Investigating Natural Shoulder Surfing Behavior in the Wild: A Research Space and Case Study

Yasmeen Abdrabou[1,7], Rivu Radiah[2], Alessia Fischer[3], Alia Saad[4], Habiba Farzand[5], Pascal Knierim[6,7], and Florian Alt[3,7]

[1] Technical University of Munich, Germany
[2] IU Hochschule, Germany
[3] LMU Munich, Germany
[4] University of Duisburg-Essen, Germany
[5] University of Bristol, UK
[6] University of Innsbruck, Austria
[7] University of the Bundeswehr Munich, Germany
yasmeen.abdrabou@tum.de

Abstract. Studying shoulder surfing in natural settings presents substantial methodological challenges, including ethical considerations and preventing mutual influence between observers and those observed. This paper contributes to understanding opportunistic shoulder surfing in public environments, particularly public transportation, a vital space for diverse user groups. We first derive a research space on shoulder surfing to identify unresolved methodological issues. Then, we introduce our methodology for studying shoulder surfing in the wild using eye tracking. In our case study, participant observers wore mobile eye trackers during public transport journeys, allowing us to capture natural, often opportunistic, shoulder surfing behavior. From this, we derived valuable lessons learned and recommendations for future research. Our efforts deepen the understanding of shoulder surfing in real-world settings and pave the way for more effective mitigation strategies.

Keywords: Shoulder Surfing · Eye Tracking · Privacy · Case Study.

1 Introduction

Shoulder surfing refers to observing the display of others' devices without their consent (e.g., smartphones) [34]. This human-centered attack has been at the focus of research in the usable security community for many years [28,83]. This act poses a security risk by potentially disclosing private information such as credentials [54]. Researchers have sought to understand who is affected by shoulder surfing [4], the contexts in which it occurs [28], the content being observed [5,28], the influence of personal relationships [31], and methods to mitigate it [20,31,55].

Prior research has employed methods such as interviews [61], self-reports [28, 53, 54], diaries [33], and user behavior studies using static 360° videos [71] or virtual reality [5]. While these approaches offer controlled environments, they fail to capture the spontaneous, socially embedded nature of shoulder surfing in real-world settings. In everyday life, this behavior is shaped by situational

factors such as spatial constraints, crowding, lighting, and subtle interpersonal dynamics—all difficult to replicate or observe accurately in the lab. Real-world observation is thus essential to understanding when, where, and how shoulder surfing occurs, its duration, and how it may go unnoticed. This knowledge is crucial for developing adaptive interfaces and real-time mitigations that reflect actual usage contexts rather than idealized conditions.

Real-world observation of shoulder surfing remains limited due to challenges such as its opportunistic nature, potential observer influence, and ethical constraints. Yet understanding factors like duration, viewing angles, and frequency is crucial for designing interfaces that raise awareness and help mitigate attacks. This methodological gap has left key questions unanswered, specifically regarding the prevalence, triggers, and characteristics of shoulder surfing in natural settings. To address these gaps, we investigate the following research questions:

RQ$_1$ How has prior work studied and countered shoulder surfing?

RQ$_2$ How can we effectively study shoulder surfing in a real-world scenario?

We begin by reviewing prior work to highlight unresolved methodological challenges. Next, we present a real-world case study using mobile eye tracking to capture natural shoulder surfing behavior in public transport. This approach enabled the collection of first-person gaze data in dynamic, real-world settings where shoulder surfing naturally occurs. We describe the study design, setting, data collection, and ethical considerations, offering practical insights for future research. While some biases remain, our methodology yields valuable lessons and recommendations for studying shoulder surfing in situ.

Contribution Statement: We present a *comprehensive research space* contextualizing shoulder surfing attacks based on existing literature, focusing on research questions and approaches. We conduct the first real-world *case study* (N=15) investigating shoulder surfing from the observer/attacker perspective. Finally, we propose a set of *best practices* and learned lessons for future research.

2 Shoulder Surfing as a Threat Model

Shoulder surfing refers to the unauthorized observation of another user's actions, potentially exposing sensitive information such as passwords, PINs, or personal messages [28]. This behavior poses risks like identity theft and financial fraud [33]. However, not all incidents are malicious—many stem from boredom or curiosity [18,28,69]. Abdrabou et al. [5] conceptualize shoulder surfing in three phases: an *idle* phase (no active observation), an *approach* phase (preparation), and an *attack* phase (direct observation). The following section reviews existing literature, focusing on the key aspects explored to date.

3 Research Space for Shoulder Surfing

To address RQ_1, we synthesized prior research on shoulder surfing, focusing on relevant studies. While many papers propose countermeasures, we included a representative sample to provide an overview rather than a detailed analysis. Our research space comprises two parts: an overview of research questions and a summary of research approaches (see Figure 1).

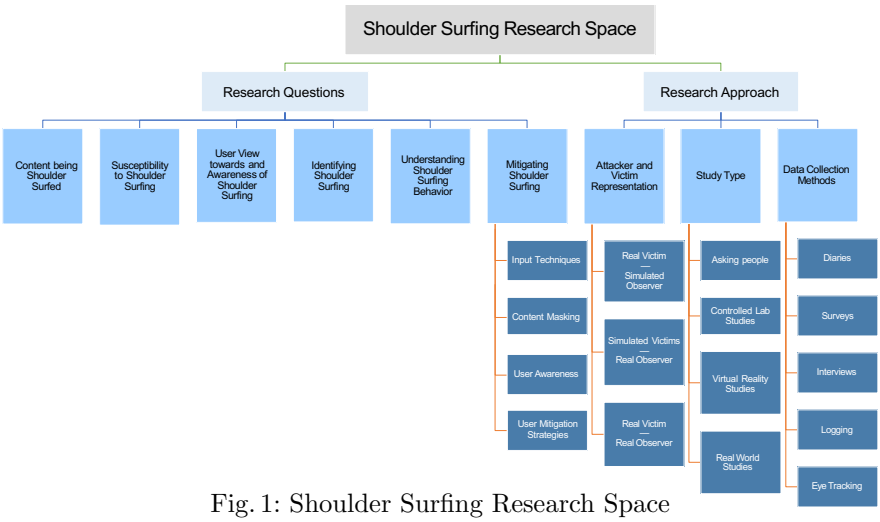

Fig. 1: Shoulder Surfing Research Space

3.1 Research Objectives

Prior work on shoulder surfing has explored a broad range of topics [18, 19, 72]. Research objectives span empirical, technical, and behavioral dimensions. Empirical work seeks to understand what content is typically shoulder surfed, who is most susceptible, and how users perceive such incidents. Technical investigations focus on detecting and mitigating shoulder surfing. Behavioral studies aim to uncover how and why shoulder surfing occurs in practice.

Content being Shoulder Surfed Researchers investigated which content is primarily subject to shoulder surfing. Analyzing self-reporting user stories, Eiband et al. [28] found that text is the most shoulder surfed content (46.6%), followed by pictures (24.1%) and games (12.6%). Regarding text, shoulder surfers mainly focused on instant messaging, e.g., WhatsApp (41.8%), followed by social networks, e.g., Facebook (17.5%), email (7.9%), and news (7.4%). In a 360° video virtual reality (VR) study, Saad et al. [71] compared four content types used by different applications: (1) WhatsApp, (2) Photo Gallery, (3) Game, and (4) Facebook. Based on a study (N=16), they found that after observing the authentication process, WhatsApp was the most shoulder-surfed application, followed by the Photo Gallery. In a VR study (N=24), Abdrabou et al. [5] reported that dynamic, engaging content (games/videos) was more often shoulder-surfed than text (chatting/articles), attributing this to the fact that motion grabbed more attention. The authors reported that attacks on games lasted the longest, followed by videos and typing. However, typing was the most frequently observed (3.7 times on average), followed by games (3.14) and video (2.4).

