Investigating the Third Dimension for Authentication in Immersive Virtual Reality and in the Real World

Ceenu George* LMU Munich
Daniel Buschek † LMU Munich
Mohamed Khamis‡ University of Glasgow LMU Munich
Heinrich Hussmann§ LMU Munich

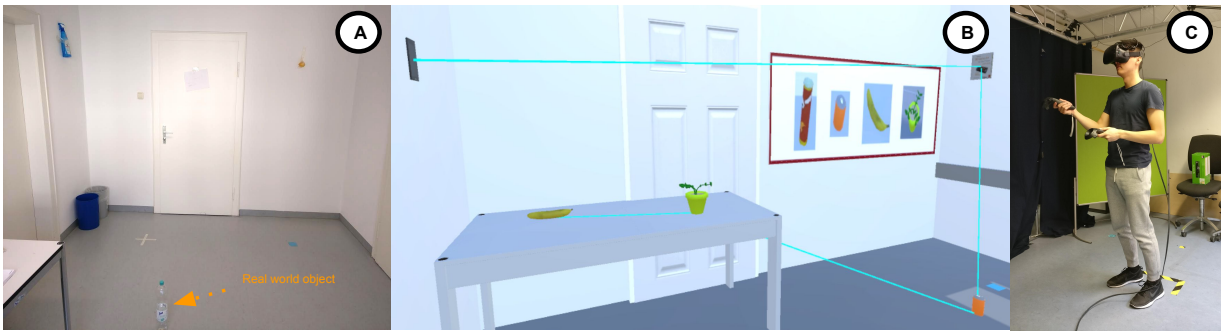

Figure 1: We study how the third dimension can be leveraged to improve the usability and security of authentication. (A) shows the sample real world room used for the study. (B) is a replica of the real world room. This screenshot depicts a view of the virtual scene from the view-point of the user during authentication process. When more than one object is selected, a blue connecting line appears. (C) is the view from the real world during user authentication in the virtual scene with a HMD, namely the HTC Vive [22].

Abstract

Immersive Virtual Reality (IVR) is a growing 3D environment, where social and commercial applications will require user authentication. Similarly, smart homes in the real world (RW), offer an opportunity to authenticate in the third dimension. For both environments, there is a gap in understanding which elements of the third dimension can be leveraged to improve usability and security of authentication. In particular, investigating transferability of findings between these environments would help towards understanding how rapid prototyping of authentication concepts can be achieved in this context.

We identify key elements from prior research that are promising for authentication in the third dimension. Based on these, we propose a concept in which users' authenticate by selecting a series of 3D objects in a room using a pointer. We created a virtual 3D replica of a real world room, which we leverage to evaluate and compare the factors that impact the usability and security of authentication in IVR and RW. In particular, we investigate the influence of randomized user and object positions, in a series of user studies (N=48). We also evaluate shoulder surfing by real world bystanders for IVR (N=75). Our results show that 3D passwords within our concept are resistant against shoulder surfing attacks. Interactions are faster in RW compared to IVR, yet workload is comparable.

Index Terms: Human-centered computing—User studies—; Human-centered computing—Virtual reality—

1 Introduction

Head mounted displays (HMD) allow users to experience immersive virtual reality (IVR) at their leisure. As users start spending more time using HMDs, storing personal data on the devices (e.g., credit card credentials), and using them for social interactions, the need for seamless authentication in IVR becomes increasingly important. Our vision for seamless authentication entails entering passwords when needed during the IVR interaction, such as upon buying a shopping item. This stands in contrast to all-in-one solutions, such as entering a password before using the HMD, as previous research has found the latter to be a poor fit for users' preferences [14]. This is especially the case when considering (1) the wireless future of HMDs as self-contained devices with no additional input hardware, such as a keyboard and external monitor, for example the Oculus Go [24], and (2) that taking the headset off leads to a break of immersion and presence, which would diminish one of the greatest strengths of this technology.

Similarly, smart homes in the real world (RW), provide an immersive environment, where users can authenticate in 3D. For example, consider a person entering a room and enabling all ubiquitous technologies within that room/house by selecting a number of tracked or digital objects. We regard this as the next step towards embedding authentication into our natural environment.

Contrary to prior work, which focused on transferring 2D concepts (e.g. PIN) to IVR [10] and smart homes in the RW [15], we investigate the third dimension for authentication with two research questions:

R1 *How can the usage of special properties offered by the 3D environment improve usability and security of authentication?*
The 3D environment provides an opportunity to increase usability, by making the authentication concept part of the immersive world, and to improve security, by utilizing virtual 3D objects as passwords, which makes it more difficult for a real world bystander to observe them. Thus, we see a clear need to investigate the third dimension for authentication.

R2 *Can the concept of using 3D objects for authentication be transferred from VR to a real world setting, for example for smart homes?*
Understanding the transferability of findings between these environments could help with realizing rapid prototyping of future authentication concepts.

*e-mail: ceenu.george@ifi.lmu.de
†e-mail: daniel.buschek@ifi.lmu.de
‡e-mail: mohamed.khamis@glasgow.ac.uk
§e-mail: heinrich.hussmann@ifi.lmu.de

For example, if a researcher wants to evaluate authentication via mid-air gestures in a smart home, they would have to prototype a smart home environment in their lab. Prototyping this setup is not only costly, but also fundamentally limited in creating the conditions necessary to understand all aspects that influence the usability and security of authentication schemes. On the other hand, by using IVR for high-level prototyping, researchers can experiment with conditions that are infeasible to replicate in the real world e.g., evaluating observation resistance from all possible angles.

