Exploring the Unprecedented Privacy Risks of the Metaverse

Vivek Nair*
UC Berkeley
vcn@berkeley.edu

Gonzalo Munilla Garrido*
Technical University of Munich
gonzalo.munilla-garrido@tum.de

Dawn Song
UC Berkeley
dawnsong@berkeley.edu

***Abstract*—Thirty study participants playtested an innocent-looking "escape room" game in virtual reality (VR). Behind the scenes, an adversarial program had accurately inferred over 25 personal data attributes, from anthropometrics like height and wingspan to demographics like age and gender, within just a few minutes of gameplay. As notoriously data-hungry companies become increasingly involved in VR development, this experimental scenario may soon represent a typical VR user experience. While virtual telepresence applications (and the so-called "metaverse") have recently received increased attention and investment from major tech firms, these environments remain relatively understudied from a security and privacy standpoint. In this work, we illustrate how VR attackers can covertly ascertain dozens of personal data attributes from seemingly-anonymous users of popular metaverse applications like VRChat. These attackers can be as simple as other VR users without special privilege, and the potential scale and scope of this data collection far exceed what is feasible within traditional mobile and web applications. We aim to shed light on the unique privacy risks of the metaverse, and provide the first holistic framework for understanding intrusive data harvesting attacks in these emerging VR ecosystems.**

I. Introduction

Through the fog of rapidly shifting consumer preferences for internet technologies, one clear trend has stood the test of time: with each new and improved medium for accessing the web comes a new and improved method for harvesting personal user data. As these technologies become more immersive and tightly integrated with our daily lives, so too do the corresponding intrusive attacks on user privacy.

In the first era of the world wide web, users primarily accessed information through static websites with limited opportunity for data-revealing interaction. The emergence of social media platforms in the early 2000s quickly changed this paradigm, generating a torrent of data on user behavior. Third-party (tracking) cookies that can uniquely identify and follow individuals [10] around the web allowed this data to be deployed for everything from surveillance advertisement [11] to pushing political agendas [55].

In the past decade, users shifted to accessing the web primarily via their mobile phones (92.1% as of 2022 [69]), simultaneously introducing a suite of newly-extractable data attributes like audio, video, and geolocation. Next, the wave of wearable devices such as smart watches added critically sensitive data like biometrics and health information into the mix [74]. Most recently, virtual home assistants have made possible pernicious intrusions into users' most private activities [15]. Overall, the tendency is clear: each new technology has gradually expanded the scope of data attributes accessible to would-be attackers.

Virtual reality (VR) is well positioned to become a natural continuation of this trend. While VR devices have been around in some form since well before the internet [2], the true ambition of major corporations to turn these devices into massively-connected social "metaverse" platforms has only recently come to light [51], [63], [71]. These platforms, by their very nature, turn every single gaze, movement, and utterance of a user into a stream of data, instantaneously broadcast to other users around the world in the name of facilitating real-time interaction.

This paper aims to shed light on the unprecedented privacy risks of the metaverse by providing the first comprehensive security and privacy framework for VR environments. We have identified over 25 examples of private data attributes that attackers can covertly harvest from VR users, which we experimentally demonstrate in our 30-person user study. Some of these attributes would be difficult, if not impossible, to observe within traditional mobile and web applications. Others mirror attacks seen elsewhere, but are demonstrated for the first time to be feasibly observable within VR. Moreover, the sheer breadth of private attribute classes revealed by VR users has scarcely been seen in other environments. We hope our results increase broad awareness of privacy concerns within VR and compel privacy practitioners to examine the challenges and solutions that lie at the intersection of privacy and the emerging VR-enhanced social internet.

The main contributions of our study are:

1) We provide the first comprehensive framework of virtual reality threat models, data sources, observable attribute classes (§II), and systematic privacy attacks (§III).
2) With our open-source VR demo [52], we illustrate how malicious game developers can design seemingly-innocuous VR environments that trick users into revealing personal information through their behavior (§IV).
3) We experimentally demonstrate how an attacker can covertly harvest from VR users over 25 unique data attributes, many of which are infeasible to obtain through traditional mobile and web applications (§V-A–V-D).
4) We further show that these VR-specific attribute sets are sufficient to accurately infer the demographics (age, gender, ethnicity, etc.) of an "anonymous" user (§V-E).

*Equal contribution.

Attacker Type	Data Sources					Observable Attribute Classes				
	Raw Sensor Data	Processed Telemetry	Rendering Pipeline & Host System APIs	Networked Telemetry	Presented Telemetry	Device	Network	Geospatial	Audio	Behavior
Privileged Attacker I	✓	✓	✓			✓		✓	✓	✓
Privileged Attacker II		✓	✓	✓		✓	✓	✓*	✓	✓
Privileged Attacker III				✓	✓		✓	✓*	✓*	✓
Non-Privileged Attacker					✓			✓*	✓*	✓

*Observable only in weaker filtered/preprocessed format

TABLE I: Virtual reality threat actor capabilities.

II. VR Threat Model

This section provides a holistic framework for attacker types and vulnerable data attributes in the context of VR, thereby framing the privacy attacks introduced in section III. First, we describe a typical information flow for a VR telepresence application. Subsequently, we consider the types of parties ("attackers") having access to data sources associated with a VR device. Finally, we identify the attribute classes observable by the attackers and their potential privacy risks. In the context of this study, we consider a state of privacy as the lack of a breach of sensitive attributes of any individual [85].

A. VR Information Flow

Users can download various games and applications from the app store their VR device's manufacturer provides (e.g., Oculus Store). One increasingly popular category of VR application is virtual telepresence (e.g., VRChat [24]), whereby users around the world interact with each other in real-time within a 3D virtual world (or "metaverse").

The typical information flow for such an application, as depicted in Fig. 1, is as follows: The VR device processes raw sensor data into useful telemetry, which it provides to the application via an API (Step 1A). The application uses this data to provide visual stimuli (frames) to the user via a graphics rendering pipeline, which the application completely controls (Step 1B). If the application involves interactions with other users, the client-side VR application streams processed telemetry data to an external server via a network to facilitate such interactions (Step 2). The server then relays this data to other users for their devices to render (Step 3).

The exact sensor readings available vary significantly depending on the device, but processed telemetry generally includes at least the position and orientation of the headset and controllers, with more data available on systems supporting advanced features such as eye tracking or full-body tracking.

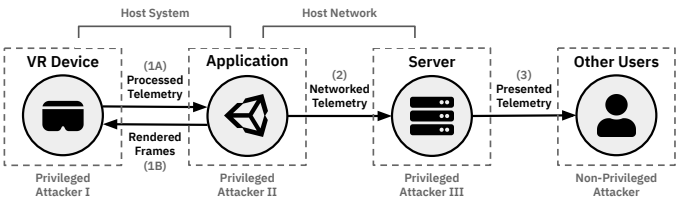


Fig. 1: Virtual reality information flow and threat model.

B. VR Attackers

Given the data flow of Fig. 1, we consider four types of VR privacy attackers which correspond to four distinct entities typically associated with VR data processing. We summarize the capabilities of each attacker in Table I. The goal of each attacker is to learn as much information as possible about the target user employing only the information naturally presented to the attacker via standard APIs (i.e., without using malware, side channels, or privilege escalation). We based this attack model on our experience with the Oculus and Steam VR ecosystems, but the exact capabilities of each attacker may vary depending on the APIs made available by different platforms.

Privileged Attacker I *(the "Hardware Adversary")*

The first privileged attacker represents the party controlling the firmware of a target user's VR device. This attacker has access to raw sensor data from the VR device, including spatial telemetry, audio/visual streams, and device specifications. There is a bi-directional information flow between the device and the local application: the device provides processed telemetry to a running application, which the attacker can manipulate arbitrarily (1A), and the device receives a stream of audio/visual stimuli from the application, which the attacker can manipulate arbitrarily before presenting to the user (1B). However, this attacker cannot read or manipulate the network communications of the application.

Privileged Attacker II *(the "Client Adversary")*

Our second privileged attacker represents the developer of the client-side VR application running on the target user's device. This attacker has full access to the APIs provided by the VR device and host system (1A) and controls a graphics rendering pipeline which the attacker can use to provide visual stimuli to the target user (1B). In the case of a multiplayer application, it can process this data arbitrarily before streaming it to an external server (2).

Privileged Attacker III *(the "Server Adversary")*

Our third privileged attacker represents the entity controlling the external server used to facilitate multiplayer functionality for the application running on the target user's device. This entity may, in practice, be the same party developing the client-side application (in the case of a "public server") or an entirely separate entity (a "private server"). Thus, Privileged Attackers II and III could often be controlled by the same entity. This attacker receives a stream of telemetry data from

the client-side application (2), which it can process arbitrarily before relaying such data to one or more other client devices (3). The data available to this attacker is generally weaker than the previous attacker; for example, a client application may receive tracking data at 120 Hz and broadcast it at 30 Hz instead [81], and audio signals are typically heavily compressed before being broadcast.

Non-Privileged Attacker *(the "User Adversary").*

A non-privileged attacker represents a second end-user of the same multiplayer application as the target user. The attacker receives low-fidelity telemetry and audio streams from the external server (3), which it uses to render a representation of the target user. They can also interact with the target user as permitted by the application, such as to provide stimuli and observe the target's response. While the audio and telemetry streams are likely highly processed and filtered by this point, they are typically still sufficient to observe the general behavior of the target user.