Susceptibility to Shoulder Surfing Eiband et al. [28] examined the *relationship between shoulder surfers and shoulder-surfees*, reporting that most observers were strangers (126/170), followed by friends (11), acquaintances (10), colleagues (8), and family members (3). Muslukhov et al. [61] found that younger users ($age \leq 17$) are at higher risk of unauthorized access.

In terms of *gender*, males most frequently shoulder surfed females (44/151), followed by male–male (38), female–female (37), and female–male (32) pairings [28]. A VR study reported minor differences in observation frequency (females: 2.43, males: 2.37), but longer observations for females (18 s vs. 13 s) [5].

Timing patterns varied: Eiband et al. [28] found most incidents occurred in the morning (57/158). Farzand et al. [33] observed more cases at night (11/23).

Location also plays a role. Public transport emerged as the most common setting (141/189) [28], followed by work or study environments (16) and other places (32). Farzand et al. [33] and Schneegass et al. [74] likewise highlighted public transport, lecture halls, crowded areas, and cafés as typical contexts.

Observer *positioning* also affects visibility: Ali et al. [7] showed that proximity influences text readability and user alertness. Abdrabou et al. [5] found that in VR, observers approach more closely when viewing smaller screens (smartphones) and maintain greater distance for larger displays (desktops).

User Views & Awareness of Shoulder Surfing Most shoulder surfing incidents go unnoticed, with observers typically acting out of curiosity or boredom [28]. Bianchi et al. [13] describe such individuals as *Casual Observers*. Harbach et al. [37] reported in an online survey ($N = 260$) that 83% of users did not consider shoulder surfing to be a realistic threat. Schneegass et al. [74] found that users perceived only 2.5% of observed incidents (11 out of 437) as threatening. Farzand et al. [33] noted that 66% of diary study participants considered their current smartphone task more important than preventing shoulder surfing.

Identifying Shoulder Surfing To detect shoulder surfing, Saad et al. [70] proposed a computer vision algorithm that estimates attackers' gaze to confirm whether they are looking at a user's screen. Similarly, PrivacyScout [11] uses facial features—including gaze direction, head pose, and eye distance—to compute a shoulder surfing risk factor. Schneegass et al. [74] explored the frequency of potential incidents by equipping participants with smartphone-attached fish-eye cameras. In 918 of 9145 recorded events (appr. 10%), another person was visible behind the user. Eye gaze was also used by Saad et al. [71] and Abdrabou et al. [5], who considered screen fixations of at least 1 s as shoulder surfing attempts.

Shoulder Surfing Behavior & Behavioral Modeling Abdrabou et al. [5] explored the use of VR to study shoulder surfing in controlled yet immersive settings. They developed two 3D environments to track observers' gaze and movement patterns. From this, they identified three distinct attack patterns and proposed a behavioral model comprising three stages: *idle* (no active observation), *approach* (preparation to observe), and *attack* (direct observation). This model provides a detailed framework for characterizing attacker behavior and serves as a foundation for developing future mitigation strategies.

Mitigating Shoulder Surfing Mitigation strategies for shoulder surfing generally fall into four categories: resistant input techniques, content masking, user awareness mechanisms, and user-initiated mitigation strategies.

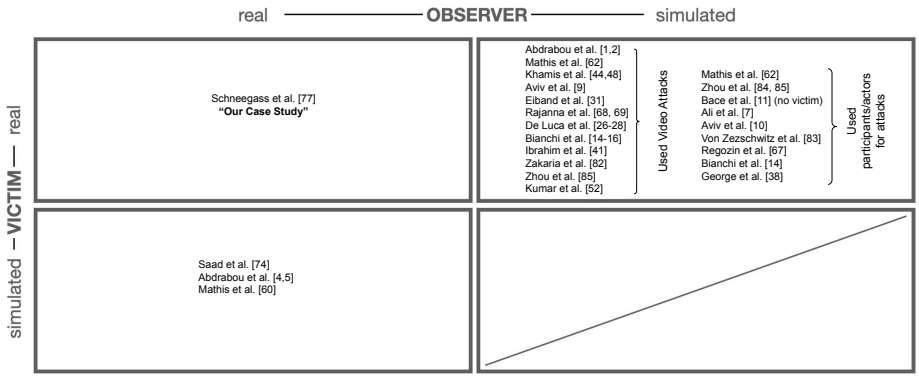

Fig. 2: Attacker and Victim Representation in the Literature

Shoulder-Surfing Resistant Input Techniques Many approaches aim to replace or obscure traditional input methods. Alternatives include gaze [21, 35, 49, 66], gestures [1], haptic/audio cues [14, 15], non-visual keypads [77], and invisible pressure input [46]. Multi-modal techniques combine modalities such as gaze and touch [41], PINs and gestures [68], or EEG and eye movements [22]. Other methods include scrambled PIN layouts [2] and graphical password schemes [17, 39], with a comprehensive review provided in [16].

Content Masking These techniques obscure screen content when a threat is detected. Examples include face-triggered blackouts [20], distance-sensitive images [76], and gaze-based content filtering [44,64]. Strategies also involve selective hiding [82], content distortion [80], or rendering text unreadable [29].

User Awareness Several systems aim to notify users when a potential shoulder surfer is detected. Saad et al. [69] used facial recognition to trigger visual and haptic alerts, including flashing borders and vibrations. Other studies explored visual and auditory cues [81], tactile feedback [52], and proxemic signals like silhouette alerts on public displays [20].

User Mitigation Strategies Users often rely on self-initiated defenses, such as shielding their screens, tilting devices, or ignoring perceived threats [48]. These measures are not always effective—Khan et al. [46] showed that tilting provides only limited protection, revealing common misconceptions about its efficacy.

3.2 Research Approaches

Attacker and Victim Representation Researchers used different combinations of real and simulated observers and victims to study shoulder surfing. Figure 2 positions prior work in this 2×2 space. Prior work mostly simulated either the victim or the observer.

Real Victim – Simulated Observer This is the most commonly used approach. Simulated observers, either researchers or study participants, observe

real users via live sessions or recorded footage. Often, fixed-position cameras capture the victim's interaction, allowing for post-hoc analysis where participants can pause, rewind, or rewatch scenes to replicate attempts.

Simulated Victim – Real Observer In this setup, researchers simulate victims and recruit real participants as observers. For example, Saad et al. [71] used 360° videos of staged shoulder surfing scenarios. VR studies by Abdrabou et al. [4, 5] and Mathis et al. [57] employed avatars to simulate victims, enabling the study of observer behavior in immersive settings. Similarly, George et al. [36] and Khamis et al. [43] investigated how real observers respond to various authentication techniques using simulated contexts.

Simulating both the victim and the observer is also possible, but it precludes the observation of natural interactions. To our knowledge, no prior work has pursued this configuration. Our focus lies in the fourth category, where both the victim and observer are real. This setting allows for the observation of truly natural behavior, including the dynamic interplay between parties, such as when an observer realizes they are being noticed.

Real Victim – Real Observer Only a few studies have explored this condition. Schneegass et al. [74] mounted fish-eye cameras on participants' smartphones to capture bystanders behind them during everyday phone use. While this approach enabled the capture of real-world interactions, it offered only limited context and could not determine observer intent, the nature of interactions, or the consequences of potential attacks.

Study Type Shoulder surfing research can be categorized into four types: self-reports, controlled lab studies, VR experiments, and real-world observations.

Asking People Surveys, interviews, and diary studies are commonly used to collect self-reported experiences and perceptions. Eiband et al. [28] surveyed users about shoulder surfing incidents, whether as victims or observers. Farzand et al. [31] conducted interviews to examine the social acceptability of mitigation strategies, emphasizing how interpersonal relationships influence user responses. They also deployed surveys [32] to assess perceptions of privacy-sensitive content, and a diary study [33] to document real-life occurrences over time.

