Privacy in Immersive Extended Reality: Exploring User Perceptions, Concerns, and Coping Strategies

Hilda Hadan
hhadan@uwaterloo.ca
Stratford School of Interaction Design and Business, University of Waterloo
Waterloo, Canada

Derrick M. Wang
dwmaru@uwaterloo.ca
Stratford School of Interaction Design and Business, University of Waterloo
Waterloo, Canada

Lennart E. Nacke
lennart.nacke@acm.org
Stratford School of Interaction Design and Business, University of Waterloo
Waterloo, Canada

Leah Zhang-Kennedy
lzhangkennedy@uwaterloo.ca
Stratford School of Interaction Design and Business, University of Waterloo
Waterloo, Canada

ABSTRACT

Extended Reality (XR) technology is changing online interactions, but its granular data collection sensors may be more invasive to user privacy than web, mobile, and the Internet of Things technologies. Despite an increased interest in studying developers' concerns about XR device privacy, user perceptions have rarely been addressed. We surveyed 464 XR users to assess their awareness, concerns, and coping strategies around XR data in 18 scenarios. Our findings demonstrate that many factors, such as data types and sensitivity, affect users' perceptions of privacy in XR. However, users' limited awareness of XR sensors' granular data collection capabilities, such as involuntary body signals of emotional responses, restricted the range of privacy-protective strategies they used. Our results highlight a need to enhance users' awareness of data privacy threats in XR, design privacy-choice interfaces tailored to XR environments, and develop transparent XR data practices.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; • **Security and privacy** → **Human and societal aspects of security and privacy**; • **Computing methodologies** → **Perception**; **Virtual reality**; **Mixed / augmented reality**.

KEYWORDS

User privacy, Privacy Perception, Privacy-Seeking Strategies, Virtual Reality, Augmented Reality, Mixed Reality, Extended Reality

ACM Reference Format:

Hilda Hadan, Derrick M. Wang, Lennart E. Nacke, and Leah Zhang-Kennedy. 2024. Privacy in Immersive Extended Reality: Exploring User Perceptions, Concerns, and Coping Strategies. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 24 pages. https://doi.org/10.1145/3613904.3642104

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642104

1 INTRODUCTION

Extended Reality (XR) is the umbrella term for immersive technologies that includes Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). XR lets users easily create, collaborate, and traverse immersive digital environments [75]. These environments offer realistic visual and auditory experiences, as well as interactive elements. XR is revolutionizing the way people live, work, and play. Following Meta's introduction of the first commercially available VR equipment to the general public in 2016 [69], the XR market has experienced rapid growth, reaching a market size of $3.31 billion USD in 2023 [3]. Since then, XR technologies have been widely used in various domains such as gaming [63], social networking [42, 43], research experiments [44], education [25, 54], and healthcare [29, 64].

XR technologies create a virtual world for people to interact with digital objects by tracking their positions, movements, and surroundings through body sensors [7]. As XR changes the way people access digital information and interact with the real world, advertisers can profile users using XR sensor data [46]. The immersive features of XR create unique privacy challenges that cannot be easily mitigated or recognized by users. For example, sensors used to operate XR devices can access behavioural pattern data that is traditionally considered non-sensitive to identify users [47, 56]. Hyperpersonalized avatars that imitate trusted loved ones may steal personal information from users [46]. Furthermore, the potential to make inferences about users based on data collected from XR systems [60] and the ambiguity surrounding multi-user access control [36, 43], are emerging issues that raise crucial privacy concerns. The data collected from the XR systems can be used to deduce personal details about people, including their location, interests, and behaviour. The data can be used for targeted advertising or criminal purposes, infringing on people's privacy rights.

Existing literature had explored privacy issues around AR browsers [45], social VR [42, 43, 51], VR learning [25], and behavioural biometrics [47, 56] from the developers' perspective. These investigations have employed various methodologies to explore the potential privacy risks posed by XR, such as analyzing

the system functional requirements [15, 45, 51], using threat modeling of internal and external vulnerabilities [25], and surveying existing privacy policies [2, 15]. However, the perceptions of XR users were rarely assessed.

The objective of our research is to *understand users' awareness and privacy concerns in XR, and to identify the contributing factors (e.g., data type, device status) in raising or inhibiting their concerns.* For users who express privacy concerns in XR, we also study their coping strategies to protect their data privacy. We use a scenario-based survey method to answer the following Research Questions (RQs):

- **RQ1:** How much are people aware of data collection through XR devices?
- **RQ2:** To what extent are people concerned about their privacy and data collection in XR?
- **RQ3:** What factors contribute to user privacy concerns in XR?
- **RQ4:** What coping strategies (if any) do people use to mitigate their privacy concerns?

Our work makes three main contributions. First, we found that XR users lack awareness that XR sensors can capture highly granular data, including involuntary body signals of emotional responses, which could potentially make XR more invasive to their privacy than other types of devices they use on the web, mobile, and Internet of Things (IoT) devices. Second, our survey identified the impacts of various factors through 18 scenarios, such as data type and data sensitivity, on XR users' privacy concerns. Although several of these privacy-related factors were also observed in non-XR environments in previous research (e.g., [37–39, 48, 72]), our work provides a more precise understanding of the effects of the factors on users in XR environments. Third, we show that participants lack sophisticated mitigation strategies to address their privacy concerns. For example, despite expressing discomfort with data collection, most of our participants opted to "give up" protecting their privacy. Qualitative feedback from participants indicates that they were very pessimistic about their privacy due to the belief that they had already lost privacy on other non-XR devices they use. Furthermore, their limited privacy-protective strategies are likely based on their experience interacting with non-XR devices. Therefore, we suggest that understanding and improving users' mental models of data privacy threats in XR is vital in supporting effective XR risk communication. To achieve transparent XR data practices, the design of privacy-choice interfaces tailored to XR environments is necessary, such as taking into consideration their immersive nature in design to minimize user disruption.

2 BACKGROUND AND RELATED WORK

Extended Reality (XR) is an umbrella term that describes the following immersive technologies [75].

Virtual Reality (VR): VR technologies bring people into completely immersive virtual reality, according to the XR Safety Initiative (XRSI) [75]. This software-generated three-dimensional (3D) environment substitutes reality with a virtual world by covering the user's eyes with equipment like head-mounted displays (HMDs) (e.g., Oculus Rift [69]). VR provides an immersive view that separates the user from the real environment. It enables the user to virtually visit remote locations far from their actual location [67].

Augmented Reality (AR): Unlike VR, AR refers to technologies that augment, rather than replace, the real world [75]. AR technology overlays digital visuals on top of real-world items to provide users with a more engaging experience. For example, the mobile game *Pokémon GO*[1] uses augmented reality to display computer-generated monsters on lawns and sidewalks as players stroll through their areas. AR is user-dependent and always takes place in the physical area surrounding users [67].

Mixed Reality (MR): MR technologies exist at the confluence of VR and AR, mixing the virtual environment with the users' real-world [67, 75]. The virtual and physical worlds coexist and interact in MR. Virtual objects behave and function much like physical items in the real world. For example, the illumination of the virtual object may match the physical light source, or the virtual objects may sound as though they are physically separated from the users. Although most MR systems currently rely on visual and acoustic feedback, researchers are working on technologies that will give haptics, taste, smell, and other cues in the future to simulate temperature, balance, and other factors [67].

Given that XR encompasses VR, AR, and MR, any devices such as HMDs, smart glasses, controllers, and projectors are considered XR [75]. These devices rely on similar sensors to gather information about the user and their environment to produce the corresponding visual, audio, and haptic effects to simulate an immersive and interactive user experience. Given the similarity in embedded sensors and data tracking capabilities across VR, AR, and MR, we do not focus on a specific XR system. Instead, our study aims to investigate user privacy concerns raised by XR sensors that are generalizable to all three XR systems.

2.1 Privacy Risks in Extended Reality

XR uses body sensors to track user movements and generate visual, audio, and haptic feedback when interacting with the virtual environment [7]. Based on these data, inferences can be made about user's physical and mental conditions [47, 51] as well as habitual movements [41, 47, 56]. Cognitive, emotional, and personality issues can be estimated [7, 46, 51, 52]. To better use XR systems, trade-offs between providing users' biometric and demographic information to enable system functionality are inevitable [43]. All this information can further be used to deanonymize the user's identity [24, 47, 56], manipulate the user toward certain behaviours and buying decisions [7, 46, 51], and derive profit for the manufacturer and third-party companies [18, 46]. Furthermore, some "always-on" XR devices enable constant surveillance without user knowledge [2, 52, 60]. Often, users are unaware of the data collection and use [1, 22, 52]. As a result, they might lose control over the information they did not intend to reveal [51]. Additionally, eavesdropping and data breaches could also endanger data privacy [25].

Concerns regarding data tracking sensors and user privacy also arise in multi-user XR environments. While XR enhances communication, it also brings forth privacy risks. Within social VR, participants find greater comfort in discussing their emotions, personal information, and experiences because they believe they remain anonymous or engage with familiar individuals [43]. However,

[1] Pokémon GO official website. https://pokemongolive.com/

users face the possibility of being deanonymized [24, 47, 56] and deceived by digital avatars impersonating their loved ones or trusted friends [7]. Furthermore, because users are always in ecosystems with other users and bystanders, they may accidentally share private information without realizing that others can see it [36]. Thus, more personal data becomes open to misuse [51].

Furthermore, many studies identified the potential erosion of privacy through bystanders in XR. As unwilling participants in the immersive experiences of others, bystanders might face problems such as being unaware of data collection [52], being recorded without consent [51], being recognized unexpectedly by XR applications [36, 52], and being manipulated virtually (i.e., altering their visual appearance or placing unwanted virtual objects on top of their head) without their authorization [24, 36, 59].

2.2 Factors Influencing User Privacy Concerns

Although previous research has investigated potential privacy risks in XR technologies, most of them focused on technical aspects of the issues through threat modelling [25, 76], system functional requirements analyses [15, 45, 51], privacy policy examinations [2, 15], and expert discussions [1].

Only a few studies focused on privacy concerns from users' perspectives. Among these, many used participants' willingness to share information (measured by the *comfort of data collection*) as an indicator of privacy concerns [39, 43, 48, 72]. For instance, Maloney et al. [43] measured user privacy concerns based on their willingness to share information. They found that social VR users feel comfortable sharing personal information when perceiving anonymity and high familiarity with friends in the virtual environment. Adams et al. [2] revealed that VR users' privacy concerns often originate from "always on" sensors and the low reputation of device manufacturers. Harborth and Pape [28] suggested that AR users' privacy concerns are directly driven by perceived permission sensitivity (i.e., perceived sensitivity of the information to which they give permissions), trust in the application, and the general feeling of being a victim of privacy invasion online. Similarly, O'Brien [50] revealed that AR users might be comfortable with data collection depending on perceived data sensitivity and the purpose of use. Gallardo et al. [22] found that AR users expressed discomfort about what data were collected and who received the data.

Beyond XR, previous studies have identified various factors that could influence participants' privacy concerns in IoT, mobile, and on the web. For instance, Naeini et al. [48] demonstrated that users' comfort with data collection through IoT devices is significantly impacted by the data type, data collection location, and the purpose of data use. Lee and Kobsa [37] found that data collections that occurred in personal space, captured photo and voice data, by unknown entities or the government, are considered unacceptable. Hadan and Patil [27] found that data collections that are irrelevant to the device's primary functionality raised concerns. Regarding privacy concerns in mobile devices and applications, Tsai et al. [72] indicated that users were more willing to share data when they were informed about data recipients and when the data was not shared. Lin et al. [39] revealed that clarifying the purpose of data use reduces the perceived uncertainty of mobile applications and thus increases users' comfort with data sharing. Other studies [4, 35, 38] identified that privacy concerns varied by perceived ownership of data, data retention time, and perceived value of data. In our study, we consider these influential factors from prior research in IoT, mobile, and web contexts to examine their potential impact on user privacy concerns in XR.

Unlike IoT and mobile devices, XR sensors are capable of collecting more granular user data such as eyeball movements [18, 51, 56], pupil dilation [15, 56], and brain activity [15, 24, 41] that enabled more privacy-invasive user surveillance and inferences, and raised the discussion around mental privacy [24] or even neuro-rights [78]. However, most people do not understand how involuntary bodily indicators of emotional responses, mental state, or health can disclose fundamentally private information, such as truthfulness, inner feelings, and sexual arousal [30]. Thus, we see the need for investigating user data collection awareness and privacy concerns in XR.

2.3 Research Gap and Connection to our study

Although previous work has explored privacy issues in XR environments, little is known about privacy concerns from the perspectives of XR users. In addition, user-centred XR privacy studies have primarily focused on the requirements of specific XR systems (e.g., AR [28, 50, 52], VR [2, 22, 43]), making it necessary to investigate other devices that fall on the Reality-Virtuality Spectrum. Our research lays the groundwork and helps to understand the baseline of user privacy concerns and factors contributing to their comfort with data collection in XR environments. Given the similarity in XR sensors and data collection capabilities, we do not focus on specific XR systems, but aim to investigate users' privacy concerns generalizable across various XR systems.

We reviewed previous studies on user privacy concerns in XR [2, 22, 50, 52, 60], IoT [37, 48], mobile [39, 72], and on the web [38] and iteratively incorporated factors that impacted user privacy perception and preference in our study. We aim to determine whether these factors remained influential on user privacy concerns across various XR systems, and we validate experts' assumptions of privacy issues within XR from users' perspectives [1]. Given that privacy norms are context-dependent [49], we used a scenario-based approach inspired by prior work [28, 37, 48]. As the first attempt to seek essential insights into users' privacy concerns in XR, we narrow our scope by focusing on personal XR systems in single-user scenarios.

3 METHODOLOGY

To understand XR users' privacy concerns, we conducted a scenario-based online survey with 464 Prolific[2] participants who had prior experience with XR technologies, but ownership of XR devices was not a requirement. Our study received Research Ethics Board approval (REB#44772) from the University of Waterloo before recruiting participants. We presented each participant with four XR data collection scenarios, randomly selected from a total of 18 predesigned scenarios. The randomized selection allowed us to explore a wider range of scenarios without overwhelming participants with a long survey.