The user study in this paper aims to show the feasibility of each attack model. We thus minimized interaction with the participants to closely emulate a realistic attack scenario. Moreover, participants reported not knowing exactly which data attributes we collected during the experiments, and thus could not directly cooperate with any of the attacks. However, insofar as the attacker's capability correlates with the quality of the received telemetry, our high-fidelity VR setups provided favorable conditions for demonstrating these attacks.

C. Observable Attribute Classes

We now shift our discussion to the broad classes of private user data observable by each of the attackers using only their corresponding data sources. Fig. 2 shows an overview of the VR data sources we consider in this paper. We categorize the collected attributes into primary (captured directly from a data source), secondary (derived deterministically from primary attributes), and inferred (derived from primary and secondary attributes using machine learning).

Geospatial Telemetry. The first major source of user data is directly from geospatial telemetry (namely, the position and orientation of the VR headset and controllers over time). Such data is useful for revealing a user's anthropometric measurements, such as height and wingspan. While all attackers can observe telemetry to some extent, less privileged attackers are likely to experience degraded precision when estimating these metrics due to the use of intermediate filtering and processing. For example, we found that privileged attackers I and II can determine a user's interpupillary distance (IPD) from telemetry to within 0.1mm, while IPD is difficult for privileged attackers III and non-privileged attackers to ascertain.

Device Specifications. Another class of attack aims to use VR-specific heuristics to determine information about the VR device and the user's host computer. Of course, privileged attackers I and II can directly query device specifications such as resolution and field of view (FOV) from available system APIs; however, we will later demonstrate how even non-privileged attackers can attempt to learn some of this

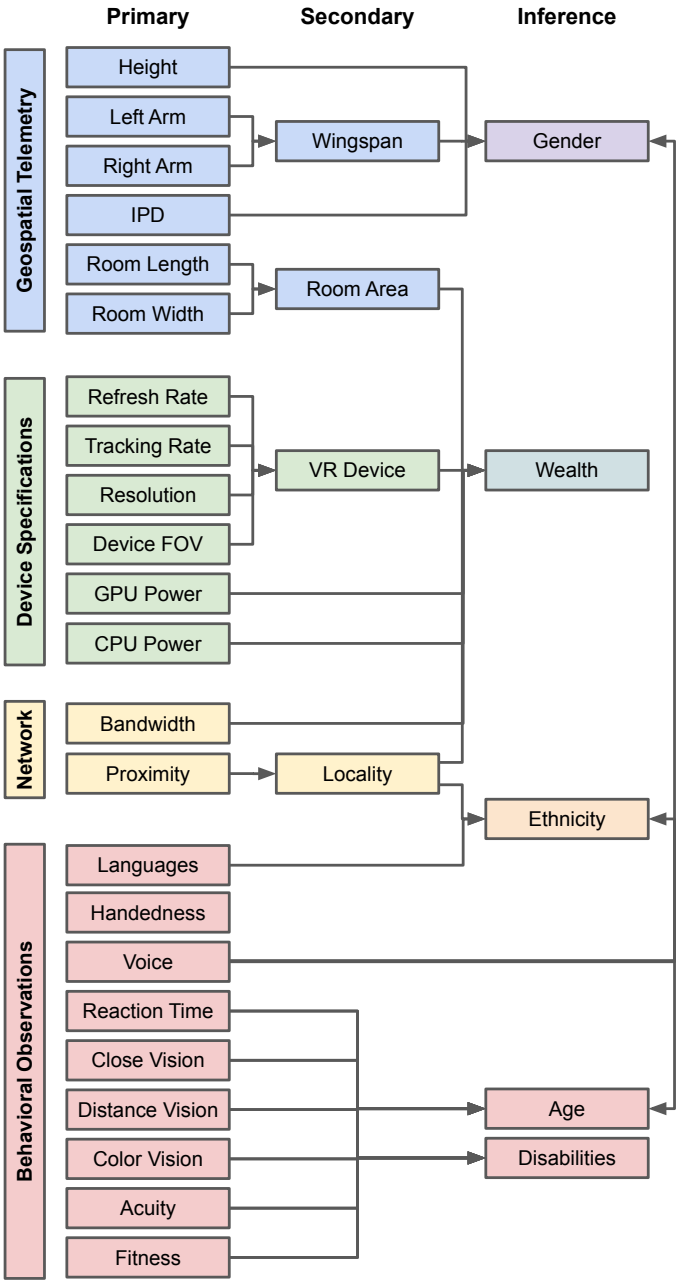


Fig. 2: Taxonomy of VR-derived data attributes.

information by creating puzzles that only users of high-fidelity devices can feasibly solve. Determining the specifications of a user's device can reveal personal information about the end-users themselves; for instance, the cost of commercially-available VR setups spans at least two orders of magnitude; thus, determining the exact hardware of a target user may reveal their level of income/wealth.

Network Observations. An additional source of information about a target user is the observation of network characteristics. While not necessarily unique to virtual reality, attacks that leverage network observations to geolocate users are a natural fit for virtual telepresence applications, which often facilitate the use of multiple game servers to minimize perceived la-

tency [81]. Thus, privileged attackers II and III can efficiently capitalize on such attacks.

Behavioral Observations. Behavioral observations are a fourth key source of private information enabled by virtual reality applications, and observing how users react to carefully chosen stimuli can reveal a wide variety of personal information. Attacks based on observing user behavior typically require less privilege than other types of attacks discussed herein, with even non-privileged attackers typically receiving enough information to observe general user interactions. We also include listening to user vocalizations in this category (audio), although one could also consider it a category.

III. VR Privacy Attacks

A. *Biometrics*

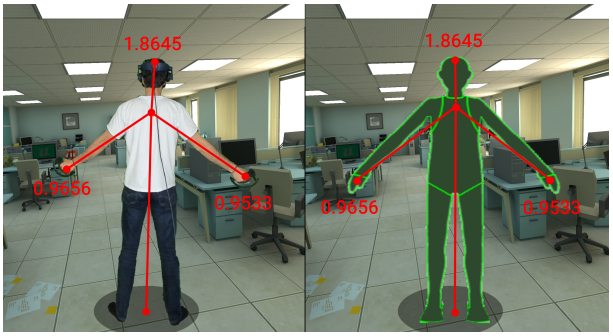


Fig. 3: Measuring user anthropometrics from telemetry.

Continuous Anthropometrics. Fig. 3 illustrates how attackers can directly measure a user's anthropometrics from VR telemetry. While basic headset-and-controller setups are sufficient to reveal height, arm lengths, and wingspan, more advanced full-body tracking systems can yield additional anthropometric measurements. Additionally, measuring the distance between the virtual cameras used to render an image for each eye can also reveal a user's interpupillary distance (IPD).

Fig. 4: Estimating handedness from behavior.

Binary Anthropometrics. An attacker can collect binary anthropometrics, which include characteristics such as longer-arm and dominant handedness, both directly from telemetry (e.g., "which hand moves more?") and from behavior (e.g., "which hand is used to press a button?"). Fig. 4 illustrates an example process of determining a user's handedness. The handedness attack is inspired, in part, by a similar method used previously in smartphones [5].

Fig. 5: VR puzzle revealing deuteranopia.

Vision. VR attackers can carefully construct interactive elements that secretly reveal aspects of a player's visual acuity, such as nearsightedness, farsightedness, or color blindness. For example, Fig. 5 shows a puzzle element of our VR game that appears innocuous to most users but is not solvable by users with red-green color blindness (deuteranopia).

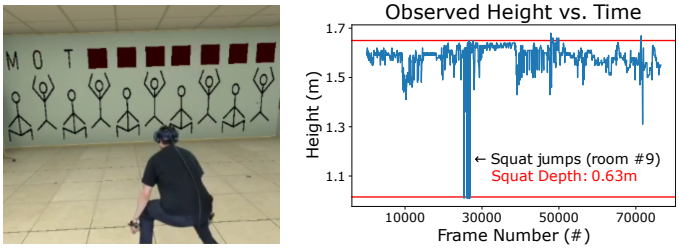


Fig. 6: Measurement of physical fitness.

Fitness. An attacker could also use behavioral and telemetric measurements to asses a subject's degree of physical fitness. Fig. 6 illustrates a virtual room designed to elicit physical activity and shows the resulting metric of physical fitness measurable on a headset position (y-coordinate) vs. time graph. We observed that a squat depth of less than 25% of height corresponded to low physical fitness. An extreme lack of fitness may reveal a participant's age or physical disabilities.

Fig. 7: VR puzzle measuring reaction time.

Reaction Time. Fig. 7 shows a VR environment constructed to reveal the participant's reaction time by measuring the time interval between a visual stimulus and motor response. Reaction time is strongly correlated with age [84].

B. Environment

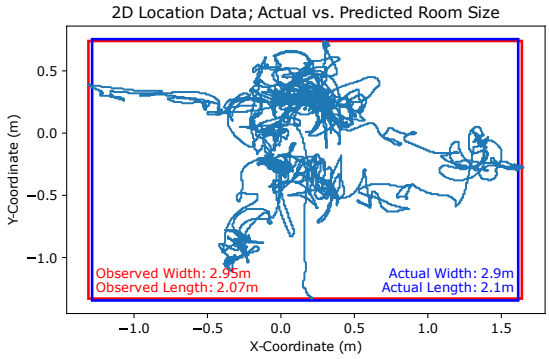


Fig. 8: Estimating room size from geospatial telemetry.

Room Size. Fig. 8 shows how an attacker could estimate the size of a user's physical environment by tracking their virtual movements. We designed virtual environments that contain interactive elements which encourage the participant to explore the boundaries of their physical environment.