Controlled Lab Studies Lab-based studies typically evaluate different input techniques designed to resist shoulder surfing [2,25,26,35,41,45,65,66,79,80], and examine the usability and effectiveness of countermeasures [20,29,44,64, 69, 70, 80–82]. Other studies explore factors like user positioning and observation angles [7, 9–11], or directly observe behavior in controlled setups [48]. While lab environments offer controlled conditions and repeatability, they often lack ecological validity, as tasks and contexts may feel artificial [81, 82]. Some researchers address this by showing pre-recorded scenarios [2, 41, 45], though this still falls short of fully replicating real-world dynamics.

Virtual Reality Studies VR has been used both as a tool to simulate real-world scenarios and as a context in which shoulder surfing itself poses a threat. On one hand, researchers have leveraged VR to explore shoulder surfing behaviors under controlled yet immersive conditions. Mathis et al. [56,58] compared shoulder surfing in VR and physical environments, enabling detailed analysis of scenarios that are difficult to replicate in the lab. However, constructing realistic VR settings remains challenging. For instance, Saad et al. [71] used 360° videos filmed in public transport with minimal bystanders, while Abdrabou et al. [5] simulated scenes like office spaces and bus stops with avatars. Both studies used partial deception to reduce bias, aiming to preserve natural behavior. While these methods offer valuable insights, their generalizability to real-world settings remains limited. On the other hand, VR itself is increasingly treated as a vulnerable environment. George et al. [36] evaluated traditional authentication methods under observation from outside VR, while others tested VR-specific schemes for resistance to shoulder surfing [**?**, 59]. These studies underscore the need to secure VR interactions against both co-located and remote observers.

Real World Studies Few studies have addressed shoulder surfing in natural environments. Schneegass et al. [74] equipped smartphones with fish-eye cameras to record surrounding activity during regular use. While this setup captured authentic interactions, it introduced several limitations. To preserve privacy, images were not shared with researchers, requiring participants to self-assess whether shoulder surfing occurred. Moreover, the visible hardware may have influenced both participant behavior and bystanders.

Data Collection Methods Shoulder surfing studies employ data collection methods like diaries, surveys, interviews, observations, logging, and eye tracking.

Diaries Diary studies provide longitudinal insights into shoulder surfing in daily life. Farzand et al. [33] used a 29-day diary to capture incidents and context. A related approach is experience sampling [12], where prompts are triggered by relevant events (e.g., phone unlocks). Harbach et al. [37] used a smartphone app to assess shoulder surfing risk at unlock moments. These methods capture rich, in-the-moment data but depend on participant compliance and awareness, risking underreporting of brief or unnoticed events.

Surveys Surveys reach large and diverse populations and are widely used to gather accounts of past shoulder surfing incidents. Eiband et al. [28] collected stories from both victims and observers. Marques et al. [53] explored unauthorized phone access among acquaintances. Farzand et al. [32] examined perceptions of privacy-sensitive content in shoulder surfing contexts.

Interviews Interviews complement other methods by uncovering participants' reasoning and perceptions. Schneegass et al. [74] used interviews post-study to interpret image-based incidents and perceived threats. Saad et al. [69] explored preferences for mitigation strategies, while Mathis et al. [57] compared VR and real-world experiences. Farzand et al. [31] investigated how personal relationships influence defense strategies.

Observations Observational methods involve directly watching and logging behavior in lab or field settings. For example, Kuhl et al. [48] observed participants in the lab to explore countermeasures used during shoulder surfing.

Logging Automated logging captures contextual data without relying on user input. Schneegass et al. [74] used fish-eye smartphone cameras and an app to log events during phone unlocks, including time, location, and visible surroundings. Bâce et al. [11] assessed risk from different viewing angles using wide-angle cameras and face detection to notify users of observers [69].

Eye Tracking Eye tracking is a precise measurement of gaze behavior. Ragozin et al. [64] used gaze data to trigger privacy-preserving techniques. Saad et al. [70] detected shoulder surfing using gaze estimation from front-facing cameras. In follow-up work [71], they combined 360° video with eye tracking to study observer attention. Abdrabou et al. [5] used gaze data in VR to analyze attack patterns. Corbett et al. [24] developed ShouldAR, an AR system that detects shoulder surfing using multimodal gaze tracking and rear-facing image capture, achieving 87.28% detection accuracy across settings.

3.3 Reflection on the Research Space

Our exploration highlights two key insights. First, while real-world studies are rare in shoulder surfing research, they offer unique insights by observing natural scenarios. Researchers must balance internal, external, and ecological validity when choosing between controlled (lab, VR) and less controlled (real-world) conditions, as observer and victim behaviors can influence each other [8].

Second, selecting the right data collection method is crucial. Modified smartphone cameras can affect observer behavior and limit contextual insights. In contrast, eye tracking, used in VR studies, provides richer data by simultaneously capturing observer and victim interactions. Exploring its use in real-world studies is worthwhile [5, 71, 74]. While eye-tracking research in lab settings is well-established, its application in real-world usable security research is limited. Mobile eye-trackers have only recently become available, offering the opportunity to conduct real-world eye-tracking research. However, it simultaneously presents a gap in understanding the challenges and pitfalls of these investigations, particularly in the context of usable security. Addressing this gap, our work takes an early step towards exploring and understanding the implications of eye tracking in real-world usable security research [27, 40, 50].

4 Case Study

The research space highlights the strengths and weaknesses of different approaches and demonstrates the importance of extending our understanding of shoulder surfing. Much of the research on shoulder surfing seeks to counter it. At the same time, work on how shoulder stuffing is triggered and how attackers behave is scarce, which makes it difficult to employ mitigation techniques effectively. Furthermore, the review of research approaches points out many methodological challenges. While the variety of approaches yielded many interesting results, the ability to observe natural, real-world shoulder surfing behavior would

be of high value to 1) confirm and extend what is so far known about shoulder surfing, as well as 2) discover and understand novel aspects.

To this end, the objective of this work is to develop a methodology capable of observing and comprehensively capturing natural shoulder surfing. To answer RQ_2, we report how we designed and conducted a case study to meet this objective. Our considerations towards coming up with a methodology to observe natural shoulder surfing included several major aspects: the study environment and task, the approach to data collection, and ethical considerations.

4.1 Threat Model

Shoulder surfing can occur in various scenarios, and the choice of the threat model greatly influences experimental design and conclusions. Wiese and Roth outline four categories based on live versus video observation and single versus multiple observations [78]. We consider opportunistic observers to be random adversaries in a public space, such as public transport. We use weak assumptions on the severity of the threat as they represent the worst-case scenario, giving us a better understanding of how vulnerable users are to shoulder-surfing in public.

4.2 Study Environment and Task

From prior work, we learned that public transport is among the most popular environments for shoulder surfing because most people primarily kill time during commutes [28]. We thus decided on a study entailing a task in this context.

It was essential for the study that participants remained naive to its true objective, which was to observe shoulder surfing events. We devised a pretext that made participants believe that our main objective was to investigate behavior change in public transport after COVID-19.

We provided participants with a task requiring them to reach a particular destination by public transport, similar to commuting to work, visiting friends, or sightseeing. More specifically, we asked them to visit a particular place using public transport, take a picture, and return. We chose the route so that shoulder surfing opportunities would occur. In particular, the route would begin in the city center with a short walk to the subway station. This included a busy crossing where people potentially would have to wait, thus creating an opportunity for shoulder surfing. The subway ride required them to ride six stops. The one-way commute would require about 20 minutes. We expected further shoulder surfing opportunities at the station, track, and on the subway.

4.3 Data Collection

Studying shoulder surfing from both the victim's and the observer's perspectives can yield complementary insights. We focus primarily on observers but also capture the victims' behavior and situational context. A promising way to log shoulder surfing incidents is to track where, when, and why observers direct their gaze. This can be done with a wearable eye tracker [40], which identifies the user's gaze location and overlays gaze data on a video of the surroundings.