These reasons underline the need to investigate how interaction in the virtual 3D environment differs compared to that of a real environment.

To this end, we first propose novel interaction concepts for authentication in the third dimension by applying findings from prior work to IVR. To test these concepts in the real world and IVR, we created a virtual 3D replica of a real world room. We evaluated and compared these concepts in a series of user studies based on their usability ($n = 48$) and memorability[1] ($n = 27$) in both environments. We also investigate observation resistance (immediate observation attacks $n = 15$, post-hoc observation attacks $n = 36$) for IVR.

Our analyses show that leveraging the third dimension for authentication in IVR increases resistance against observation attacks. While interactions took significantly longer in the virtual world compared to the real world, we found no differences in workload, which implies that users' perception of difficulty is not influenced by the environment, whereas their performance is. Our findings are valuable for researchers and practitioners who design authentication concepts for virtual reality and for the real world.

To summarize, we investigate (R1) how using the third dimension that IVRs offer can be leveraged when designing interactions for authentication purposes and (R2) whether the concept to authenticate with 3D objects can be transferred to real world settings.

2 Related Work

We build on: 1) Authentication Concepts that can be suitable for IVR, and 2) IVR features that are relevant to authentication.

2.1 Authentication Concepts that are Promising for IVR

Virtual reality environments and the usage of HMDs provide distinctive features (e.g., limitless space for 3D interaction) that have not been investigated before in the context of authentication. In this section we focus on a) authentication concepts that are promising to apply in virtual reality, and b) features of IVR that can be leveraged to improve authentication.

2.1.1 Knowledge-based Authentication

Knowledge-based authentication is based on something the user "knows", and can be classified into recall-based and recognition-based authentication [26]. These authentication concepts present a promising solution for authentication in virtual reality: In the real world, observation by attackers ("shoulder surfing") has been identified as one of the main drawbacks of such methods [3,8]. In contrast, IVR offers a 'secret channel between the user and the system' [10] that is not visible to attackers in the real world. This channel thus seems ideal for password cues that are only visible to the user. We leverage this advantage to improve resistance to shoulder surfing in IVR, to improve over previous work on authentication using PINs and patterns in IVR by George et. al [10].

Yadav et. al [34] explored knowledge based systems (e.g. PIN) in AR systems which were perceived to be usable but entry times were high (8s-14s on Google Glass). In contrast to AR, where for example attackers obtain visual cues from glass reflections, the combination of non-observable visual cues and mid-air interactions has not been explored in prior research to our knowledge.

[1] Memorability refers to how well users can memorize passwords [26]

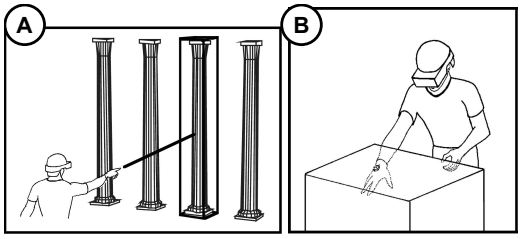

Figure 2: Adapted from Mine et al. [20], interaction in IVR can be performed in two different ways: Controllers can act as as a virtual pointer (A), whereby a laser beam is projected into the virtual scene, and as virtual hands (B), imitating touch interaction. Previous work revealed pointer interaction to be preferable for authentication [10].

Another way to overwhelm observers when authenticating via knowledge-based schemes is to randomize the input cues. This was leveraged in some previous systems such as PassFace [5], V-Go [17] and keyboard scrambling [32]. To the best of our knowledge, we are the first to report on an authentication randomization scheme for 3D environments.

2.1.2 Biometric Authentication

Android and iOS require setting a fall back knowledge-based authentication method (e.g., PIN or pattern) when using biometric authentication such as Touch ID or Face ID. This is due to the limitations of biometric authentication in terms of not being feasible all the time e.g., improper lighting conditions for face detection or wet/greasy fingers that distort the fingerprint [11]. Moreover, not all users want to share biometric data with third parties [23], especially after it has been shown multiple times that they can be remotely stolen from companies [29]. Hence, although biometric authentication concepts are a valid alternative for authentication in IVR and RW, this paper focuses on knowledge-based solutions that will continue to exist despite the popularity of biometric ones.

2.2 Features of IVR that are Relevant to Authentication

Many features of IVR, such as high memorability, plethora of input and design possibilities, immersion and presence, can be leveraged for designing usable and secure authentication schemes.

2.2.1 Memorability in Virtual Reality

Taiabul et al. [12] tested two memory techniques in virtual environments: the method of loci and the link method. Both leverage spatial memory which supports humans in recording their surroundings, forming spatial orientation and awareness [4, 19, 21]. In contrast to prior work, we empower the user to apply the method of loci in an immersive 3D environment, where users' are placed in the center of the room and surrounded by (virtual and real) 3D objects.

2.2.2 Input and Output Features

HMD controllers offer a vast amount of interaction modalities. Several buttons and a touchpad may be used on the hardware itself. They may be programmed to act as virtual hands, imitating touch interaction, and as a virtual pointer, projecting a "laser beam" into the virtual scene (Fig. 2) [20]. Previous work revealed pointer interaction to be preferable for authentication [10], which is the method of choice for our test environment.