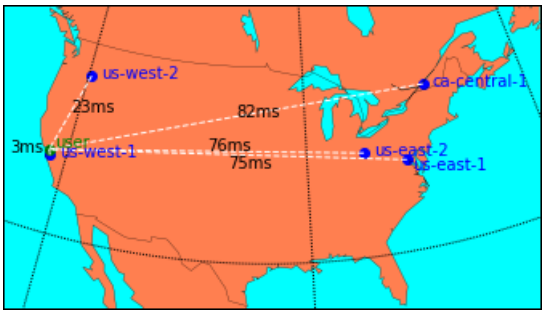
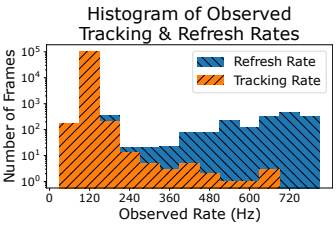


Fig. 9: Estimating user location from network latency.

Geolocation. Fig. 9 shows how observing the round-trip delay between a client device and multiple game servers (proximity) can reveal an end user's location (locality) via multilateration. A non-privileged attacker could use the round trip delay of audio signals as an approximate measure of latency.

C. Device Specifications



(a) Measuring headset and tracking refresh rate from device API throughput (log scale).

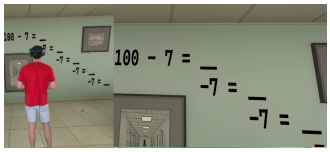
(b) Environment designed to reveal headset refresh rate via perceived motion differences.

Fig. 10: Methods of attaining VR device metrics.

VR Device. We assume that privileged attackers I and II have intrinsic knowledge of the VR device specifications via direct API interaction. Fig. 10a shows how privileged attacker III may use telemetry throughput to determine the refresh rate of a target user's head-mounted display (HMD) and controller tracking. Further, Fig. 10b shows how even a non-privileged attacker can construct a virtual environment that users perceive differently depending on their devices' refresh rate (see puzzle 15 in Appendix A). Determining several specifications such as refresh rate, resolution, and field of view, is sufficient to reveal the exact make and model of the VR device.

Host Device. Privileged attackers can also embed a variety of standardized benchmarks in their source code to assess the quality of the target user's host device (gaming computer). An attacker can use metrics such as CPU power, GPU power, and network bandwidth to reveal the age and price tier of the system and, thus, potentially the wealth of the target user.

(a) abstraction



(b) attention

(c) naming

(d) orientation

Fig. 11: Methods of measuring cognitive acuity.

D. Acuity (MoCA)

A number of standardized cognitive, diagnostic, and aptitude tests can be adapted for VR environments. Fig. 11 illustrates VR environments designed to covertly asses four categories of the Montreal Cognitive Assessment (MoCA): abstraction (11a), attention (11b), naming (11c), and orientation (11d).

E. Demographics

Fig. 12: Determining language from user behavior.

Language. There are a number of ways to ascertain a user's spoken language(s) in VR, including via speech recognition. Fig. 12 illustrates how a non-privileged attacker can observe a user's direction of gaze while solving a puzzle to reveal the languages they speak.

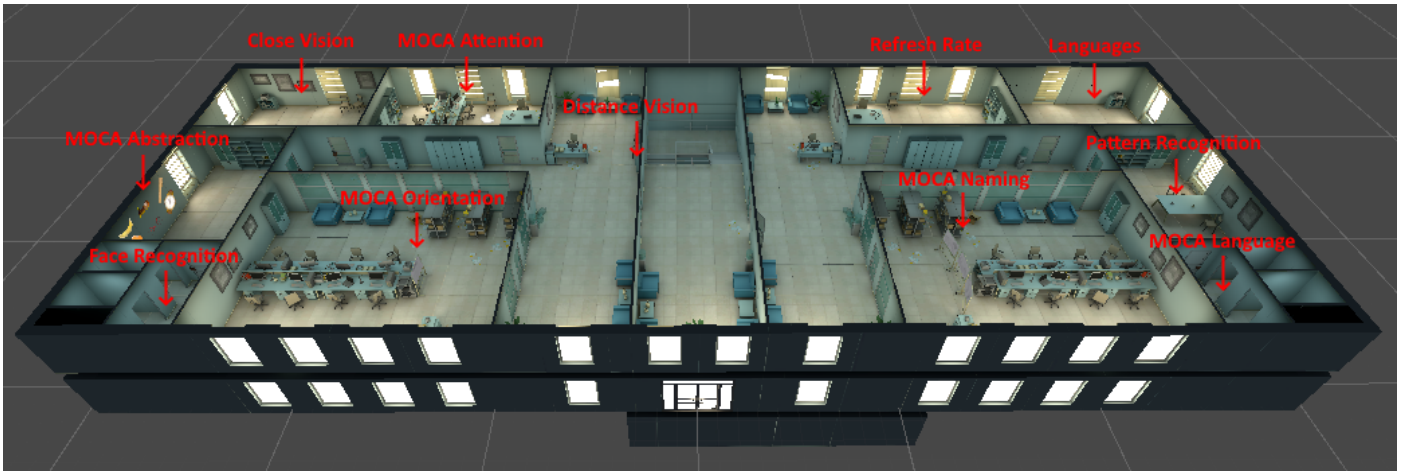


Fig. 13: Virtual office building hosting the puzzle rooms.

Vocal Characteristics. Listening to the voice of a user may reveal key demographic attributes such as age, gender, and ethnicity [6], [18]. Shared VR environments with voice streaming provide a strong opportunity to capitalize upon such a feature, as attackers can cue target users to speak certain words or phrases that reveal more information.

Inferred Attributes. While most demographic attributes cannot be observed directly from VR data, an attacker could often accurately infer them from primary data attributes. For example, height, wingspan, and IPD correlate strongly with gender, while eyesight, reaction time, and fitness correlate with age.

IV. Experimental Design

The question we aim to answer is whether, and to what degree, an attacker can use data collected from consumer-grade VR devices to accurately extract and infer users' private information. This section details the experimental design, technical setup, and protocol used to answer this important question. We begun by identifying a number of privacy-sensitive variables we believed to be uniquely accessible within VR. We designed and implemented systematic methods to collect and analyze these variables from within VR applications, as summarized in section III. To test the efficacy of these attacks, we designed an "escape room"-style VR game themed as an office building (see Fig. 13). We then disguised the attacks as a set of puzzles within the game, which users were highly motivated to solve to the best of their ability in order to unlock a sequence of doors and win the game. We describe the exact puzzles in detail in Appendix A. We endeavored to design the experiment such that it did not bluntly reveal the ulterior goal (namely facilitating the measurement of the variables in Fig. 2), thereby illustrating how other VR applications could also accomplish the same goal covertly. To this end, we also added innocuous (i.e., "noisy") rooms which did not necessarily collect meaningful personal information, but instead served to camouflage the data-harvesting puzzles.

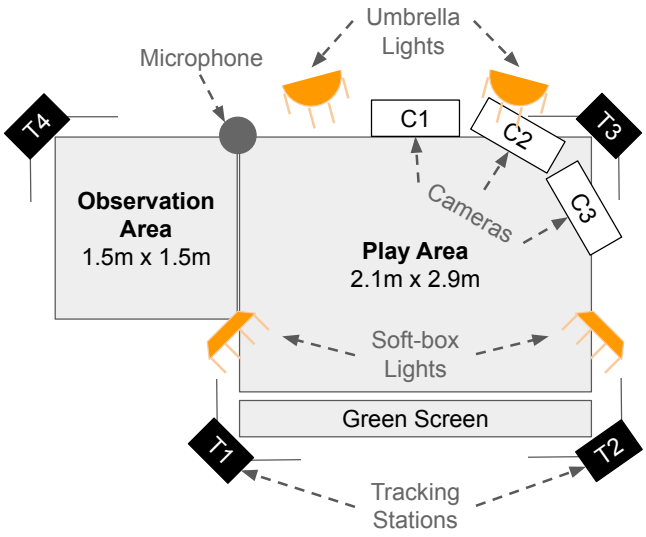


Fig. 14: VR laboratory room layout.

A. Setup and Protocol

We recruited 30 individuals for the experiments (6 female and 24 male, 18–64 years old, with $\bar{x} = 27.3$ yrs and $s = 11.1$ yrs). The recruiting channels we had access to were predominantly department-specific; as such, the demographics of our participants mirror those of our own department within our institution.

After completing a thorough informed consent and orientation process, we helped the participants don a VR headset (HTC Vive, Vive Pro 2, or Oculus Quest 2) and its handheld controllers (Vive Controllers, Valve Index Controllers, or Oculus Quest Controllers, respectively), after which the participant proceeded to play the VR game (see the laboratory room layout in Fig. 14 and the primary VR setup in Fig. 15). Each headset was paired with a gaming computer sufficiently powerful to run it at full fidelity; the main experimental setup

We tested three devices to determine if there were any noteworthy differences, which we did not observe other than in IPD (see §V-A).

had 64 GB of RAM, an AMD Ryzen 9 5950X CPU, and an Nvidia RTX 3090 GPU. Finally, the participants completed a post-game survey to collect the "ground truth" values for attributes of interest.

