[DC] VirSec: Virtual Reality as Cost-Effective Test Bed for Usability and Security Evaluations

Florian Mathis*
University of Glasgow & University of Edinburgh

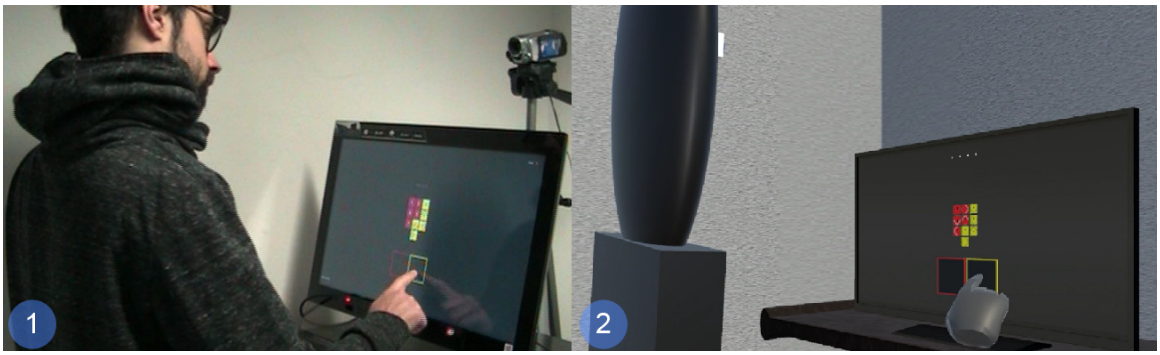

Figure 1: This doctoral thesis evaluates the suitability of using virtual reality for human-centred usability and security evaluations. The picture shows a real-world study of a recently introduced authentication scheme called CueAuth [5] (❶) and our replication study [7] in where we evaluated the system's usability and observation resistance through two VR studies (❷).

Abstract

In this doctoral thesis, we explore how virtual reality (VR) can better support the development and evaluation of prototype systems, with a focus on usable privacy and security (USEC) as a sub-domain of Human-computer Interaction (HCI) research. We argue that VR, as a study platform, can augment existing research paradigms in usable security research by 1) enabling researchers to study systems and user behaviour in contexts that are otherwise challenging due to ethical or legal constraints and 2) improving evaluations that are constrained to conditions that can be physically replicated in the lab.

Keywords: Virtual Reality, Usable Security, HCI.

Index Terms: Human-centered computing—Virtual reality; Security and privacy—Usability in security and privacy;

1 Introduction & Background

Evaluating the usability and security, e.g., when looking over the user's shoulder, of usable security systems can be challenging. Building physical prototypes can be expensive, lab-based studies of corresponding evaluations are often not capable of representing real-world conditions, and conducting studies of that type with considerable large sample sizes can be time-consuming. Although online studies (e.g., surveys) are well-respected and a fundamental part of usable security research, they are often not capable of providing comprehensive usability and security evaluations of physical prototypes. Finding other means to evaluate security systems is vital to be capable of handling the fast pace of emerging technologies that come with novel challenges [1]. Through this doctoral thesis, we first identify the unique challenges faced by researchers who prototype and evaluate usable security systems. We then explore the suitability of VR as a test bed for human-centred evaluations. Leveraging VR as a test bed is deemed promising beyond USEC research; for example, recent suggestions to run user studies that involve virtual and augmented reality amidst COVID-19 include collaborations across labs to provide participants for each other's experiments, and building an infrastructure that provides equipment to a pool of participants [8]. That being said, the idea of leveraging virtual reality as a test bed is not new in human-centred research. George et al. [3], for example, stressed that VR could serve as a feasible research tool to simulate and evaluate complex security systems. Mäkelä et al. [6] argued that digital deployments are cheaper and faster to build than real deployments. Virtual environments can be used for several studies without the need to, for example, clear physical rooms or maintain real-world setups. Mäkelä et al. [6] also emphasised that VR can enable researchers to conduct studies even when physical spaces are under construction or not accessible (e.g., due to COVID-19). Weiß et al. [9] further argued that the use of surrogate empirical methods, such as VR studies, can be particularly promising when evaluating systems within safety-critical or expensive scenarios.