Privacy norms are dependent on context [49]. Scenarios are hypothetical situations in specific circumstances [21] that have been

[2]Prolific. https://www.prolific.co/

frequently used in prior studies to provide participants with contexts when evaluating privacy concerns [28, 48]. For this reason, we developed scenarios to support our investigation of user privacy concerns in XR. We incorporated factors that were found to influence user privacy perception and preference in XR [2, 50, 60], IoT [37, 48], mobile [39, 72], and on the web [38]. Our aimed to determine whether these factors remained influential in XR environments.

In the following sections, we describe the design of the scenarios, the development of the survey questions, and our participant recruitment process.

3.1 Scenarios Development

Drawing from previous studies on factors that potentially influence users' privacy perception in XR [2, 22, 50, 52, 60], IoT [37, 48], mobile [39, 72], and on the web [38], and discussions with our research team of privacy experts, we iteratively identified five high-level factors that are most likely to affect people's privacy concerns in XR technologies. Three are static factors (i.e., `USER_TYPE`, `LOCATION` and `DEVICE_TYPE`) that remain unchanged throughout all scenarios, and two are dynamic factors (i.e., `DATA_TYPE` and `DEVICE_STATUS`) that varied between scenarios.

Static Factors:

(1) **User Type** (`USER_TYPE`): the type of user being involved in data collection; Always set to the user (i.e., "you")
(2) **Location** (`LOCATION`): the location where the data collection occurs; Always set to a private space (i.e., "home / personal room")
(3) **Device Type** (`DEVICE_TYPE`): the type of technology that collects data; Always set to XR devices that includes VR, AR, and MR (i.e., "XR devices")

Dynamic Factors:

(4) **Data Type** (`DATA_TYPE`): The type of data collected by XR devices;
(5) **Device Status** (`DEVICE_STATUS`): Whether the data collection happens actively (i.e., when the user is interacting with the device) or passively (i.e., when the device is running in the background).

Two researchers with expertise in usable privacy reviewed and discussed these factors in multiple iterations, and developed more granular levels for each of the five factors identified above. We illustrate our process below.

3.1.1 User Type, Location, and Device Type (factors #1–3). Our aims to understand the baseline of user privacy concerns and factors that contribute to their comfort with data collection across various XR environments. While previous studies found privacy issues from both users' and bystanders' perspectives [24, 36, 51, 52] and in multi-user environments [43, 59], as discussed in subsection 2.3, we narrow our scope by focusing on personal XR systems in single-user scenarios. Thus, our data collection `LOCATION` was "Home/ personal room" for all scenarios, and the `USER_TYPE` was always "You", representing the user themselves. Furthermore, our study described data collection `DEVICE_TYPE` as "XR devices" as a whole and did not differentiate between specific XR systems in our scenarios. The decision was made to keep the number of scenarios manageable for the participants and to take into consideration that, in most cases, VR, AR, and MR sensors share similar data collection sensors. Although some privacy concerns, such as bystander effects, may be more likely to occur in AR, users with virtual avatar representations may become bystanders for other users in VR [14], and disconnecting from physical surroundings can cause VR users to be monitored by bystanders in the physical world without their knowledge [17]. Therefore, by treating XR systems as a cohesive unit, our study focuses on shared privacy concerns raised by common XR sensors and data practices, rather than focusing on the device type and manufacturer.

3.1.2 Device Status (factor #4). Previous research found that user privacy concerns could originate from their interaction with XR devices and devices with continuously operating sensors [2]. Thus, our factor `DEVICE_STATUS` encompasses two levels: *in-use* representing situations when the user actively interacts with the device, and *background-running* representing situations when the device is running in the background.

3.1.3 Data Type (factor #5). Several studies have tried to classify the data types that XR technology collects [15, 22, 52, 70]. The primary framework we chose for scenario design is a XR data classification from an article by Dick [15] because it clearly described each category of XR data, taking user privacy into account. Based on the framework, we incorporated data types previously examined in XR studies to ensure the inclusion of a diverse range of data types in our scenarios.

We initially compiled 26 data types that XR sensors could capture based on the framework from Dick [15]. These data types comprise user account information (e.g., demographics, billing address, phone number) [2, 15, 52], device information (e.g., crash report, model information, system logs) [15], user geo-location [25, 45, 51], user audio data [2, 18, 45], infrared camera data or user images [2, 18, 51, 52], user movements [13, 41, 47, 52], user visual attention [12, 18, 56], physical body dimensions [2, 18, 52], user surroundings [15, 18, 45, 52], and digital communication messages and friend list. Since our research aims to investigate XR-related concerns, we further refined the list of data types by excluding common data types such as user geo-location, facial features, payment information, and IP address that were already investigated in previous mobile, web, and IoT privacy studies. Thus, in our scenarios, we focused only on nine data types that are frequently collected by XR devices, but are less prevalent in other technologies. To enhance participants' comprehension of each data type, descriptive examples enclosed in "()" are also presented to them in the scenarios.

Overall, our `DATA_TYPE` factor includes nine levels in four data categories [15]:

a *Observable data:* data about an individual that third-parties can observe and replicate, which includes:
- Physical appearance (e.g., body dimensions) [2, 18, 51, 52];
- Identifying in-app/in-world assets (e.g., personal virtual objects, personal avatars) [15, 52];
- Environment information and dimensions (e.g., users' surroundings, room layout, device/user position in relation to the environment, user position in relation to device) [15, 18, 45, 52].

b *Observed data:* data that individual provides or generates, which third-parties cannot replicate, including:
- Physical movements and characteristics (e.g., head/hand/body/eye motion, orientation, position, gestures, posture, fitness information) [7, 13, 15, 24, 36, 41, 47, 52, 75];
- Physiological data (e.g., pupil and cornea reflections, brainwaves, skin signals) [7, 24, 36, 52, 75].

c *Computed data:* new data inferred by manipulating observable and observed data, including:
- Visual attention (e.g., eye gaze, area of interest, fixation, heatmaps, time to first fixation, time spent on a certain point, fixation sequences) [7, 12, 18, 52, 56, 75];
- Mannerisms (e.g., gait, habitual movements) [15, 41, 47, 56];
- Cognitive, emotional, and personality estimates (e.g., cognitive load, stress, depression, excitement, identity traits, etc.) [7, 46, 47, 51, 52]

d *Associated data:* data that, on its own, does not provide descriptive details about an individual, such as:
- Non-identifying virtual assets (e.g., in-app achievements) [15].

Finally, the combination of three static factors and two dynamic factors with nine data type categories and two device statuses resulted in a total of 18 scenarios shown in Table 1, developed using a scenario template:

*"You are at home/personal room, and your XR device (e.g., smart headset, touch controller, 3D projector) is keeping track of your [*DATA_TYPE*] when [*DEVICE_STATUS*]."*

3.2 Survey Design

3.2.1 Screening Questionnaire. The survey began with a study information sheet and consent form, followed by a screening questionnaire. We set our screening criteria to recruit XR users from North America and Europe, the leading regions in the XR market [32]. The North American region has been a market leader in early XR device adoption, with many well-known vendors, and the European Commission actively supports research and innovation in XR technologies. To participate, individuals had to be at least 18 years old and possess some familiarity using XR devices, but ownership of XR devices was not mandatory.

3.2.2 XR Understanding. To ensure that participants' privacy concerns are specifically related to their experience in XR environments, we first evaluated participants' understanding of VR, AR, and MR technologies. In a drag and drop pair-matching question, participants were asked to match the different descriptions of XR technologies to the corresponding definitions. After this exercise, we provided participants with the correct descriptions of VR, AR, and MR definitions from the XR Safety Initiative (XRSI) [75] to ensure that have an appropriate baseline understanding of different types XR technologies to answer the remaining questions in the survey.

3.2.3 Perceptions and Comfort Towards Data Collection. To assess participants' awareness of XR's data collection capability, they were asked to share what data they thought were being observed by XR devices based on their experience. Then, the participants were presented with four randomly selected scenarios from 18 scenarios (see Table 1). For each scenario, we adapt the methodology of Naeini et al. [48], where participants were asked to rate how likely the scenario is to happen today, within 2 years, and within 10 years, to assess how *realistic* participants think the scenarios are, with the scenarios perceived to happen today being the most realistic. We coded their responses as LIKELY_TODAY, LIKELY_IN_2YRS, and LIKELY_IN_10YRS. These questions were answered on 5-point Likert scales from *1-extremely unlikely* to *5-extremely likely*.

To prevent biasing participants' perceptions, we did not explicitly mention "privacy" or related terms when measuring their privacy concerns. Instead, we followed the approach used in previous studies (see subsection 2.2) to measure participants' comfort level with data collection (COMFORT_LEVEL) using a 5-point Likert-scale from *1-very uncomfortable* to *5-very comfortable*). Participants were also asked to rate the perceived sensitivity of the data type (DATA_SENSITIVITY) on a 5-point Likert-scale from *1-not sensitive* to *5-very sensitive*). For participants who expressed discomfort with the situation described in the scenario, we asked them to share their mitigation strategy in an open-ended question.

3.2.4 Experience and Privacy Concern. Finally, we inquired about participants' personal experience with XR devices, including how many years they have been using XR devices (EXP_YEARS), the type of devices used (EXP_DEVICE), the frequency of use (EXP_FREQUENCY), and the purpose for using these devices (EXP_PURPOSE). We measured participants' general level of privacy concern on the Internet using the Internet Users' Information Privacy Concerns Scale (IUIPC) [40], which consists of three dimensions of user privacy concerns: collection (IUIPC_COLLECTION), control (IUIPC_CONTROL), and awareness (IUIPC_AWARENESS). To ensure data quality, we included an attention check question mixed in the scenarios. We closed the survey with demographic questions. Figure 1 summarizes the survey flow. A complete set of questions is included in the Supplementary Materials.

3.3 Participants Recruitment and Demographics

We piloted the survey questionnaire with 15 students. Using their feedback, we improved the questions' wording and order to enhance understandability. A power analysis [19] for a Wilcoxon-signed rank test determined that we need a minimum of 57 participants for each pair of scenarios to achieve a power of $\geq 95\%$, allowing a margin for random error of $\leq 5\%$ with an estimated effect size=0.50. We deployed our survey on Prolific in December 2022 and received 549 responses with an average completion time of 13 minutes. We rejected 49 incomplete responses. The remaining 500 participants were remunerated US $5.8 each[3]. We excluded 23 responses that failed the attention check question and another 13 responses that were automatically identified as bots by Qualtrics. Our final analysis was based on the remaining $N = 464$ valid responses, with $n \approx 64$ for each pair of scenarios, providing sufficient data for assessing significant differences on participants' privacy concerns (measured as COMFORT_LEVEL) both within subjects.

[3]US $5.8 is about CAD $8. This remuneration rate is calculated based on Ontario local minimum hourly wage. See: https://www.immigrationwaterlooregion.ca/en/study-and-work/salary-standards-and-minimum-wage.aspx

Table 1: Summary of the 18 scenarios used in the study. ID= Scenario identifier, n = number of participants presented with each scenario. Each scenario was presented to approximately 103 participants on average.

ID		Scenarios
	n	*All scenarios begin with*:
Scenario:Type-Status		"You are at home/personal room, and your XR device (e.g., smart headset, touch controller, 3D projectors) is keeping track of your.."
S:Appearance-Use	99	Physical appearance (e.g., body dimensions) when you use the device.
S:Appearance-Bg	107	Physical appearance (e.g., body dimensions) when the device is running in the background.
S:Identifying-Use	105	Identifying in-app/in-world assets (e.g., personal virtual objects, personal avatars) when you use the device.
S:Identifying-Bg	103	Identifying in-app/in-world assets (e.g., personal virtual objects, personal avatars) when the device is running in the background.
S:Environment-Use	102	Environment information and dimensions (e.g., users' surroundings, room layout, device/user position in relation to the environment, user position in relation to device) when you use the device.
S:Environment-Bg	107	Environment information and dimensions (e.g., users' surroundings, room layout, device/user position in relation to the environment, user position in relation to device) when the device is running in the background.
S:Movements-Use	106	Physical movements and characteristics (e.g., head/hand/body/eye motion, orientation, position, gestures, posture, fitness information) when you use the device.
S:Movements-Bg	103	Physical movements and characteristics (e.g., head/hand/body/eye motion, orientation, position, gestures, posture, fitness information) when the device is running in the background.
S:Physiological-Use	104	Physiological data (e.g., pupil and cornea reflections, brainwaves, skin signals) when you use the device.
S:Physiological-Bg	102	Physiological data (e.g., pupil and cornea reflections, brainwaves, skin signals) when the device is running in the background.
S:Visual-Use	99	Visual attention (e.g., eye gaze, area of interest, fixation, heatmaps, time to the first fixation, time spent on a certain point, fixation sequences) when you use the device.
S:Visual-Bg	101	Visual attention (e.g., eye gaze, area of interest, fixation, heatmaps, time to the first fixation, time spent on a certain point, fixation sequences) when the device is running in the background.
S:Mannerisms-Use	100	Mannerisms (e.g., gait, habitual movements) when you use the device.
S:Mannerisms-Bg	106	Mannerisms (e.g., gait, habitual movements) when the device is running in the background.
S:Cognitive-Use	104	Cognitive, emotional, and personality estimates (e.g., cognitive load, stress, depression, excitement, identity traits, etc.) when you use the device.
S:Cognitive-Bg	102	Cognitive, emotional, and personality estimates (e.g., cognitive load, stress, depression, excitement, identity traits, etc.) when the device is running in the background.
S:Non-identifying-Use	102	Non-identifying virtual assets (e.g., in-app achievements) when you use the device.
S:Non-identifying-Bg	107	Non-identifying virtual assets (e.g., in-app achievements) when the device is running in the background.