Each experiment lasted approximately 10–20 minutes within VR, plus around 10 minutes for completing the survey. Throughout the experiments, we minimized the interactions with the participants and ensured their safety by intervening when they approached a wall in the room. The experiments remained the same for all participants; we did not alter the game play-through or logic. The game collected the targeted data points in CSV format during the play-through. Furthermore, the researchers manually annotated data points for data collection that required game development beyond what is reasonable for this study, e.g., automating voice recognition to register the escape room "passwords" (solutions) the participants articulated aloud. The researchers pressed keys on a keyboard to trigger animations in the virtual environment and teleport the player between rooms. These elements could be automated in a production-ready VR game.

Fig. 15: Experimental setup.

Once the experiment ended, the participants filled out a form with their ground truth, which we used to validate the accuracy of the proposed privacy attacks. To collect the ground truth unknown to the participants themselves, we performed onsite measurements, e.g., we annotated the VR device and VR-room area, tested their reaction time with a desktop app, and measured their height and wingspan with a metric tape. Furthermore, knowing that researchers have studied the use of cognitive assessments in the diagnosis of attention disorders [60], autism [25], PTSD [40], and dementia [83], we chose the Montreal cognitive assessment (MoCA) [27] as a simple example of what advanced, immersive VR games could hide in their play-throughs. We randomized the order of the VR experiment and paper MoCA test (with half the participants taking the MoCA before and with the other half after the experiment) to neutralize potential biases in either direction. Once we collected the ground truth, we ran our analysis scripts (privacy attacks) over the collected data to compile and infer data points, which we compared to the ground truth to assess the attacks' accuracy.

B. Ethical considerations

We identified three primary ethical risks in our protocol: (i) the risk of discomfort using a VR device, (ii) the risk of a confidentiality breach of participant data, and (iii) the risk that participants might not have wished to disclose certain information about themselves during the course of the study.

To address the first risk (i), we used high-fidelity VR devices and appropriately powerful gaming computers for all participants, together capable of consistently providing 120 frames per second, well above the minimum specifications recommended to mitigate the risk of VR sickness [67]. We designed our VR game to avoid distressing elements such as horror, claustrophobia, or flickering/strobing lights. Furthermore, a researcher was present to ensure participants did not collide with real-world objects during each play-through.

To address the second risk (ii), we anonymized all collected data using random alphanumeric codes that we could not reasonably trace back to a participant's identity. Moreover, we avoided collecting any highly-sensitive data that could potentially damage participants in a breach. Lastly, we normalized biometric measurements on a scale of 0 to 1 to avoid revealing exact measurements in this paper (e.g., in Fig. 16).

To address the third risk (iii), we made sure participants clearly understood the nature of the study. We emphasize that this is not a deception study. Our claims about the non-obviousness of the presented attacks should not be construed to imply that participants were unaware that their data was being collected during the study. Participants were informed that their data was being collected, including a description of the categories of data being observed. After completing the VR portion of the study, participants were made aware of the exact attributes being collected. They were explicitly given the opportunity to withdraw consent without penalty at any point in the process, including after having detailed knowledge of the data attributes involved, in which case their data would not have been included in the results.

In light of these considerations, the study was deemed a minimal-risk behavioral intervention and was granted an IRB exempt certification under 45 C.F.R. § 46.104(d)(3) by an OHRP-registered institutional review board.

Attribute	Type / Source	Precision	Accuracy	Attackers
Height	Primary Telemetry	1 cm	70% within 5 cm 100% within 7 cm	Privileged I-III Non-Privileged*
Longer Arm	Primary Telemetry	boolean	64% for ≥ 1 difference 100% for ≥ 3 cm difference	Privileged I-III Non-Privileged*
Interpupillary Distance	Primary Telemetry	0.1 mm	96% within 0.5 mm (Vive Pro 2) 87% within 0.5 mm (All Devices)	Privileged I-II
Wingspan	Secondary Telemetry	1 cm	86% within 7 cm 100% within 12 cm	Privileged I-III Non-Privileged*
Room Size	Secondary Telemetry	1 m^2	78% within 2 m^2 97% within 3 m^2	Privileged I-III Non-Privileged*
Geolocation	Primary Network	100 km	50% within 400 km 90% within 500 km	Privileged II-III
HMD Refresh Rate	Primary Device	1 Hz	100% within 3 Hz (Privileged Attacker) 81% wtihin 60 Hz (Unprivileged Attacker)	Privileged I-II Privileged III* Non-Privileged*
Controller Tracking Rate	Primary Device	1 Hz	100% within 2.5 Hz	Privileged I-II Privileged III* Non-Privileged*
Device Resolution (MP)	Primary Device	0.1 MP	100% within 0.1 MP	Privileged I-II
Device FOV	Primary Device	10°	100% within 10°	Privileged I-II Privileged III* Non-Privileged*
Computational Power	Primary Device	0.1 GHz 10 Mh/s	CPU: 100% within 0.4 GHz GPU: 100% within 20 Mh/s	Privileged I-II
VR Device	Secondary Device	N/A	100%	Privileged I-III Non-Privileged*
Handedness	Primary Behavior	boolean	97%†	Privileged I-III Non-Privileged
Eyesight	Primary Behavior	boolean	70% (Hyperopia) 81% (Myopia)	Privileged I-III Non-Privileged
Color Blindness	Primary Behavior	boolean	100%	Privileged I-III Non-Privileged
Languages	Primary Behavior	boolean	88%	Privileged I-III Non-Privileged
Physical Fitness	Primary Behavior	boolean	90%	Privileged I-III Non-Privileged
Reaction Time	Primary Behavior	17 ms	88%	Privileged I-II Privileged III* Non-Privileged*
Acuity (MoCA)	Primary Behavior	1 point	81% within 1 point 90% within 2 points 100% diagnostic accuracy	Privileged I-III Non-Privileged
Gender	Inferred Classification	boolean	100%	Privileged I-III Non-Privileged
Age	Inferred Regression	1 yr	100% within 1 yr	Privileged I-III Non-Privileged
Ethnicity	Inferred Classification	categorical	100%	Privileged I-III Non-Privileged
Income	Inferred Regression	$1k	100% within $25k	Privileged I-III Non-Privileged
Disability Status‡	Inferred Classification	boolean	100%	Privileged I-III Non-Privileged

* With degraded accuracy. † Only $1/30$ were left-handed. ‡ No physical disabilities observed, see §V-E.

TABLE II: Selected attributes collected and analyzed during the experiment.

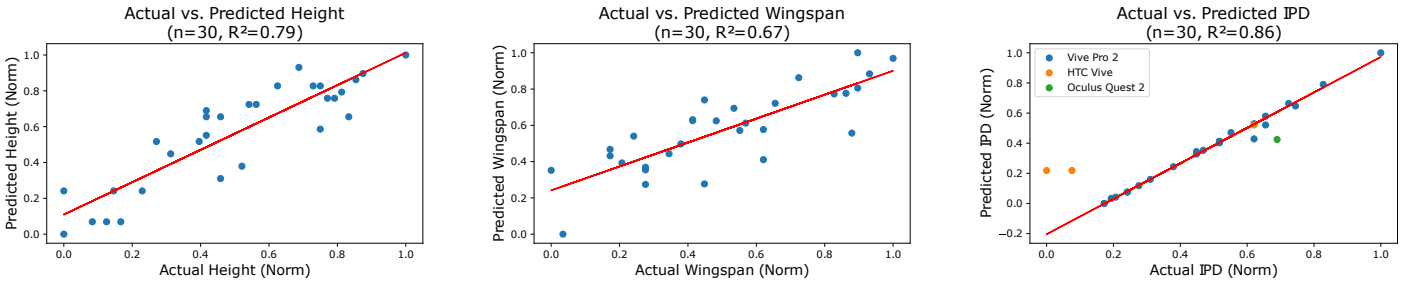


Fig. 16: Actual and predicted user anthropometrics.

V. Results

In this section, we present the empirical effectiveness of the privacy attacks introduced in § III, as summarized in Table II.

A. Biometrics

Continuous Anthropometrics. Fig. 16 shows (scaled) actual and predicted values for *height* ($R^2 = 0.79$), *wingspan* ($R^2 = 0.67$), and *interpupillary distance* (IPD) ($R^2 = 0.86$). IPD measurements were most accurate on the Vive Pro 2, with $R^2 = 0.99$ when excluding other devices. In general, we could accurately determine these three metrics for most users from just a few seconds of telemetry. We were not, however, able to accurately predict the individual lengths of the left and right arms ($R^2 = 0.01$ and $R^2 = 0.08$ respectively), due to the lack of a reliable center point from which to measure.

Binary Anthropometrics. Although absolute arm lengths were not discernible, relative lengths were accurate enough that we could usually identify which of the participant's arms was longer. We observed increasing accuracy for participants with greater differences in length, reaching 100% accuracy for the 13% of participants with a difference of at least 3 cm. We believe that handedness can also be determined accurately from certain behavioral observations; we note, however, that 97% of our participants were right-handed.

Vision. Our vision tests achieved diagnostic accuracies for *hyperopia* (farsightedness), *myopia* (nearsightedness), and *deuteranopia* (red-green color blindness) of 70%, 81%, and 100% respectively. The overall accuracy of detecting a visual deficiency was 81%, in part because some users of contact lenses could not remove their contacts for the experiment.

Fitness. Using squat depth as a correlate of *physical fitness* discriminated "low" fitness with an accuracy of 90%; our tests were not able to differentiate between "moderate" and "high" fitness.

Reaction Time. We measured *reaction time* to a precision of one recorded frame (16.6 ms). We were able to detect whether a participant's reaction time was above or below 250 ms (the approximate median reaction time) with an accuracy of 88%.