The options to display visual elements and cues are limitless in IVR. Feedback can be displayed within the virtual scene in any size and shape. It was found that authentication by selecting targets that are at a far distance is easier than when targets are close by [10]. Attention should be paid to the amount of information being displayed in order to not negatively affect cognitive load.

Table 1: Comparison of post-hoc vs. intermediate attack for the security studies.

	Post-hoc vs. Immediate Threat Model	
	Post-hoc Attack	**Immediate Attack**
Attack mode	Video Attack	Live Attack
Tools	Video of victim authenticating	Access to VR device and software
	Pen and paper	Pen and paper
Attack model	1. View video of RW authentication	1. Interact with *RoomLock* in VR
	2. View video of VR authentication	2. Observe user live in-situ
	3. Guess password on pen and paper	3. Guess password on pen and paper
Attack opportunity	Unlimited view of authentication videos	View authentication three times

2.2.3 Immersion and Presence in Virtual Reality

Experiences in virtual reality aim to imitate interactions in the real world, which is also reflected in the way these environments are evaluated. Immersion and Presence are two of the key terms used when describing how well a virtual environment is perceived by users. This paper follows Slater et al. [25, 28]: immersion is quantified by quality of the technology, and presence is the users' subjective perception of how real the virtual world is compared to the real world. We consider these aspects in several ways: First, high-end HMDs (e.g., HTC Vive [22]) establish a high level of immersion in our studies. As we established earlier, this can serve as an advantage, as it provides a secret channel between the user and the system. However, based on prior work, which established that users' generally do not notice when being shoulder surfed [8], being immersed may further increase users unawareness of attackers. Second, presence is measured to understand how well the virtual world is perceived [2]. Finally, we study the interactions both in IVR and the real world.

Based on Legge et al. [18], this allows us to gain a deeper understanding of how spatial memory, awareness and interaction in these two environments relate to each other. Furthermore, we argue that findings from both worlds are necessary to understand whether virtual environments can serve as a testbed for usable security research in the real world, specifically for rapid prototyping purposes.

2.3 Threat Model

We illustrate the addressed threat models with two scenarios, whereby both start with the victim using their self-contained HMD, for example the Oculus Go [24], at home while friends and acquaintances are close-by. A possible scenario where the user needs to authenticate is when, for example, confirming a purchase made in a virtual store or an in-game purchase, or verifying the user's identity when logging into a virtual social network or a player's account.

2.3.1 Post-hoc Attack

The user cannot see the real world, hence they do not notice that a bystanders (the attacker) is recording the user as they authenticate. The recording covers the whole room in the RW, including the authentication process. The attacker watches the recordings later to recreate the password. Once the password is found, the attacker could use the observed password to make in-app purchases by, for example, logging into the user's account from a different HMD, or getting hold of the user's HMD while it is unattended.

2.3.2 Immediate Attack

Shortly after authenticating, the user takes off the HMD and leave it temporarily unattended e.g., to grab a glass of water from the next door room. One of the bystanders in the room picks up the headset and continues with the game. When prompted to authenticate in order to do an in-app purchase, attacker enters the password they had just observed.

3 An Example 3D Authentication Concept

To test the above mentioned concepts from prior work that would benefit from the virtual reality setting, we developed an authentication scheme called *RoomLock*.

We implemented it using Unity 3D with C#. A HTC Vive controller (Fig. 3) is used for virtual pointing [2].

3.1 Overview

To meet the needs of a knowledge-based authentication concept, users authenticate by pointing at a pre-defined number of stationary objects in a virtual room, in a specific order (Fig. 1). In our prototype and study, a password consists of a list of 4 objects, and a total of 9 objects were available for selection (Fig. 4). The limitless 3D space in virtual reality, allows both password length and set of objects to be easily extended. Due to the novelty of the authentication concept in immersive 3D environments and to limit the time required for introducing the study to participants, we chose objects based on the assumption that they are well-known to anyone joining our study, rather than choosing digital products for smart homes.

3.2 Input Method and Feedback

Object selection requires pointing with a laser and a button press to avoid unwanted selections of objects placed in the same visual path. Upon selection of two or more objects, visual feedback in the form of a blue connecting line appears in the virtual scene. The concept allows for the same objects to be selected multiple times but not

[2] Note, at the point of the study, self-contained devices, such as the Oculus Go [24], were not yet available, thus the study used the HTC Vive but assumes the interaction without the desktop and keyboard being available.

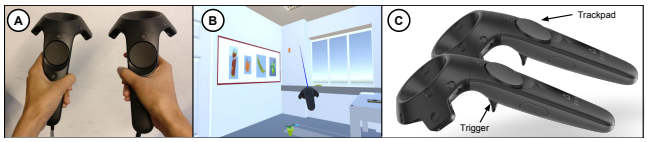

Figure 3: (A) displays the view of the HTC Vive controller [22] in the real world, whereas (B) shows how the controllers are displayed in the virtual world. There is a one to one mapping between the real and virtual world, hence all movements are observable in real-time from the real world. The details of which buttons are used for interaction during password entry can be seen in (C).

Figure 4: The first row displays 3D models of the real objects in the second row. Objects in the first row were used for authentication in $Environment_{virtual}$ whereas the once in the second made up the password in $Environment_{real}$.

consecutively. Further, the blue connecting line changes its colour when an authentication attempt is completed, turning green if it was successful, or red otherwise.

