Perceptions of Tracking and Inference in VR

Wentao Liu, Kevin Guo, Meixiu Li
University of Maryland, Baltimore

Abstract

Kevin worked on creating the VR application, populating it with tracked objects, and implementing the cat minigame. Yuxin worked on implementing both forms of tracking (gaze and reaction time), integrating them with the VR application, and exporting data. Wentao worked on overall integration (including adding more tracked objects and debugging the VR application), the R script for data analysis, and designing/running the pilot user studies.

1 Introduction

Virtual reality (VR) applications expose users to potentially unfamiliar forms of data collection, including gaze tracking and a wide range of motion and position sensors. As VR becomes more popular, this data will likely be used by companies to make inferences about users, as is common nowadays across the Internet; however, given the relative novelty of these new forms of tracking, VR users may not know of or understand the changes and implications involved. We designed and piloted a user study designed to answer two research questions:

1. What are people's perceptions of tracking and inference in VR when reflecting on examples drawn from their own activity? How does perceived acceptability differ by context?
2. How does awareness of possible tracking and inference affect people's attitudes towards VR, such as their expressed level of comfort participating in VR?

USENIX Symposium on Usable Privacy and Security (SOUPS) 2023. August 6–8, 2023, Anaheim, CA, USA

In the rest of this report, we briefly discuss related work; explain our study method, including the design of the VR environment we created to expose participants to tracking; highlight some interesting findings from two pilot studies; and discuss how we plan to refine our method in preparation for a full user study.

2 Background

We first present background information on tracking and inference in VR, before reviewing related studies on user perceptions of tracking and inference both in VR and in other contexts.

2.1 Tracking and inference in VR

Previous articles and papers have extensively discussed tracking and inference in VR. O'Brolcháin et al. categorized threats posed by VR social networks to privacy and autonomy [12]. Kelly et al. evaluated seven popular VR devices (covering all major VR companies) and five popular third-party VR apps, analyzing what data companies collect and what they intend to do with personal information [6]. They found that all appear to track and monetize data, whether by selling it to third parties or by using it to drive marketing and target ads. Others have raised alarms about continuing threats to privacy [4,14], focused on specific forms of tracking as particularly new and concerning, including tracking eyes [2], facial expressions [5], and digital bodies (realistic avatars) [7].

In addition to expanding the forms of tracking, VR is also immersive in a way that is distinctive from most other applications. While a fitness tracker may collect a user's physiological responses, such as heart rate, a VR application can know exactly what is in the user's environment that might prompt certain responses, and it can even alter the environment to test those responses. This raises additional concerns for user privacy in VR.

Some researchers have implemented and tested proofs of concept for novel forms of tracking and inference in VR, such as identifying users from movement [8,13]. Still others have developed privacy-enhancing tools to protect users from tracking [9,15].

In a user study similar to our own, Nair et al. designed a

VR escape room designed to secretly harvest personal data from users in order to make inferences; they found that they could make accurate inferences without tipping off users, who were surprised by the breadth of data collection [11]. While their study shows that bespoke VR environments can be designed to make accurate inferences about users, ours aims to add passive tracking in more natural VR contexts, and our research goals are primarily concerned with users' reactions rather than accuracy.

2.2 User perceptions of tracking

Limited prior work on user perceptions suggests that novel forms of tracking and inference in VR may not be top of mind for users. VR users interviewed by Adams et al. in 2018 were primarily concerned about data collection from microphones and infrared sensors, and some felt that privacy would become a greater issue in the future [1]. In 2015, Motti et al. analyzed comments posted online and found minimal privacy concern around VR headsets [10]. Through our study, we aim to make privacy issues more tangible for users in the present so as to explore their reactions and perceptions.

Other studies outside of the VR context have taken a similar approach of showing participants how they were tracked in order to elicit reactions. Weinshel et al. created a browser extension that mimicked third-party trackers and visualized tracking data for users [18]; in a longitudinal field study, they found that this tool helped participants understand tracking better and led to increased intent to take privacy-preserving actions. Wei et al. showed participants their Twitter ad targeting data, identifying practices that felt more invasive and suggesting improvements to ad targeting explanations [17]. We take a similar approach, simulating tracking in VR for our participants and asking questions to understand their perceptions.