Below, we describe how we selected and evaluated hardware solutions.

(a) (b) (c)

Fig. 3: Comparison between AdHawk MindLink (a), Tobii Pro Glasses 3 (b), and PupilLabs Invisible (c). The MindLink was most comfortable and unobtrusive.

Hardware Considerations We strive to sustain natural behavior to record unbiased victim-observer interactions. Hence, the choice of hardware is crucial. The used hardware requires being 1) unobtrusive, 2) lightweight, 3) with high resolution, and 4) adaptable to different lighting conditions. This enables users to behave naturally without feeling observed. Based on our experience and a market review, we selected three candidate devices: (a) AdHawk MindLink[8], (b) Tobii Pro Glasses 3[9], and (c) PupilLabs Pupil Invisible[10] (cf. Figure 3).

Assessment of Hardware Based on our considerations for the study environment and task, we conducted a pilot test to evaluate the suitability of various eye-tracking glasses for our study environment and task. Key considerations included the ability to record data continuously without user interaction, capture high-quality data in diverse lighting conditions, and minimize obtrusiveness.

The pilot study involved two female participants, both aged 25, with one wearing contact lenses. They completed the anticipated route three times, each time using a different eye tracker. Upon their return, we assessed several factors: the obtrusiveness of the eye tracker, comfort, and its impact on wearing a face mask. We also analyzed data and video streams for quality, focusing on resolution, light adaptation, readability of different contexts, and focus clarity.

Obtrusiveness Participants found the Tobii Pro Glasses 3 conspicuous, as they attracted attention from people nearby, resulting in frequent glances in their direction. In contrast, participants wearing the Pupil Labs and AdHawk glasses felt less noticeable, allowing them to behave more naturally. This may be attributed to the less visible world camera of the AdHawk glasses. While the Pupil Labs glasses were also less noticeable initially, some bystanders eventually took notice. Therefore, while Pupil glasses may be suitable for brief exposures, they may not be ideal for longer durations where observers receive prolonged attention, such as during extended public transport rides.

[8] AdHawk: https://www.adhawkmicrosystems.com/eye-tracking
[9] Tobii Pro Glasses 3: https://www.tobiipro.com/product-listing/tobii-pro-glasses-3/
[10] Pupil Invisible: https://pupil-labs.com/products/invisible/

Additionally, the AdHawk glasses have a disadvantage due to the hard-to-hide cable attached to the middle of the temples.

Wearing Comfort The Pupil Labs eye tracker frame was difficult to fit into people's head forms. The long temples caused the glasses to slip up and down frequently. In contrast, both the Tobii Pro Glasses 3 and the AdHawk were easy to fit. The nose piece was useful: it kept glasses in place and made them comfortable to wear.

Condensation For safety reasons, our participants wore face masks during the study. This introduces an additional challenge, as the eye-tracking glasses become foggy. The AdHawk glasses came with an anti-fog cleaning cloth, substantially reducing this issue.

Resolution and Content Visibility We tested various font sizes and smartphone orientations in a controlled lab to evaluate each eye tracker's resolution and content visibility. We used font sizes 8, 12, and 16 pt, tested black and white screen backgrounds, and assessed text visibility at display angles of 0°, 30°, and 60°. We also examined different hand postures (one hand vs. two hands) and distances from the eye tracker (15, 30, and 60 cm). The Tobii Pro Glasses 3 provided the best resolution, with content clearly visible.

Summary We reviewed various wearable eye trackers, finding that the Tobii Pro Glasses 3 and AdHawk eye tracking provided comparable results. We chose AdHawk as it is less obtrusive, making participants more comfortable wearing it in public. Recording on a smartphone also enhances discretion. The AdHawk tracker features an 82-degree field of view, a 16:9 aspect ratio, 1080 p resolution, and a 30 Hz recording rate.

4.4 Ethical Considerations

Our case study was guided by several ethical considerations and approved by the university's ethical review board.

Partial Deception Our study required partial deception. Deception studies are commonly used in Human-Computer Interaction research [6] to reduce participant bias and ensure the collection of ecologically valid data. They also enable researchers to gain insights that might not be attainable through other methods [30]. We informed participants about the data that would be collected throughout the study, yet did not reveal our interest in shoulder surfing to not influence their natural behavior. Participants were debriefed after data collection. Participants were explicitly told that at any point, they could decide without any disadvantage to have data and recordings deleted if they felt uncomfortable sharing them for research purposes.

Observer's Gaze Data Researchers need to consider how to preserve participants' anonymity (in this case, observers). In particular, recordings of participants' eyes might reveal their identities. We addressed this by collecting gaze paths only. The data was used to calculate fixations and saccades (to automatically identify shoulder surfing) and to visualize gaze points in videos (to support visual inspection).

Recording Video Data in Public In the jurisdiction where the study was conducted, recording video data in public spaces is legal, even with visible individuals. However, publishing this data requires explicit consent from identifiable individuals. Presence in the recording does not make someone a study participant, so no consent is needed for recording. Note that laws may differ elsewhere, and researchers should consult experts before their studies.

Public Transport Permission We acquired permission from the local transit authority to conduct our study and record videos on their premises. Our study adhered to their general transport regulations, and they were generally interested in our research. Yet, we were only permitted to conduct the study during off-peak hours.

Visibility of Bystanders' Faces In case researchers plan to release video data from the world camera in publications or datasets, bystanders' faces must be anonymized. During our study, mandatory face masks on public transport naturally anonymized the data. However, if full faces are visible, researchers must employ anonymization techniques like face blurring.

Visibility of Content on Observer and Victim's Devices Recording the view of a mobile eye tracker's world camera may capture sensitive content on both observers' and bystanders' devices. We informed participants about this during the debriefing and obtained their consent, though it may have influenced their smartphone use. To protect bystanders, we blurred screens before data analysis. Future research could develop methods to abstract screen content, such as overlays indicating the type of app in use.

4.5 Design

We collected the following data: participants' fixations on others' screens, video recordings, post-study questionnaires reflecting on personality traits [75] and privacy perceptions questionnaire [23], and answers to the post-study interview.

4.6 Recruiting & Demographics

We recruited 15 participants (7 females) through university mailing lists and word of mouth. Participants were 21 to 34 years old ($M = 25.5; SD = 3.5$). Thirteen participants were students, and two were full-time workers. Participants were rather inexperienced with eye tracking (5-Point Likert scale; 1=no experience at all; 5=strong experience; $M = 2.57; SD = 1.13$). No participant wore glasses. One had contact lenses. Our sample size is comparable with previous studies using cameras and eye tracking in real-world settings. For instance, Khamis et al. [42] recruited 11 participants, and Lebreton et al. [51] included 16 participants in their study.

4.7 Procedure

As participants arrived at our lab, we explained to them that the study investigates how people's behavior changed in public transport after COVID-19 and that we would record gaze data. After obtaining consent and demographic information, participants were equipped with the eye tracker and assisted with fitting and calibration.

Calibration was performed before each session using AdHawk's calibration software, and participants were instructed to behave naturally and keep the glasses on to minimize hardware intrusiveness over time. We also provided them with transportation tickets and face masks if needed. We maintained communication for assistance during the journey. Upon return, recordings were transferred and reviewed for errors. Participants provided feedback on 1) their familiarity with the route, 2) the eye tracker, 3) how they felt wearing the glasses in public, and 4) how COVID-19 changed their behavior and distance from others in public transport. Then, we asked them to fill in the personality traits and the privacy attitudes questionnaires. Finally, we conducted a semi-structured interview covering route description, observations during the ride, and debriefing. Participants had the option to delete their gaze data, though none chose to do so. Compensation of 15 EUR was provided for the 1.5-hour study.

4.8 Limitations

Participants' unfamiliarity with the route may have affected their commuting behavior, potentially reducing shoulder surfing opportunities. However, their behavior likely aligns with patterns in unfamiliar locations.