B. Environment

Room Size. The *length* and *width* of each of three testing rooms was determined to within 1.0 m with accuracies of 90% and 100% respectively. This allowed true room area to be found within 3 m^2 in 97% of trials. Taking the average estimated area for each tested room vs. the true accessible room area yields $R^2 = 0.97$.

Geolocation. Using the server latency multilateration (hyperbolic positioning) technique for *geolocation* yielded a mean longitudinal error of 2.58° and mean latitudinal error of 2.50° across three tested locations. This was sufficient to locate the test subject to within 500 km in 94% of cases, and within the correct U.S. state in 100% of cases.

C. Device Specifications

VR Device. We found that privileged attackers could determine various VR Device specifications (namely, *display refresh rate*, *display resolution*, *field of view*, and *tracking rate*) with 100% accuracy. This allows privileged attackers to determine the type of VR device with 100% accuracy. We also found that non-privileged attackers could determine the refresh rate to within 30 Hz with an accuracy of 38% and to within 60 Hz with an accuracy of 81%; however, this was not sufficient to accurately determine the type of device.

Host Device. We found that an attacker benchmarking host device specifications can determine *GPU power* with 100% accuracy to within 20 Mh/s (daggerhashimoto) and *CPU clock speed* to within 0.4 GHz, allowing them to estimate the price tier of the host device.

D. Acuity (MoCA)

Table III summarizes the numerical (continuous, i.e., the score of each category) and diagnostic (binary, i.e., passing or failing a category) accuracy of the *Montreal Cognitive Assessment* (MoCA) we conducted in the VR experiments. We achieved a diagnostic accuracy of 90% or greater for 5 of the 7 scored MoCA categories (excluding visuospatial/executive and delayed recall), with an overall diagnostic accuracy of 100%.

MoCA Category	Accuracy (Numerical)	Accuracy (Diagnostic)
Executive	N/A	N/A
Naming	100%	100%
Memory	75%	75%
Serial 7	90%	100%
Attention	88%	100%
Repetition	75%	94%
Language	72%	94%
Abstraction	100%	100%
Recall	53%	84%
Orientation	100%	100%
Overall	81% within 1 point 90% within 2 points	100%

TABLE III: Accuracy of each MoCA category.

E. Demographics

Language. The visual focus method of *language* determination identified a spoken language (other than English) with at least conversational proficiency in 88% of multilingual participants.

Vocal Characteristics. We used existing machine learning models to determine the gender [6] and ethnicity [18] of participants from their voice with an accuracy of 97% and 63% respectively; these accuracy values improved to 100% when combined with other attributes such as height and wingspan as described in "Inferred Attributes" below.

Inferred Attributes. We used Azure Automated Machine Learning [1] to determine the optimal preprocessor, model type, and input metrics for inferring several demographic attributes. Table IV summarizes the results of this meta-analysis. Using the identified optimal models, we determined the participant's gender, ethnicity, disability status, age (within one year), and income (within $\$25,000$) with 100% accuracy across several Monte Carlo cross-validations. Participants using a VR device other than their own were excluded from consideration for the income attribute. In each case, the model far outperformed any individual attribute; for example, ethnicity was 100% accurate despite its most significant input (voice) being only 63% accurate on its own.

Attribute (Prediction)	Inputs	Preprocessing / Model
Gender (Classification)	Voice, Height, Wingspan, Interpupillary Distance (IPD)	TruncatedSVDWrapper SVM
Age (Regression)	Close Vision, Reaction Time, Height, Test Duration, Acuity	MaxAbsScaler ExtremeRandomTrees
Ethnicity (Classification)	Voice, Language, Height	StandardScalerWrapper LightGBM
Income (Regression)	VR Device, GPU Power, CPU Power	MaxAbsScaler XGBoostRegressor
Disabilities (Classification)	Vision, Fitness, Acuity	MaxAbsScaler NaiveBayes

TABLE IV: Inputs and methodology of inferred attributes.

With respect to disability, we did not observe any physical disabilities in our 30 participants; instead, we expanded the scope of disabilities to include, for example, visual and cognitive impairments, for the purposes of this study.

VI. Discussion

We now return to the question of whether an attacker can use data collected from consumer-grade VR devices to extract and infer users' private information accurately. In this study, we have shown that this is indeed possible, with moderate to high accuracy values for most of the aggregated and inferred data points presented in Table II. We found that an attacker could uniquely and consistently identify a participant among the pool of 30 within a few minutes of gameplay and based on as few as two data points: *height* and *wingspan*. Moreover, we have collected more than 25 granular data points, well above the 15 necessary to uniquely identify every individual in the United States [61]. While we were required to condense this data collection into a concise 20-minute experiment for logistical reasons, real-world attackers could gain increased accuracy and covertness by aggregating data collected over much longer periods of time.

In sections III and IV, we argued that a developer could design VR environments and games to facilitate the covert collection of targeted data points disguised as normal game elements. Indeed, after the experiment, all 30 participants reported not knowing exactly which attributes were being collected and inferred during the game. Many participants expressed surprise during both the initial consent process and the later debrief at the breadth of information that could be collected within VR, but none expressed particular shock at the existence of some degree of data harvesting (perhaps having already grown accustomed to these practices in other environments).

While for ethical reasons we limited our attacks to relatively benign data points, an attacker could potentially track and infer additional information about other more critically sensitive personality traits, like sexual, religious, or political orientation, educational level, and illnesses, among others, to enhance practices such as surveillance advertisement [11] or pushing political agendas [55]. Given how immersive and emotionally engaging VR environments can be [54], [28], [82], [38], such practices could become more pernicious and effective than with current mobile and desktop applications.

Although we use the terms "attack" and "attacker" throughout this paper, to the best of our knowledge, there is nothing strictly illegal about the methods described herein. It is therefore possible that in the future, many VR users would knowingly or unknowingly consent to this form of data collection via clauses contained in platform terms of service or end-user license agreements. In fact, major VR device manufacturers have been observed selling headsets at a loss [58] of up to $10 billion per year [59], and it is evident that these corporations will aim to recoup said losses with some form of after-sales revenue.

Limitations. We would like to note that our sample of participants was unfortunately not perfectly representative of the general population; for example, there were more men than women (due to the demographics of the department population from which we recruited), we had only one left-handed participant, and none had physical disabilities (but other types of disabilities were also considered). The majority of our participants were students. For logistical reasons, we were unable to tamper with VR device firmware and thus could not

consider hardware-level attacks in this paper. Therefore, privileged attackers I and II, while different in theory, had identical capabilities within the scope of our experiment. Furthermore, we obtained our results through high-fidelity latest-generation VR devices; as such, they may not be applicable to lower-fidelity devices. Lastly, the researchers were forced to interact with participants outside of VR on some occasions, such as to warn of nearby obstacles. While we did attempt to minimize such occurrences, these interactions could nevertheless have biased certain results (particularly behavior).

Future work. Given the early stage of research on privacy and VR, there are many outstanding questions in this field for researchers to tackle. Among them is the question of how developers can design VR games or applications that make privacy attacks even more stealthy, including by integrating these attacks into daily tasks in future VR/AR environments. On the other hand, researchers could also study analysis techniques for revealing hidden data collection mechanisms (where possible) to make these attacks harder to achieve. Furthermore, studying what additional data attributes an attacker could leverage from data sources we did not consider (including eye tracking and full-body tracking) will expand our overall awareness of VR-related vulnerabilities. Additionally, future work dedicated to how an attacker could not just observe but actually change users' opinions will shed light on the implications of future immersive and impactful metaverse applications. Above all, we think it most important to study the potential countermeasures to these VR privacy attacks, such as by adding noise to raw VR device data without compromising the user experience.

VII. Related Work

Virtual reality (VR) is an interdisciplinary field of research used in a multitude of contexts such as education [31], [19], healthcare [73], transportation [45], [22], [35], [44], work environments [34], [78], productivity [29], [64], and entertainment [66], [24]. Additionally, new fields open as the current market trends push VR to become an extension of the social internet (where security and privacy are critical) in the form of the so-called "metaverse". There are many related works on privacy and the web, e.g., privacy attacks on web browsers [37], leveraging social media data [49], [20] and searchable personal information [41], and on internet privacy policies [33]. There is also research on privacy attacks on mobile location data [79], smart wearables [26], and across mobile and desktop browsers and mobile applications [46]. As VR environments become an increasingly prevalent part of the social internet [51], these attacks may overlap and expose VR users' private information.

Regarding related work on VR and privacy specifically, top searches of studies related to the "*metaverse*", or "*virtual reality*" and "*privacy*" in digital libraries such as IEEE [23], ACM [3], ScienceDirect [68], or Springer Link [70] and references cited thereof revealed high-level literature reviews related to privacy in VR [56], [9], [32], [16], [51], [77], [42], [13], [12]. Notably, O'Brolcháin et al. [56] highlights the ethical concerns of converging social networks with VR, and Falchuk et al. [16] qualitatively discusses potential privacy protections in the metaverse, e.g., producing virtual clones to mask the authentic user. These works highlight the importance of addressing privacy in VR, which is a sensitive environment and may emotionally charge users as VR emulates the real world [54], [28], [82], [38]. However, these studies lack technical implementations and practical demonstrations of VR-specific attacks.