3 Method

In this section, we first describe how we designed a VR application to collect data from participants. We then describe our study method, in which participants interacted in this VR environment and then answered questions about their experience and perceptions. We note that the primary purpose of the VR activity was not to make accurate inferences about participants, but rather to simulate a company's process for making inferences plausibly enough that the experience would help participants explore new perspectives.

3.1 Tracking participants in a VR application

As the backbone of our user study, we created a VR application in Unity, starting from a free apartment asset in the Unity Asset Store. We chose an apartment because visiting a friend's virtual space is a common activity in VR social networks. To track user preferences from gaze, we added three categories of objects to the apartment:

1. music album covers on the walls of a bedroom, 2. an ad for different virtual cars on the living room TV,
3. and recipe books on the kitchen counter.

To increase the likelihood that participants would engage with these objects, we use a cat-chasing minigame to guide them to the right locations. The goal of this minigame is for the user to point the VR controller at the cat and click; the cat switches between various locations near the objects of interest after being found, which subtly guides the user to those locations. Furthermore, there is a spatialized meow sound effect that plays on loop to help cue the user to where they should look for the cat.

Since our current headset lacks the capability to provide accurate gaze-tracking data, our tracking system relies on the assumption that the user is looking at the center of the screen. For each frame, we project a ray from the center of the camera pointing forward and determine the intersection point of the ray with an object. By measuring the distance between the intersection point and all objects of interest, we identify the nearest neighbor. If the distance falls below a predefined threshold, we infer that the person is looking at that object. This threshold is carefully selected to minimize false positives.

Using a similar gaze tracking approach, we also monitor the user's reaction time. When the user initially enters the virtual world, we position the cat from the minigame to their left. When the cat meows for the first time, we start a timer, and the timer continues until the user looks at the cat (presumably guided by the spatialized audio). If this is under a threshold of about three seconds, we count it as a valid measurement of the user's reaction time.

Tracking data can be exported from the VR application in the form of a JSON file encoding the amount of time a user spent looking at each object of interest, as well as their reaction time if one was measured. We developed an R script that automatically analyzes this data to make inferences about the user. The types of tracking and inference we implemented were chosen after a review of related work based on two criteria: they explored forms of tracking that are relatively novel to VR, and they represented

data that we believe VR companies could plausibly monetize. The R script generated a visualization of each inference using ioslides. We made three different kinds of inferences:

1. For each category of objects for which we tracked a user's gaze, we inferred a product interest based what they spent the most time looking at.
2. Several of the tracked objects were associated with a liberal or conservative political leaning, based on prior work [3]; based on the average amount of time a user spent looking at objects associated with either side, we made an extremely naïve inference of their political leaning (an ad-relevant trait other than product interest).
3. Based on a user's reaction time as compared to a madeup baseline, we inferred whether their cognitive processing speed (a health-related trait) was below, around, or above average.

3.2 User study design

Our user study comprises three main portions, summarized in
Figure 1. First, participants spent around 10 minutes in our VR environment, while data about their activities was automatically collected in the background. We told participants to imagine they were entering a large virtual world designed for fun and socializing by a company called Total VR. We explained that the specific space contained in our VR application was designed by Total VR to help users get the hang of moving around in VR, and we instructed them to play the cat minigame and also spend some time familiarizing themself with the world and the controls.

Second, after exiting the VR environment, participants reviewed up to five examples of tracking and inference with us and answered questions about their sense of surprise and fairness. We also asked about whether they considered it acceptable for Total VR to keep these data and inferences to itself and use them to personalize ads and suggest content, and whether they considered it acceptable for Total VR to sell these data and inferences to other companies, who might combine them with other data about the participant to build a more complete marketing profile.

Thirdly and finally, we asked participants questions about their perceptions of VR more broadly, in a semi-structured interview format. These included questions about whether they perceived VR social media as a place where they would feel comfortable acting freely and being themselves, and about what they would want from a tool or setting like private browsing mode for VR (if anything).

At the beginning of the study, participants also filled out a questionnaire on demographic information and their experience with VR and social media; at the end, they filled out a questionnaire containing 11 true-or-false questions about online tracking and inference in the U.S., selected from a survey run by Turow et al. [16]. Participants volunteered to take part but were assured of baked goods in the future.

Analysis. We recorded audio of the second and third portions of our user studies. When we conduct our full study, we will transcribe and qualitatively code these interview portions, but for the purposes of reporting on our pilot studies, Wentao listened to the recordings and pulled out segments corresponding to our research questions.