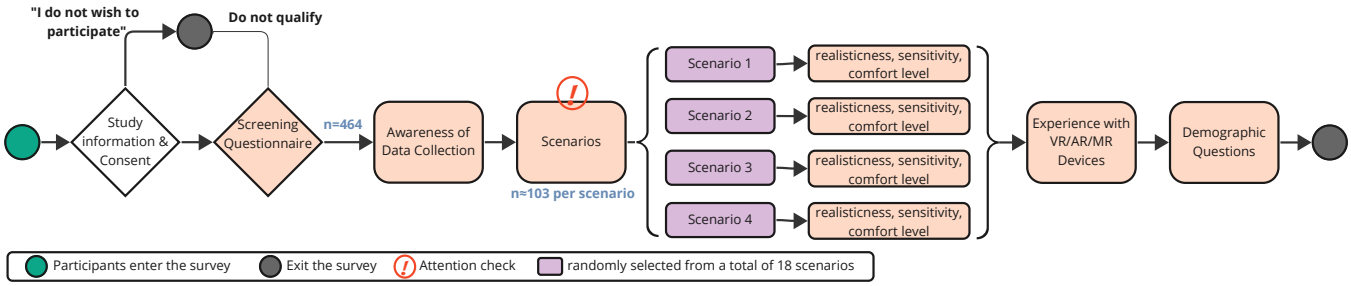

Figure 1: Survey flow: Upon passing the screening questionnaire, 464 participants were asked about their understanding of XR and awareness of data collection. Next, they were presented with four randomly selected scenarios from a total of 18 scenarios from Table 1, resulting in approximately 103 responses per scenario. For each scenario, participants answered questions relating to comfort with data collection, perceived data sensitivity, and how realistic they perceived the scenario to be. Finally, the participants shared their experience with XR devices and demographic information.

We summarize participants' demographic background in Table 2. Our 464 participants consist of 50% men, 48% women, 2% non-binary/third gender, and less than 1% chose not to disclose their gender. The majority of our participants, aged between 20 and 29, possessed an Associate degree or higher (Associate degree = 6%, Bachelor's degree = 34%, Graduate or professional degree = 20%), with an annual income of less than $50,000 USD (Less than $25,000 = 39%, $25,000 to $49,999 = 32%) and displayed relevantly high IUIPC scores (see Table 2). The participants came from 21 countries in North America and Europe (see Table 6 in Supplementary Materials).

The majority of the participants (94%) used XR devices for entertainment purposes. Regarding their experience with XR devices, most of the participants had less than five years of experience using XR devices (Less than 6 months = 40%; 6 months to 1 year = 26%; 1 to 5 years = 36%). In terms of frequency, a small number of participants engaged with XR devices at least once a day (5%), while others used them once a week (25%) or once a month (26%). When it comes to specific types of XR devices, the majority of participants (85%) had experience with smart headsets or glasses, and most (58%) had experience with handheld devices (see Table 6 in Supplementary Materials).

Table 2: Breakdown of our participants' (N = 464) gender, age, education level, income level, and three IUIPC dimensions. A detailed overview of participants' demographic is located in Table 6 in Supplementary Materials.

Gender		Age		Education Level		Income		IUIPC factors	
Man	230 (50%)	Range	18-56	No high school	4 (<1%)	Less than $25,000	182 (39%)	*IUIPC_Collection*	
Woman	222 (48%)	Mean (SD)	27.78 (7.95)	High school	173 (37%)	$25,000-$49,999	151 (33%)	Range	(1.25 - 7)
Non-binary	11 (2%)			Associates	30 (7%)	$50,000-$99,999	86 (19%)	Mean (SD)	5.69 (1.14)
Prefer not to say	1 (<1%)			Bachelors	162 (35%)	$100,000-$199,999	19 (4%)	*IUIPC_Control*	
				Professional	94 (20%)	More than $200,000	5 (1%)	Range	(3 - 7)
				No answer	1 (<1%)	Prefer not to say	21 (5%)	Mean (SD)	5.78 (0.92)
								IUIPC_Awareness	
								Range	(3 - 7)
								Mean (SD)	6.29 (0.74)

3.4 Data Analysis

We demonstrate the details of our quantitative data analysis while introducing our results in section 4. For the open-ended question, we analyzed the responses using inductive thematic analysis with two researchers [6]. We started by familiarizing ourselves by scanning through the data and excluding blank or incoherent responses. We kept responses such as "none" or "nothing" because having no way to mitigate privacy concerns was still a valid response (see numbers of included responses for each question in Figure 4). The two researchers first coded approximately 10% of the data individually, then discussed and resolved conflicts. This process was repeated twice until the codebook was finalized and all remaining data were coded. Using the codebook, the researchers developed and refined the themes. A summary of themes and codes is shown on Table 5.

4 FINDINGS

4.1 RQ1: How much are people aware of the data collected through XR devices?

Figure 2 illustrates the overall response distribution among the N = 464 participants regarding the types of data they believed XR devices could capture. The majority of participants (84%) acknowledged that XR devices can collect *Physical movements and characteristics*, while 74% recognized the capture of *Environment information and dimensions* through XR devices. Furthermore, a large proportion of participants realized that their *Visual attention* (65%), *Identifying in-app/in-world assets* (48%), and *Non-identifying virtual assets* (46%) are getting observed by the XR devices. Some participants also believed that XR devices observe their *Mannerisms* (32%), *Physical appearance* (26%), and *Physiological data* (18%). Only a small number of participants (11%) thought that their *Cognitive, emotional, and personality estimates* could be observed. A few participants also mentioned *Other* information, such as "purchases and utilization of assets" (≤ 1%), while a small percentage (1%) were uncertain about what data their devices might observe.

4.2 RQ2: To what extend are people concerned about their privacy regarding data collection in XR?

The distribution of participants' `COMFORT_LEVEL` across various data types and device statuses are shown in Figure 3. To validate the observed differences in participants' comfort level, we performed a set of pairwise Wilcoxon-signed-rank comparisons [74] with Bonferroni correction [8] (see Table 7 and Table 8 in Supplementary Materials).

Among all scenarios, participants expressed relevantly strong discomfort when the data type associated with the scenario was *Cognitive, emotional, and personality estimates*, with 29% of participants selected very uncomfortable and 35% selected uncomfortable. This difference in participants' perceived comfort level was significant ($P \leq .007$) across all data types (see Table 7). Moreover, this particular data type was considered the most sensitive and the most unlikely to happen today, with 47% of participants selecting very sensitive and 40% selecting extremely unlikely (see Figure 3). This difference in participants' perceived data sensitivity and realisticness were significant ($P < .001$) across all data types as well. Compared to other data types, participants exhibited a significant ($P \leq .005$) higher level of comfort when the scenario involved *Non-identifying virtual assets*, as only 9% of participants selected unco'mfortable and 3% selected very uncomfortable. This data type was also perceived as the least sensitive among the others, with 43% of participants indicating that it was not sensitive. This difference is also significant ($P < .001$) across all data types.

Furthermore, a slightly greater number of participants expressed discomfort, when the device operates passively in the background (18% selected very uncomfortable, and 27% selected uncomfortable). On the other hand, a slightly greater number of participants perceived higher level of data sensitivity and realism when scenarios involving active device interaction, as 27% and 21% of participants selected sensitive and very sensitive, and 30% and 17% of participants selected likely and extremely likely (see Figure 3). These differences are also significant ($P \leq .005$) based on our pairwise analyses (see Table 8).

4.3 RQ3: What factors contribute to user privacy concerns in XR?

This section summarizes participants' perceptions of how realistic the scenarios are, their general privacy concerns on the Internet, and their experience with and frequency of using XR devices. We further compare the influence of these factors on participants' comfort levels towards data collection in XR.

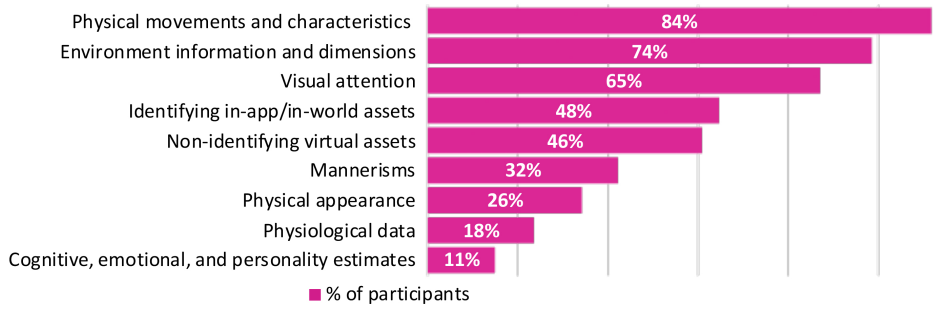

Figure 2: Percentage of participants' (N = 464) perceived data collection of various data types through XR devices.

Data Type	1 - Very uncomfortable	2 - Uncomfortable	3 - Neutral	4 - Comfortable	5 - Very comfortable	1 - Not at all sensitive	2 - Slightly sensitive	3 - Somewhat sensitive	4 - Sensitive	5 - Very sensitive	1 - Extremely unlikely	2 - Unlikely	3 - Neutral	4 - Likely	5 - Extremely likely
Physical appearance	14%	34%	25%	21%	5%	11%	16%	21%	32%	21%	18%	26%	21%	24%	10%
Identifying in-app/in-world assets	7%	17%	30%	32%	13%	25%	27%	23%	18%	8%	11%	14%	22%	30%	23%
Environment info & dimensions	10%	27%	35%	22%	6%	7%	22%	23%	30%	17%	13%	17%	20%	34%	15%
Physical movements/characteristics	14%	23%	26%	28%	9%	12%	20%	25%	25%	18%	11%	16%	21%	36%	16%
Physiological data	21%	30%	33%	14%	2%	6%	13%	20%	33%	28%	30%	28%	19%	19%	4%
Visual attention	11%	28%	30%	23%	9%	14%	24%	28%	23%	12%	13%	28%	19%	27%	14%
Mannerisms	14%	31%	32%	19%	4%	10%	19%	29%	27%	15%	19%	24%	19%	33%	5%
Cognitive/emotional/personality	29%	35%	18%	15%	4%	2%	7%	13%	31%	47%	40%	30%	17%	12%	2%
Non-identifying virtual assets	3%	9%	33%	32%	24%	43%	20%	19%	11%	8%	10%	7%	17%	27%	39%
Devices Status															
Running in the background	18%	27%	27%	19%	8%	15%	21%	23%	24%	17%	21%	24%	21%	24%	11%
In-use	9%	24%	32%	27%	9%	14%	16%	22%	27%	21%	15%	18%	18%	30%	17%

Figure 3: Percentage of participants' (n = 464) comfort level, perceived data sensitivity and perceived realism of a scenario between the data type and the device status. Cells with a higher percentage have a darker background colour. For instance, 29% of participants selected very uncomfortable when their *cognitive, emotional, and personality estimates* were being observed by the device.

4.3.1 Factors impacting participants' comfort level. To evaluate the relationship between participants' comfort levels and various factors, we employed Cumulative Link Mixed Model (CLMM) regression[4], incorporating a random intercept for each participant as part of our models. CLMM regression proves particularly beneficial for analyzing repeated measures experiments with ordinal dependent variable such as our study, where participants were presented with multiple parallel scenarios [9].

We first conducted a series of Univariate CLMM regressions. Each regression employed the participants' `COMFORT_LEVEL` as the dependent variable (DV) and one of the factors as the predictor. The purpose was to ascertain the existence of relationships and select predictors for our multivariate model. As shown in Table 3, the results indicated significant relationships between participants' `COMFORT_LEVEL` (DV) and predictors such as `DATA_TYPE` (P<.001), `DEVICE_STATUS` (P<.001), perceived `DATA_SENSITIVITY` (P<.001), `LIKELY_TODAY/IN_2YRS/IN_10YRS` (P<.001), general privacy concerns on the Internet (IUIPC scores P<.001). No significance was found regarding participants' demographic information or previous XR experience.

We further conducted a multivariate CLMM regression using participants' `COMFORT_LEVEL` as the dependent variable (DV), with significant predictors identified in the univariate CLMM regressions (see Table 3). To obtain the best-fit model, we utilized the backwards elimination method, which involved starting with a full model containing all variables and sequentially eliminating the variable with the highest p-value in each step until reaching the global minimum Akaike information criterion (AIC) [33]. The final multivariate model, presented in Table 4, was derived from this process. We arranged the predictors in descending order based on their contribution to the `COMFORT_LEVEL`, as determined by the AIC values obtained in the univariate CLMM (where each model contained only one predictor and the random intercept). Predictors were ranked with the highest contribution (lowest AIC in univariate CLMM) appearing first. Positive estimates (or OR>1) indicate a tendency towards comfort, while negative estimates (or OR<1) indicate a tendency towards discomfort.

[4]We used the Ordinal R-package (https://cran.r-project.org/web/packages/ordinal/) for modeling participants' comfort level.

Table 3: Univariate Cumulative Linked Mixed Model analyses of factors impacting participants' comfort level (5-point Likert data, from 1-very uncomfortable to 5-very comfortable), with a random intercept per participant. Ordinal data are used as is.