Moreover, to meet privacy and ethical guidelines, participants were informed about the eye tracker's world camera, which may have influenced their smartphone use. Yet, reduced phone usage could have shifted attention to the environment, increasing the chances of shoulder surfing.

Although new hardware can influence participants' behavior, extended use minimizes this effect as they become accustomed to it [62, 63].

Finally, the AdHawk MindLink's 10K EUR cost restricted participation to university students and staff due to financial and insurance constraints. While this limited our sample, we believe it still fulfilled the study's purpose.

5 Results and Discussion

5.1 Data Overview & Preprocessing

We collected approximately 600 minutes of video and gaze data (40 minutes per participant, $N = 15$). To ensure data integrity, we visually inspected the recordings using the AdHawk MindLink software[11].

Despite extensive precautions to reduce data loss, such as pilot testing and careful planning, issues still emerged. Because participants moved naturally throughout the study, occasional data corruption occurred, primarily due to cable friction and movement. This was more common when participants carried the tracking device in a pocket, whereas placing it in a bag yielded more stable recordings. Such data loss is a well-documented challenge in real-world eye tracking [51], often caused by motion, lighting variability, or hardware limitations. While standard mitigation strategies include multiple trials and route validation, data degradation remains difficult to eliminate entirely. Due to these issues, we report results from 7 participants whose data met quality standards.

[11] AdHawk MindLink: https://www.adhawkmicrosystems.com/adhawk-mindlink

We identified gaze fixations using the Dispersion-Threshold Identification (I-DT) algorithm [73], applying default parameters (dispersion threshold = 25, duration threshold = 100 ms). Future work could improve robustness by incorporating automated gaze metrics and inter-rater reliability checks to support consistent labeling and interpretation.

5.2 Shoulder Surfing Attempts

To assess shoulder surfing behavior, we employed a multi-step analysis pipeline. First, we used YOLOv3 object detection [67] to identify gaze intersections with areas of interest (AOIs) such as smartphones, tablets, laptops, books, and newspapers. Gaze fixations overlapping with these AOIs were extracted and grouped into consecutive intervals, which we labeled as potential shoulder surfing events. To improve detection accuracy, particularly in cases where YOLOv3 missed or misclassified AOIs due to occlusions, we manually annotated the video data using MAXQDA[12]. Manual annotations also included environmental features such as direct light sources (e.g., spotlights, flashing doors) and crowd density, categorized as low (fewer than 2 people nearby), medium (2–5), or high (more than 5). We additionally annotated participants' own devices, people, and visible displays in the field of view. To protect bystander privacy, we applied YOLOv3-based blurring to detected AOIs on other users' screens, which were occasionally visible in low resolution through the world camera.

We identified 35 shoulder surfing instances, averaging five per participant. These lasted 588 ms (SD=419) on average, with half being brief (100 ms) and two exceeding 1300 ms. All but one targeted smartphones, aligning with prior findings [28,33]. VR studies suggest gaze duration varies with content type, with immersive media attracting longer glances [5].

Next, we analyzed observer-victim positioning. Contrary to Saad et al. [71], who found most incidents occur while both are sitting, 22 of 35 cases (62.8%) in our study happened while both were standing, 7 with the observer standing and victim sitting (20%), 3 with the observer sitting and victim standing (8.5%), and 3 with both sitting (8.5%).

Regarding the observation duration, sitting–sitting positions resulted in the longest average duration ($M = 450\,ms; SD = 100$), while standing–standing and sitting–standing had similar durations ($M = 300\,ms; SD = 70$). The longest individual observations occurred in the standing–standing position. Overall, durations in our study were shorter than prior work (¡ 1 s) [5, 71], suggesting more casual observations [13].

We also analyzed the effects of lighting and crowd density on participants' behavior during these attempts. In 85% of the cases, the environment was very crowded, with only 5 attempts occurring in low-crowd. Lighting conditions remained consistent throughout the observations, with no notable variation.

Inferential statistics were unsuitable due to the small sample and our exploratory aim. Still, our descriptive and behavioral analyses offer meaningful insights into participants' behavior.

[12] MAXQDA https://www.maxqda.com/

5.3 Post Study Questionnaires

Upon their return, participants rated their ride experience and eye tracker usage on a scale from 1 to 5. They reported frequent use of public transport ($M = 4.1; SD = 1.3$) and were initially unfamiliar with eye tracking ($M = 2.5; SD = 1.3$). Despite varying attempts to maintain distance from others ($M = 3.3; SD = 1.3$), participants felt comfortable wearing the eye tracker ($M = 4.3; SD = 0.7$) and experienced few issues with it during the study ($M = 4.1; SD = 1.2$).

In terms of privacy attitudes, participants were willing to be exposed and monitored by others, but reluctant to share personal information online. They reported using protection techniques such as two-factor authentication (2FA). Regarding personality traits, participants showed average to strong levels of extraversion ($M = 3.2, SD = 0.7$), agreeableness ($M = 3.8, SD = 0.7$), conscientiousness ($M = 3.4, SD = 1.0$), and low negative emotionality ($M = 2.6, SD = 0.8$). They exhibited high levels of open-mindedness ($M = 4.5, SD = 0.5$).

5.4 Interview Analysis

We conducted semi-structured interviews to explore participants' activities during rides and notable observations. Most reported observing people, listening to music, messaging, and navigating, while two avoided phone use for privacy. Many noticed others on their phones; for instance, P3 saw a video, P4 an article, and P1 an unidentified screen.

During debriefing, all but one admitted to past shoulder surfing, often unconsciously. P5, an IT professional, was particularly interested in phone interactions. Participants cited boredom, distraction, and curiosity as motives [28]. Longer commutes encouraged screen observation, especially colorful content like videos and puzzles [5]. To avoid onlookers, they limited phone use or used subtle cues like turning their heads. While they acknowledged shoulder surfing, they framed it as casual rather than malicious.

5.5 Discussion and Implications

This case study explored shoulder surfing in real-world public settings, with an emphasis on preserving natural observer behavior and protecting data privacy. While the in-situ approach offers high ecological validity, several challenges impacted our ability to study shoulder surfing without bias. First, participant awareness of being monitored may have led to altered behavior, potentially suppressing shoulder surfing attempts. Additionally, participants' unfamiliarity with the chosen routes may have diverted attention away from their surroundings. Future studies could mitigate this by leveraging participants' regular commute paths, reducing distraction and enhancing ecological realism.

The study also required significant effort in design, execution, and ethics management. Despite attempts to automate data collection, the duration and complexity of the sessions still necessitated manual annotation and analysis.

Although the dataset was limited, our findings demonstrate the feasibility of capturing naturalistic shoulder surfing behavior. Notably, we observed both alignment and divergence with previous research. Compared to VR-based studies, participants exhibited fewer and shorter shoulder surfing attempts—likely due to the real social consequences present in public settings, which may increase self-consciousness and suppress overt observation behavior. Interview data further revealed nuanced interactions between observers and victims, with some victims displaying heightened awareness and actively adopting defensive behaviors. This highlights the importance of capturing such dynamics, which are often absent from simulated studies.

From a practical standpoint, our findings suggest that privacy-aware interfaces should adapt to contextual signals such as crowd density, posture, or gaze direction. For example, smartphones could trigger subtle alerts or activate privacy-enhancing display modes when nearby gaze is detected [44]. Similarly, gaze estimation and ambient sensing could enable proactive, context-sensitive defenses that operate independently of user action [11, 44]. Given the variation in lighting and spatial conditions, adaptive display technologies—responsive to light levels, proximity, or user movement—may further reduce accidental screen exposure. Crucially, these solutions should prioritize passive protection mechanisms that minimize user burden while maximizing real-world effectiveness.