Among technical works in the field of privacy and VR [21], [80], [43], [48], [38], [47], [72], [39], [8], notable studies investigate the consequences of traditional security and privacy attack vectors on VR learning environments, e.g., packet sniffing, shoulder surfing, or network attacks, and create a risk assessment framework thereof [21], [80]. Furthermore, Martinovic et al. [43] study shared similarities to our work's goal and method despite studying privacy in brain-computer interfaces instead of in VR. Moreover, given that eye movement can reveal users' gender, age, and interest in a scene, Steil et al. [72] and Ao et al. [39] employed differential privacy to protect users' eye-tracking data without significantly compromising the utility of heatmaps for inferring, e.g., document types or reading speed. Other works are less involved, for instance, claiming privacy preservation by just capturing data in "short" time frames without a threat model [8]. Aware of these privacy attacks, Lim et al. [38] created a privacy tutorial for users to navigate VR safely. Lastly, researchers have also focused on full-body tracking data in VR environments. Many works revolve around identifying body motions for user authentication [50], [57], [65], [36], which led Miller et al. [48] to investigate the privacy implications in user identification with VR tracking body movements. However, Miller et al. and the rest of the literature did not include other critical VR data such as device specifications, network, and behavioral observations in combination to identify users.

Overall, we are the first to provide a holistic taxonomy of VR attackers, data sources, vulnerable attributes, and corresponding attacks that practically demonstrate how attackers may accurately harvest sensitive user information.

VIII. Conclusion

In this study, we shed light on the unprecedented privacy risks of the metaverse by showing how VR can be turned against its users. Specifically, we provided a comprehensive security and privacy framework for VR environments that classifies (i) attackers, (ii) data sources, (iii) vulnerable attributes, and their corresponding (iv) attacks. We demonstrated the practicality and accuracy of these attacks by designing and conducting experiments with 30 participants using consumer-grade VR devices. The participants played our "escape room" VR game, which was secretly designed to collect personal information, like biometrics, demographics, and VR device and network details, among numerous other data points. The results demonstrate high information leakage with moderate to high accuracy values over most identified vulnerable attributes, with just a handful of these attributes being sufficient to uniquely identify a user [62], [75], [53], [17], [30], [14], [76], [4].

The alarming accuracy and covertness of these attacks and the push of data-hungry companies towards metaverse technologies indicate that data collection and inference practices in VR environments will soon become more pervasive in our daily lives. Furthermore, the breadth of possible VR applications, increasing quality of VR devices, and relative

simplicity of our demonstration, all suggest that more sophisticated attacks with a higher success rate are possible and perhaps on the horizon. Therefore, we hope our work encourages other privacy practitioners to advance research at the intersection of privacy and VR, in particular to propose countermeasures for new and existing privacy attacks in the metaverse.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation, by the National Physical Science Consortium, and by the Fannie and John Hertz Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the supporting entities. We sincerely thank our study participants for making this work possible.

AVAILABILITY

The Unity (C#) source code and compiled binaries for the "escape room" VR game we designed for our experiments are available for download from our public repository [52]. The repository also contains the data collection instruments and data analysis scripts used to determine the primary attributes and sample data with which researchers can experiment.

https://github.com/MetaGuard/MetaData

REFERENCES

[1] Azure Automated Machine Learning - AutoML | Microsoft Azure.

[2] The Very Real History of Virtual Reality (+A Look Ahead).

[3] ACM. ACM Digital Library. https://dl.acm.org/. Online; accessed 20 May 2022.

[4] Maryam Archie, Sophie Gershon, Abigail Katcoff, and Aileen Zeng. Who′s watching? de-anonymization of netflix reviews using amazon reviews. 2018.

[5] Jeff Avery, Daniel Vogel, Edward Lank, Damien Masson, and Hanae Rateau. Holding patterns: detecting handedness with a moving smartphone at pickup. In *Proceedings of the 31st Conference on l'Interaction Homme-Machine - IHM '19*, pages 1–7, Grenoble, France, 2019. ACM Press.

[6] Kory Becker. primaryobjects/voice-gender, May 2022. original-date: 2016-06-09T14:30:44Z.

[7] Blur Busters. UFO Motion Tests. https://www.testufo.com/. Online; accessed 30 April 2022.

[8] Efe Bozkir, David Geisler, and Enkelejda Kasneci. Person independent, privacy preserving, and real time assessment of cognitive load using eye tracking in a virtual reality setup. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1834–1837, 2019.

[9] Kent Bye, Diane Hosfelt, Sam Chase, Matt Miesnieks, and Taylor Beck. The ethical and privacy implications of mixed reality. In *ACM SIGGRAPH 2019 Panels*, SIGGRAPH '19, New York, NY, USA, 2019. Association for Computing Machinery.

[10] Aaron Cahn, Scott Alfeld, Paul Barford, and S. Muthukrishnan. An empirical study of web cookies. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, page 891–901, Republic and Canton of Geneva, CHE, 2016. International World Wide Web Conferences Steering Committee.

[11] Matthew Crain. *Profit Over Privacy*. Minneapolis: University of Minnesota Press, 2021.

[12] Jaybie A. De Guzman, Kanchana Thilakarathna, and Aruna Seneviratne. Security and privacy approaches in mixed reality: A literature survey. 52(6):1–37.

[13] Ellysse Dick. Balancing user privacy and innovation in augmented and virtual reality. page 28.

[14] Cynthia Dwork, Adam Smith, Thomas Steinke, and Jonathan Ullman. Exposed! a survey of attacks on private data. 4(1):61–84. Publisher: Annual Reviews.

[15] Jide S. Edu, Jose M. Such, and Guillermo Suarez-Tangil. Smart Home Personal Assistants: A Security and Privacy Review. *ACM Computing Surveys*, 53(6):1–36, November 2021. arXiv:1903.05593 [cs].

[16] Ben Falchuk, Shoshana Loeb, and Ralph Neff. The social metaverse: Battle for privacy. *IEEE Technology and Society Magazine*, 37(2):52–61, 2018.

[17] Xianyi Gao, Bernhard Firner, Shridatt Sugrim, Victor Kaiser-Pendergrast, Yulong Yang, and Janne Lindqvist. Elastic pathing: your speed is enough to track you. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '14 Adjunct*, pages 975–986, Seattle, Washington, 2014. ACM Press.

[18] Yatharth Garg. Speech-Accent-Recognition, May 2022. original-date: 2018-06-21T07:55:52Z.

[19] Joy Gisler, Valentin Holzwarth, Christian Hirt, and Andreas Kunz. Work-in-progress-enhancing training in virtual reality with hand tracking and a real tool. In *2021 7th International Conference of the Immersive Learning Research Network (iLRN)*, pages 1–3, 2021.

[20] Neil Zhenqiang Gong and Bin Liu. Attribute inference attacks in online social networks. *ACM Trans. Priv. Secur.*, 21(1), jan 2018.

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

[22] Philipp Hock, Sebastian Benedikter, Jan Gugenheimer, and Enrico Rukzio. Carvr: Enabling in-car virtual reality entertainment. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 4034–4044, New York, NY, USA, 2017. Association for Computing Machinery.

[23] IEEE. IEEE Xplore. https://ieeexplore.ieee.org. Online; accessed 20 May 2022.

[24] VRChat Inc. Vrchat. https://hello.vrchat.com/. Online; accessed 17 May 2022.

[25] W. Jarrold, P. Mundy, M. Gwaltney, J. Bailenson, N. Hatt, N. McIntyre, K. Kim, M. Solomon, S Novotny, and L. Swain. Social attention in a virtual public speaking task in higher functioning children with autism. *Autism Res.*, 2013.

[26] Carlos Jensen, Chandan Sarkar, Christian Jensen, and Colin Potts. Tracking website data-collection and privacy practices with the iwatch web crawler. In *Proceedings of the 3rd Symposium on Usable Privacy and Security*, SOUPS '07, page 29–40, New York, NY, USA, 2007. Association for Computing Machinery.

[27] Parunyou Julayanont and Ziad S. Nasreddine. Montreal cognitive assessment (MoCA): Concept and clinical review. In A. J. Larner, editor, *Cognitive Screening Instruments*, pages 139–195. Springer International Publishing.

[28] Orin S Kerr. Criminal law in virtual worlds. page 17, 2008.

[29] Pascal Knierim and Albrecht Schmidt. The virtual office of the future: Are centralized workplaces obsolete?

[30] Daniel Kondor, Behrooz Hashemian, Yves-Alexandre de Montjoye, and Carlo Ratti. Towards Matching User Mobility Traces in Large-Scale Datasets. *IEEE Transactions on Big Data*, 6(4):714–726, December 2020.

[31] Jun Lee, PhamSy Quy, Jee-In Kim, Lin-Woo Kang, Anna Seo, and HyungSeok Kim. A collaborative virtual reality environment for molecular biology. In *2009 International Symposium on Ubiquitous Virtual Reality*, pages 68–71, 2009.

[32] Ronald Leenes. Privacy in the metaverse. In Simone Fischer-Hübner, Penny Duquenoy, Albin Zuccato, and Leonardo Martucci, editors, *The Future of Identity in the Information Society*, pages 95–112, Boston, MA, 2008. Springer US.

[33] Stephen E. Levy and Carl Gutwin. Improving understanding of website privacy policies with fine-grained policy anchors. In *Proceedings of the 14th International Conference on World Wide Web*, WWW '05,

page 480–488, New York, NY, USA, 2005. Association for Computing Machinery.

[34] Jingyi Li, Ceenu George, Andrea Ngao, Kai Holländer, Stefan Mayer, and Andreas Butz. An exploration of users' thoughts on rear-seat productivity in virtual reality. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2020.