Predictor						**AIC**
data_type	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5185.92
Non-identifying virtual assets	Reference					
Identifying in-app/in-world assets	-0.84	0.20	-4.13	<.001***	0.43 (0.29, 0.64)	
Visual attention	-1.67	0.21	-8.05	<.001***	0.19 (0.13, 0.28)	
Physical movements and characteristics	-1.45	0.20	-7.14	<.001***	0.23 (0.16, 0.35)	
Mannerisms	-2.09	0.21	-10.10	<.001***	0.12 (0.08, 0.19)	
Environment information and dimensions	-1.72	0.20	-8.57	<.001***	0.18 (0.12, 0.27)	
Physical appearance	-2.10	0.21	-10.13	<.001***	0.12 (0.08, 0.18)	
Physiological data	-2.54	0.21	-12.08	<.001***	0.08 (0.05, 0.12)	
Cognitive, emotional, and personality estimates	-3.17	0.22	-14.41	<.001***	0.04 (0.03, 0.07)	
device_status	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5442.24
In-use	Reference					
Running in the background	-0.55	0.09	-5.94	<.001***	0.58 (0.48, 0.69)	
data_sensitivity	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	4442.67
1 - Not sensitive	Reference					
2	-2.25	0.20	-11.09	<.001***	0.11 (0.07, 0.16)	
3	-3.57	0.22	-16.52	<.001***	0.03 (0.02, 0.04)	
4	-5.12	0.24	-21.64	<.001***	0.01 (0.00, 0.01)	
5 - Very sensitive	-6.74	0.27	-24.72	<.001***	0.00 (0.00, 0.00)	
likely_today	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	4984.37
1 - Extremely unlikely	Reference					
2 - Unlikely	0.40	0.18	2.26	.02*	1.49 (1.05, 2.11)	
3 - Neutral	1.21	0.19	6.46	<.001***	3.34 (2.32, 4.82)	
4 - Likely	2.22	0.19	11.85	<.001***	9.17 (6.36, 13.23)	
5 - Extremely likely	4.12	0.23	17.61	<.001***	61.44 (38.85, 97.15)	
likely_in_2yrs	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5097.70
1 - Extremely unlikely	Reference					
2 - Unlikely	0.61	0.34	1.81	.07	1.84 (0.95, 3.58)	
3 - Neutral	1.21	0.34	3.59	<.001***	3.37 (1.73, 6.53)	
4 - Likely	2.37	0.33	7.18	<.001***	10.68 (5.59, 20.40)	
5 - Extremely likely	3.75	0.34	10.94	<.001***	42.31 (21.62, 82.73)	
likely_in_10yrs	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5305.95
1 - Extremely unlikely	Reference					
2 - Unlikely	-0.26	0.51	-0.50	0.62	0.77 (0.28, 2.11)	
3 - Neutral	0.58	0.49	1.18	0.24	1.78 (0.69, 4.62)	
4 - Likely	1.10	0.47	2.35	.02*	3.01 (1.20, 7.54)	
5 - Extremely likely	2.31	0.47	4.91	<.001***	10.04 (4.00, 25.22)	
IUIPC scores	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5409.70
IUIPC_CONTROL	-0.34	0.08	-4.45	<.001***	0.71 (0.61, 0.83)	
IUIPC_AWARENESS	-0.36	0.10	-3.80	<.001***	0.70 (0.58, 0.84)	
IUIPC_COLLECTION	-0.50	0.06	-8.37	<.001***	0.61 (0.54, 0.68)	
Age	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5477.09
	0.01	0.01	0.85	0.40	1.01 (1.00, 1.03)	
Education	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5479.58
Less than high school	Reference					
High school	1.89	0.82	2.32	0.12	6.65 (1.34, 32.94)	
Associate's degree	2.11	0.86	2.46	0.28	8.26 (1.53, 44.40)	
Bachelor's degree	1.68	0.82	2.06	0.12	5.38 (1.08, 26.65)	
Graduate's degree	1.77	0.82	2.15	0.25	5.86 (1.17, 29.42)	
exp_years	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5479.80
Less than 6 months	Reference					
6 months to 1 year	0.37	0.19	1.96	0.05	1.45 (1.00, 2.11)	
1 year to 5 years	0.09	0.17	0.53	0.60	1.09 (0.79, 1.51)	
5 to 10 years	0.25	0.54	0.46	0.65	1.28 (0.45, 3.67)	
More than 10 years	0.29	1.11	0.26	0.80	1.33 (0.15, 11.75)	
exp_frequency	**Estimates**	**Std. Error**	**z**	***P* value**	**OR (95%CI)**	5470.92
Once a year	Reference					
At least once every six months	-0.25	0.27	-0.92	0.36	0.78 (0.46, 1.32)	
At least once every three months	0.27	0.26	1.04	0.30	1.31 (0.79, 2.16)	
At least once a month	0.13	0.23	0.57	0.57	1.14 (0.73, 1.78)	
At least once a week	0.26	0.23	1.15	0.25	1.30 (0.83, 2.03)	
At least once a day	1.12	0.36	3.08	0.07	3.06 (1.50, 6.23)	

*Note. Significance are displayed as: *** P<.001, ** P<.01, * P<.05. OR=Odds Ratio, CI=Confidence Interval. The Reference categories were selected to enhance result interpretability. For OR, a value greater than 1 indicates a positive relationship, and a value less than 1 indicates a negative relationship.*

As illustrated in Table 4, the DATA_SENSITIVITY described in the Scenarios had the greatest negative effect on participants' COMFORT_LEVEL with data collection (*P<.001*). Participants' COMFORT_LEVEL was positively influenced by the perceived likelihood of happening today (LIKELY_TODAY, *P<.001*), indicating a negative association between perceived data sensitivity and participants' comfort, as well as a positive association between the perceived realisticness of the scenario and their comfort with data collection.

In addition, not all data types yielded statistically significant results. For instance, compared to Scenarios where *Non-identifying virtual assets* were being observed, participants had lower COMFORT_LEVEL when Scenarios collecting *Visual attention* (*P=.007*), *Mannerisms* (*P=.002*), *Physical appearance* (*P=.02*), and *Physiological data* (*P=.02*). That is, participants were more likely to express discomfort when these data types were being collected. On the other hand, we did not identify significance from Scenarios in which *Identifying in-app/in-world assets*, *Physical movements and characteristics*, and *Environment information and dimensions* were observed. This is in line with our results in Figure 3, where participants demonstrated a relevantly neutral COMFORT_LEVEL when the scenario involved the collection of these three data types, possibly due to their unclear (i.e., evenly distributed) DATA_SENSITIVITY and their lack of understanding of the associated privacy implication. We found evidence from our qualitative data, where more participants sought detailed information about the data collection and surrounding privacy laws ("comprehend and control what happens to my data") when the scenarios involved these three data types (see Figure 4: IDENTIFYING-USE, IDENTIFYING-BG, MOVEMENTS-USE, MOVEMENTS-BG, ENVIRONMENT-USE, and ENVIRONMENT-BG).

Participants expressed a strong discomfort with scenarios involving *cognitive, emotional, and personality estimates* and this particular type of data was also considered the most sensitive. In the participants' qualitative responses, they also mentioned that the collection of these types of data is *"too much intrusion"* compared to other types of data. The discomfort may also be related to the trust issue with data recipients and perceived purpose of data use [22, 52], as some participants sought to understand data practices and expressed trust issues with the manufacturer and the device system in their qualitative responses (see Figure 4). However, we did not find a significant difference from scenarios involving these data types. We believe this because while *cognitive, emotional and personality estimates* are perceived as invasive, scenarios involving the data types were also perceived as the least likely to occur today, providing a possible explanation for the lack of significance in our CLMM. This suggest that if the participants perceive the scenario is unlikely to happen, it may contribute to downplaying the perceived privacy risk [22, 48].

Lastly, we found that participants' COMFORT_LEVEL was negatively affected by their IUIPC_COLLECTION score (*P<.001*), suggesting that their concerns about the fairness of data collection (i.e., cost and benefits) was negatively impacting their comfort level. We also identified a negative relationship between participants' COMFORT_LEVEL and the DEVICE_STATUS (*P<.001*), suggesting that participants were more likely to express discomfort when the device is running in the background. These results correspond to our qualitative analysis, where many participants expressed concerns about perceived privacy risks, such as fear of surveillance, use of data for other than the stated purpose (misuse), and lack of trust in how sensitive information would be handled by entities collecting or processing it (see Section 4.4).

4.4 RQ4: What coping strategies (if any) do people use to mitigate their privacy concerns?

This section further discusses the themes derived from the participants' qualitative responses (Figure 4) and how they relate to the corresponding quantitative responses. For clarity, themes are placed within quotations and participants' quotes are set in *italics*. We use a percentage range (e.g., 7% ∼ 28%) to represent diverse proportions of participants that mentioned a same theme across different scenarios, and we use a single percentage (e.g., 3%) for themes mentioned in a specific scenario or by a specific participant group. The questions and themes can be found in the Supplementary Materials (Q5 and Table 5).

Across all 18 Scenarios, a large portion of the participants expressed an inability to mitigate discomfort towards potential data collection and stated their intention to continue using the device ("ignore or give up"). From their qualitative responses, we identified that these participants were pessimistic about their privacy within the XR environment and had become accustomed to the absence of privacy.

> *"Technology will overcome us. I think I would feel strange at first, but eventually I would get used to, even tho I would rather not share this information with unknowned entities. We don't have privacy anymore, I realize that."*
>
> *— Participant 313*

Notably, for Scenarios that involve the same data type, a higher proportion of participants indicated their willingness to disregard potential data collection when actively using the device. This result aligns with our findings in Figure 3, where a higher number of participants expressed discomfort when the device operated passively in the background. In addition, a significantly lower number of participants expressed discomfort when *Non-identifying virtual assets* were collected in the Scenario ($n = 11$ in S:NON-IDENTIFYING-USE, $n = 14$ in S:NON-IDENTIFYING-BG). Most of these participants (73% in S:NON-IDENTIFYING-USE and 71% in S:NON-IDENTIFYING-BG) expressed the tendency to "ignore and give up". This finding could be explained by our quantitative results, where *Non-identifying virtual assets* were perceived as the least sensitive.

While the majority of participants opted to *"give up"*, some simultaneously claimed that nothing can be done *"unless I'm aware of data collection."* This suggests that raising awareness and transparency about XR data collection might encourage participants to safeguard their privacy in XR. In addition, 29% of participants mentioned about ignoring the discomfort when their *Environment info & dimensions* were collected by an XR device running in the background in S:ENVIRONMENT-BG because, in contrast to other data types, participants perceived the devices that track environmental data as easily containable. Most of them (51%) suggested isolating the device by placing it in a designated and non-personal room (*"limit the data collection"*).

Table 4: Multivariate Cumulative Linked Mixed Model analyses of factors impacting participants' comfort level (1-very uncomfortable to 5-very comfortable), with a random intercept per participant. Ordinal data are used as is.

Predictor	Estimates	Std. Error	z	*P* value	OR (95%CI)
data_sensitivity					
1 - Not sensitive	Reference				
2	-1.83	0.21	-8.66	<.001***	0.16 (0.11, 0.24)
3	-2.92	0.23	-12.93	<.001***	0.05 (0.04, 0.08)
4	-4.27	0.25	-17.13	<.001***	0.01 (0.01, 0.02)
5 - Very sensitive	-5.66	0.29	-19.54	<.001***	0.00 (0.00, 0.01)
likely_today					
1 - Extremely unlikely	Reference				
2 - Unlikely	0.29	0.19	1.54	.13	1.33 (0.92, 1.92)
3 - Neutral	0.81	0.20	4.09	<.001***	2.25 (1.53, 3.33)
4 - Likely	1.44	0.20	7.22	<.001***	4.22 (2.85, 6.23)
5 - Extremely likely	2.52	0.25	10.10	<.001***	12.44 (7.63, 20.31)
data_type					
Non-identifying virtual assets	Reference				
Identifying in-app/in-world assets	-0.25	0.22	-1.17	.24	0.78 (0.51, 1.19)
Visual attention	-0.61	0.23	-2.71	.007**	0.54 (0.35, 0.84)
Physical movements and characteristics	-0.25	0.22	-1.10	.27	0.78 (0.51, 1.21)
Mannerisms	-0.72	0.23	-3.17	.002**	0.49 (0.31, 0.76)
Environment information and dimensions	-0.38	0.22	-1.71	.09	0.69 (0.45, 1.06)
Physical appearance	-0.55	0.23	-2.42	.02*	0.58 (0.37, 0.90)
Physiological data	-0.56	0.23	-2.38	.02*	0.57 (0.36, 0.91)
Cognitive, emotional, and personality estimates	-0.47	0.25	-1.85	.06	0.63 (0.38, 1.03)
IUIPC scores					
IUIPC_COLLECTION	-0.38	0.09	-4.26	<.001***	0.68 (0.57, 0.82)
IUIPC_CONTROL	-0.22	0.11	-2.02	.06	0.80 (0.65, 0.99)
device_status					
In-use	Reference				
Running in the background	-0.40	0.10	-3.86	<.001***	0.67 (0.55, 0.82)

*Note. final AIC = 4244.43. Significance are displayed as follows: *** P<.001, ** P<.01, * P<.05. OR=Odds Ratio. CI=Confidence Interval. The Reference categories were selected to enhance result interpretability. For OR, a value greater than 1 indicates a positive relationship, and a value less than 1 indicates a negative relationship.*

In fact, 7% ∼ 51% of participants proposed the idea of *"limiting the data collection"* as the strategy to safeguard their privacy. This encompassed actions like disabling device sensors, placing devices in designated rooms with low privacy concerns, disconnecting or unplugging the device, uninstalling privacy-invasive applications, adjusting privacy settings, or installing specialized applications to intervene in the data collection processes. Some participants engaged in self-censorship behaviours, such as using the device only when necessary, avoiding regular mannerisms, or suppressing genuine emotions to prevent monitoring.

> *"I would probably change my mannerism and try to limit my displays of any sort of emotions from time to time to feel like my 'real me' is not being monitored."*
>
> *— Participant 215*

Moreover, some participants (7% ∼ 28%) expressed their intention to stay "away from the device," either by discontinuing its use or disposing of the device. A few participants (3%) also indicated that they would have refrained from purchasing such devices had they been aware of the data collection beforehand. Conversely, six participants (1%) weighed the device's benefits and the potential compromise to their privacy when deciding to purchase or use it.

Furthermore, ∼ 9% of the participants wanted greater and more effective control over their data. Some mentioned actively reading user agreements, privacy policies, data privacy laws, and online articles to seek detailed information about the collected data. Others reported exploring options to reduce the amount and type of data being collected and methods to delete their data from the device. They emphasized that the data should be anonymized and only used to benefit users' lives and experiences, and deemed data use for targeted advertising or behaviour manipulation inappropriate. In addition, ∼ 8% of participants emphasized the importance of informed consent. In general, participants expressed a desire to understand what is being collected, how it is used, and with whom their data is shared.

> *"I don't like the idea of my personal data being collected by other companies for their own profit. Show the users that data is only being used for their interest and not with second intentions."*
>
> *— Participant 163*

Lastly, ∼ 18% of the participants prioritized their privacy over all other XR features and stated their intention to switch to alternative models or manufacturers that value and protect their privacy. On the contrary, ∼ 2% leaned towards completely trusting the device manufacturer and are willing to accept any data-related processes.

