bubbles [18]. These solutions are important when considering marginalized populations as well. For instance, various users (e.g., trans users, women, non-native speakers) report that concealing their voice can help avoid harassment [30]. However, harassment mitigation strategies can unintentionally marginalize the users they are intended to help, such as reports of personal space bubbles creating barriers to positive interactions [46, 49]. Further, some privacy-enhancing strategies are less effective in immersive VR environments that not only display users' avatars, but also their voice and body language [46]. Our results suggest additional complexity: balancing personal privacy with interpersonal security, and the need for both authenticity and protections, in avatar-based interactions.

Our findings also reveal that users want to represent themselves flexibly, with customization options that allow for accurate and non-stereotypical expressions. These findings highlight the need for design that promotes inclusive identity representation, such as affording more granular customization options that allow accurate representations of multiple identities [25], indirect representation for users with "invisible" identities [60], and inclusion of neopronouns and androgynous avatars [9]. However, we also found that some people felt uncomfortable with other users concealing or changing social identities in avatars, especially when this could pose threats to safety online. Grappling with these safety threats, such as by mitigating harassment and harm from users who might maliciously adopt marginalized identities, is critical. A complex question posed by our work is how to balance restrictions of inappropriate use with the need for both accurate identity representation options and the ability to maintain privacy online. Our results highlight fundamental challenges to designing equitable and safe avatar-based worlds, which research should continue to explore.

We also found that people with marginalized identities felt more identity threat than non-marginalized people, or were more concerned about being devalued based on their identities, which was associated with a greater desire for privacy via avatars. There is clear cause for concern; people with marginalized identities face continued harassment in social VR [2, 10, 29, 47, 59]. Although some users may conceal marginalized identities to avoid harassment, this choice itself comes with costs. Concealing one's social identity can be related to reduced feelings of belonging and authenticity [37], reduced embodiment of one's avatar [46], more negative affect [1], and lowered working memory [50]. Ideally, users should not have to face such a tradeoff. Although people may adopt various strategies to navigate marginalizing experiences, this structural problem cannot be solved via only individualized solutions [3, 7, 16]. In short, the onus should not fall entirely on users facing harassment. Although design solutions should afford privacy and flexibility in identity representations, a great effort should also be put toward creating virtual worlds where users can feel their identities are valued and welcomed.

Finally, our studies were not designed to enable us to propose design recommendations for specific groups, though other research highlighted above explores just these questions. Instead, this work provides a foundation for understanding the issues that need to be fully considered before final designs are determined. For instance, the VR design space should grapple with these tensions between privacy and transparency. Design research has begun to explore related questions, including considerations for handling identity and fraud (e.g., identity verification in some virtual contexts) [32]. Our work suggests designers should consider societal and structural systems of marginalization (which will differ across societies, regions, and peoples) that exist beyond specific virtual worlds to ensure technical solutions do not unintentionally reinforce marginalization. Although users should have agency over their virtual representations, the option to reveal or conceal identities is not a complete solution. That marginalized users feel less comfortable revealing their identities and safely interacting with other avatars suggests design can reinforce marginalization. Grappling with these issues involves explicitly considering who is being included in design considerations and who is being pushed out [48], anticipating how the harms and benefits of a system will be distributed (un)equally across groups [42], and including communities who are most impacted by design in the design process itself [5].

7.3 Conclusion

As avatars move toward richer representations of users, it is imperative to understand how avatars can afford or revoke privacy, and how users from marginalized backgrounds may have unique privacy-related needs in these contexts. Applying psychological theories of identity, we find that marginalization can surface important concerns about privacy in avatars. Although people may conceal marginalized identities in avatars to avoid devaluation, this self-protective strategy can come with costs. Such asymmetrical needs, wanting privacy for oneself but to feel informed about others, present important considerations for designing equitable and safe virtual worlds.