[35] Jingyi Li, Ceenu George, Andrea Ngao, Kai Holländer, Stefan Mayer, and Andreas Butz. Rear-seat productivity in virtual reality: Investigating vr interaction in the confined space of a car. *Multimodal Technologies and Interaction*, 5(4), 2021.

[36] Sugang Li, Ashwin Ashok, Yanyong Zhang, Chenren Xu, Janne Lindqvist, and Macro Gruteser. Whose move is it anyway? authenticating smart wearable devices using unique head movement patterns. In *2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 1–9, 2016.

[37] Bin Liang, Wei You, Liangkun Liu, Wenchang Shi, and Mario Heiderich. Scriptless timing attacks on web browser privacy. In *2014 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, pages 112–123, 2014.

[38] Junsu Lim, Hyeonggeun Yun, Auejin Ham, and Sunjun Kim. Mine yourself!: A role-playing privacy tutorial in virtual reality environment. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, CHI EA '22, New York, NY, USA, 2022. Association for Computing Machinery.

[39] Ao Liu, Lirong Xia, Andrew Duchowski, Reynold Bailey, Kenneth Holmqvist, and Eakta Jain. Differential privacy for eye-tracking data. ETRA '19, New York, NY, USA, 2019. Association for Computing Machinery.

[40] L. Loucks, C. Yasinski, SD. Norrholm, J. Maples-Keller, L. Post, L. Zwiebach, D. Fiorillo, M. Goodlin, T. Jovanovic, AA. Rizzo, and BO. Rothbaum. You can do that?!: Feasibility of virtual reality exposure therapy in the treatment of ptsd due to military sexual trauma. *Anxiety Disord.*, 2019.

[41] Ruxia Ma, Xiaofeng Meng, and Zhongyuan Wang. Preserving privacy on the searchable internet. iiWAS '11, page 238–245, New York, NY, USA, 2011. Association for Computing Machinery.

[42] Divine Maloney, Samaneh Zamanifard, and Guo Freeman. Anonymity vs. familiarity: Self-disclosure and privacy in social virtual reality. In *26th ACM Symposium on Virtual Reality Software and Technology*, VRST '20, New York, NY, USA, 2020. Association for Computing Machinery.

[43] Ivan Martinovic, Doug Davies, Mario Frank, Daniele Perito, Tomas Ros, and Dawn Song. On the feasibility of side-channel attacks with brain-computer interfaces. page 16.

[44] Mark McGill and Stephen Brewster. Virtual reality passenger experiences. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*, AutomotiveUI '19, page 434–441, New York, NY, USA, 2019. Association for Computing Machinery.

[45] Mark McGill, Julie Williamson, Alexander Ng, Frank Pollick, and Stephen Brewster. Challenges in passenger use of mixed reality headsets in cars and other transportation. 24(4):583–603.

[46] Maryam Mehrnezhad. A cross-platform evaluation of privacy notices and tracking practices. In *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*, pages 97–106, 2020.

[47] Liang Men and Danqi Zhao. *Designing Privacy for Collaborative Music Making in Virtual Reality*, page 93–100. Association for Computing Machinery, New York, NY, USA, 2021.

[48] Mark Roman Miller, Fernanda Herrera, Hanseul Jun, James A. Landay, and Jeremy N. Bailenson. Personal identifiability of user tracking data during observation of 360-degree VR video. 10(1):17404.

[49] Tehila Minkus, Yuan Ding, Ratan Dey, and Keith W. Ross. The city privacy attack: Combining social media and public records for detailed profiles of adults and children. In *Proceedings of the 2015 ACM on Conference on Online Social Networks*, COSN '15, page 71–81, New York, NY, USA, 2015. Association for Computing Machinery.

[50] Tahrima Mustafa, Richard Matovu, Abdul Serwadda, and Nicholas Muirhead. Unsure how to authenticate on your vr headset? come on, use your head! In *Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics*, IWSPA '18, page 23–30, New York, NY, USA, 2018. Association for Computing Machinery.

[51] Stylianos Mystakidis. Metaverse. 2(1):486–497.

[52] Vivek Nair and Gonzalo Munilla Garrido. MetaData Study. https://github.com/MetaGuard/MetaData. Online; accessed 22 May 2022.

[53] Arvind Narayanan and Vitaly Shmatikov. Robust De-anonymization of Large Sparse Datasets. In *2008 IEEE Symposium on Security and Privacy (sp 2008)*, pages 111–125, Oakland, CA, USA, May 2008. IEEE. ISSN: 1081-6011.

[54] John William Nelson. A virtual property solution: How privacy law can protect the citizens of virtual worlds. page 24, 2010.

[55] UK's Information Commissioner's Office. Audits of data protection compliance by uk political parties. https://ico.org.uk/media/action-weve-taken/2618567/audits-of-data-protection-compliance-by-uk-political-parties-summary-report.pdf. Online; accessed 17 May 2022.

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

[57] Ken Pfeuffer, Matthias J. Geiger, Sarah Prange, Lukas Mecke, Daniel Buschek, and Florian Alt. 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*, CHI '19, page 1–12, New York, NY, USA, 2019. Association for Computing Machinery.

[58] Alan Dexter published. Oculus will sell you a Quest 2 headset that doesn't need Facebook for an extra $500. *PC Gamer*, April 2021.

[59] Michael L. Hicks published. Despite Quest 2 sales success, Meta lost $10.2 billion on VR/AR last year, February 2022.

[60] A.A. Rizzo, T. Bowerly, C. Shahabi, J.G. Buckwalter, D. Klimchuk, and R. Mitura. Diagnosing attention disorders in a virtual classroom. *Computer*, 37(6):87–89, 2004.

[61] Luc Rocher, Julien M. Hendrickx, and Yves-Alexandre de Montjoye. Estimating the success of re-identifications in incomplete datasets using generative models. 10(1):3069.

[62] Luc Rocher, Julien M. Hendrickx, and Yves-Alexandre de Montjoye. Estimating the success of re-identifications in incomplete datasets using generative models. 10(1):3069.

[63] Black Rock. The metaverse: Investing in the future now. https://www.blackrock.com/us/individual/insights/metaverse-investing-in-the-future. Online; accessed 17 May 2022.

[64] C.C. Rodrigues and K.M. Pavlosky. An industrial application of telepresence technology: productivity improvements in material handling tasks. In *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*, volume 3, pages 2115–2120 vol.3, 1995.

[65] Cynthia E. Rogers, Alexander W. Witt, Alexander D. Solomon, and Krishna K. Venkatasubramanian. An approach for user identification for head-mounted displays. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, ISWC '15, page 143–146, New York, NY, USA, 2015. Association for Computing Machinery.

[66] Sylvia Rothe, Alexander Schmidt, Mario Montagud, Daniel Buschek, and Heinrich Hußmann. Social viewing in cinematic virtual reality: a design space for social movie applications. 25(3):613–630.

[67] NATO Science and Technology Organization. Guidelines for mitigating cybersickness in virtual reality systems.

[68] ScienceDirect. ScienceDirect Digital Library. https://www.sciencedirect.com/. Online; accessed 20 May 2022.

[69] We Are Social. Digital 2022: Another year of bumper growth. Online; accessed 17 May 2022.

[70] Springer Link. Springer Link Digital Library. https://link.springer.com/. Online; accessed 20 May 2022.

[71] Morgan Stanley. Metaverse: more evolutionary than revolutionary. Online; accessed 17 May 2022.

[72] Julian Steil, Inken Hagestedt, Michael Xuelin Huang, and Andreas Bulling. Privacy-aware eye tracking using differential privacy. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research &*

amp; Applications, ETRA '19, New York, NY, USA, 2019. Association for Computing Machinery.

[73] Robert Stone, Charlotte Small, James Knight, Cheng Qian, and Vishant Shingari. *Virtual Natural Environments for Restoration and Rehabilitation in Healthcare*, pages 497–521. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014.

[74] Labaton Sucharow. Record-breaking $650 million settlement of biometric privacy lawsuit reached by labaton sucharow, edelson, robbins geller and facebook. Online; accessed 23 May 2022.

[75] Latanya Sweeney. Simple demographics often identify people uniquely. page 34, 2000.

[76] Latanya Sweeney, Akua Abu, and Julia Winn. Identifying Participants in the Personal Genome Project by Name. *SSRN Electronic Journal*, 2013.

[77] Philipp Sykownik, Divine Maloney, Guo Freeman, and Maic Masuch. Something personal from the metaverse: Goals, topics, and contextual factors of self-disclosure in commercial social vr. In *CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.

[78] Krsna Das Thoondee and Andreas Oikonomou. Using virtual reality to reduce stress at work. In *2017 Computing Conference*, pages 492–499, 2017.

[79] Zhen Tu, Fengli Xu, Yong Li, Pengyu Zhang, and Depeng Jin. A new privacy breach: User trajectory recovery from aggregated mobility data. *IEEE/ACM Transactions on Networking*, 26(3):1446–1459, 2018.

[80] Samaikya Valluripally, Aniket Gulhane, Reshmi Mitra, Khaza Anuarul Hoque, and Prasad Calyam. Attack trees for security and privacy in social virtual reality learning environments. In *2020 IEEE 17th Annual Consumer Communications Networking Conference (CCNC)*, pages 1–9, 2020.

[81] VRChat. Network specs and tips.

[82] Ian Warren and Darren Palmer. Crime risks of three-dimensional virtual environments. page 7, 2010.