5 DISCUSSION

We first compare our study with privacy studies on other types of devices. We then discuss the implications of our findings to future

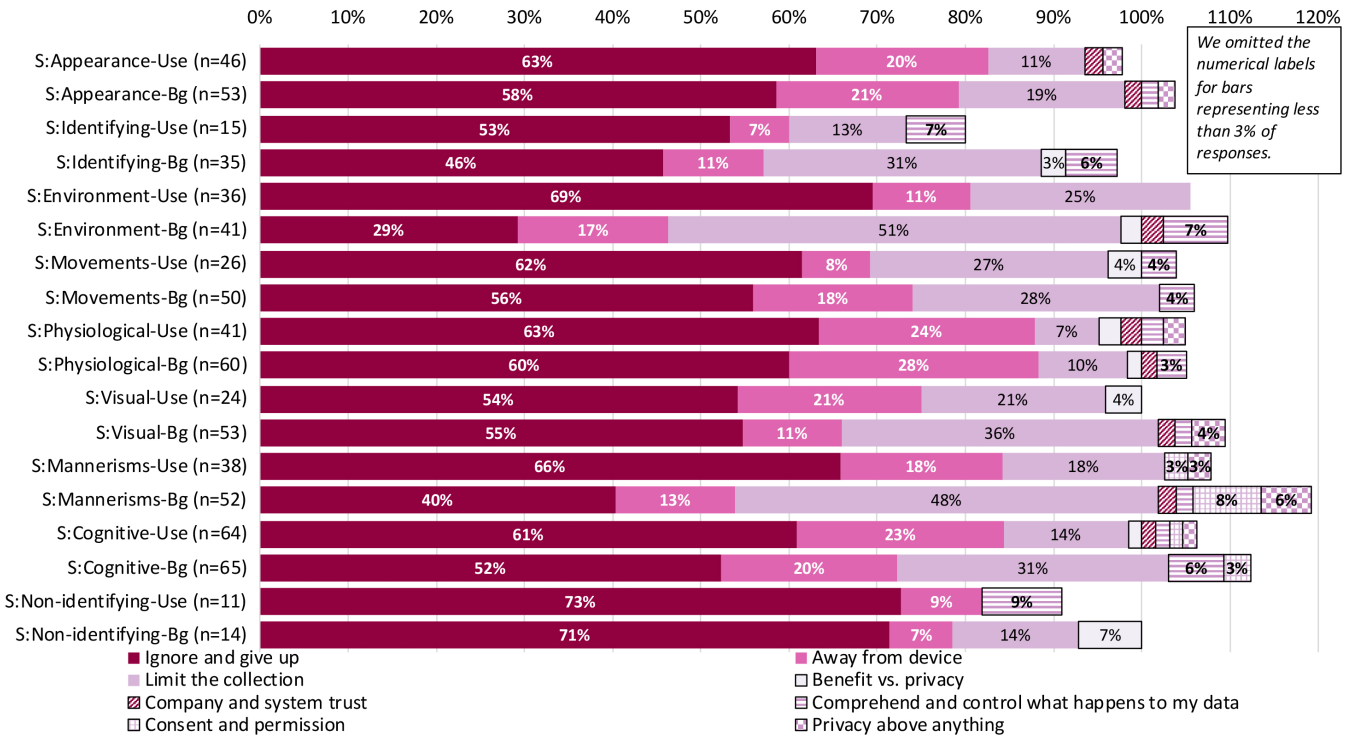

Figure 4: n= number of participants who responded to this question. Summary of participants' responses to the conditional open-ended question "What would you do (if any) to mitigate the discomfort raised?" (Q5) for each Scenario. The question was only shown to participants who selected *1-very uncomfortable* or *2-uncomfortable* regarding data collection (Q3) in each Scenario. Total percentage > 100% since a participant might mention multiple coping strategies.

work, including the importance of addressing users' unawareness of XR data privacy threats and passive XR data collection processes, the need for designing privacy-choice interfaces tailored to XR environments and implementing privacy-friendly default settings to reduce user burden, and the necessity to enforce and develop transparent XR data practices through XR product design and new regulations.

5.1 Similarities and Differences Between XR Privacy and Other Types of Devices

Many of our findings supported the results from prior work in IoT (e.g., [37, 48, 53]), mobile (e.g., [39, 72]), and web privacy studies (e.g., [38]), suggesting a convergence in users' privacy concerns regardless of settings. For instance, our regression analysis revealed that users' privacy concerns (comfort with data collection) were driven by their perceived data sensitivity (e.g., [28, 50]), data type (e.g., [48]), perceived equity of data collection (i.e., the IUIPC_collection [28]), and their perceived realism of the scenario (e.g., [48]). Our study also found a significant association between users' privacy concern and device status, aligning with concerns regarding "always-on" devices from Adams et al. [2] and Roesner et al. [60]. Although our participants' demographic information and experience with XR devices do not affect their privacy concerns, we suspect it is primarily due to their lower exposure with new commercial XR technologies.

In addition, our qualitative analysis showed that participants mainly agreed on several privacy-seeking strategies frequently mentioned in non-XR environments. For example, our participants reported self-censorship in response to the presence of nearby XR devices, such as a "chilling effect" of being more conscious of their speech and emotions [36, 51]. Their reported strategy of removing the device or blocking sensors echoes previous findings on personal space control challenges in AR [36], VR threats of recording devices in private places [51], AR bystanders' interests in physical or technical measures [14], and users' strategies in dealing with IoT surveillance [53].

Pessimistically, there was a strong agreement among our participants on the difficulty of protecting their privacy in XR. Many participants felt a sense of privacy resignation because they believed that they had already lost their privacy on other devices. On the other hand, stopping using of the device was embraced by a notable group of participants who value personal privacy more than the innovative functionality that XR provides. The underlying core agreement was the assumption that their data is collected to support other commercial activities, such as targeted advertising. This assumption was confirmed by a network traffic analysis on Oculus VR [71]. While infrequently mentioned, our qualitative responses identified participants' trust issues with the manufacturer and the XR system, and their desire to comprehend XR data use and to give explicit consent on XR data collection, echoing findings in IoT studies [28, 48].

While we found that XR privacy has some of the same concerns as other types of technology, we suggest that interaction paradigms and privacy interfaces users hold in 2D applications will need to be adapted and expanded.

5.2 Implications for Future Work

5.2.1 Data Transparency and Awareness in XR. Unlike web, mobile, and some IoT devices, XR collects users' physical, physiological, environmental, and even emotional estimates to produce the immersive experience [11]. Data are often collected without explicit notice or consent [71]. Although our participants were generally comfortable with XR data collection, which was slightly higher than the comfort level with IoT devices in previous studies (e.g., [48]), we believe it is in part due to their limited awareness of the types, uses, and management of their XR data [22, 52, 73]. For example, only a few participants recognized that physiological and physical data could be observed and mannerisms and cognitive, emotional, and personality estimates could be inferred from other data (see Figure 2). These data types could serve as a behavioural identifier [24, 47, 56] and could make users vulnerable to privacy risks from emotional manipulations [46]. The lack of awareness further leads to misconceptions between perceived risks and benefits (i.e., privacy calculus [16]).

Since many users are new to XR technologies, it is crucial for product and policymakers to ensure that people are aware of various types of data used by XR technologies in 3D environments that are not generally required for 2D web and mobile-based applications. In fact, since we found the types data that are least understood (e.g., cognitive, emotions, and personality estimates) are not generally used on non-XR devices, we suspect that many users rely on their privacy experiences on non-XR devices to make privacy decisions in XR environments. While users' experience with other types of technology may make them more familiar with some types of data collection—such as location tracking and voice recording—data types like movement tracking may be less understood, particularly when the scale and nature of their use may change in XR technologies to adapt to new use cases.

While some participants showed indifference to XR data collection, many still expressed discomfort towards data sharing in XR—especially regarding passive data collection and inferences made from sensors when a person is not intentionally sharing their data (e.g., emotional inferences from facial expressions and physiological sensor inputs). One possible explanation could be that it is difficult for users to assess data collected from unconscious action and behaviours (e.g., system 'reads' users' facial expressions to make product recommendations) because it is invisible to the users and often operating in the background, compared to more active data collection (e.g., users intentionally smile to signal agreement). This is evident by our qualitative data that a slightly higher proportion of participants mentioned to "comprehend and control what happens to my data" when the scenario involved a background running device (see Figure 4). This suggests that in addition to *what* novel data types are collected, the *way* the data are collected can also influence people's comfort with the data. Therefore, XR designers should help users be aware of the data they provide to the device, especially those data collected from unconscious processes when users have little awareness. As XR continues to innovate, we encourage future research to explore users' perceptions on emerging data types and explore visual and creative ways to communicate data collection beyond 2D interfaces. The results from the research could inform XR designers about the types of data that are considered more intrusive and should be minimized.

5.2.2 Informed Data Collection without Sacrificing Enjoyment and Immersion. Many participants wanted to be informed when data collection happens. Some mentioned about trying to limit XR device data collection by blocking sensors physically, practicing self-censorship, and discontinuing its use. However, these strategies often meant sacrificing the enjoyment and functionality of their XR devices. One way to achieve informed data collection and allow users to opt-out when necessary is through notifications in XR. However, creating efficient notifications in XR environments is a challenging task. To date, various forms of XR notifications have been developed to inform users through head-up display, on-body haptic feedback, floating display, and in-situ display on a surrounding virtual object [61]. On the one hand, the rich information and feedback that XR provides often leads to disconnection from the real world and causes important notifications being missed [61]. On the other hand, notifications that distract users' attention in XR are often found to disrupt the user experience [23, 34]. Thus, future XR designers and researchers should explore ways to integrate XR notifications that better balance their effectiveness and their disruptiveness to the immersive experience. 3D interactions facilitated by XR technology are naturally different than those enabled by 2D screen-based interfaces. Therefore, the way users interact with and in 3D environments provides an opportunity for privacy researchers to study how these 3D interactions (e.g., gestural and spatial controls) can help users understand how their data are collected and processed. For example, digital objects can become privacy interfaces and user consent might be communicated through a range of physical gestures and embodied actions that feel more natural in XR environments than text, icons, and menu-based consent notices. As the design space for XR privacy expands, we encourage designers to prioritize functionalities that are less intrusive to user privacy. This requires future research to explore alternative designs that achieve similar functionality but involve less sensitive data types. Our findings revealed participants' perceived data sensitivity and comfort level for a comprehensive list of data types.

Another way to achieve immersion is for XR designers to implement privacy-friendly default settings [58]. For instance, default settings can restrict the sharing of user information [58, 77]. While having a privacy-friendly default can be effective in reducing users' burden in making the 'better' privacy decisions by offering a path of least resistance, researchers worry that nudging users with these defaults may threaten their autonomy. A person's decision to take an action or change a setting might become biased outside of their awareness [65, 66]. To ensure user autonomy and resolve the dilemma of privacy decision-making, we recommend future research to explore ways to personalize the privacy defaults based on user preferences. AI-supported XR systems could automatically predict user preferences based on prior use and apply relevant settings as the new default that is more likely to fit with user expectations [31]. However, users should also be able to easily

override the default settings when they want other options [68] or if they feel that the system misinterpreted their preferences.

5.2.3 Customizable Consent and Data Flow Tailored to XR. Many participants desired more effective control and explicit consent for specific types of data collection through XR software configurations. While privacy notifications are essential for enhancing informed data collection, research suggests that users feel greater privacy violations when they cannot manage their data collection and use [5, 55]. Their frustration on the absence of privacy choices may further lead to privacy resignation [10]. Since privacy controls in XR could move beyond screen-based paradigms, we suggest providing more natural and intuitive privacy controls tailored to 3D environments to make users feel more empowered. For example, using gestures like a "thumbs up" for consent may be more intuitive than interacting with a screen-based "I agree" button. Emerging controls through gestures and object-oriented interfaces provide valuable opportunities for XR designers to help users make informed privacy decisions and consent to their data use. Future work could explore the best suited interaction types in XR for privacy notifications and control mechanisms.

Currently, there is a lack of guidance to support XR system designers and privacy engineers when designing privacy notices. The traditional design dimensions for the web and mobile interfaces (i.e., timing, channel, modality, control [62]) and the extended design space for IoT (i.e., type, functionality, timing, channel, modality) [20] need to be expanded further for XR environments. For example, the modality design dimension can extend beyond visual, auditory, and haptic to include motion, neurofeedback (e.g., emotion), and even user-defined olfaction (e.g., scent), and gustation (e.g., taste) [57] to convey users' privacy decisions to XR systems. The effectiveness of existing usability evaluation frameworks [26] needs assessment, and new evaluation frameworks tailored to XR environments need to be developed.

5.2.4 Enforcing and Developing Transparent XR Data Practices. Many participants reported strategies such as trusting and relying on the manufacturer to do what had been stated in their privacy policy, and understanding XR data practices by reading user agreements, privacy policies, or legal documentation. They also expressed the need to mandate data collection and that the data should primarily be used to serve user interests due to the fact that VR application data policies are often vague [2], and most data are used for commercial purposes without user consent [71]. These problems highlight the need for transparent data practices associated with XR devices. Since many data types explored in this study are not used in most other types of consumer technology, current data privacy regulations are inadequate to address the privacy challenges posed by the proliferation of sensors and uncertainties in XR data practices [15]. The immersive nature of XR makes it difficult to mitigate risks by applying existing privacy policies and practices for other types of digital media. Future research could help product makers design transparent controls in their products and help policymakers identify and develop new regulations around sensitive data types and mandate appropriate consent mechanisms and controls. When possible, data minimization practices should be implemented around personal and potentially sensitive data types, whether the data is required for the functionalities of the service or to enhance personalization and user experience. Our study provides some indication of the data types that users consider sensitive and suggests that particular attention should be paid to data types that are currently considered non-personally identifiable in non-XR technologies (e.g., behavioural identifier) [24, 47, 56]). Our study also raises the need to design intuitive notices and consent around passive data collection and make use of personalized default settings to help people manage their privacy and data in immersive XR experiences.