Acknowledgments

reviewers for their feedback that further improved this paper. This work was done as part of the Center for Privacy and Security for Marginalized and Vulnerable Populations (PRISM), supported by the National Science Foundation under Awards SaTC-2205171 and 2207019.

References

[1] Manuela Barreto, Naomi Ellemers, and Serena Banal. 2006. Working under cover: Performance-related self-confidence among members of contextually devalued groups who try to pass. *European Journal of Social Psychology* 36, 3 (2006), 337–352. https://doi.org/10.1002/ejsp.314

[2] Lindsay Blackwell, Nicole Ellison, Natasha Elliott-Deflo, and Raz Schwartz. 2019. Harassment in Social Virtual Reality: Challenges for Platform Governance. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 3. 1–25. https://doi.org/10.1145/3359202

[3] Eduardo Bonilla-Silva. 2021. *Racism without Racists: Color-Blind Racism and the Persistence of Racial Inequality in America* (6th ed.). Rowman & Littlefield, Lanham, MD. https://doi.org/10.2307/40264552

[4] Nyla R. Branscombe, Naomi Ellemers, Russell Spears, and Bertjan Doosje. 1999. The Context and Content of Social Identity Threat. In *Social Identity: Context, Commitment, Content*, Naomi Ellemers, Russell Spears, and Bertjan Doosje (Eds.). Blackwell Science, Oxford, UK, 35–58.

[5] Sasha Costanza-Chock. 2018. Design justice: Towards an intersectional feminist framework for design theory and practice. In *Proceedings of the Design Research Society*. 528–540. https://doi.org/10.21606/drs.2018.679

[6] Jennifer Crocker and Brenda Major. 1989. Social stigma and self-esteem: The self-protective properties of stigma. *Psychological Review* 96, 4 (1989), 608–630. http://dx.doi.org/10.1037/0033-295x.96.4.608

[7] Richard Delgado and Jean Stefancic. 2011. *Critical Race Theory: An Introduction* (2nd ed.). New York University Press, New York, NY.

[8] Katrina Fong, Joshua A. Quinlan, and Raymond A. Mar. 2023. Select your character: Individual needs and avatar choice. *Psychology of Popular Media* 12, 1 (2023), 30–39. https://psycnet.apa.org/doi/10.1037/ppm0000384

[9] 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 *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–15. https://doi.org/10.1145/3491102.3502082

[10] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Dane Acena. 2022. Disturbing the peace: Experiencing and mitigating emerging harassment in social virtual reality. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 6. 1–30. https://doi.org/10.1145/3512932

[11] Jin X. Goh, Daniel N. Kort, Avery M. Thurston, Lina R. Benson, and Cheryl R. Kaiser. 2019. Does concealing a sexual minority identity prevent exposure to prejudice? *Social Psychological and Personality Science* 10, 8 (2019), 1056–1064. https://doi.org/10.1177/1948550619829065

[12] Brendan John, Sophie Jörg, Sanjeev Koppal, and Eakta Jain. 2020. The security-utility trade-off for iris authentication and eye animation for social virtual avatars. In *IEEE transactions on visualization and computer graphics*, Vol. 26. 1880–1890. https://doi.org/10.1109/TVCG.2020.2973052

[13] Cheryl R. Kaiser, S. Brooke Vick, and Brenda Major. 2006. Prejudice expectations moderate preconscious attention to cues that are threatening to social identity. *Psychological Science* 17, 4 (2006), 332–338. https://doi.org/10.1111/j.1467-9280.2006.01707.x

[14] Sonia K. Kang, Katherine A. DeCelles, András Tilcsik, and Sora Jun. 2016. Whitened résumés: Race and self-presentation in the labor market. *Administrative Science Quarterly* 61, 3 (2016), 469–502. https://doi.org/10.1177/0001839216639577

[15] Priyanka Kesarwani. 2023. Meta Unveils New Avatar Customization Options Across Platforms. *Marketing Mind* (2023). https://www.marketingmind.in/meta-unveils-new-avatar-customization-options-across-platforms/.