[83] P. Werner, S. Rabinowitz, E. Klinger, AD. Korczyn, and N. Josman. Use of the virtual action planning supermarket for the diagnosis of mild cognitive impairment: a preliminary study. *Dement. Geriatr. Cogn. Disord.*, 2009.

[84] David L. Woods, John M. Wyma, E. William Yund, Timothy J. Herron, and Bruce Reed. Age-related slowing of response selection and production in a visual choice reaction time task. *Frontiers in Human Neuroscience*, 9, 2015.

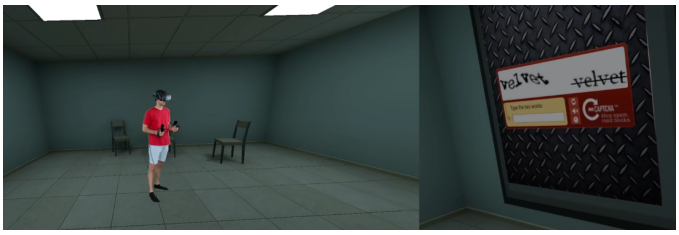
[85] Felix T Wu. Defining privacy and utility in data sets. *84 University of Colorado Law Review 1117 (2013); 2012 TRPC*, pages 1117–1177, 2012.

APPENDIX

This section describes the experiment design in detail. Our experiment consists of puzzles located in VR rooms that the participants visit. The puzzles are artifacts that facilitate collecting privacy-sensitive variables that might not otherwise be evident. The rooms are themed as a virtual office. Before initiating the game, we explained to the participants that they would find the password by solving a puzzle, thereby "escaping" the room. As developing a full-fledged game with voice recognition or virtual password pads is out of scope, the participants spoke the passwords aloud so that the researchers could press a key and "teleport" them to the next room. We include five "noisy" rooms, i.e., rooms that do not serve the purpose of facilitating the measurement of sensitive information but help to mask the rooms that do. Nonetheless, noisy rooms habituate the player to the game mechanics, e.g., looking around the room or immersing the player further in the game. If the player gets stuck in one room, we press a key to teleport the participant to the next room. We request the users to remove their glasses or contact lenses for puzzles 23 and 24, measuring eyesight. While influencing players in such a way is not possible in a real scenario, these puzzles could at least identify the players who do not have good eyesight, i.e., they do not wear glasses/contacts when playing.

Puzzle 1: The first room introduces the player to the dynamics of the game, containing only a door and a poster with the word "*hello*", which is the password. Upon instinctively reading the word aloud, the player is teleported to the next room.

Puzzle 2: The second room contains a poster with the password "*face*". The player spawns facing the opposite wall of the poster; thus, we accustom the player to turn and explore the virtual environment to find the password and reinforce finding and speaking the password aloud.



Puzzle 3: Similarly, a poster depicts a captcha with the word "*velvet*."

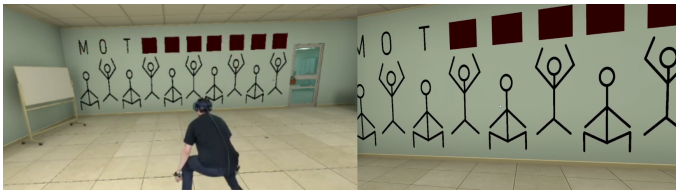
Puzzle 4: The room contains several tables with monitors, on whose screens are letters spelling "*church*" appropriately ordered from left to right.

Puzzle 5: This room tests for color blindness. Similarly to puzzle 4, monitors display letters on Ishihara color test plates. Without colorblindness, the player would read "*daisy*"; with colorblindness, the player would read "*as*" instead. Each of these passwords unlocks the room.

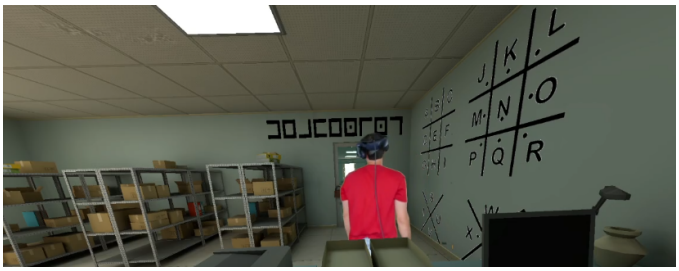
Puzzle 6: There is a button on a table; upon pressing it three times, the three balloons next to the opposite wall pop sequentially, revealing the password "*red*".

Puzzle 7: The puzzle tests the short-term memory of the participants (MoCA memory). A whiteboard displays seven rows arranged vertically, each with fill-in blanks. The first two rows contain the already filled-in words "*VR*" and "*hello*", respectively. The last five rows correspond to the previous passwords from puzzles 2 to 6. Connecting the highlighted letters sequentially from up to bottom, the participant reveals the password "*recluse*".

Puzzle 8: To measure wingspan, we depict on a wall four human stick figures with different poses. The participant must mimic the poses on the wall to uncover the four letters of the password "*cave*". One of the poses is a T-pose, which facilitates wingspan measurement.

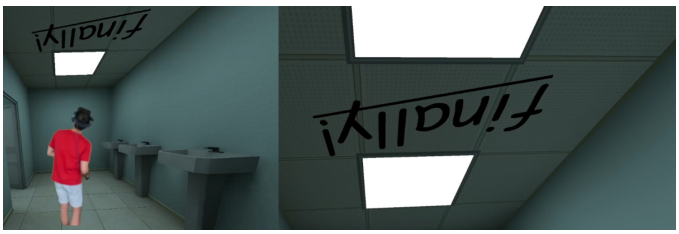


Puzzle 9: The participant must mimic the sequence of poses on the wall, a set of squats. For every squat, the participant uncovers two letters of the password "*motivation*". We correlate the distance traveled during the squats to fitness.



Puzzle 10: The (noisy) room depicts on a wall a pigpen cipher hiding the password "*deafening*".

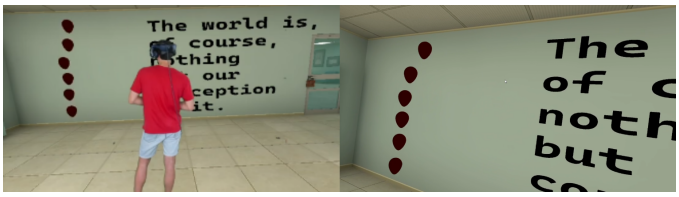
Puzzle 11: The player presses a button on a table in time with a visual input, thereby revealing their reaction time.



Puzzle 12: The (noisy) room presents the password "*finally*" on the ceiling, habituating the user to look also upwards.

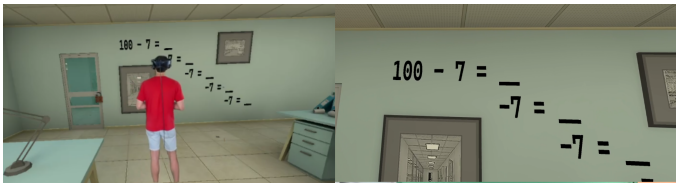
Puzzle 13: The room depicts the word "*apple*" in Hindi, Mandarin, French, Japanese, Russian, Spanish, Portuguese, and Arabic. The direction of gaze of the player when speaking the password reveals which language the participant recognizes.

Puzzle 14: This (noisy) room presents the sentence "*Everything you can do, I can do meta*" broken down vertically into five rows. To the left of each row, there is a shape. The last three shapes are the same (circles). To solve the puzzle, the participant must read aloud the words next to the first instance of the repeated shape "*I can*."



Puzzle 15: Similarly to puzzle 14 and inspired by screen refresh rate tests [7], we present a number of balloons moving at different refresh rates. Depending on the refresh rate of the VR device, users cannot distinguish between some balloons.

Puzzle 16: To deploy the "naming" MoCA task, the room presents three whiteboards depicting three animals.



Puzzle 17: To measure an "attention" task from MoCA, we present a serial seven subtraction starting at 100, the password is the sequence of numbers that lead to the final answer: "65."

Puzzle 18: This room contains puzzle 7, thereby measuring delayed recall from the MoCA test.

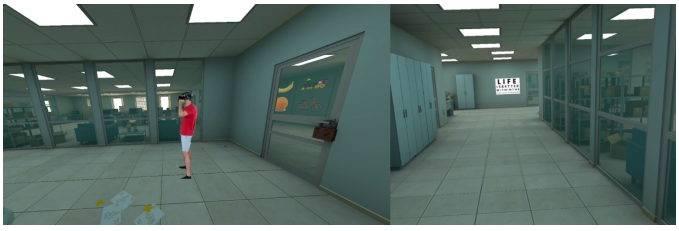
Puzzle 19: This room pictographically recreates the MoCA abstraction test.

Puzzle 20 (no image): To complete this audio-only room, the participant must repeat aloud two recorded sentences after listening to them once, thereby measuring one of the language tests of the MoCA.

Puzzle 21: The (noisy) room depicts three pictures of a famous physicist—"*Albert Einstein*" is the password.

Puzzle 22: The room presents calendar days on a whiteboard with "*Today?*" as the header and without disclosing the year, month, weekday, or date, which prompts the participant to identify the date of the experiment, thereby measuring one variable of the orientation task in MoCA.

Puzzle 23: We measure whether a participant can read the text at a close distance. We write the sentence "*The code is equal to three times four*" in four lines on the screen of a monitor, each line becoming more diminutive than the above line.



Puzzle 24: Similarly, we measure whether a participant can read the sentence "*Life is better within the digital playground*" at a long distance.