Ethics. This study was reviewed and approved by the University of Maryland Institutional Review Board. To protect participant privacy, we have deleted the raw data collected by our VR application, as well as the inferences made from them. We told participants that they could take a break or withdraw from our study at any point for any reason, including concern about privacy and nausea from being in VR. We have taken care to not release identifying data about participants.

4 Findings

We conducted two pilot studies, which we present here as case studies. Both participants were recruited from Wentao's contacts and have highly technical backgrounds, which is a significant and acknowledged limitation to even these very preliminary findings. P1 took approximately 40 minutes, had significant experience with VR, and was highly knowledgeable about tracking in the U.S. (10/11 questions answered correctly, 1/11 answered unsure). P2 took approximately 30 minutes, had minimal experience with VR, and was moderately knowledgeable (above average, based on the original survey results [16]) about tracking in the U.S. (6/11 questions answered correctly, 2/11 unsure, 3/11 incorrectly).

Our VR application worked reasonably well for P1, who spent a while walking around and looking at many objects, including the tracked ones, although a reaction time was not registered due to P1 not turning quickly to look for the source of the noise. In contrast, our VR application did not work very well for P2, who played the cat minigame very efficiently without going close to the cat or looking around more than necessary, resulting in minimal and noisy data; we did, however, successfully register a reaction time for P2.

In the rest of this section, we highlight some findings relevant to our research questions.

4.1 RQ1: perceptions of tracking and inference

Neither participant was surprised that a company could or would do the kinds of tracking and inference we presented. Interestingly, although P1 knew companies can "draw shockingly similar data" from how people use websites, they said this "does feel more invasive . . . maybe because looking is a more unconscious activity than purposely moving around a webpage." Similarly, they later said, "It was really interesting to just see that data collected and charted in front of me. That definitely felt different from just conceptually knowing that this is a thing that could be done."

P1 perceived the accuracy of inferences differently based on the situation. In one instance, they expressed measured skepticism, saying, "I don't feel like it makes much sense to draw much information from [what I was absent-mindedly gazing at on the TV], but hey, maybe psychologists would disagree"; in another, they acknowledged that the inference was fair, saying, "I remember I did distinctly pause to read those, and I kind of did glaze over the [other cookbook] . . . But I did make a point to read the title of the spicy food one." P2 perceived most inferences as inaccurate, as expected given the minimal amount of data we were able to collect.

When asked about the acceptability of tracking and inference in different contexts, P1 said they consider it "quite invasive" for a company to use or sell this kind of data. P2, however, said using data is fine if it is disclosed to users in a privacy policy and the data is anonymized, but it is not acceptable if the data is non anonymized or if it is sold without the user making some profit.

4.2 RQ2: effect on attitudes towards VR

Both of our participants said the types of tracking and inference they observed in the study were not new or surprising to them; therefore, they expressed little change in attitude towards VR. Though this was not a newly developed opinion, P1 offered, "I find that when this kind of information is being tracked about me, I find it very off-putting and it makes me kind of hyper-conscious about how my behavior's being
interpreted."

Both participants said they use privacy-enhancing tools outside of VR such as private browsing mode, and they both expressed interest in such tools for VR. P2 suggested that they would like a tool to either block tracking or to perturb it by giving the tracking algorithm wrong information. P1 expressed that it is important that such a tool doesn't interfere with the "cool" aspects of VR that allow users to see each other's body language and make eye contact.