[16] Anne M. Koenig and Alice H. Eagly. 2014. Evidence for the Social Role Theory of Stereotype Content: Observations of Groups' Roles Shape Stereotypes. *Journal of Personality and Social Psychology* 107, 3 (2014), 371–392. http://dx.doi.org/10.1037/a0037215

[17] Kathryn M. Kroeper, Heidi E. Williams, and Mary C. Murphy. 2020. Counterfeit diversity: How strategically misrepresenting gender diversity dampens organizations' perceived sincerity and elevates women's identity threat concerns. *Journal of Personality and Social Psychology* 3, 122 (2020), 399–426. https://dx.doi.org/10.1037/pspi0000348

[18] Karina LaRubbio, Ethan Wilson, Sanjeev Koppal, Sophie Jörg, and Eakta Jain. 2023. Give me some room please! Personal space bubbles for safety and performance. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. 897–898. https://doi.org/10.1109/VRW58643.2023.00291

[19] Joel M. Le Forestier, Elizabeth Page-Gould, and Alison L. Chasteen. 2023. Which Identities Are Concealable? Individual Differences in Concealability. *Personality and Social Psychology Bulletin* (2023), 1–19. https://doi.org/10.1177/01461672231198162

[20] Jong-Eun Roselyn Lee. 2014. Does virtual diversity matter? Effects of avatar-based diversity representation on willingness to express offline racial identity and avatar customization. *Computers in Human Behavior* 36 (2014), 190–197. https://doi.org/10.1016/j.chb.2014.03.040

[21] Jim Lewis and Jeff Sauro. 2020. Are Sliders Better Than Numbered Scales? *MeasuringU* (2020). https://measuringu.com/numbers-versus-sliders/.

[22] Calvin A. Liang, Sean A. Munson, and Julie A. Kientz. 2021. Embracing Four Tensions in Human-Computer Interaction Research with Marginalized People. *ACM Transactions on Computer-Human Interaction (TOCHI)* 28, 2 (2021), 1–47. https://doi.org/10.1145/3443686

[23] Hsin Lin and Hua Wang. 2014. Avatar creation in virtual worlds: Behaviors and motivations. *Computers in Human Behavior* 34 (2014), 213–218. https://doi.org/10.1016/j.chb.2013.10.005

[24] Jinghuai Lin and Marc Erich Latoschik. 2018. Digital body, identity and privacy in social virtual reality: A systematic review. *Frontiers in Virtual Reality* 3, 974652 (2018), 1–25. https://doi.org/10.3389/frvir.2022.974652

[25] Kelly Mack, Rai Ching Ling Hsu, Andrés Monroy-Hernández, Brian A. Smith, and Fannie Liu. 2023. Towards Inclusive Avatars: Disability Representation in Avatar Platforms. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–13. https://doi.org/10.1145/3544548.3581481

[26] Cameron D. Mackey, Christopher F. Silver, Kimberly Rios, Colleen M. Cowgill, and Ralph W. Hood Jr. 2021. Concealment of nonreligious identity: Exploring social identity threat among atheists and other nonreligious individuals. *Group Processes & Intergroup Relations* 24, 5 (2021), 860–877. https://doi.org/10.1177/1368430220905661

[27] Moira Maguire and Brid Delahunt. 2017. Doing a Thematic Analysis: A Practical, Step-by-Step Guide for Learning and Teaching Scholars. *All Ireland Journal of Teaching and Learning in Higher Education* 8, 3 (2017), 3351–33514.