Our study sheds light on the complexity of studying shoulder surfing in the wild. It offers key insights into real-world behavior, challenges current assumptions derived from lab and VR-based research, and lays the groundwork for refining methodologies and designing more context-aware mitigation strategies. These contributions support future work with stronger ecological grounding.

6 Best Practices and Lessons Learned

Based on our case study design and results, we present best practices and lessons learned. While our focus was on using eye tracking to study shoulder surfing, many findings apply to other technologies. We group our best practices into three categories: study environment, study design, and study ethics and privacy.

6.1 Study Environment

Select Locations That Maximize Shoulder Surfing Opportunities The study environment should reflect real-world conditions where shoulder surfing naturally occurs. Public transport, waiting areas, and busy cafés provide ideal settings due to high interaction rates with mobile devices [5, 28]. Researchers should pilot test multiple locations to identify where participants exhibit the most opportunistic observations.

Manage Environmental Distractions Common distractions, such as digital billboards or background noise, can reduce participants' likelihood of engaging in shoulder surfing. Previous studies highlight the importance of accounting for environmental factors that may divert attention and bias behavior [33, 71]. Document major distractions to account for their influence in the analysis.

Minimize Observer Influence Equipment should be discreet to prevent altering natural behavior. In our study, a visible world camera sometimes raised bystanders awareness. Abdrabou et al. [4] similarly found that subtle VR cues influenced social behavior, highlighting the need to reduce hardware visibility and observer presence in situated studies. Future research could explore alternative hardware placement or clothing accessories to conceal devices while maintaining data integrity.

6.2 Study Design

Create Natural Shoulder Surfing Opportunities Design tasks that blend seamlessly into the environment and participant routines. For example, asking participants to navigate a new area or wait for transport naturally encourages observational behavior. Avoid artificial scenarios that make participants overly aware of being observed [28].

Reduce Awareness of Study Objectives Participants aware of being studied may alter their behavior. Using deployment-based research, such as embedding the study within a routine task, can minimize this effect [38,60]. If full deception is necessary, ensure ethical safeguards, including thorough debriefing.

Limit Technology Interaction to Preserve Natural Behavior Frequent user interactions with research equipment can disrupt natural behavior. Ensure minimal setup adjustments and automate data collection as much as possible. Abdrabou et al. [5] emphasize minimizing intrusions in situated studies to preserve natural behavior, especially in mobile contexts. Testing the equipment in real-world scenarios before the study can help refine the process.

Use Delayed Debriefing to Avoid Bias Revealing study objectives too early may influence participant behavior. Withholding study intent until after participation reduces social desirability bias while maintaining autonomy through post-hoc consent [61]. Conducting debriefing only after data collection ensures more natural responses while providing participants the option to withdraw their data if needed.

Account for Novelty Effects New or unfamiliar technology, such as eye trackers, can make participants self-conscious. Farzand et al. [33] discuss the importance of acclimating participants to new devices in diary and deployment studies, as novelty can skew initial behaviors. Where possible, use common devices or support familiarization before starting the study to reduce artificial behavior.

6.3 Study Ethics and Privacy

Anonymize Gaze Data to Protect Participants Gaze data can reveal personal habits and attention patterns [3,47]. To preserve anonymity, only necessary gaze path information must be collected and processed into aggregated metrics rather than raw gaze recordings.

Follow Legal Guidelines for Public Video Recording nderstand and adhere to the legal requirements of recording video data in public settings. Farzand et al. [33] highlight the need for ethical and legal compliance in public in-the-wild studies, stressing the distinction between data collection and broader use. Differentiate between permissible recording and publishing, and obtain explicit consent for the latter to respect jurisdictional regulations and participant rights.

Implement Bystander Anonymization Techniques When using a world camera, anonymizing bystanders is essential. Saad et al. [70] illustrate techniques such as real-time face detection and masking to protect the identity of surrounding individuals captured in studies. Employ real-time face-blurring where feasible or process the data post-capture to remove identifying features.

Obtain Explicit Consent for Recording Device Screens Since shoulder surfing studies involve viewing sensitive content, participants should provide informed consent regarding the visibility of their screens. Implementing methods to abstract sensitive details, such as overlaying dummy text [44], can help balance ethical concerns with data accuracy.

Ensure a Thorough Debriefing Process If deception is used to maintain study integrity, researchers must conduct a structured debriefing to explain the true study purpose, address concerns, and allow participants to opt-out. This follows best practices in deceptive field research, highlighting the need for transparent, participant-centered debriefing [61].

By implementing these best practices, researchers can better capture authentic user behavior, leading to more valid findings and the development of effective shoulder surfing mitigation strategies.

7 Conclusion

Introducing a framework for shoulder surfing research, our case study investigates real-world exploration using eye tracking. Through our case study, we demonstrate this approach, detailing study considerations, findings, and lessons learned. While our methodology can be further enhanced, it lays the groundwork for future research aiming to devise innovative mitigation strategies.

References

1. Abdrabou, Y., Khamis, M., Eisa, R.M., Ismael, S., Elmougy, A.: Engage: Resisting shoulder surfing using novel gaze gestures authentication. In: Proc. MUM'17. ACM, New York, NY, USA (2018). https://doi.org/10.1145/3282894.3289741
2. Abdrabou, Y., Khamis, M., Eisa, R.M., Ismail, S., Elmougy, A.: Just gaze and wave: Exploring the use of gaze and gestures for shoulder-surfing resilient authentication. In: Proc. ETRA'19. ACM, New York, NY, USA (2019)
3. Abdrabou, Y., Omelina, T., Dietz, F., Khamis, M., Alt, F., Hassib, M.: Where do you look when unlocking your phone?: A field study of gaze behaviour during smartphone unlock. In: CHI EA'24. ACM, New York, NY, USA (2024)