5.3 Limitations

We note several limitations of our study. Our sample does not reflect the general population distribution of North America and Europe, as we did not include non-users in our study. Our survey was conducted only in English, non-English speaking XR users are inevitably excluded. Despite these limitations, our results offer valuable insights into the perceptions of those with experience using XR. Further studies could consider expanding the scope to understand the perspectives of non-users, the non-English speaking XR users, and the difference between users and non-users.

Our survey-based study methodology relies on the participants' self-reported privacy concerns, preferences, and behaviour, which might not match their actual behaviour of using XR devices. Additionally, our study can suffer from bias when the scenarios describe situations that the participants have never encountered, affecting the precision of their perceptions and concerns. Although we recognize that mentioning data collection could influence participants' privacy concerns, efforts were made to minimize its impact. For example, we promoted our study as research about XR user experience and explicitly avoided using words like "privacy," "risk," and "concern" in the recruitment materials and questionnaire. We also acknowledge that the factors we examined in our scenarios were not an exhaustive list. Currently, much information regarding XR data practices is unclear [70]. We hope that our study can provide valuable insights for future discussions and encourage researchers to explore user privacy concerns in more complex situations, such as bystander involvement, multiple user scenarios, and other related factors.

6 CONCLUSION

Our paper reported a scenario-based study on privacy concerns related to XR technologies. We surveyed 464 participants and identified their (lack of) awareness of XR data collection, their levels of privacy concerns, factors that influence their concerns, and coping strategies that mitigate their discomfort caused by XR data collection. Our results indicated that participants' comfort regarding XR data collection is directly driven by type of data that are collected, perceived data sensitivity, perceived realism of the scenarios, and perceived equity of data collection. Furthermore, many reported coping strategies similar to web, mobile and IoT scenarios. Based on our findings, we discussed the need to address users' lack of awareness and correct privacy misconceptions originated from previous interactions with non-XR devices, the design of privacy interfaces tailored to immersive XR technologies, and the need to develop and mandate transparent data practices. We hope that our work can provide insight for future user-centred XR privacy studies. XR

technologies are still in early stages, which provides an opportunity to proactively develop solutions and countermeasures to address potential privacy issues. Additionally, we hope that our finding will encourage and facilitate future privacy research for other XR user groups, settings, and usage scenarios, such as bystanders and multi-user environments.

ACKNOWLEDGMENTS

The research was funded by the Games Institute Seed Grant from the University of Waterloo. L. Zhang-Kennedy (#RGPIN-2022-03353) and L. Nacke (#RGPIN-2023-03705) also acknowledge support from the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Games Institute, the University of Waterloo, or NSERC.

REFERENCES

[1] Melvin Abraham, Pejman Saeghe, Mark Mcgill, and Mohamed Khamis. 2022. Implications of XR on Privacy, Security and Behaviour: Insights from Experts. In *Nordic Human-Computer Interaction Conference* (Aarhus, Denmark) *(NordiCHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 30, 12 pages. https://doi.org/10.1145/3546155.3546691

[2] Devon Adams, Alseny Bah, Catherine Barwulor, Nureli Musabay, Kadeem Pitkin, and Elissa M. Redmiles. 2018. Ethics Emerging: The Story of Privacy and Security Perceptions in Virtual Reality. In *Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security* (Baltimore, MD, USA) *(SOUPS '18)*. USENIX Association, USA, 443–458.

[3] Thomas Alsop. 2022. XR: AR, VR, and the metaverse - statistics & facts. Last accessed on November 27, 2022.

[4] Debjanee Barua, Judy Kay, and Cécile Paris. 2013. Viewing and Controlling Personal Sensor Data: What Do Users Want?. In *Persuasive Technology*, Shlomo Berkovsky and Jill Freyne (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 15–26.

[5] Florian Bemmann, Maximiliane Windl, Jonas Erbe, Sven Mayer, and Heinrich Hussmann. 2022. The influence of transparency and control on the willingness of data sharing in adaptive mobile apps. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (2022), 1–26.

[6] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological*. American Psychological Association, Washington, DC, US, 57–71. https://doi.org/10.1037/13620-004

[7] Lauren Buck and Rachel McDonnell. 2022. Security and Privacy in the Metaverse: The Threat of the Digital Human. In *Proceedings of CHI Conference on Human Factors in Computing Systems (CHI EA '22, Proceedings of the 1st Workshop on Novel Challenges of Safety, Security and Privacy in Extended Reality*. ACM, New York, USA, 4 pages.

[8] Shi-Yi Chen, Zhe Feng, and Xiaolian Yi. 2017. A general introduction to adjustment for multiple comparisons. *Journal of thoracic disease* 9, 6 (2017), 1725.

[9] Rune Haubo B Christensen. 2019. A Tutorial on fitting Cumulative Link Mixed Models with clmm2 from the ordinal Package. *Tutorial for the R Package ordinal* 1 (2019), 10 pages.

[10] Jessica Colnago, Yuanyuan Feng, Tharangini Palanivel, Sarah Pearman, Megan Ung, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. 2020. Informing the Design of a Personalized Privacy Assistant for the Internet of Things. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376389

[11] James J Cummings and Jeremy N Bailenson. 2016. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media psychology* 19, 2 (2016), 272–309.

[12] Jaybie Agullo de Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. 2019. SafeMR: Privacy-aware Visual Information Protection for Mobile Mixed Reality. In *2019 IEEE 44th Conference on Local Computer Networks (LCN)*. IEEE, Osnabrueck, Germany, 254–257. https://doi.org/10.1109/LCN44214.2019.8990850

[13] Jaybie Agullo de Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. 2019. Security and privacy approaches in mixed reality: A literature survey. *ACM Computing Surveys (CSUR)* 52, 6 (2019), 1–37.

[14] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. 2014. In Situ with Bystanders of Augmented Reality Glasses: Perspectives on Recording and Privacy-Mediating Technologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2377–2386. https://doi.org/10.1145/2556288.2557352

[15] Ellysse Dick. 2021. *Balancing user privacy and innovation in augmented and virtual reality*. Technical Report. Information Technology and Innovation Foundation.

[16] Tamara Dinev and Paul Hart. 2006. An extended privacy calculus model for e-commerce transactions. *Information systems research* 17, 1 (2006), 61–80.

[17] Youngwook Do, Frederik Brudy, George W Fitzmaurice, and Fraser Anderson. 2023. Vice VRsa: Balancing Bystander's and VR user's Privacy through Awareness Cues Inside and Outside VR. In *Graphics Interface 2023-second deadline*. Canadian Human-Computer Communications Society, Victoria, BC, Canada, 10 pages.

[18] Ben Egliston and Marcus Carter. 2023. Examining visions of surveillance in Oculus' data and privacy policies, 2014–2020. *Media International Australia* 188, 1 (2023), 52–66. https://doi.org/10.1177/1329878X211041670

[19] Franz Faul, Edgar Erdfelder, Axel Buchner, and Albert-Georg Lang. 2009. Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior research methods* 41, 4 (2009), 1149–1160.

[20] Yuanyuan Feng, Yaxing Yao, and Norman Sadeh. 2021. A Design Space for Privacy Choices: Towards Meaningful Privacy Control in the Internet of Things. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 64, 16 pages. https://doi.org/10.1145/3411764.3445148

[21] Janet Finch. 1987. The vignette technique in survey research. *Sociology* 21, 1 (1987), 105–114.

[22] Andrea Gallardo, Chris Choy, Jaideep Juneja, Efe Bozkir, Camille Cobb, Lujo Bauer, and Lorrie Cranor. 2023. Speculative Privacy Concerns About AR Glasses Data Collection. *Proceedings on Privacy Enhancing Technologies* 4 (2023), 416–435.

[23] Ceenu George, Manuel Demmler, and Heinrich Hussmann. 2018. Intelligent Interruptions for IVR: Investigating the Interplay between Presence, Workload and Attention. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188686

[24] Jan Gugenheimer, Wen-Jie Tseng, Abraham Hani Mhaidli, Jan Ole Rixen, Mark McGill, Michael Nebeling, Mohamed Khamis, Florian Schaub, and Sanchari Das. 2022. Novel Challenges of Safety, Security and Privacy in Extended Reality. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 108, 5 pages. https://doi.org/10.1145/3491101.3503741

[25] Aniket Gulhane, Akhil Vyas, Reshmi Mitra, Roland Oruche, Gabriela Hoefer, Samaikya Valluripally, Prasad Calyam, and Khaza Anuarul Hoque. 2019. Security, privacy and safety risk assessment for virtual reality learning environment applications. In *2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, Las Vegas, NV, USA, 1–9.

[26] Hana Habib and Lorrie Faith Cranor. 2022. Evaluating the usability of privacy choice mechanisms. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS)*. USENIX Association, Boston, MA, 273–289.

[27] Hilda Hadan and Sameer Patil. 2020. Understanding perceptions of smart devices. In *Financial Cryptography and Data Security: FC 2020 International Workshops, AsiaUSEC, CoDeFi, VOTING, and WTSC, Kota Kinabalu, Malaysia, February 14, 2020, Revised Selected Papers 24*. Springer International Publishing, Kota Kinabalu, Sabah, Malaysia, 102–121.

[28] David Harborth and Sebastian Pape. 2021. Investigating privacy concerns related to mobile augmented reality apps–A vignette based online experiment. *Computers in Human Behavior* 122 (2021), 106833.

[29] M Brandon Haworth, Melanie Baljko, and Petros Faloutsos. 2012. PhoVR: a virtual reality system to treat phobias. In *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*. ACM, New York, USA, 171–174.

[30] Brittan Heller. 2020. Watching Androids Dream of Electric Sheep: Immersive Technology, Biometric Psychography, and the Law. *Vand. J. Ent. & Tech. L.* 23 (2020), 1.

[31] Yue Huang, Borke Obada-Obieh, and Konstantin Beznosov. 2020. Amazon vs. my brother: How users of shared smart speakers perceive and cope with privacy risks. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1–13.

[32] Fortune Business Insights. 2021. Extended Reality Market Size, Share & Industry Analysis, By Type (Consumer Engagement, Business Engagement), By Application (Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR)), By Industry Vertical (BFSI, Healthcare & Life Sciences, Telecommunications & IT, Government & Public Sector, Manufacturing, Consumer Goods & Retail, Media & Entertainment, Others) And Regional Forecast 2022-2029. Last access on October 21st, 2022.

[33] Joseph B Kadane and Nicole A Lazar. 2004. Methods and criteria for model selection. *Journal of the American statistical Association* 99, 465 (2004), 279–290.

[34] Kan_G3. 2021. Popup notification while playing VR. Reddit. https://www.reddit.com/r/HPReverb/comments/lrfcq0/popup_notification_while_playing_vr/. Last accessed on January 26, 2023.

[35] Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. 2009. Exploring privacy concerns about personal sensing. In *International Conference on Pervasive Computing*. Springer, Berlin, Heidelberg, 176–183.

[36] Kiron Lebeck, Kimberly Ruth, Tadayoshi Kohno, and Franziska Roesner. 2018. Towards security and privacy for multi-user augmented reality: Foundations with end users. In *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, San Francisco, CA, USA, 392–408.

[37] Hosub Lee and Alfred Kobsa. 2016. Understanding user privacy in Internet of Things environments. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*. IEEE, Reston, VA, USA, 407–412.

[38] Pedro Giovanni Leon, Blase Ur, Yang Wang, Manya Sleeper, Rebecca Balebako, Richard Shay, Lujo Bauer, Mihai Christodorescu, and Lorrie Faith Cranor. 2013. What Matters to Users? Factors That Affect Users' Willingness to Share Information with Online Advertisers. In *Proceedings of the Ninth Symposium on Usable Privacy and Security* (Newcastle, United Kingdom) *(SOUPS '13)*. Association for Computing Machinery, New York, NY, USA, Article 7, 12 pages. https://doi.org/10.1145/2501604.2501611

[39] Jialiu Lin, Shahriyar Amini, Jason I Hong, Norman Sadeh, Janne Lindqvist, and Joy Zhang. 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM conference on ubiquitous computing*. Association for Computing Machinery, New York, NY, USA, 501–510.

[40] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.

[41] Divine Maloney, Guo Freeman, and Andrew Robb. 2021. Social virtual reality: ethical considerations and future directions for an emerging research space. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, Lisbon, Portugal, 271–277.

[42] Divine Maloney, Guo Freeman, and Andrew Robb. 2021. Stay Connected in An Immersive World: Why Teenagers Engage in Social Virtual Reality. In *Proceedings of the 20th Annual ACM Interaction Design and Children Conference* (Athens, Greece) *(IDC '21)*. Association for Computing Machinery, New York, NY, USA, 69–79. https://doi.org/10.1145/3459990.3460703

[43] Divine Maloney, Samaneh Zamanifard, and Guo Freeman. 2020. Anonymity vs. familiarity: Self-disclosure and privacy in social virtual reality. In *26th ACM Symposium on Virtual Reality Software and Technology*. Association for Computing Machinery, New York, NY, USA, 1–9.

[44] Florian Mathis. 2022. Moving Usable Security and Privacy Research Out of the Lab: Adding Virtual Reality to the Research Arsenal.

[45] Richard McPherson, Suman Jana, and Vitaly Shmatikov. 2015. No Escape From Reality: Security and Privacy of Augmented Reality Browsers. In *Proceedings of the 24th International Conference on World Wide Web* (Florence, Italy) *(WWW '15)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 743–753. https://doi.org/10.1145/2736277.2741657

[46] Abraham Hani Mhaidli and Florian Schaub. 2021. Identifying Manipulative Advertising Techniques in XR Through Scenario Construction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 296, 18 pages. https://doi.org/10.1145/3411764.3445253

[47] Mark Roman Miller, Fernanda Herrera, Hanseul Jun, James A Landay, and Jeremy N Bailenson. 2020. Personal identifiability of user tracking data during observation of 360-degree VR video. *Scientific Reports* 10, 1 (2020), 1–10.

[48] Pardis Emami Naeini, Sruti Bhagavatula, Hana Habib, Martin Degeling, Lujo Bauer, Lorrie Faith Cranor, and Norman Sadeh. 2017. Privacy expectations and preferences in an IoT world. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*. USENIX Association, Santa Clara, CA, 399–412.

[49] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Washington Law Review* 79 (2004), 119.