[28] Brenda Major and Laurie T. O'Brien. 2005. The Social Psychology of Stigma. *Annual Review of Psychology* 56 (2005), 393–421. https://doi.org/10.1146/annurev.psych.56.091103.070137

[29] Divine Maloney, Guo Freeman, and Andrew Robb. 2020. A virtual space for all: Exploring children's experience in social virtual reality. In *In Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 472–483. https://doi.org/10.1145/3410404.3414268

[30] Divine Maloney, Guo Freeman, and Donghee Yvette Wohn. 2020. "Talking without a voice": Understanding Non-Verbal Communication in Social Virtual Reality. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 4. 1–25. https://doi.org/10.1145/3415246

[31] 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. https://doi.org/10.1145/3385956.3418967

[32] Rachel McAmis, Betül Durak, Melissa Chase, Kim Laine, Franziska Roesner, and Tadayoshi Kohno. 2024. Handling Identity and Fraud in the Metaverse. *IEEE Security & Privacy* (2024), 2–12. https://doi.org/10.1109/MSEC.2024.3399699

[33] Willajeanne F. McLean. 2020. Who are you wearing? Avatars, blackface and commodification of the other. *The Law Review of the Franklin Pierce Center for Intellectual Property* 61, 3 (2020), 455–506.

[34] Jen Mills. 2020. Paedophile abused 51 boys by pretending to be teenage girl online. *Metro* (2020). https://metro.co.uk/2020/11/23/paedophile-abused-51-boys-by-pretending-to-be-teenage-girl-online-13641246/.

[35] Catherine Mooney and Brett A. Becker. 2020. Sense of Belonging: The Intersectionality of Self-Identified Minority Status and Gender in Undergraduate Computer Science Students. In *United Kingdom & Ireland Computing Education Research conference (UKICER '20)*. 24–30. https://doi.org/10.1145/3416465.3416476

[36] Lisa Nakamura. 2005. Race In/For Cyberspace: Identity Tourism and Racial Passing on the Internet. *Works and Days* (2005). https://api.semanticscholar.org/CorpusID:17777456

[37] Anna-Kaisa Newheiser and Manuela Barreto. 2014. Hidden costs of hiding stigma: Ironic interpersonal consequences of concealing a stigmatized identity in social interactions. *Journal of Experimental Social Psychology* 52 (2014), 58–70. https://doi.org/10.1016/j.jesp.2014.01.002

[38] Jaeyoung Park and Seongcheol Kim. 2022. How do people with physical disabilities want to construct virtual identities with avatars? *Frontiers in Psychology* 13 (2022), 994786. https://doi.org/10.3389/fpsyg.2022.994786

[39] Kristin Pauker, Nicholas O. Rule, and Nalini Ambady. 2010. Ambiguity and social perception. In *Social Psychology of Visual Perception*, E. Balcetis and D. Lassiter (Eds.). Psychology Press, New York, NY, 7–26.

[40] Andreas Pfitzmann and Marit Hansen. 2005. Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management – A Consolidated Proposal for Terminology. v0.28 (2005), 1–54. https://doi.org/10.1007/3-540-44702-4_1

[41] Diane M. Quinn and Valerie A. Earnshaw. 2011. Understanding concealable stigmatized identities: The role of identity in psychological, physical, and behavioral outcomes. *Social Issues and Policy Review* 5, 1 (2011), 160–190. https://doi.org/10.1111/j.1751-2409.2011.01029.x

[42] Harshini Sri Ramulu, Helen Schmitt, Dominik Wermke, and Yasemin Acar. 2024. Security and Privacy Software Creators' Perspectives on Unintended Consequences. In *33rd USENIX Security Symposium (USENIX Security 24)*. 3259–3276.

[43] Emma Roth. 2023. Apple's Vision Pro headset will turn you into a digital avatar when FaceTiming. *The Verge* (2023). https://www.theverge.com/2023/6/5/23750096/apple-vision-pro-headset-persona-facetime.

[44] Abhinaya S.B., Aafaq Sabir, and Anupam Das. 2024. Enabling Developers, Protecting Users: Investigating Harassment and Safety in VR. In *33rd USENIX Security Symposium (USENIX Security 24)*. 6561–6578.

