# Understanding Shoulder Surfer Behavior Using Virtual Reality

Yasmeen Abdrabou[1] Radiah Rivu[1] Tarek Ammar[2] Jonathan Liebers[3] Alia Saad[3]
Carina Liebers[3] Uwe Gruenefeld[3] Pascal Knierim[1] Mohamed Khamis[4] Ville Mäkelä[5]
Stefan Schneegass[3] Florian Alt[1]

[1]Bundeswehr University Munich, Germany, {yasmeen.essam, sheikh.rivu, pascal.knierim, florian.alt}@unibw.de
[2]LMU Munich, Germany, tarek.ammar@campus.lmu.de [3]University of Duisburg-Essen, Germany, {firstname.lastname}@uni-due.de
[4]University of Glasgow, United Kingdom, mohamed.khamis@glasgow.ac.uk [5]University of Waterloo, Canada, ville.makela@uwaterloo.ca

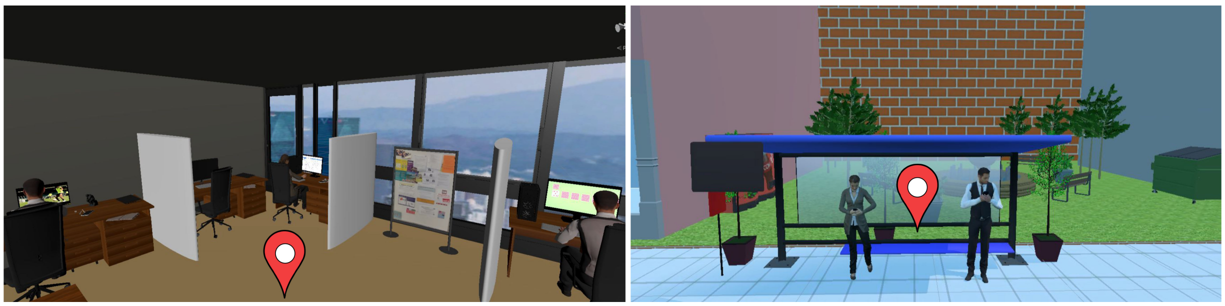

Figure 1: We explore shoulder surfers' behavior using a virtual reality environment. In particular, we investigate factors influencing shoulder surfers' behavior in two different situations (a) open office scene, and (b) bus stop. Initial user position is market with a pin.

## Abstract

We explore how attackers behave during shoulder surfing. Unfortunately, such behavior is challenging to study as it is often opportunistic and can occur wherever potential attackers can observe other people's private screens. Therefore, we investigate shoulder surfing using virtual reality (VR). We recruited 24 participants and observed their behavior in two virtual waiting scenarios: at a bus stop and in an open office space. In both scenarios, avatars interacted with private screens displaying different content, thus providing opportunities for shoulder surfing. From the results, we derive an understanding of factors influencing shoulder surfing behavior.

**Index Terms:** H.5.2 [Security and privacy]; H.5.m [Human-centered computing]: Human computer interaction (HCI)

## 1 Introduction and Related Work

Investigating shoulder surfers' behavior in the real world is challenging because it often happens in complex situations. In addition, conducting such in-the-wild studies is highly challenging from an ethical point-of-view [5]. Researchers tried to address this in several ways: Marques et al. collected stories from people to learn how they feel about unauthorized access to their smartphones [2]. Eiband et al. *asked people* about their shoulder surfing behavior in an online survey [1]. The study revealed attackers' motivations and provided an understanding of contexts in which shoulder surfing occurs – yet it could not capture the shoulder surfers' actual behavior. Saad et al. exposed participants to *360-degree videos* [3] of photo-realistic, pre-recorded public transport situations, including passengers whose smartphones could be shoulder surfed. However, the study was limited as the videos were recorded from a static viewpoint, hence not allowing participants to move around and position themselves freely. The results show that Virtual Reality (VR) in combination with eye tracking acts as a next step towards obtaining a more nuanced understanding of shoulder surfers' behavior. This is also supported by recent research, demonstrating that VR can serve as a means to observe behavior as it occurs in real-world settings.