[50] Kyle O'Brien. 2019. Coca-Cola embraces augmented reality with interactive experience. The Drum. https://www.thedrum.com/news/2019/09/10/coca-cola-embraces-ar-interactive-experience. Last accessed on January 6, 2023.

[51] Fiachra O'Brolcháin, Tim Jacquemard, David Monaghan, Noel O'Connor, Peter Novitzky, and Bert Gordijn. 2016. The convergence of virtual reality and social networks: threats to privacy and autonomy. *Science and engineering ethics* 22, 1 (2016), 1–29.

[52] Joseph O'Hagan, Pejman Saeghe, Jan Gugenheimer, Daniel Medeiros, Karola Marky, Mohamed Khamis, and Mark McGill. 2023. Privacy-Enhancing Technology and Everyday Augmented Reality: Understanding Bystanders' Varying Needs for Awareness and Consent. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–35.

[53] Antti Oulasvirta, Aurora Pihlajamaa, Jukka Perkiö, Debarshi Ray, Taneli Vähäkangas, Tero Hasu, Niklas Vainio, and Petri Myllymäki. 2012. Long-term effects of ubiquitous surveillance in the home. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. Association for Computing Machinery, New York, NY, USA, 41–50.

[54] Zhigeng Pan, Adrian David Cheok, Hongwei Yang, Jiejie Zhu, and Jiaoying Shi. 2006. Virtual reality and mixed reality for virtual learning environments. *Computers & graphics* 30, 1 (2006), 20–28.

[55] Sameer Patil, Roberto Hoyle, Roman Schlegel, Apu Kapadia, and Adam J Lee. 2015. Interrupt now or inform later? Comparing immediate and delayed privacy feedback. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1415–1418.

[56] Ken Pfeuffer, Matthias J Geiger, Sarah Prange, Lukas Mecke, Daniel Buschek, and Florian Alt. 2019. Behavioural biometrics in vr: Identifying people from body motion and relations in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12.

[57] Ismo Rakkolainen, Ahmed Farooq, Jari Kangas, Jaakko Hakulinen, Jussi Rantala, Markku Turunen, and Roope Raisamo. 2021. Technologies for multimodal interaction in extended reality—a scoping review. *Multimodal Technologies and Interaction* 5, 12 (2021), 81.

[58] REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. 2016. Art. 25 GDPRData protection by design and by default.

[59] Jan Ole Rixen, Teresa Hirzle, Mark Colley, Yannick Etzel, Enrico Rukzio, and Jan Gugenheimer. 2021. Exploring augmented visual alterations in interpersonal communication. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–11.

[60] Franziska Roesner, Tadayoshi Kohno, and David Molnar. 2014. Security and privacy for augmented reality systems. *Commun. ACM* 57, 4 (2014), 88–96.

[61] Rufat Rzayev, Sven Mayer, Christian Krauter, and Niels Henze. 2019. Notification in VR: The Effect of Notification Placement, Task and Environment. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) *(CHI PLAY '19)*. Association for Computing Machinery, New York, NY, USA, 199–211. https://doi.org/10.1145/3311350.3347190

[62] Florian Schaub, Rebecca Balebako, Adam L Durity, and Lorrie Faith Cranor. 2015. A design space for effective privacy notices. In *Eleventh Symposium on Usable Privacy and Security (SOUPS)*. USENIX Association, Ottawa, Canada, 1–17.

[63] Daniel M Shafer, Corey P Carbonara, and Michael F Korpi. 2019. Factors affecting enjoyment of virtual reality games: a comparison involving consumer-grade virtual reality technology. *Games for health journal* 8, 1 (2019), 15–23.

[64] Joon-Ho Shin, Si Bog Park, and Seong Ho Jang. 2015. Effects of game-based virtual reality on health-related quality of life in chronic stroke patients: a randomized, controlled study. *Computers in biology and medicine* 63 (2015), 92–98.

[65] N Craig Smith, Daniel G Goldstein, and Eric J Johnson. 2013. Choice without awareness: Ethical and policy implications of defaults. *Journal of Public Policy & Marketing* 32, 2 (2013), 159–172.

[66] Daniel J Solove. 2012. Introduction: Privacy self-management and the consent dilemma. *Harv. L. Rev.* 126 (2012), 1880.

[67] Maximilian Speicher, Brian D Hall, and Michael Nebeling. 2019. What is mixed reality?. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1–15.

[68] Cass R Sunstein. 2014. *Why nudge?: The politics of libertarian paternalism*. Yale University Press, New Haven, Connecticut, USA.

[69] Meta Oculus Team. 2015. FIRST LOOK AT THE RIFT, SHIPPING Q1 2016. https://www.oculus.com/blog/first-look-at-the-rift-shipping-q1-2016/. Last accessed on November 21, 2022.

[70] The XR Safety Initiative (XRSI). 2021. *1st XR Data Classification Roundtable Report*. XR Safety Week 2021. The XR Safety Initiative (XRSI). Last accessed on October 20th, 2022.

[71] Rahmadi Trimananda, Hieu Le, Hao Cui, Janice Tran Ho, Anastasia Shuba, and Athina Markopoulou. 2022. {OVRseen}: Auditing Network Traffic and Privacy Policies in Oculus {VR}. In *31st USENIX security symposium (USENIX security 22)*. USENIX Association, Boston, MA, 3789–3806.

[72] Janice Y Tsai, Patrick Kelley, Paul Drielsma, Lorrie Faith Cranor, Jason Hong, and Norman Sadeh. 2009. Who's viewed you? The impact of feedback in a mobile location-sharing application. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2003–2012.

[73] Max Van Kleek, Ilaria Liccardi, Reuben Binns, Jun Zhao, Daniel J Weitzner, and Nigel Shadbolt. 2017. Better the devil you know: Exposing the data sharing practices of smartphone apps. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 5208–5220.

[74] Robert F Woolson. 2007. Wilcoxon signed-rank test. , 3 pages.

[75] XR Safety Initiative (XRSI). 2020. *The XRSI Definition of Extended Reality (XR)*. XR Safety Initiative Standard Publication XR 001. The XR Safety Initiative (XRSI). Last accessed on October 16th, 2022.

[76] Toshihiko Yamakami. 2020. A privacy threat model in xr applications. In *International Conference on Emerging Internetworking, Data & Web Technologies*.

Springer, Berlin, Heidelberg, 384–394.
[77] Yaxing Yao, Justin Reed Basdeo, Smirity Kaushik, and Yang Wang. 2019. Defending my castle: A co-design study of privacy mechanisms for smart homes. In *Proceedings of the 2019 chi conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1–12.
[78] Rafael Yuste, Jared Genser, and Stephanie Herrmann. 2021. It's time for neurorights. *Horizons* 18 (2021), 154–164.

A SUPPLEMENTARY MATERIALS — SURVEY QUESTIONS

A.1 User Understanding of XR Technologies

For the introductory purpose, the participants were first presented with a drag and drop pair-matching exercise, where they were asked to match the following descriptions (from XR Safety Initiative (XRSI) [75]) with corresponding technologies: VR, AR, MR. After this exercise, we further reinforced participants' understanding by providing them with the correct VR, AR, and MR definitions [75]:

Extended Reality (XR) is a fusion of all the realities — including **Virtual Reality (VR)**, **Augmented Reality (AR)**, and **Mixed Reality (MR)** — which consists of technology-mediated experiences enabled via a wide spectrum of hardware and software, including sensory interfaces, applications, and infrastructures.

- **VR**
  - It is a fully immersive software-generated, artificial digital environment.
  - It is a simulation of three-dimensional images, experienced by users via special electronic equipment, such as a headset.
  - It places the experiencer in another environment entirely. Whether that environment has been generated by a computer or captured by video, it entirely occludes the experiencer's natural surroundings.
- **AR**
  - It overlays digital-created content on top of the user's real-world environment, viewed through a device (such as a smartphone) that incorporates real-time inputs to create an enhanced version of reality.
  - Digital and virtual objects (e.g., graphics, sounds) are superimposed on an existing environment to create an immersive experience.
- **MR**
  - It seamlessly blends the user's real-world environment with digitally-created content, where both environments can coexist and interact with each other.
  - The virtual objects behave in all aspects as if they are present in the real-world, e.g., they are occluded by physical objects, they sound as though they are in the same space as the user. As the user interacts with the real and virtual objects, the virtual objects will reflect the changes in the environment as would any real object in the same space.

A.2 User Awareness of XR Data Operation

Q1: You mentioned that you have experience using XR devices. What data do you think are being observed by the device?
- Identifying in-app/in-world assets (e.g., personal virtual objects, personal avatars)
- Physical appearance (e.g., body dimensions)
- Environment information and dimensions (e.g., users' surroundings, room layout, device/user position in relation to environment, user position in relation to device)
- Physical movements and characteristics (e.g., head/hand/body/eye motion, orientation, position, gestures, posture, fitness information)
- Physiological data (e.g., pupil and cornea reflections, brainwaves, skin signals)
- Visual attention (e.g., eye gaze, area of interest, fixation, heatmaps, time to first fixation, time spent on a certain point, fixation sequences)
- Mannerisms (e.g., gait, habitual movements)
- Cognitive, emotional, personality estimates (e.g., cognitive load, stress, depression, excitement, identity traits, etc.)
- Non-identifying virtual assets (e.g., in-app achievements)

A.3 Scenarios and Questions

Sample Scenario:

> *You are at home / personal room, and your XR device (e.g., smart headset, touch controller, 3D projects) is keeping track of your **Identifying in-app / in-world assets (e.g., personal virtual objects, personal avatars)** when **you use the device**.*

Questions below are repeated with each Scenario.

Q2: "I think the Scenario like this (will) happen..." (each answered on 5-point Likert scales from "1-extremely unlikely" to "5-extremely likely")
- Today
- Within 2 years
- Within 10 years

Q3: How comfortable are you about the data collection described in the Scenario? (answered on a 5-point Likert scale from "1-very uncomfortable" to "5-very comfortable")

Q4: Please indicate how sensitive you consider the data being collected in the Scenario. **Sensitive Information** is any information which could cause serious mental, physical, or financial harm if it's lost, misused, or shared without permission. (answered on a 5-point Likert scale from "1-not sensitive" to "5-very sensitive")

Q5: (if "1-very uncomfortable" or "2-uncomfortable" in Q3) If you had no choice on data collection, what would you do (if any) to mitigate the discomfort raised? [open-ended question]

Table 5: Codebook — This codebook presents the codes and themes we synthesized from participants' responses to the open-ended question "If you had no choice on data collection, what would you do (if any) to mitigate the discomfort raised?" (Q5).

Theme	Frequent codes	Example responses
Away from device	Avoid at all cost Cease using equipment Get rid of device Wouldn't buy the device	"It can be very dangerous. Avoid at all cost." "Most likely the only way to reduce data collection would be to not use the device." "I would get rid of device." "I'd sell the VR headset." "I'd not purchase device which is capable of this kind of data collection."
Benefit vs. Privacy	Accept the data collection if its beneficial to me Evaluate if the device worth the value I should be paid if they use my data	"If them obtaining that data benefited me in some way, improved my health for example." "ask myself is it worth it.If the reasons are strong enough I [will] convince myself to only focus on the price." "at least reward people for their own data." "[I should] getting paid for the collection and use of my data."
Company and System Trust	Bring concern to the company Choose the device/brands I trust Must comply with user privacy settings Trust the system	"I'd try to raise concerns to such companies." "choose the devices/brands I feel most comfortable sharing my data with." "allow users to set privacy settings and actually comply with them." "I'd trust a good system, so in the end I'd be comfortable sharing my data."
Consent and Permission	Be able to opt out Data collection should with my permission Access and control data usage post-collection	"I'd like to be able to opt out." "Informing that it is doing this and when it is collecting data, how it is being used, and viewing it if necessary." "I'd not partake at all if sensitive data was collected from me without my explicit permission." "Having access on how and from whom my personal data are used for."
Comprehend and control what happen to my data	Ensure data is used only for my interests Search option from the Internet Seek privacy protection regulations Understand data operation	"show the users that data is only used for their interests and not with second intentions." "data is not used for malicious actions." "search for third-party ways to cheat that device so it would process wrong data [not my personal data]." "I'd check the data privacy laws surrounding the product." "make sure to have some sort of data collection protection contract." "try research at maximum about the data collection to prevent discomfort."
Ignore and Give up	Don't mind data collection but can't be 24/7 Everyone's being monitored I can mitigate the discomfort Keep using device and ignore the data collection None	"I don't mind the data it collects, but every waking moment in my day would be uncomfortable." "everyone is being monitored anyway. I'm not special." "Technology will overcome us." "We have no privacy." "I believe I have the ability to mitigate the discomfort." "just not think about the data being collected." "ignore the discomfort." "None." "Nothing can be done."
Limit the collection	Disconnect / turnoff /uninstall when not use Have separate devices for different activities Limit the device usage only when necessary Not register personal info Physically block device sensors Physically contain the tracking ability Seek data protection apps Seek data protection configuration options Self-censor privacy disclosure	"unplug the device when I'm not using it." "disconnect the device from the Internet." "having a separate device for activities that wouldn't mix with my main personal computer or devices." "Use the device a lot less." "only use it for certain things and games." "I can lie about my digital self and not have a very personal experience." "Not register the personal avatar on the device." "obscure the view of the sensor that the device is using to operate." "cover the device when not in use." "find a way to totally block any data that is sent." "put the device in a room that is less personal." "I'd download some apps to stop it [data collection]." "find if there's a setting that disables data collection for commercial purposes." "there should be options to reduce/control the amount and type of data being collected." "avoid my usual mannerism so that the machine can't learn them." "try to not demonstrate my feelings to that device."
Privacy above anything	Be careful and not support it Concerned about impersonation Personalized advertisements and manipulation Ensure that I can't be identified from my data Switch to alternatives that respect privacy	"have more attention when i use it." "I'd refuse to support it [data collection] in any way." "I'd not agree for this kind of monitoring." "it would be real easy to impersonate someone if those details were to be gathered." "it will be used to sell me something in the future." "it can be used for manipulate my behavior through the shopping, voting and important life decisions habits." "assure me somehow data is anoymized." "something like a VPN would exist for those circumstances in the future, where my data is collected by third parties are unable to connect that data to your identity." "I'd try other companies who doing device who respect privacy." "I'd search for alternative headsets without data collection."