4. Abdrabou, Y., Rivu, R., Ammar, T., Liebers, J., Saad, A., Liebers, C., Gruenefeld, U., Knierim, P., Khamis, M., Mäkelä, V., Schneegass, S., Alt, F.: Understanding shoulder surfer behavior using virtual reality. In: Proc. VRW'22) (2022)
5. Abdrabou, Y., Rivu, S.R., Ammar, T., Liebers, J., Saad, A., Liebers, C., Gruenefeld, U., Knierim, P., Khamis, M., Makela, V., Schneegass, S., Alt, F.: Understanding shoulder surfer behavior and attack patterns using virtual reality. In: Proc. AVI'22. ACM, New York, NY, USA (2022). https://doi.org/10.1145/3531073.3531106
6. Adar, E., Tan, D.S., Teevan, J.: Benevolent deception in human computer interaction. In: Proc. CHI'13. ACM, New York, NY, USA (2013)
7. Ali, M.E., Anwar, A., Ahmed, I., Hashem, T., Kulik, L., Tanin, E.: Protecting mobile users from visual privacy attacks. In: Adj. Proc. UbiComp'14. ACM, New York, NY, USA (2014). https://doi.org/10.1145/2638728.2638788
8. Alt, F., Schneegaß, S., Schmidt, A., Müller, J., Memarovic, N.: How to evaluate public displays. In: Proc. PerDis '12. ACM, New York, NY, USA (2012)
9. Aviv, A.J., Davin, J.T., Wolf, F., Kuber, R.: Towards baselines for shoulder surfing on mobile authentication. In: Proc. ACSAC'17. ACM, New York, NY, USA (2017)
10. Aviv, A.J., Wolf, F., Kuber, R.: Comparing video based shoulder surfing with live simulation. In: Proc. ACSAC'18. ACM, New York, NY, USA (2018)
11. Bâce, M., Saad, A., Khamis, M., Schneegass, S., Bulling, A.: Privacyscout: Assessing vulnerability to shoulder surfing on mobile devices. In: Proc. PETs'22 (2022)
12. van Berkel, N., Ferreira, D., Kostakos, V.: The experience sampling method on mobile devices. ACM CSUR 50(6) (2017). https://doi.org/10.1145/3123988
13. Bianchi, A., Oakley, I.: Multiplexed input to protect against casual observers. In: Proc. HCIK'15. Hanbit Media, Inc., Seoul, KOR (2014)
14. Bianchi, A., Oakley, I., Kwon, D.S.: Spinlock: A single-cue haptic and audio pin input technique for authentication. In: Haptic and Audio Interaction Design. Springer, Berlin, Heidelberg (2011)
15. Bianchi, A., Oakley, I., Kwon, D.S.: Counting clicks and beeps: Exploring numerosity based haptic and audio PIN entry. Interacting with Computers 24(5) (07 2012)
16. Binbeshr, F., Siong, K.C., Yee, L., Imam, M., Al-Saggaf, A.A., Abudaqa, A.A.: A systematic review of graphical password methods resistant to shoulder-surfing attacks. Intl. Journal of Information Security 24(1), 1–22 (2025)
17. Bostan, H., Bostan, A.: Shoulder surfing resistant graphical password schema: Randomized pass points (rpp). Multimedia Tools and Applications 82(28) (2023)
18. Bošnjak, L., Brumen, B.: Shoulder surfing: From an experimental study to a comparative framework. IJHCS 130, 1–20 (2019)
19. Bošnjak, L., Brumen, B.: Shoulder surfing experiments: A systematic literature review. Computers & Security 99, 102023 (2020)
20. Brudy, F., Ledo, D., Greenberg, S., Butz, A.: Is anyone looking? mitigating shoulder surfing on public displays through awareness and protection. In: Proc. PerDis'14. ACM, New York, NY, USA (2014). https://doi.org/10.1145/2611009.2611028
21. Bulling, A., Alt, F., Schmidt, A.: Increasing the security of gaze-based cued-recall graphical passwords using saliency masks. In: Proc. CHI '12. ACM, NY, USA (2012)
22. Cheng, S., Wang, J., Sheng, D., Chen, Y.: Identification with your mind: A hybrid bci-based authentication approach for anti-shoulder-surfing attacks using eeg and eye movement data. IEEE Trans. on Instrumentation and Measurement 72 (2023)
23. Chignell, M.H., Quan-Haase, A., Gwizdka, J.: The privacy attitudes questionnaire (paq): initial development and validation. In: Proc. Human Factors and Ergonomics Society annual meeting. vol. 47. SAGE Publications (2003)
24. Corbett, M., David-John, B., Shang, J., Ji, B.: Shouldar: Detecting shoulder surfing attacks using multimodal eye tracking and ar. Proc. ACM IMWUT 8(3) (Sep 2024)

25. De Luca, A., Denzel, M., Hussmann, H.: Look into my eyes! can you guess my password? In: Proc. SOUPS'09. ACM, New York, NY, USA (2009)
26. De Luca, A., von Zezschwitz, E., Nguyen, N.D.H., Maurer, M.E., Rubegni, E., Scipioni, M.P., Langheinrich, M.: Back-of-device authentication on smartphones. In: Proc. CHI'13. ACM, New York, NY, USA (2013)
27. Duchowski, A.: Eye tracking methodology: Theory and practice. Springer (2017)
28. Eiband, M., Khamis, M., Von Zezschwitz, E., Hussmann, H., Alt, F.: Understanding shoulder surfing in the wild: Stories from users and observers. In: Proc. CHI'17. ACM, New York, NY, USA (2017)
29. Eiband, M., von Zezschwitz, E., Buschek, D., Hußmann, H.: My scrawl hides it all: Protecting text messages against shoulder surfing with handwritten fonts. In: CHI EA'16. ACM, New York, NY, USA (2016)
30. Fan, J., Shen, X.: New progress in the paradigm of elicited deception: Application of hci in deception detection. In: Proc. ICISE-IE'21'. IEEE (2021)
31. Farzand, H., Bhardwaj, K., Marky, K., Khamis, M.: The interplay between personal relationships & shoulder surfing mitigation. In: MuC'21. ACM, NY, USA (2021)
32. Farzand, H., Marky, K., Khamis, M.: I hate when people do this; there'sa lot of sensitive content for me": A typology of perceived privacy-sensitive content in shoulder surfing scenarios. In: Proc. SOUPS'22 (2022)
33. Farzand, H., Marky, K., Khamis, M.: Shoulder surfing through the social lens: A longitudinal investigation & insights from an exploratory diary study. In: Proc. EuroUSEC'22 (2022)
34. Farzand, H., Marky, K., Khamis, M.: Out-of-device privacy unveiled: Designing and validating the out-of-device privacy scale (odps). In: Proc. CHI'24. ACM, New York, NY, USA (2024)
35. Forget, A., Chiasson, S., Biddle, R.: Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In: Proc. CHI'10. ACM, NY, USA (2010)
36. George, C., Khamis, M., von Zezschwitz, E., Burger, M., Schmidt, H., Alt, F., Hussmann, H.: Seamless and secure vr: Adapting and evaluating established authentication systems for virtual reality. In: Proc. USEC'17. NDSS (2017)
37. Harbach, M., Von Zezschwitz, E., Fichtner, A., De Luca, A., Smith, M.: {It's} a hard lock life: A field study of smartphone (un) locking behavior and risk perception. In: Proc. SOUPS'14) (2014)
38. Henze, N., Rukzio, E., Boll, S.: 100,000,000 taps: Analysis and improvement of touch performance in the large. In: Proc. MobileHCI'11. ACM, NY, USA (2011)
39. Jermyn, I., Mayer, A., Monrose, F., Reiter, M., Rubin, A.: The design and analysis of graphical passwords. In: Proc. Security'99. USENIX, Berkeley, CA (1999)
40. Katsini, C., Abdrabou, Y., Raptis, G.E., Khamis, M., Alt, F.: The role of eye gaze in security and privacy applications: Survey and future hci research directions. In: Proc. CHI'20. ACM, New York, NY, USA (2020)
41. Khamis, M., Alt, F., Hassib, M., von Zezschwitz, E., Hasholzner, R., Bulling, A.: Gazetouchpass: Multimodal authentication using gaze and touch on mobile devices. In: CHI EA'16. ACM, New York, NY, USA (2016)
42. Khamis, M., Baier, A., Henze, N., Alt, F., Bulling, A.: Understanding face and eye visibility in front-facing cameras of smartphones used in the wild. In: Proc. CHI'18. ACM, New York, NY, USA (2018)
43. Khamis, M., Bandelow, L., Schick, S., Casadevall, D., Bulling, A., Alt, F.: They are all after you: investigating the viability of a threat model that involves multiple shoulder surfers. In: Proc. MUM'17. ACM, New York, NY, USA (2017)
44. Khamis, M., Eiband, M., Zürn, M., Hussmann, H.: Eyespot: Leveraging gaze to protect private text content on mobile devices from shoulder surfing. MTI (2018)