We thus contribute an investigation of shoulder surfing behavior by means of a VR environment. We implemented two waiting scenarios, motivated by a previously conducted survey that identified waiting as one of the primary contexts in which shoulder surfing occurs [1]. The two scenarios simulate a bus stop and an office environment, in which a diverse set of avatars are interacting with different devices. The bus stop serves to investigate shoulder surfing behavior towards smartphones, and the office scene allows shoulder surfing towards larger screens (desktops) to be investigated.

## 2 Investigating Shoulder Surfing Behavior

We conducted a within-subjects design lab study to answer our research question: *What is the influence of (a) gender, (b) screen content, and (c) distance between attacker and victim on shoulder surfers' behavior?* Each participant experienced two environments (bus stop, open office space), counterbalanced using a Latin-square.

The study involved the following *independent variables*: environment (bus stop vs. office), avatar gender (male vs. female), and screen content (news article/video vs. chat/email vs. game). The avatar screen content was counterbalanced for every participant and scene. For the bus stop scene, as there were two screens but three types of content, we ensured that each content was shown equally often (again using a Latin square). As *dependent variables* we collected 1) participants' gaze direction (x, y, z), 2) participants' position in the virtual environment, and 3) Euclidean distance between the participant and the screen. Participants were initially placed 3 meters away from all avatars in the *open office scene* and 1 meter away from all the avatars in the *bus scene*.

### 2.1 Apparatus and Participants

We used the HTC Vive Pro Eye headset with a resolution of 1440 $\times$ 1600 pixels per eye and at a frame rate of 90 Hz. The headset was attached to an Alienware Core i7 and 16 GB RAM and was developed using Unity. We recruited 24 participants (9 female) through mailing lists and social networks. Participants' average age was 27.25 years

(SD = 6.66). Participants had diverse backgrounds and they rated their VR experience as average (M=2.4) on a 5-Point Likert item.

### 2.2 Study Procedure

When participants came to our lab, we briefed them that the objective of the study was to better understand how they perceive day-to-day situations in virtual reality. Then we asked the participants to fill in the consent and demographics questionnaire. We calibrated the VR headset using the integrated Tobii routine. Participants saw the scenes randomly balanced on a Latin Square. The task was to wait in both scenes, for the bus in the bus stop and for the boss in the office scene. Each scene lasted two minutes. Then participants were asked to fill in a Presence questionnaire [6]. During the debriefing, we explained the true purpose of the study. Each session lasted for around 30 minutes and participants were compensated with 5 Euros.

## 3 Results and Discussion

We analyzed participants' behavior in both settings. More specifically, we investigated gaze behavior, participants' movement, and influencing factors such as avatar gender, screen content, and users' distance to the screen. Our data was not normally distributed (confirmed by Shapiro-Wilk and Anderson-Darling tests); thus, we performed non-parametric tests and report the mean values (M).

### 3.1 Gaze Behavior

To explore how often and for how long participants looked at the screens in the different scenarios, we identified fixations using the unmodified Dispersion-Threshold Identification algorithm [4].

**Number of Eye Contacts with Screens.** Each participant looked at least once on an avatar's screen. On average, participants fixated 5.7 times on the screens in each the scenarios. For the *open office scene*, participants looked on average 8.04 times in the two minute time frame (2–25 fixations). In the *bus stop scene*, we found that participants looked 3.36 times on average (2–22 fixations).

**Duration of Eye Contact with Screens.** We also analyzed the duration of eye contact towards screens. Participants looked on average for 1.61 s at the screens. For the *open office scene*, we found that participants looked on average for 2.1 s on the screens (between 1.16 s and 7.53 s). For the *bus scene*, participants looked on average for 1.12 s (between 1.01 s and 2.96 s). These findings show that shoulder surfing happened in our scenarios as participants looked into the avatar screens and for periods of 1.6 s.

### 3.2 Positioning and Distance

We looked at which average distance people positioned themselves from the screens. Therefore, we considered cases in which people moved away from their initial position and calculated the distance to the screen for each time, people stopped. People on average positioned themselves at 17.6 cm from the screen in the *bus stop scene* and at 150.79 cm in the *open office scene*. This means that participants positioned themselves almost halfway through the distance between their initial position and the avatar's screen in the office scene. However, when the screen became smaller, participants positioned themselves closer (two-thirds of the distance) to the screen.