*e-mail: florian.mathis@glasgow.ac.uk

2 Research Arc

Our research arc comprises three stages: **(1)** Identifying unique challenges faced by usable security researchers when developing and evaluating prototype systems; **(2)** Validating the use of virtual reality for usability and security evaluations; **(3)** Exploring unique characteristics of virtual reality to augment the development and evaluation of usable security systems. Our research is roughly guided by following three research questions that are segmented into smaller research objectives over the course of time.

R_{Q1}: *What are the challenges researchers face when designing, implementing, and evaluating human-centred security prototypes?*

R_{Q2}: *What characteristics of security systems fully evaluated in VR match findings from a corresponding real-world evaluation?*

R_{Q3}: *What are the characteristics of VR that can augment current research paradigms to evaluate usable security prototypes?*

2.1 Challenges faced by USEC Researchers

We conducted a series of in-depth interviews with established and nascent usable privacy and security researchers, ranging from full professors to senior PhD candidates who have published work in top-tier HCI and usable security venues. Through twelve in-depth interviews, we identified unique challenges faced by USEC researchers

such as **the high costs of conducting usability and security evaluations of hardware prototypes** and the **difficulties of conducting security field studies** (R_{Q1}). Experts voiced that novel technologies, such as 3D printing and VR, could augment existing research paradigms by replacing study-specific hardware with 3D printed items and/or digital artefacts, and could enable researchers to create *"[virtual] replications of everyday life scenarios"* (P8).

To scope our research appropriately to a PhD thesis, we focused in the next step specifically on the above-mentioned challenges and investigate **(1)** the transferability of findings collected in VR to the real world and **(2)** how VR can augment existing human-centred research paradigms such as lab and field studies.

2.2 Validating the Use of Virtual Reality Studies

While our interviews unveiled many existing bottlenecks in the development and evaluation of usable security systems, it remains unclear whether findings from VR-based usability and security studies match those previously reported in real-world studies. We therefore conducted two user studies to answer R_{Q2} (see also [7]). We first conducted an in-VR usability study where we replicated an authentication scheme that comes with three input methods (touch, mid-air gestures, eye gaze) fully in VR and validated our findings through a comparison between our results and those previously reported in the real-world evaluation [5]. We then conducted a follow-up remote security study where we showed participants VR recordings of authentications and compared their observation performance to the observation performance on real-world authentications.

Our analysis indicates that VR can serve as a suitable test bed for usability and security evaluations in many aspects, but the used technology (e.g., hand tracking) can have a significant impact on the evaluation. While we found many similarities between our VR security study and the equivalent real-world security study (e.g., similar attack rate), authentications using touch input were significantly faster in the real world study than in our in-VR usability study, mainly due to the virtual artefact (the virtual hand) that is required when providing touch input in VR. Gaze-based input was significantly faster in our VR study compared to the real-world study, indicating that VR studies could help mitigate limitations of hardware used in the real world, but could also be misleading. The differences between our VR and the real-world studies suggest that transferability of quantitative results from VR to the real world highly depends on how well reality and its limitations are emulated. Findings from an additional user study suggest that an abstract avatar (Fig. 1,❷) is capable of providing the same information to bystanders when providing input with touch, mid-air, and eye gaze compared to a more realistic avatar by Microsoft Research [4]. This finding is particularly interesting as an abstract avatar can reduce the required effort and expertise to use VR as an evaluation platform; thus, can make VR studies more accessible to the broader HCI community.

2.3 Exploring the Strengths of VR as a Test Bed

Building upon our recent works that suggest VR security studies through virtual avatars match results collected in real-world security studies, we further explore to what extent we can use VR to conduct more realistic security studies. This is motivated by existing research that evaluated a system's security through 2D recordings of authentications (e.g., [2]). While assessing the observation resistance of a system through 2D recordings is a well-respected approach in usable privacy and security, previous research by Aviv et al. [2] showed that 2D recordings of authentications cannot replace evaluation in live settings; yet, studying live observations is often costly and time-consuming, and therefore often results in security evaluations based on 2D recordings. Expanding upon 2D recordings than can be used as a baseline measurement when assessing a system's observation resistance, we look into the additional value of 3D video recordings in VR. This allows us to explore the impact of the third-dimension provided by an immersive virtual environment on security evaluations and contributes to answering R_{Q3}. Furthermore, we plan to study user authentications in contexts that are otherwise challenging to study due to legal and ethical concerns. For example, one direction could be to study users' authentication behaviour at ATMs by replicating an ATM authentication scenario fully in VR.