Haptic feedback is included to meet the needs of the previously proposed password space: The controller vibrates briefly two times on correct input; and slowly two times on false input.

3.3 Error Handling

Error handling followed pattern entry on smartphones: Users could not correct individual object selections [10,30], but had to start anew if they made wanted to change a selection.

3.4 Real versus Virtual World

In order to test the 3D authentication concept in the real world, for example for smart home purposes, the virtual 3D room used for *RoomLock* in our study replicated a real world room at our institute. The concept and system are flexible and could be easily extended with different virtual rooms that do not require a RW equivalent.

A physical laser pointer was used for pointing in the real world. Due to the limitations of the real world, there was no connecting line between objects upon selection. However, similar to the virtual version, passwords consisted of 4 real objects from a total of 9 possible options (Fig.4).

4 User Studies

RoomLock was evaluated in three parts: (1) a lab study ($N = 48$) to test the usability of *RoomLock*, and to understand how the interaction compares to the real world; (2) a follow-up questionnaire ($N = 27$) to gain insights into password memorability; and (3) two security studies ($N = 15$ and $N = 36$). Security was solely assessed for the IVR environment, in two separate lab studies to test the immediate vs. the posterior threat models (Table 1) for shoulder surfing. All studies adhered to ethical research standards within our institution.

Table 2: Overview of studies completed as part of the main study in chronological order.

Type of study	Participants
Usability study (RW vs. IVR)	$N = 48$
Memorability study	$N = 27$
Security study I - Post-hoc attack	$N = 15$
Security study II - Immediate attack	$N = 36$

4.1 Variables

Independent and dependent variables are consistent for all studies evaluating *RoomLock*. 2.

4.1.1 Independent Variables

Position explores the impact of randomization of the authentication procedure: either no randomization ($Position_{\text{baseline}}$), or varying users' starting position ($Position_{\text{user}}$), or varying the position of the objects ($Position_{\text{object}}$). In $Position_{\text{baseline}}$, users start at a fix point in the virtual scene. In $Position_{\text{user}}$, the starting position is randomized for every authentication. But in all three conditions the users' field of view is always on the starting object (albeit from varying angles.)

Repetition was another variable as participants entered passwords in each position three times.

Environment differentiates between participants that completed the study in the real world (Control group: $Environment_{\text{real}}$) and virtual reality (Test group: $Environment_{\text{virtual}}$).

4.1.2 Dependent Variables

Entrytime measured the time taken to enter a password. In $Environment_{\text{virtual}}$ the time was tracked from selection of the first object until the last one. In $Environment_{\text{real}}$ the end of a password entry was verbally communicated upon selection of the last object to imitate the virtual setup.

Error counted the number of times a password was entered incorrectly. It was tracked by visual inspection in $Environment_{\text{real}}$ and automatically in $Environment_{\text{virtual}}$.

We also measured cognitive load with a NASA TLX questionnaire [13] and conducted a focus group to capture the perception of presence in VR [27].

4.2 Usability Study

We used a mixed model design with a between subjects variable *Environment* and a within subjects variable position.

4.2.1 Procedure

We recruited 48 participants (14 female) through our University mailing list. The average age was 23.71 (SD = 3.26). 33% had no prior experience with VR. Participants were compensated with a 10 EUR voucher for an online shop or institute internal credit.

Passwords were randomly generated from the list of available objects prior to the start of the study. However, the same set of passwords was used for each *Environment*. Each participant entered each password three times.

Virtual world: Participants were then introduced to the project and hardware (HTC Vive [22]). They put on the HMD with controllers, placed in the middle of the virtual room, facing a whiteboard (Fig.1). This starting position was the same for $Environment_{\text{real}}$ and $Environment_{\text{virtual}}$.

The interaction in the virtual world was introduced by a training session, which consisted of entering a pre-defined password. This password appeared on the whiteboard in the virtual room and participants were verbally directed by the experimenter on how to enter it with the HTC Vive controllers. To complete the training session they had to enter the password three times correctly.

Subsequently, participants completed three different rounds of password entry with three repetitions each (*position* x *Repetition*), whereby *position* was counterbalanced. Errors were tracked and participants had to repeat their entry until they had provided three correct password repetitions, before moving on to the next condition. Therefore, each participant entered 9 correct passwords (3 repetitions $\times$ 3 repetitions). A NASA TLX questionnaire and one for demographics concluded the session.

Real world: Participants were given the same introduction to the project as in $Environment_{\text{virtual}}$ and started at the same position in front of the real world whiteboard.

They were instructed on how to use the physical laser pointer and asked to complete a training session by correctly entering a pre-defined password three times. The password was provided on the real whiteboard.

As there was no feedback available from the physical laser pointer, we asked participants to count out loud whenever they chose an object with the laser pointer. Similarly, they were asked to say "finished" at the end of their entry. This allowed us to track errors and check that the correct objects were selected in the correct order. The whole procedure was demonstrated prior to starting the study.

The study was completed as in the virtual world: Participants entered given passwords for all three conditions of *position* with three repetitions. Finally, they completed the NASA TLX and demographics questionnaires.