45. Khamis, M., Hassib, M., Zezschwitz, E.v., Bulling, A., Alt, F.: GazeTouchPIN: protecting sensitive data on mobile devices using secure multimodal authentication. In: Proc. ICMI'17. ACM, New York, NY, USA (2017)
46. Khan, H., Hengartner, U., Vogel, D.: Evaluating attack and defense strategies for smartphone pin shoulder surfing. In: Proc. CHI'18. ACM, NY, USA (2018)
47. Kröger, J.L., Lutz, O.H.M., Müller, F.: What Does Your Gaze Reveal About You? On the Privacy Implications of Eye Tracking. Springer Intl. Publishing, Cham (2020)
48. Kühn, R., Korzetz, M., Schlegel, T.: User strategies for mobile device-based interactions to prevent shoulder surfing. In: Proc. MUM'19. ACM, NY, USA (2019)
49. Kumar, M., Garfinkel, T., Boneh, D., Winograd, T.: Reducing shoulder-surfing by using gaze-based password entry. In: Proc. SOUPS'07. ACM, NY, USA (2007)
50. Land, M., Tatler, B.: Looking and acting: vision and eye movements in natural behaviour. Oxford University Press (2009)
51. Lebreton, P., Hupont, I., Mäki, T., Skodras, E., Hirth, M.: Eye tracker in the wild: studying the delta between what is said and measured in a crowdsourcing experiment. In: Proc. 4th Intl. Workshop on Crowdsourcing for Multimedia (2015)
52. Luo, W., Lan, B., Wan, X., Liu, Z., Zeng, Y., Ma, J.: Feel vibration: Challenge-response mobile authentication with covert channel. In: Proc. ICCT'20 (2020)
53. Marques, D., Guerreiro, T., Carriço, L., Beschastnikh, I., Beznosov, K.: Vulnerability & blame: Making sense of unauthorized access to smartphones. In: Proc. CHI'19. ACM, New York, NY, USA (2019)
54. Marques, D., Muslukhov, I., Guerreiro, T., Carriço, L., Beznosov, K.: Snooping on mobile phones: Prevalence and trends. In: Proc. SOUPS'16. USENIX Association, Denver, CO (Jun 2016)
55. Masuzawa, R., Aburada, K., Yamaba, H., Katayama, T., Okazaki, N.: Development of keypads which use colors or shapes to prevent shoulder surfing. Artificial Life and Robotics 28(4), 710–717 (2023)
56. Mathis, F., O'Hagan, J., Vaniea, K., Khamis, M.: Stay home! conducting remote usability evaluations of novel real-world authentication systems using virtual reality. In: Proc. AVI'22. ACM, New York, NY, USA (2022)
57. Mathis, F., O'Hagan, J., Khamis, M., Vaniea, K.: Virtual reality observations: Using vr to augment lab-based shoulder surfing research. In: IEEE VR'22 (2022)
58. Mathis, F., Vaniea, K., Khamis, M.: Can i borrow your atm? using virtual reality for (simulated) in situ authentication research. In: IEEE VR'22 (2022)
59. Mathis, F., Williamson, J., Vaniea, K., Khamis, M.: Rubikauth: Fast and secure authentication in virtual reality. In: CHI EA'20. ACM, New York, NY, USA (2020)
60. Müller, J., Walter, R., Bailly, G., Nischt, M., Alt, F.: Looking glass: A field study on noticing interactivity of a shop window. In: Proc. CHI'12. ACM, NY, USA (2012)
61. Muslukhov, I., Boshmaf, Y., Kuo, C., Lester, J., Beznosov, K.: Know your enemy: The risk of unauthorized access in smartphones by insiders. In: Proc. MobileHCI'13. ACM, New York, NY, USA (2013). https://doi.org/10.1145/2493190.2493223
62. Peng, W., Li, L., Kononova, A., Cotten, S., Kamp, K., Bowen, M., et al.: Habit formation in wearable activity tracker use among older adults: qualitative study. JMIR mHealth and uHealth 9(1), e22488 (2021)
63. Pires, D., Gayet-Ageron, A., Guitart, C., Robert, Y.A., Fankhauser, C., Tartari, E., Peters, A., Tymurkaynak, F., Fourquier, S., et al.: Effect of wearing a novel electronic wearable device on hand hygiene compliance among health care workers: a stepped-wedge cluster randomized clinical trial. JAMA Network Open 4(2) (2021)
64. Ragozin, K., Pai, Y.S., Augereau, O., Kise, K., Kerdels, J., Kunze, K.: Private reader: Using eye tracking to improve reading privacy in public spaces. In: Proc. MobileHCI'19. ACM, New York, NY, USA (2019)

65. Rajanna, V., Malla, A.H., Bhagat, R.A., Hammond, T.: Dygazepass: A gaze gesture-based dynamic authentication system to counter shoulder surfing and video analysis attacks. In: Proc. ISBA'18' (2018)
66. Rajanna, V., Polsley, S., Taele, P., Hammond, T.: A gaze gesture-based user authentication system to counter shoulder-surfing attacks. In: CHI EA '17. ACM, New York, NY, USA (2017)
67. Redmon, J., Farhadi, A.: Yolov3: An incremental improvement. arXiv (2018)
68. Riyadh, H., Bhardwaj, D., Dabrowski, A., Krombholz, K.: Usable authentication in virtual reality: Exploring the usability of pins and gestures. In: Intl. Conf. on Applied Cryptography and Network Security. Springer (2024)
69. Saad, A., Chukwu, M., Schneegass, S.: Communicating shoulder surfing attacks to users. In: Proc. MUM'17. ACM, New York, NY, USA (2018)
70. Saad, A., Elkafrawy, D.H., Abdennadher, S., Schneegass, S.: Are they actually looking? identifying smartphones shoulder surfing through gaze estimation. In: Adj. Proc. ETRA '20. ACM, New York, NY, USA (2020)
71. Saad, A., Liebers, J., Gruenefeld, U., Alt, F., Schneegass, S.: Understanding bystanders' tendency to shoulder surf smartphones using 360-degree videos in virtual reality. In: Proc. MobileHCI'21. ACM, New York, NY, USA (2021)
72. Saad, A., Liebers, J., Schneegass, S., Gruenefeld, U.: "they see me scrollin"—lessons learned from investigating shoulder surfing behavior and attack mitigation strategies. In: Human Factors in Privacy Research, pp. 199–218. Springer (2023)
73. Salvucci, D.D., Goldberg, J.H.: Identifying fixations and saccades in eye-tracking protocols. In: Proc. ETRA'00. ACM, New York, NY, USA (2000)
74. Schneegaß, S., Saad, A., Heger, R., Delgado, S., Poguntke, R., Alt, F.: An investigation of shoulder surfing attacks on touch-based unlock events. Proc. ACM HCI (2022)
75. Soto, C.J., John, O.P.: Short and extra-short forms of the big five inventory–2: The bfi-2-s and bfi-2-xs. Journal of Research in Personality 68, 69–81 (2017)
76. Tang, B.J., Shin, K.G.: Eye-Shield: Real-Time protection of mobile device screen information from shoulder surfing. In: Proc. USENIX Security'23 (2023)
77. Varma, M., Watson, S., Chan, L., Peiris, R.: Vibroauth: Authentication with haptics based non-visual, rearranged keypads to mitigate shoulder surfing attacks. In: HCI for Cybersecurity, Privacy and Trust. Springer Intl. Publishing, Cham (2022)
78. Wiese, O., Roth, V.: See you next time: a model for modern shoulder surfers. In: Proc. MobileHCI'16. ACM, New York, NY, USA (2016)
79. Zakaria, N.H., Griffiths, D., Brostoff, S., Yan, J.: Shoulder surfing defence for recall-based graphical passwords. In: Proc. SOUPS'11. ACM, New York, NY, USA (2011)
80. von Zezschwitz, E., Ebbinghaus, S., Hussmann, H., De Luca, A.: You can't watch this! privacy-respectful photo browsing on smartphones. In: Proc. CHI'16. ACM, New York, NY, USA (2016). https://doi.org/10.1145/2858036.2858120
81. Zhou, H., Ferreira, V., Alves, T., Hawkey, K., Reilly, D.: Somebody is peeking! a proximity and privacy aware tablet interface. In: CHI EA'15. ACM, NY, USA (2015)
82. Zhou, H., Tearo, K., Waje, A., Alghamdi, E., Alves, T., Ferreira, V., Hawkey, K., Reilly, D.: Enhancing mobile content privacy with proxemics aware notifications and protection. In: Proc. CHI'16. ACM, New York, NY, USA (2016)
83. Zurita, B., Bosque, S., Fuertes, W., Macas, M.: Social engineering shoulder surfing attacks (ssas): A literature review. lessons, challenges, and future directions. In: Intl. Conf. on Advanced Research in Technologies, Information, Innovation and Sustainability. Springer (2023)