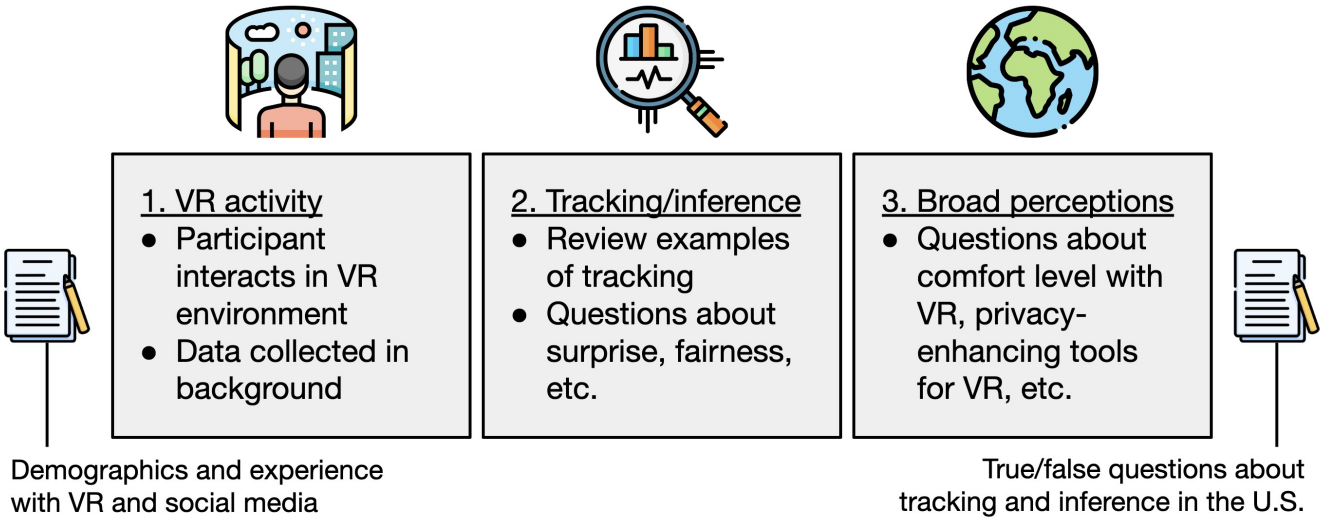

Figure 1: A summary of our user study design.

5 Discussion

Based on promising results from our pilot studies, we aim to continue this project after making some changes to the VR application and to the user study protocol.

5.1 Changes to VR application

Gaze tracking is perhaps the most interesting "new" form of tracking in VR to us; our naïve implementation worked somewhat convincingly in the first pilot study, but we would like to use a headset with eye tracking to get more accurate data.

As P2 demonstrated, participants can easily avoid tracking while in the VR environment, making it difficult to reflect on plausible inferences in the next portion of the user study. We can mitigate this by adding more tracking (e.g., more tracked objects, as well as more types of tracking) to the environment. We may also want to redesign the activity; rather than have participants look for a cat, for instance, we could have them talk to pre-programmed avatars that stand in for other users in an actual social VR world. This particular solution would have the double benefit of mimicking actual social VR more closely and potentially giving participants' eyes time to roam freely (incurring tracking) as they listen to these avatars speak.

Reaction time did not produce convincing inferences; P1 did not even turn immediately to look at the cat when the sound was played, and P2 expressed skepticism about whether it was fair to compare different users' reaction times. We may try to make a different health-related inference, such as height or posture.

Finally, both participants did experience motion sickness due in part to the joystick-based controls. Different locomotion methods might help reduce discomfort.

5.2 Changes to user study protocol

While conducting pilot studies, we observed that some questions need refining; for example, it is unclear what exactly the word *fair* means in "How fair do you think this inference is, based on your activity?"

The true-or-false questions that we ask at the end were designed for a survey of Americans' knowledge specifically, and some contain U.S.-specific statements, such as "The Health Insurance Portability and Accountability Act (HIPAA) prevents apps that provide information about health from selling data collected about app users to marketers." This makes it a poor assessment of knowledge for non-American participants; P2, who is not from the U.S., noted this as a reason for not knowing the answer to some questions. We will need to select questions that are less biased by nationality.

6 Conclusion

For our class project, we designed a VR application and user study to elicit people's perceptions of tracking and inference in VR. We ran two pilot studies and have identified changes to make before moving forward with a full user study.

References

[1] Devon Adams, Alseny Bah, Catherine Barwulor, Nureli Musaby, Kadeem Pitkin, and Elissa M. Redmiles. Ethics emerging: The story of privacy and security perceptions in virtual reality. In *Proceedings of the Fourteenth Symposium on Usable Privacy and Security*, pages 443– 458, August 2018. https://www.usenix.org/conference/soups2018/presentation/adams.

[2] Avi Bar-Zeev. The eyes are the prize: Eyetracking technology is advertising's holy grail. https://www.vice.com/en/article/bj9ygv/the-eyes-arethe-prize-eye-tracking-technology-is-advertisingsholy-grail, May 2019.

[3] Marc Hetherington and Jonathan Weiler. *Prius or Pickup? How the Answers to Four Simple Questions Explain America's Great Divide*. 2018. https://www.google.com/books/edition/Prius_Or_Pickup/Z6lvDwAAQBAJ.