[45] Toni Schmader and Michael Johns. 2003. Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology* 84 (2003), 440–452. https://psycnet.apa.org/doi/10.1037/0022-3514.85.3.440

[46] 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. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 7. 1–29. https://doi.org/10.1145/3610027

[47] 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)*. 225–226. https://doi.org/10.1109/VR.2017.7892258

[48] Mattea Sim, Kurt Hugenberg, Tadayoshi Kohno, and Franziska Roesner. 2023. A Scalable Inclusive Security Intervention to Center Marginalized & Vulnerable Populations in Security & Privacy Design. In *Proceedings of the 2023 New Security Paradigms Workshop*. 102–115. https://doi.org/10.1145/3633500.3633508

[49] Jaisie Sin, Rachel L. Franz, Cosmin Munteanu, and Barbara Barbosa Neves. 2021. Digital Design Marginalization: New Perspectives on Designing Inclusive Interfaces. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11. https://doi.org/10.1145/3411764.3445180

[50] Laura Smart and Daniel M. Wegner. 1999. Covering up what can't be seen: Concealable stigma and mental control. *Journal of Personality and Social Psychology* 77, 3 (1999), 474–486. https://psycnet.apa.org/doi/10.1037/0022-3514.77.3.474

[51] Steven J. Spencer, Claude M. Steele, and Diane M. Quinn. 1999. Stereotype threat and women's math performance. *Journal of Experimental Social Psychology* 35, 4 (1999), 4–28. https://doi.org/10.1006/jesp.1998.1373

[52] Claude M. Steele. 1997. A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance. *American Psychologist* 52, 6 (1997), 613–629. https://psycnet.apa.org/doi/10.1037/0003-066X.52.6.613

[53] Claude M. Steele, Steven J. Spencer, and Joshua Aronson. 2002. Contending with group image: The psychology of stereotype and social identity threat. In *Advances in Experimental Social Psychology*, M. P. Zanna (Ed.). Vol. 34. Academic Press, 379–440. https://doi.org/10.1016/S0065-2601(02)80009-0

[54] 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. https://doi.org/10.1145/3491102.3502008

[55] Henry Tajfel and John C. Turner. 1986. The social identity theory of intergroup behavior. In *The Psychology of Intergroup Relations* (2nd ed.), S. Worchel and W.G. Austin (Eds.). Hall Publishers, Chicago, IL, 7–24.

[56] Ruth van Veelen, Jenny Veldman, Colette Van Laar, and Belle Derks. 2020. Distancing from a stigmatized social identity: State of the art and future research agenda on self-group distancing. *European Journal of Social Psychology* 50, 6 (2020), 1089–1107. https://doi.org/10.1002/ejsp.2714

[57] Ari Voutilainen, Taina Pitkäaho, Tarja Kvist, and Katri Vehviläinen-Julkunen. 2016. How to ask about patient satisfaction? The visual analogue scale is less vulnerable to confounding factors and ceiling effect than a symmetric Likert scale. *Journal of Advanced Nursing* 72, 4 (2016), 946–957. https://doi.org/10.1111/jan.12875

[58] Cheng Yao Wang, Sandhya Sriram, and Andrea Stevenson Won. 2021. Shared Realities: Avatar Identification and Privacy Concerns in Reconstructed Experiences. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 5. 1–25. https://doi.org/10.1145/3476078

[59] Queenie Wong. 2021. As Facebook Plans the Metaverse, It Struggles to Combat Harassment in VR. *CNET* (2021). https://www.cnet.com/tech/gaming/features/as-facebook-plans-the-metaverse-it-struggles-to-combat-harassment-in-vr//.

[60] 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 *The 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22)*. 1–16. https://doi.org/10.1145/3517428.3544829

[61] Daniel Zimmermann, Anna Wehler, and Kai Kaspar. 2022. Self-representation through avatars in digital environments. *Current Psychology* (2022), 1–15. https://doi.org/10.1007/s12144-022-03232-6

A Study 1 Materials

[Social Identity] Before you begin the study, we would like to know a little bit more about you. On the next page, you will be asked to respond to some demographic questions.