4.2.2 Results

Entry Time in VR: We confirmed that our data was normally distributed within the two different environments by visual inspection of the normal distribution curve for both. There were no outliers. Mauchlys Test of Sphericity did not indicate that it had been violated. A repeated measures ANOVA determined that *position* had a significant effect on entry time ($F_{2,40} = 14.44$, $p < 0.05$). Post-hoc analysis showed a significantly higher entry time in $Position_{\text{object}}$($Mean = 14.33$, $SD = 0.83$) compared to the other two

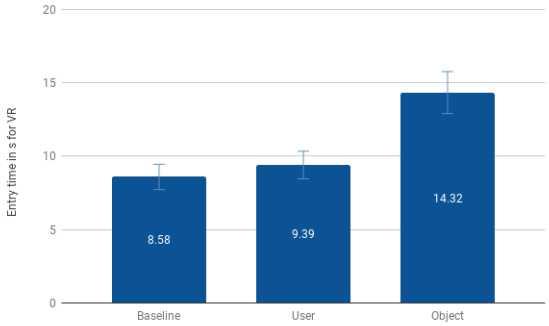

Figure 5: A significantly higher entry time in $Position_{\text{object}}$ was found compared to the other two *position* conditions, namely $Position_{\text{baseline}}$ and $Position_{\text{user}}$.

position conditions ($Position_{\text{baseline}}(Mean = 8.58, SD = 1.24)$ and $Position_{\text{user}}(Mean = 9.4, SD = 0.71)$) (Fig. 5).

Results showed that entry time differed significantly depending on $Repetition(F_{1.6,36.5} = 7.94, p < 0.05)$ for all *position*. Post-hoc analysis revealed that the first round $(Mean = 12.71, SD = 1.1)$ was significantly slower $(p < 0.05)$ than in the second $(Mean = 10.21, SD = 0.73)$ or third $(Mean = 9.39, SD = 0.64)$ try.

However, a closer look at *Repetition* for each individual $position(F_{3.45,79.37} = 2.9, p < 0.05)$ revealed that participants only showed a positive learning curve for $Position_{\text{baseline}}$ and $Position_{\text{user}}$, on average they were 43% slower in round 1 and 2 than 3. In contrast, participants in $Position_{\text{object}}$ showed no learning effect.

Error per Entry in VR: We also logged error, which is a binary value that captured whether users entered a correct (true) or incorrect (false) password. Only 2 errors occurred in 72 entries.

Subjective Feedback in VR: Based on additional questions asked at the end of the usability study, we found that the connecting line (Fig.1) was not perceived to be distracting (Median=2, 1="not distracting at all") confirm this.

Participants were asked to speak out loud if they had feedback after every password entry. This qualitative feedback revealed that participants perceived the interactions to be intuitive. Moreover, 42% remarked that objects which were not in their FOV, especially the once on the ground or close to the ceiling, were harder to find.

Comparison VR and Real World: The between subjects analysis with a Shapiro-Wilk test on *Environment* revealed that our data was not normally distributed $(p < 0.05)$. A Mann-Whitney test indicated that the entry time was greater in $Environment_{\text{virtual}}$ (Mean = 10.77s) than in $Environment_{\text{real}}$ (Mean = 6.97s, U = 14145, $(p < 0.05)$). When comparing the paired conditions between both variables, all pairs show a significant $(p < 0.05)$ higher entry time in $Environment_{\text{virtual}}$ than in $Environment_{\text{real}}$ (Table3).

There were five errors out of 144 password entries. Two occurred in $Environment_{\text{virtual}}$ and three in $Environment_{\text{real}}$. One error in $Environment_{\text{real}}$ was out of frustration, as the participant could not find an object hence a wrong password was forcefully entered. Three of the remaining errors were due to incorrect order entries. One participant in$Environment_{\text{virtual}}$ mistook the spray bottle for the water, which shares a similar colour scheme (Fig.4). According to the conducted Sign Test no significant $(p > 0.05)$ difference was found in the NASA-TLX scores between both world conditions.

4.3 Memorability Study

To gain insights into how memorable our passwords were, we asked participants from our usability study to complete a follow-up questionnaire one week after the first study (resulting in $N = 27$, female = 12). Participants were not informed about the nature of the follow up questionnaire, indicating that there was no memorability bias. As compensation, they took part in a raffle for a 10 EUR online shop voucher.

Table 3: A Mann-Whitney test indicated that the entry time was greater in $Environment_{\text{virtual}}$ than in $Environment_{\text{real}}$.

Position	Mean RW	Mean VR	U	p
Baseline1	7.57s	11.79s	176	0.021
Baseline2	5.47s	7.57s	182	0.029
Baseline3	4.84s	6.39s	154	0.006
User1	7.17s	12.15s	156	0.006
User2	5.4s	8.8s	136	0.002
User3	5.26s	7.25s	151	0.005
Object1	9.74s	14.19s	137	0.002
Object2	8.42s	14.27s	117	0.001
Object3	8.87s	14.53s	117	0.001

11 participants from $Environment_{\text{virtual}}$ and 16 from $Environment_{\text{real}}$ completed the questionnaire.

4.3.1 Procedure

The questionnaire consisted of three parts: Firstly, participants were asked to remember their passwords and to enter the names of the four objects in the right order without any cues. Secondly, they were provided with the 9 possible options (Fig.4) and given the opportunity to change their initial entry or to confirm the input. Finally, they had to create their own password out of the available objects, to investigate what type of passwords they would create based on their experience so far.