3 Conclusion & Contribution

Results from our works show that many findings from VR usability and security studies match corresponding real-world evaluations. We found great similarities between our online security study through virtual avatars and the original real-world security study (see also [7]). Through our research, we hope to augment current study paradigms and complement lab and field studies by enabling large-scale and cost-efficient evaluations through VR.

To further shape our research arc and explore the strengths and weaknesses of VR for usable privacy and security research, we would like to discuss following three questions during the doctoral consortium: **(1)** Does our research arc make sense and flow logically? **(2)** What additional steps are required to establish VR studies as a well-respected research paradigm in (and beyond) usable privacy and security research? **(3)** How can we scientifically measure the efficiency in terms of resources, time, and effort when conducting VR studies (vs real-world studies)?

Acknowledgments

The author thanks his supervisors Dr. Kami Vaniea and Dr. Mohamed Khamis. This work is supported by the University of Edinburgh and the University of Glasgow jointly funded PhD studentships.

References

[1] F. Alt and E. von Zezschwitz. Emerging trends in usable security and privacy. *Journal of Interactive Media*, 18(3):189–195, 2019. doi: 10.1515/icom-2019-0019

[2] A. J. Aviv, F. Wolf, and R. Kuber. Comparing video based shoulder surfing with live simulation. In *Proceedings of the 34th Annual Computer Security Applications Conference*, ACSAC '18, p. 453–466. ACM, New York, NY, USA, 2018. doi: 10.1145/3274694.3274702

[3] C. George, M. Khamis, D. Buschek, and H. Hussmann. Investigating the third dimension for authentication in immersive virtual reality and in the real world. In *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 277–285, 2019. doi: 10.1109/VR.2019.8797862

[4] M. Gonzalez-Franco, E. Ofek, Y. Pan, A. Antley, A. Steed, B. Spanlang, A. Maselli, D. Banakou, N. Pelechano, S. Orts-Escolano, V. Orvalho, L. Trutoiu, M. Wojcik, M. V. Sanchez-Vives, J. Bailenson, M. Slater, and J. Lanier. The rocketbox library and the utility of freely available rigged avatars. *Frontiers in VR*, 2020. doi: 10.3389/frvir.2020.561558

[5] M. Khamis, L. Trotter, V. Mäkelä, E. v. Zezschwitz, J. Le, A. Bulling, and F. Alt. Cueauth: Comparing touch, mid-air gestures, and gaze for cue-based authentication on situated displays. *Proc. of ACM IMWUT*, 2(4), Dec. 2018. doi: 10.1145/3287052

[6] V. Mäkelä, S. R. R. Rivu, S. Alsherif, M. Khamis, C. Xiao, L. M. Borchert, A. Schmidt, and F. Alt. Virtual Field Studies: Conducting Studies on Public Displays in Virtual Reality. In *Proceedings of the 38th Annual ACM Conference on Human Factors in Computing Systems*, CHI '20. ACM, New York, NY, USA, 2020. doi: 10.1145/3313831.3376796

[7] F. Mathis, K. Vaniea, and M. Khamis. Replicueauth: Validating the use of a lab-based virtual reality setup for evaluating authentication systems. In *Proceedings of the 39th Annual ACM Conference on Human Factors in Computing Systems*, CHI '21. ACM, New York, NY, USA, 2021. doi: 10.1145/3411764.3445478

[8] A. Steed, F. Ortega, A. Williams, E. Kruijff, W. Stuerzlinger, A. Batmaz, A. Won, E. Rosenberg, A. Simeone, and A. Hayes. Evaluating immersive experiences during covid-19 and beyond. *Interactions*, 2020.

[9] M. Weiß, K. Angerbauer, A. Voit, M. Schwarzl, M. Sedlmair, and S. Mayer. Revisited: Comparison of empirical methods to evaluate visualizations supporting crafting and assembly purposes. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2020. doi: 10.1109/TVCG.2020.3030400