A social identity is a sense of who we are based on the social groups we belong to. Some people experience marginalization based on these social identities, or experiences of stigma, disrespect, or discrimination based on one or more of their social identities. We are interested in the extent to which you believe you have experienced (or believe you may experience in the future) stigma, disrespect, or discrimination based on one or more of your social identities that are visible to other people. In other words, people know you belong to these groups just by looking at you.

Please indicate whether you have a visible stigmatized identity.
○ Yes, I have experienced or believe I may experience marginalization based on one or more visible social identities.
○ No, I have not experienced nor do I believe I will experience marginalization based on one or more visible social identities.

If you answered yes to the question above, indicate below the social category or categories for which you have experienced stigma, disrespect, or discrimination (select all that apply).
☐ Race
☐ Ethnicity
☐ Sex
☐ Gender identity
☐ Age
☐ Ability/Disability
☐ Body weight, shape, or size
☐ Height
☐ Pregnancy status
☐ Religion
☐ Sexual orientation
☐ Social class
☐ Another identity or multiple identities not listed here (please specify) [text entry box]

[Primary Instructions] In this study, we want to understand how people create avatars for online worlds and networks. Avatars are artificial persons or graphic figures that are increasingly used to represent the user (you) in online environments, instead of actual photos of the user. We are conducting preliminary research on how people want to represent themselves through avatars. In the future, people might have the choice to have dramatically different avatars across different contexts (e.g., work, school, socializing), or more like real life, people might have a single avatar to use across all contexts. In this research, we want to understand how you might create an avatar if you had to use a single avatar across a variety of contexts. This avatar would represent yourself to other users who you would interact with online. Therefore, the characteristics and physical features of your avatar would be visible to other users. Once you continue, you will respond to questions about your expectations for the virtual world and how you want your avatar to look.

[Identity Threat] Respond to the following questions about your expectations for interactions in the virtual world if your avatar looked exactly like you look in the real world. [Likert scale: 1 = Strongly Disagree, 5 = Strongly Agree]

- I worry I won't belong
- I worry that others won't respect me
- I worry that others won't value my opinions or contributions
- I worry that I can't be my true self
- I worry I will be left out or marginalized
- I worry I will be stereotyped based on my identities

[Identity Revealing] Respond to the following questions about your intentions for creating an avatar. [Likert scale: 1 = Strongly Disagree, 5 = Strongly Agree]

- I am willing to reveal all my identities through my avatar
- I am eager to reveal all my identities through my avatar
- I feel uncomfortable revealing all my identities through my avatar [reverse coded]
- I will create an avatar that resembles my actual self
- I will create an avatar different from my actual self [reverse coded]
- I want my avatar to look exactly like I look in the real world

For each of the following characteristics, rate the extent to which you want the appearance of your avatar to perfectly match your actual self. [Slider scale: 1 = Not at all like my actual self, 100 = Exactly like my actual self]

- Biological sex
- Race
- Age
- Hairstyle
- Facial features
- Skin tone
- Body shape
- Body size
- Body weight
- Height
- Ability/Disability
- Pregnancy status (skip if not applicable)

[Open-Ended Questions] Based on the previous question, provide more info below as applicable.