4.3.2 Results

The results of the memorability study ($N = 27$) indicate that *RoomLock* passwords are memorable, especially with cued-recall. One participant had to be excluded as their results clearly indicated that they did not read the instructions. Results are ordered by cued versus non-cued recall. The questionnaire was organized in the same order.

Non-cued recall: Twentyone participants memorized their passwords 100% correctly. Four participants failed to remember their passwords one week after completing the usability study. Three of these participants had the majority of objects correct but not in the correct order and only one participant failed to remember any objects. Another two remembered the password with only one object each not being in the correct order.

Cued recall: One participant changed their input after the images were displayed, which resulted in an incorrect entry as the initial password was correct. Another participant that had given a wrong password in the first non-cued recall round, changed their passwords to the correct one. They changed their input from 'something green' to the 'plant' object (Fig. 4). Thus, after the cue-recall round, there were still twentyone participants who memorized their passwords correctly.

Observations on Memory Techniques and Password Choices: Based on additional questions that we asked at the end of the questionnaire we made the following observations: (1) Without nudging participants to use specific memory techniques, they naturally used the link technique [12]. Overall passwords showed the tendency to tell a story, such that the choice of 'chips' and 'water' objects can be interpreted as 'I am eating chips, I am thirsty.'. (2) None of the participants chose duplicate objects in their own passwords.

4.4 Security Study I - Post-hoc Attack

We recruited 15 Participants (7 female) based on their participation in the usability study ($N = 7$) and from a university mailing list ($N = 8$). Their average age was 23.13 (SD=3.58). Participants were asked to provide their experience with VR on a likert scale (5=no

experience), the median was 4. Participants were rewarded with a 5 EUR voucher for an online shop or an institute internal credit.

4.4.1 Procedure

Participants answered a demographics questionnaire. Then they were provided with a sheet of paper that showed multiple printed images of the 3D room.

We randomly chose three people recorded in the usability study in condition $Position_{\text{baseline}}$. Each of these recordings was used for an "attack" in the following three steps:

First, participants watched the video that showed the interaction viewed from the real world, from the best possible angle for a video attacker. The angle was the result of a trial and error validation by the experimenters and showed the victims gestures from the view point of someone hovering over and slightly to the left of the victim (e.g. drone) . They could re-watch it as many times as they wanted.

They then watched a video of the virtual scene, to help them understand where objects were placed in the virtual room.

In the third round participants were allowed to re-watch both videos from the first two rounds.

After each of these three rounds, participants used the printout of the 3D room to indicate their guess of the pointing locations of the observed user. In addition, they provided a Likert-scale rating on perceived difficulty.

Our plan was to repeat the the above mentioned procedure for all conditions of our independent variable *position*. However, after the first iterations, participants were demotivated and started giving random guesses to complete the study. They commented that it was too difficult to guess the passwords in $Position_{\text{baseline}}$ and did not see any value in trying for $Position_{\text{user}}$ and $Position_{\text{object}}$.

4.4.2 Results

Taking the best guess from all three rounds into consideration, in $Position_{\text{baseline}}$, without changing user and object position, the video attackers were overall not able to guess the password correctly: One participant was able to identify two objects of the attacked password but not in the correct order and one participants guessed two objects in the correct order. Half of the participants were able to guess one object without being able to put it in the correct order.

When changing $Position_{\text{object}}$ and $Position_{\text{user}}$, the video attackers were not able to provide a guess. The qualitative feedback revealed that they found it impossible to imagine themselves in the correct position in the virtual scene: (P3) "I have no idea in which direction she is looking at now." (P7) "Do I have to provide a guess? I have no clue but I can pick random points in the scene and hope that I am lucky." (P13) "Can you help me out by telling me what view of the room she has right now?"

The experimenter also observed that 2 attackers tried to observe the head movements of the victim in order to guess the password. (P3) "It would be great to see where they are looking at (...) maybe I can just look at the head (...) don't think that is really helpful either."

4.5 Security Study II - Immediate Attack

We recruited 36 participants for a within-subjects lab study from our institution ($N = 25$) and random selection on the street ($N = 11$). They had an average age of 26.16 (SD=7.22, 15 female).

4.5.1 Procedure

The study started with an introduction, which was concluded by signing the consent form. Subsequently, they completed a demographics questionnaire and were introduced to the hardware. They were asked to do a training round in IVR to familiarize themselves with the hardware and *RoomLock*. Training was deemed successful once they entered a pre-defined password correctly three times in a row. For the main part of the study participants had to act as attackers in the real world whilst observing an expert user entering passwords in *RoomLock*.

To test the effect of a *position* change, the first password was entered in $Position_{\text{baseline}}$, the second and third from two random $Position_{\text{user}}$. (Due to the negative feedback for $Position_{\text{object}}$ in the usability and posterior security study, it was not included as a variable.) Thus participants observed three pre-defined passwords and after each observation, up to three guessing trials were possible.

to simulate a prepared attack, the attacker was provided with a pen and paper and encouraged to (1) draw the virtual room and (2) take notes during observation. They were also told the start and end of the authentication gesture. After each attack, participants had to rate the difficulty of observation. Finally, they were asked to put on the headset again in order to set their own password of choice. The study was concluded by a semi structured questionnaire. Participants were rewarded with an Amazon voucher worth 7.50 EUR.

4.5.2 Results

Overall, in contrast to the post-hoc attacks (0% success rate), this study showed that immediate attacks are more successful (12.5% success rate).