[4] Tatum Hunter. Surveillance will follow us into 'the metaverse,' and our bodies could be its new data source. *Washington Post*, January 2022. https://www.washingtonpost.com/technology/2022/01/ 13/privacy-vr-metaverse/.

[5] Khari Johnson. Meta's new Quest Pro VR headset harvests personal data right off your face. *Wired*, October 2022. https://www.wired.com/story/metas-vr-headsetquest-pro-personal-data-face/.

[6] Girard Kelly, Jeff Graham, Jill Bronfman, and Steve Garton. Privacy of virtual reality: Our future in the metaverse and beyond. Technical report, Common Sense Media, 2022. https://www.commonsensemedia.org/research/privacyof-virtual-reality-our-future-in-the-metaverse-andbeyond.

[7] Jinghuai Lin and Marc Erich Latoschik. Digital body, identity and privacy in social virtual reality: A systematic review. *Frontiers in Virtual Reality*, 3, November 2022. https://doi.org/10.3389/frvir.2022.974652.

[8] 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. *Scientific Reports*, 10(1), 2020. https://doi.org/10.1038/s41598-020-74486-y.

[9] Alec G. Moore, Ryan P. McMahan, Hailiang Dong, and Nicholas Ruozzi. Personal identifiability and obfuscation of user tracking data from VR training sessions. In *Proceedings of the 2021 IEEE International Symposium on Mixed and Augmented Reality*, pages 221–228, October 2021. https://doi.org/10.1109/ISMAR52148.2021. 00037.

[10] Vivian Genaro Motti and Kelly Caine. Users' privacy concerns about wearables: Impact of form factor, sensors and type of data collected. In *Financial Cryptography and Data Security*, Lecture Notes in Computer Science, pages 231–244, 2015. https://doi.org/10.1007/9783-662-48051-9_17.

[11] Liu, Fuxiao, Kevin Lin, Linjie Li, Jianfeng Wang, Yaser Yacoob, and Lijuan Wang. "Mitigating hallucination in large multi-modal models via robust instruction tuning." *arXiv preprint arXiv:2306.14565* 1, no. 2 (2023): 9.

[12] Liu, Fuxiao, Yinghan Wang, Tianlu Wang, and Vicente Ordonez. "Visual news: Benchmark and challenges in news image captioning." *arXiv preprint arXiv:2010.03743* (2020).

[13] Liu, Fuxiao, Tianrui Guan, Zongxia Li, Lichang Chen, Yaser Yacoob, Dinesh Manocha, and Tianyi Zhou. "Hallusionbench: You see what you think? or you think what you see? an image-context reasoning benchmark challenging for gpt-4v (ision), llava-1.5, and other multi-modality models." *arXiv preprint arXiv:2310.14566* (2023).

[14] Liu, Fuxiao, Yaser Yacoob, and Abhinav Shrivastava. "COVID-VTS: Fact Extraction and Verification on Short Video Platforms." *arXiv preprint arXiv:2302.07919* (2023).

[15] Liu, Fuxiao, Hao Tan, and Chris Tensmeyer. "DocumentCLIP: Linking Figures and Main Body Text in Reflowed Documents." *arXiv preprint arXiv:2306.06306* (2023).

[16] Li, Zongxia, Paiheng Xu, Fuxiao Liu, and Hyemi Song. "Towards understanding in-context learning with contrastive demonstrations and saliency maps." *arXiv preprint arXiv:2307.05052* (2023).

[17] Liu, Fuxiao, Xiaoyang Wang, Wenlin Yao, Jianshu Chen, Kaiqiang Song, Sangwoo Cho, Yaser Yacoob, and Dong Yu. "MMC: Advancing Multimodal Chart Understanding with Large-scale Instruction Tuning." *arXiv preprint arXiv:2311.10774* (2023).

[18] 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. *Science and Engineering Ethics*, 22(1):1–29, 2016. https://doi.org/10.1007/ s11948-014-9621-1.

[19] 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, pages 1–12, May 2019. https://doi. org/10.1145/3290605.3300340.

[20] Rory Mir and Katitza Rodriguez. If privacy dies in VR, it dies in real life. https://www.eff.org/deeplinks/2020/ 08/if-privacy-dies-vr-it-dies-real-life, August 2020.