Please explain more about how you want your avatar to look different from yourself and why. [essay text box]
Please explain more about how you want your avatar to look similar to yourself and why. [essay text box]

[Avatar Experiences]

- I worry that avatar creation tools won't let me accurately represent myself through an avatar [Likert scale: 1 = Strongly Disagree, 5 = Strongly Agree]
- How much experience do you have with creating avatars? [Likert scale: 1 = Not at all, 2 = A little, 3 = A moderate

amount, 4 = A lot, 5 = A great deal]

Respond to the following questions about your previous experiences with creating avatars. If you have never created an avatar before, you can skip to the next page. [Likert scale: 1 = Strongly Disagree, 5 = Strongly Agree]

- I have struggled to make an avatar that looks like me because there are not enough options
- I've noticed that my identities are not always represented in avatar creation tools
- Please select "Strongly disagree" for this question [attention check]
- I am usually satisfied with the avatar creation process

Name some platforms in which you have created avatars in the past (for example, BitMoji, Wii, Second Life, Horizon World, Sims, etc.). [essay text box]

[Attention Check] The color test is simple, when asked to enter a color you must enter the world teal in the text box below.
Based on the previous instruction, what color have you been asked to enter? [text entry box]

B Study 2 Materials

[Primary Instructions] In this study, we want to understand how people perceive other avatars in online worlds and networks. Avatars are artificial persons or graphic figures that are increasingly used to represent the user (you and others) in online environments, instead of actual photos of the user. In the future, people might have the choice to have dramatically different avatars across different contexts (e.g., work, school, socializing), or more like real life, people might have a single avatar to use across all contexts. We want to understand how you might evaluate avatars if you had to use a single avatar across a variety of contexts.

We are conducting research to understand how people react to the ways in which others represent themselves through avatars. Sometimes, people might choose to create an avatar that looks exactly like themselves in the real world, but other times, people might choose to create an avatar that looks different from themselves. Imagine you have been getting to know someone in a virtual world where both of you are represented by avatars. We will ask for your reaction to two different scenarios where you learn that this person's avatar looks different than they do in the real world.

[Participants are randomly assigned to both the conceal and adopt conditions in random order]

[Conceal Condition Manipulation] You learn that the person you've been getting to know online has an avatar with identities that are different from their identities in the real world. In actuality:

- Their real identity is typically marginalized in society
- Their avatar has an identity that is different from their marginalized identity

[Adopt Condition Manipulation] You learn that the person you've been getting to know online has an avatar with identities that are different from their identities in the real world. In actuality:

- Their avatar has an identity that is typically marginalized in society
- Their real identity is NOT marginalized

[Comfort Questions in Each Condition] For each of the following characteristics, rate how uncomfortable you would be if this person's avatar looked different from their actual self. [Slider scale: 1 = Not at all uncomfortable, 100 = Very uncomfortable]

- Biological sex
- Race
- Age
- Hairstyle
- Facial features
- Skin tone
- Body shape
- Body size
- Body weight
- Height
- Ability/Disability
- Pregnancy status

[Open-Ended Questions]
For the characteristic above that you were the most *uncomfortable* with, explain what you think their motives would be for changing this characteristic in their avatar. [essay text box]

For the characteristic above that you were the most *comfortable* with, explain what you think their motives would be for changing this characteristic in their avatar. [essay text box]

[Avatar Experiences] How many avatars have you created in the past? [Likert scale: None at all, A few, A moderate amount, A lot]

[Attention Check] The color test is simple, when asked to enter a color you must enter the world teal in the text box below.
Based on the previous instruction, what color have you been asked to enter? [text entry box]

[Social Identity] Finally, we would like to know a little bit more about you. On the next page, you will be asked to respond to some demographic questions.

A social identity is a sense of who we are based on the social groups we belong to. Some people experience marginalization based on these social identities, or experiences of stigma, disrespect, or discrimination based on one or more of their social identities. We are interested in the extent to which you believe you have experienced (or believe you may experience in the future) stigma, disrespect, or discrimination based on one or more of your social identities that are visible to other people. In other words, people know you belong to these groups just by looking at you.

Please indicate whether you have a visible stigmatized identity.

○ Yes, I have experienced or believe I may experience marginalization based on one or more visible social identities.
○ No, I have not experienced nor do I believe I will experience marginalization based on one or more visible social identities.