Visual inspection confirmed that all data was normally distributed. Binary results highlighted that after the first attack in $Position_{\text{baseline}}$ 18.5% of the passwords were guessed correctly and only 19.4% were not guessed at all. Randomizing user position within the virtual room led to a shoulder surfing rate of 12.5% after the second attack.

A one-way ANOVA ($F_{2,105} = 5.3$, $p < 0.05$) revealed a significant effect depending on changes in $Position_{\text{user}}$ on attack success. This was supported by the results from the semi structured questionnaire, where 31% stated that changing position made observation more difficult. An additional 19.45% noted that closeness of virtual objects increased observation difficulty.

In regard to observation tactics, 35 participants took notes, whereby 28 drew a 2D model of the virtual room with the pen and paper that were provided (Fig.6). Further analysis of the drawings revealed that virtual objects were represented in form of numbers (71.4%) or names (68.6%) and the order was noted with arrows (57.1%).

The security study was completed by asking participants to set their own passwords within *RoomLock*. The analysis revealed that two objects (banana and plant) which were placed on a table close to each other, rather than the floor or wall, were chosen 38% of the time. Similarly, we found that object proximity and placement within the same FOV also influenced individual password setup. 48% of participants chose to repeat a maximum of 2 objects - not consecutively - and the average length was 6.5. All participants changed their FOV at least once to select objects during password setup.

4.6 Limitations

The lack of feedback in the real world during pointing made it difficult to create the exact same experience in both worlds. In the real world the laser is only visible on the hit object, rather than continuously as in virtual reality. Moreover, there was no confirm button on the real world laser pointer. Nonetheless, it can be argued that these inconsistencies between the worlds did not affect the authentication experience, as results revealed the real world interaction to be more usable despite the lack of feedback.

Training effects were not tested as part of our study but they were already visible during the repetitions in our usability study. Thus, we believe future work, for example in form of a long lasting field study, will reveal further improvements to entry time.

We used a university mailing list for recruitment – our sample thus only reflects a certain demographic. Students are believed to be more technology-savvy compared to the general population. However this is one of the main target groups for this concept.

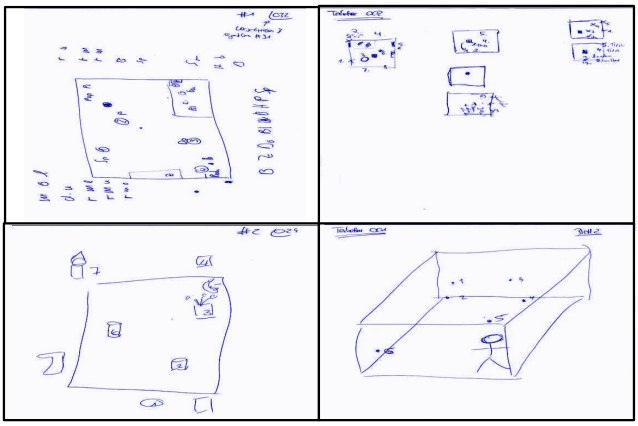

Figure 6: Four samples of participants' notes, which were taken during the security study investigating the immediate threat of *RoomLock*. The majority of participants drew a top view of the virtual room.

5 Discussion

We adapted prior concepts from authentication research that are promising for interaction in a real and virtual immersive 3D environment and investigated their usability, memorabilty and security. In addtion, we evaluated the transferrability of findings between IVR and RW and compared the results.

5.1 Leveraging the Third Dimension

5.1.1 Object Placement affects Entry Times

Placement and size of objects affects entry times, as we observed that objects placed in the center of the FOV were found quicker than ones at the edges. This confirms previous results by Arthur et al. [1], who evaluated the perception of spatial layouts in VR.

5.1.2 User Preferences during Password Setup

During password setup users preferred objects that were close to each other, which is also advantageous from a security perspective, as close proximity decreases susceptibility to shoulder surfing. Additionally, the majority of users changed their FOV at least once to find and select more objects within the virtual room which increases the practical password space which in turn improves security.

5.1.3 Password Space and Representation

The design of our concept is closely aligned with the authentication concept of *patterns* [30], as *RoomLock* passwords can, most straightforwardly, be defined as a list of object identifiers. However, they could also be described as gestures, via the lines drawn between objects. Moreover, we could use a list of object coordinates (x, y, z). Although, there are similarities to patterns, the third dimension adds another level of complexity, which results in an increased password space. In this initial investigation, we designed the password to have a total length of 4 objects, however, to increase the the password space [31] this should and can be easily extended.

It would be interesting to observe participants' behaviour regarding password creation in a 3D environment and whether differences exist compared to 2D pattern creation. Password creation in the third dimension might inspire additional concepts (e.g. users memorise object locations instead of specific objects), and might increase the (theoretical) password space. A closer treatment is beyond the scope of this paper, yet presents an interesting direction for future research.

5.1.4 Single vs. Shared Usage

We believe *RoomLock* has the potential also to act as a group authentication scheme by leveraging the notion of shared rooms and objects. Das et al. [6] highlight the need for socially-inclusive group authentication to access shared resources (e.g. equipment such as commonly used HMDs), instead of using individual secrets which may be perceived as 'rude or inappropriate'.