Furthermore, we analyzed the influence of distance between attackers and screens on the duration of the attack. We used the average distance as a threshold of "near" and "far". With regard to the *duration per attack*, we found that if participants stood closer to the display, the attack duration was longer (18 s vs. 16 s open office scene and 16 s vs 15 s bus scene near and far respectively). A Wilcoxon test confirmed a statistically significant effect of distance on attack duration – both in the open office scenario ($Z = -2.16, p < .01$) as well as in the bus stop scenario ($Z = -3.23, p < .01$).

The findings above should be interpreted with care. Our data shows that in some cases, participants moved rather close towards the screens. While this might also be the case in the real world (for example, in a crowded subway), another reason in our case (where the space was not particularly crowded) could have been that participants moved closer to better observe text-based content. Hence, while our findings hint at distance indeed having an influence on attack duration, this should be confirmed by future work.

### 3.3 Gender

We looked if gender had an influence on the frequency and duration of the attack. Here, we only considered the bus stop scene as, here, gender was equally distributed. Regarding *number of attacks*, we found that males (2.37 times) and females (2.43) were attacked almost equally often. On the other side, for the *duration per attack*, we found that attacks on the screens of female's were about 30% longer (18 s) as opposed to males (13 s). Differences in both cases were not statistically significant, $p > .05$, according to a Wilcoxon test. Prior work suggest that females were more often shoulder surfed than males [1]. We can neither confirm nor refute this finding. Previous work did not report on the duration of shoulder surfing for different genders. Here, our numbers suggests that females might be shoulder surfed longer – yet this should be further investigated.

### 3.4 Content

Finally, we then compared three different screen contents regarding how they influence shoulder surfing. For this analysis we focus on the open office space environment, as here all three different types of content were visible at the same time. First, we looked, again, at the *duration per attack*. We found that attacks on games lasted longest (on average 6.6 s), followed by videos (6 s) and typing (5.6 s). Second, we looked at the *number of attacks*. Where we found that typing was attacked most frequently (3.7 times on average), followed by games (3.14) and video (2.4). Differences in both cases were not statistically significant, $p > .05$, according to a Wilcoxon test. From this we learn that text – as it is being typed – seems to be attacked more often but for shorter time spans. The reason might be that attackers repeatedly check back on new text (as reading text is usually faster than typing it). In contrast, video and games seem to be attacked less often, but for longer time windows.

## 4 Conclusion

In this paper, we investigated shoulder surfing using virtual reality. We implemented two virtual waiting scenarios and analyzed participants' gaze and movement behavior. We assessed the influence of different aspects, namely avatar gender, screen content, and distance. We report insights on shoulder surfer's behavior that could serve as a starting point for further investigations, both on attacker and victim behavior, and for designing novel mitigation concepts.

## References


[1] M. Eiband, M. Khamis, E. Von Zezschwitz, H. Hussmann, and F. Alt. Understanding shoulder surfing in the wild: Stories from users and observers. In *Proc. CHI'17*, pp. 4254–4265. ACM, 2017.

[2] D. Marques, T. Guerreiro, L. Carriço, I. Beschastnikh, and K. Beznosov. Vulnerability & blame: Making sense of unauthorized access to smartphones. In *Proc. CHI'19*. ACM, 2019.

[3] A. Saad, J. Liebers, U. Gruenefeld, F. Alt, and S. Schneegass. Understanding Bystanders' Tendency to Shoulder Surf Smartphones Using 360-degree Videos in VR. In *Proc. MobileHCI'21*. ACM, 2021.

[4] D. D. Salvucci and J. H. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Proc. ETRA'00*, p. 71–78. ACM, 2000.

[5] O. Wiese and V. Roth. Pitfalls of shoulder surfing studies. In *NDSS Workshop on Usable Security*, pp. 1–6, 2015.

[6] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoper. Virtual Environ.*, 7(3):225–240, 1998.