Table 6: Overview of Participants' demographics (N = 464)

Gender	n (% of total)
Man	230 (50%)
Woman	222 (48%)
Non-binary / third gender	11 (2%)
Prefer not to say	1 (<1%)
Age	
18-19	12 (3%)
20-29	313 (68%)
30-39	88 (19%)
40-49	39 (8%)
50-56	12 (3%)
Mean (SD)	27.78 (7.95)
Education level	
Less than a high school diploma	4 (<1%)
High school degree or equivalent	173 (37%)
Associate's degree (e.g., AA, AS)	30 (7%)
Bachelor's degree (e.g., BA, BS)	162 (35%)
Graduate or professional degree (e.g., MA, MD, Ph.D.)	94 (20%)
Prefer not to say	1 (<1%)
Income	
Less than $25,000	182 (39%)
$25,000-$49,999	151 (33%)
$50,000-$99,999	86 (19%)
$100,000-$199,999	19 (4%)
More than $200,000	5 (1%)
I prefer not to respond	21 (5%)
Country of residence	
Belgium	5 (1%)
Canada	5 (1%)
Czech Republic	8 (2%)
Greece	25 (5%)
Hungary	20 (4%)
Italy	43 (9%)
Netherlands	5 (1%)
Poland	66 (14%)
Portugal	123 (27%)
Slovenia	7 (2%)
Spain	28 (6%)
United Kingdom of Great Britain and Northern Ireland	57 (12%)
United States of America	52 (11%)
Other (9 countries)*	20 (4%)
Experience with XR devices (EXPERIENCE_YEARS)	
Less than 6 months	185 (40%)
6 months to 1 year	103 (22%)
1 year to 5 years	165 (36%)
5 to 10 years	9 (2%)
More than 10 years	2 (<1%)
Type of devices used (EXPERIENCE_DEVICE)	
Smart headsets, smart glasses	396 (85%)
Touch controllers, wired gloves, 3D mouse	121 (26%)
Handheld or Mobile devices (e.g., handheld viewers, mobile phone, tablet)	268 (58%)
Projectors and display walls	51 (11%)
Frequency of use (EXPERIENCE_FREQUENCY)	
At least once a day	23 (5%)
At least once a week	117 (25%)
At least once a month	121 (26%)
At least once every three months	69 (15%)
At least once every six months	60 (13%)
Once a year	74 (16%)
Purpose of use (EXPERIENCE_PURPOSE)	
Arts (e.g., virtual gallery, design & prototyping, virtual graffiti)	79 (17%)
Education or training (e.g., students, sports, military, medical procedures)	80 (17%)
Entertainment (e.g., gaming, socializing)	437 (94%)
Healthcare (e.g., mental health diagnoses and treatment)	11 (2%)
Workplace requirement (e.g., remote work, staff connection, workplace functionality)	31 (7%)
Other (i.e., research, curiosity, getting directions, fitness)	10 (2%)
Internet Users' Information Privacy Concerns (IUIPC)	
Control factor (IUIPC_AWARENESS)	
Range (Min - Max)	(3 - 7)
Mean (SD)	5.78 (0.92)
Awareness factor (IUIPC_CONTROL)	
Range (Min - Max)	(3 - 7)
Mean (SD)	6.29 (0.74)
Collection factor (IUIPC_COLLECTION)	
Range (Min - Max)	(1.25 - 7)
Mean (SD)	5.69 (1.14)

*Note. * "Other" includes 9 countries, each containing fewer than 5 participants: Austria, Estonia, Finland, France, Germany, Ireland, Lativa, Norway, Sweden.*

Table 7: Wilcoxon Signed-Rank Tests with Bonferroni adjustment [8] on participants' comfort level and perceived data sensitivity across data types. A same pair of scenarios is presented to $n \approx 64$ participants.

	Descriptive Statistics		Wilcoxon-signed rank test							
data_type	**Median (SD)**	**Range (Min-Max)**	Physical appearance	Identifying in-app/in-world assets	Environment information and dimensions	Physical movements and characteristics	Physiological data	Visual attention	Mannerisms	Cognitive, emotional, and personality estimates
comfort_level										
Physical appearance	3 (1.10)	(1 - 5)								
Identifying in-app/in-world assets	3 (1.10)	(1 - 5)	*W=248.5* *P=.002*** *r=-0.39*							
Environment information and dimensions	3 (1.06)	(1 - 5)	*W=323.5* *P<.15* *r=-0.17*	*W=804.5* *P=.001**** *r=-0.46*						
Physical movements and characteristics	3 (1.14)	(1 - 5)	*W=320.0,* *P=.02** *r=-0.28*	*W=502* *P=.11* *r=-0.20*	*P=.65* *P=.65* *r=-0.06*					
Physiological data	2 (1.03)	(1 - 5)	*W=612.5* *P=.81* *r=-0.03*	*W=1473.5* *P<.001**** *r=-0.48*	*W=738.5* *P=.003*** *r=-0.37*	*W=838.5* *P<.001**** *r=-0.46*				
Visual attention	3 (1.12)	(1 - 5)	*W=395.5* *P=.64* *r=-0.06*	*W=410.5* *P=.048** *r=-0.25*	*W=465* *P=.46* *r=-0.10*	*W=367* *P=.40* *r=-0.10*	*W=181.5* *P=.04** *r=-0.28*			
Mannerisms	3 (1.07)	(1 - 5)	*W=345.0* *P=.62* *r=-0.07*	*W=827* *P<.001**** *r=-0.49*	*W=539.5* *P=.25* *r=-0.14*	*W=574.5* *P=.12* *r=-0.18*	*W=231.5* *P=.06* *r=-0.24*	*W=584.5* *P=.29* *r=-0.13*		
Cognitive, emotional, and personality estimates	2 (1.14)	(1 - 5)	*W=1052.5* *P<.001**** *r=-0.59*	*W=1031.5* *P<.001**** *r=-0.61*	*W=751* *P<.001**** *r=-0.61*	*W=1002* *P<.001**** *r=-0.50*	*W=317* *P=.007*** *r=-0.39*	*W=988* *P=.002*** *r=-0.39*	*W=567.5* *P=.001*** *r=-0.42*	
Non-identifying virtual assets	4 (1.01)	(1 - 5)	*W=143.5* *P<.001**** *r=-0.58*	*W=189* *P=.001*** *r=-0.40*	*W=285* *P<.001**** *r=-0.53*	*W=252* *P=.005*** *r=-0.35*	*W=0* *P<.001**** *r=-0.80*	*W=181* *P<.001**** *r=-0.57*	*W=146.5* *P<.001**** *r=-0.62*	*W=27.5* *P<.001**** *r=-0.78*

Note. W=Test Statistic, r=Effect size, Significance are displayed as: *** *P<.001,* ** *P<.01,* * *P<.05. Table continued on next page.*

Table 7: continued. Wilcoxon Signed-Rank Tests with Bonferroni adjustment [8] on participants' comfort level and perceived data sensitivity across data types. A same pair of scenarios is presented to $n \approx 64$ participants.

	Descriptive Statistics		Wilcoxon-signed rank test							
data_type	**Median (SD)**	**Range (Min-Max)**	Physical appearance	Identifying in-app/in-world assets	Environment information and dimensions	Physical movements and characteristics	Physiological data	Visual attention	Mannerisms	Cognitive, emotional, and personality estimates
data_sensitivity										
Physical appearance	4 (1.28)	(1 - 5)								
Identifying in-app/in-world assets	2 (1.24)	(1 - 5)	*W=782.5* *P=.008*** *r=-0.34*							
Environment information and dimensions	3 (1.20)	(1 - 5)	*W=469* *P=.42* *r=-0.10*	*W=164.5* *P<.001**** *r=-0.49*						
Physical movements and characteristics	3 (1.26)	(1 - 5)	*W=594* *P=.23* *r=-0.15*	*W=181* *P<.001**** *r=-0.49*	*W=322* *P=.67* *r=-0.06*					
Physiological data	4 (1.16)	(1 - 5)	*W=503.5* *P=.68* *r=-0.05*	*W=234* *P<.001**** *r=-0.58*	*W=327.5* *P=.006*** *r=-0.34*	*W=151* *P<.001**** *r=-0.46*				
Visual attention	3 (1.22)	(1 - 5)	*W=547.5* *P=.12* *r=-0.20*	*W=276.5* *P=.07* *r=-0.24*	*W=369* *P=.57* *r=-0.07*	*W=526.5* *P=.006*** *r=-0.34*	*W=406* *P=.002*** *r=-0.44*			
Mannerisms	3 (1.20)	(1 - 5)	*W=411* *P=.56* *r=-0.08*	*W=262* *P=.002*** *r=-0.38*	*W=695.5* *P=.40* *r=-0.10*	*W=286* *P=.09* *r=-0.20*	*W=742.5* *P<.001**** *r=-0.48*	*W=290* *P=.16* *r=-0.18*		
Cognitive, emotional, and personality estimates	4 (1.00)	(1 - 5)	*W=17* *P<.001**** *r=-0.73*	*W=113* *P<.001**** *r=-0.67*	*W=88* *P<.001**** *r=-0.61*	*W=75* *P<.001**** *r=-0.71*	*W=53* *P<.001**** *r=-0.61*	*W=74.5* *P<.001**** *r=-0.69*	*W=104* *P<.001**** *r=-0.58*	
Non-identifying virtual assets	2 (1.28)	(1 - 5)	*W=1109* *P<.001**** *r=-0.62*	*W=525.5* *P<.001**** *r=-0.49*	*W=1723.5* *P<.001**** *r=-0.56*	*W=1012.5* *P<.001**** *r=-0.61*	*W=1469.5* *P<.001**** *r=-0.79*	*W=831* *P<.001**** *r=-0.64*	*W=934.5* *P<.001**** *r=-0.52*	*W=1263* *P<.001**** *r=-0.81*

Note. W=Test Statistic, r=Effect size, Significance are displayed as: *** *P<.001,* ** *P<.01,* * *P<.05.* *Table continued on next page.*

Table 7: continued. Wilcoxon Signed-Rank Tests with Bonferroni adjustment [8] on participants' comfort level and perceived data sensitivity across data types. A same pair of scenarios is presented to $n \approx 64$ participants.

	Descriptive Statistics		Wilcoxon-signed rank test								
data_type	**Median (SD)**	**Range (Min-Max)**	Physical appearance	Identifying in-app/in-world assets	Environment information and dimensions	Physical movements and characteristics	Physiological data	Visual attention	Mannerisms	Cognitive, emotional, and personality estimates	
likely_today											
Physical appearance	3 (1.27)	(1 - 5)									
Identifying in-app/in-world assets	4 (1.28)	(1 - 5)	*W=226.5* *P<.001**** *r=-0.46*								
Environment information and dimensions	4 (1.26)	(1 - 5)	*W=217* *P=.009*** *r=-0.31*	*W=598* *P=.06* *r=-0.23*							
Physical movements and characteristics	4 (1.23)	(1 - 5)	*W=339.5* *P=.002*** *r=-0.37*	*W=683* *P=.11* *r=-0.20*	*W=217* *P=.37* *r=-0.12*						
Physiological data	2 (1.17)	(1 - 5)	*W=520.5* *P=.24* *r=-0.14*	*W=1402* *P<.001**** *r=-0.66*	*W=846.5* *P<.001**** *r=-0.57*	*W=797* *P<.001**** *r=-0.54*					
Visual attention	3 (1.25)	(1 - 5)	*W=226* *P=.03** *r=-0.28*	*W=471* *P=.14* *r=-0.19*	*W=690.5* *P=.02** *r=-0.30*	*W=361* *P=.27* *r=-0.14*	*W=62* *P<.001**** *r=-0.51*				
Mannerisms	3 (1.20)	(1 - 5)	*W=372* *P=.35* *r=-0.12*	*W=985.5* *P<.001**** *r=-0.56*	*W=637* *P=.007*** *r=-0.33*	*W=656* *P=.001*** *r=-0.40*	*W=97.5* *P<.001**** *r=-0.54*	*W=487* *P=.46* *r=-0.09*			
Cognitive, emotional, and personality estimates	2 (1.08)	(1 - 5)	*W=1016.5* *P<.001**** *r=-0.59*	*W=1387* *P<.001**** *r=-0.67*	*W=982.5* *P<.001**** *r=-0.70*	*W=968* *P<.001**** *r=-0.70*	*W=259* *P=.09* *r=-0.24*	*W=1006* *P<.001**** *r=-0.54*	*W=811* *P<.001**** *r=-0.53*		
Non-identifying virtual assets	4 (1.30)	(1 - 5)	*W=108* *P<.001**** *r=-0.64*	*W=225* *P=.09* *r=-0.21*	*W=485.5* *P=.03** *r=-0.25*	*W=195.5* *P=.02** *r=-0.30*	*W=10* *P<.001**** *r=-0.77*	*W=212* *P=.001*** *r=-0.42*	*W=111.5* *P<.001**** *r=-0.57*	*W=61* *P<.001**** *r=-0.74*	

*Note. W=Test Statistic, r=Effect size, Significance are displayed as: *** P<.001, ** P<.01, * P<.05.*

Table 8: Wilcoxon Signed-Rank Tests on participants' comfort level across device statues (N = 464).

device_status	Descriptive Statistics		Wilcoxon Signed-Rank Test
	Median (SD)	Range (Min-Max)	
comfort_level			In-use
In-use	3 (1.10)	(1 - 5)	
Running in the background	3 (1.20)	(1 - 5)	*W=38588, P<.001***, r=-0.23*
data_sensitivity			In-use
In-use	3 (1.12)	(1 - 5)	
Running in the background	3 (1.10)	(1 - 5)	*W=24216, P=.005**, r=-0.14*
likely_today			In-use
In-use	3 (1.14)	(1 - 5)	
Running in the background	3 (1.11)	(1 - 5)	*W=38028, P<.001***, r=-0.24*

*Note. W=Test Statistic, r=Effect size, Significance are displayed as:*** P<.001, ** P<.01, * P<.05.*