If you answered yes to the question above, indicate below the social category or categories for which you have experienced stigma, disrespect, or discrimination (select all that apply).
□ Race
□ Ethnicity
□ Sex
□ Gender identity
□ Age
□ Ability/Disability
□ Body weight, shape, or size
□ Height
□ Pregnancy status
□ Religion
□ Sexual orientation
□ Social class
□ Another identity or multiple identities not listed here (please specify) [text entry box]

C   Demographic Questions

Participants responded to the following demographic questions in both Studies 1 and 2.

We'd like to know just a little bit more about you. Please respond to the following demographic questions.

[Gender Identity] What is your gender identity? (Select all that apply)
□ Woman
□ Man
□ Non-binary
□ Genderqueer
□ Agender
□ A gender not listed
□ Prefer to self-describe [text entry box]
□ Prefer not to say

[Sex] What was your sex at birth?
○ Female
○ Male
○ Prefer not to say

[Sexual Orientation] What is your sexual orientation?
○ Straight/heterosexual
○ Gay or lesbian
○ Bisexual
○ Pansexual
○ Queer
○ Asexual
○ Prefer to self-describe [text entry box]

○ Prefer not to say

[Age] What is your age? Type value: [text entry box]

[Language] What is your primary language? [text entry box]

[Politics] Where would you place yourself on this scale from extremely liberal to extremely conservative? [Likert scale: 1 = Extremely Liberal, 5 = Moderate, 9 = Extremely Conservative]

[Race] What is your race? (Select all that apply)
□ Black/African American
□ South Asian
□ East Asian
□ American Indian/Alaskan Native
□ Native Hawaiian/Pacific Islander
□ White/Caucasian
□ Another race, ethnicity, or origin (please specify): [text entry box]
□ Prefer to self-describe [text entry box]
□ Prefer not to say

[Ethnicity] Are you of Hispanic, Latino/a/x, or Spanish origin? (Select all that apply)
□ No, not of Hispanic, Latino/a/x, or Spanish origin
□ Yes, Mexican, Mexican American, Chicano/a/x
□ Yes, Puerto Rican
□ Yes, Cuban
□ Yes, another Hispanic, Latino/a/x, or Spanish origin [text entry box]
□ Prefer to self-describe [text entry box]
□ Prefer not to say

[Perceived Socioeconomic Status] Think of this ladder as showing where people stand in the United States. At the top of the ladder are people who have the most money, most education, and most respected jobs. At the bottom of the ladder are people who have the least money, least education, and least respected jobs. Where would you place yourself on the ladder? [Likert scale: 1 = Worst off, 10 = Best off]

[Body] Thinking of your bodyweight, which of the following would you say you are?
○ Very underweight
○ Underweight
○ Slightly underweight
○ Neither underweight nor overweight
○ Slightly overweight
○ Overweight
○ Very overweight

[Disabilities] Do you have long-term health conditions or disabilities? (Select all that apply)
□ Vision impairment
□ Hearing impairment
□ Psychological disorder/mental health condition
□ Intellectual or learning disability

□ Neurological disability
□ Autism spectrum disorder
□ Physical disability/reduced mobility
□ Another condition not listed (please specify): [text entry box]
□ None
□ Prefer not to answer

[Pregnancy Status] Are you currently pregnant?
○ No
○ Yes
○ Prefer not to say

[Religion] What is your religious affiliation? (Select all that apply)
□ Baha'i
□ Buddhism
□ Candomblé
□ Christianity (e.g., Baptist, Church of England, Roman Catholic, Methodist, Jehovah Witness, etc.)
□ Hinduism
□ Islam
□ Jainism
□ Judaism
□ Non-religious (e.g., Agnostic, Atheist, No religion)
□ Paganism
□ Rastafari
□ Santeria
□ Shinto
□ Sikhism
□ Spiritualism
□ Taoism
□ Unitarianism
□ Zoroastrianism
□ Another religion not listed [text entry box]
□ Prefer not to answer