From our point of view the future vision for HMDs is that they are accessible both as single- and multi-user devices, similar to the usage pattern for tablet devices or apps, such as Netflix. The concept of shared authentication allows for example parents to restrict certain content from their children or companies to enable shared access to specific content.

In the first example, a common room, such as the family living room may be the room in which the objects are placed. The objects itself could be chosen by all family members who may need access, whereby the children use a subset of the objects enabling the unique identification of individual users whilst still maintaining a shared password. Although the concept of *RoomLock* allows this extension, further studies need to evaluate whether users perceive this to be usable when tested in a group context.

5.1.5 Improving Usability through the Third Dimension

At first glance, our results indicate that users take longer to complete a 4-object password in the third dimension - compared to 2D password in IVR [10] or 2D patterns in the real world [30]. However, the repetition data highlights that they were nearly twice as quick in the third round, even when changing user position, which implies a steep learning curve for authentication in 3D. Additionaly, error rates are lower than 2D real world results [30]. To summarize, our data suggests that usability in 3D is comparable to established authentication systems, however, further studies over longer periods of time, will have to confirm whether entry times can be improved.

5.2 Using the Virtual World to Aid Memorability

5.2.1 *RoomLock* Passwords are Memorable

Our memorability study indicates that *RoomLock*'s passwords are well remembered. After one week, the majority of participants remembered their objects without wearing an HMD and being immersed in the virtual room.

5.2.2 Personalisation Opportunities

The opportunity to increase memorability through personalisation *RoomLock* is a key strength of our concept and presents an avenue for future work: We believe that the entry time can be further optimized through habituation as well as by personalizing *RoomLock*. Users could not only create their own individual passwords by choosing personal objects but also authenticate in familiar rooms (e.g a 3D model of their living room), rather than the one we tested in our study. Similarly, these objects may be adjusted to suit the use case. For example, smart home authentication may be tested better with digital objects, such as smart speakers.

5.3 Randomization to Increase Security

5.3.1 Third Dimension Decreases Shoulder Surfing Risk

The security study highlighted the value of seamless authentication in VR: Leveraging the third dimension for password entry drastically decreases the risk of shoulder surfing, compared to 2D authentication in VR (compare to [10]).

5.3.2 Third Dimension Hinders Immediate Attacks

The study on the immediate threat model investigated whether bystanders are able to attack the victim in real-time when entering the same virtual room. This may be the case with shared devices, where bystanders use the same HMD after the victim has taken it off. Shoulder surfing was proven to be less successful in our 3D concept *Roomlock* (12.5%) compared to previous 2D VR authentication studies by George et al. [10] (18%). Even without randomizing user position, whereby user's starting position is changed for each password entry, binary attacking success was comparable (18.5%).

5.3.3 Post-hoc Attacks Unsuccessful even without Randomisation

In the posterior threat model, even without randomising user and object positions, video attackers were not able to guess the passwords correctly. Since we thus reached "perfect" security for our tasks already in the baseline, we could not confirm whether changing user and object position increases security even further. But the usability study revealed that changing user position is as usable as keeping it constant, whilst changing objects was found to be cumbersome.

5.3.4 Methodology: Multiple Threat Models for IVR Auth.

Comparing both attack studies, immediate attacks were much more successful than post-hoc attacks. From a methodological perspective, this finding sheds light on the importance of testing VR authentication concepts for multiple threat models. Our findings suggest that the difference in shoulder surfing rate between posterior vs immediate threat is due to attackers having experienced *RoomLock* in IVR for the latter. This enabled them to form a more sophisticated mental model of the virtual room, which is confirmed by the drawings they made during the security study testing the immediate threat model.

Another interesting finding was the use of head movements as gestures that reveal additional information about the virtual interactions [16]. This idea was raised by the attackers during the security study rather than the users of *RoomLock*. Future attack studies could thus focus on head movements to infer the placement of objects within the virtual room.

5.4 Design Recommendations for RoomLock

Based on our results and discussion, we summarize the following recommendations:
(a) Objects should be placed in the FOV of the user and close to each other to decrease input times and increase security.
(b) Leveraging depth in 3D environments provides similar security and usability as established authentication schemes, but user randomization drastically improves security without reducing usability.
(c) To maintain similar entry times to prior work, use 4 out of 9 objects to make up the 3D password.
(d) Objects should be familiar to the user and it should be possible to create a story with them to support memorability of passwords.

5.5 IVR as a Testbed for Usable Security Research

5.5.1 Transferability between RW and VR

Although, there were differences with regards to entry times, we found no significant differences in workload between the real and virtual world for *RoomLock*. This suggests that users' perception of difficulty is not influenced by the environment, whereas their performance is.

We argue that the workload similarities are due to the intuitiveness of interactions in *RoomLock*. This is supported by the qualitative feedback provided by participants during the usability study, which revealed that the interactions were perceived to be natural without needing further training. Prior work refers to this phenomenon as the theory of intuitive interactions, which also states that it will improve in usability over time [7].

Although, entry times were different in both environments, we argue that the advantages that IVR as a testbed for real world smart home authentication concepts offers, outweighs this difference. In addition to the points mentioned above, the study setup and procedure took drastically longer for the real world than IVR and considering time for completion is a main factor of rapid prototyping, we believe IVR to be a valid testbed for authentication concepts.

5.5.2 Alternative Approaches for Smart Home Authentication

There are two approaches to transfer our findings to a working prototype for smart home authentication based on *RoomLock*: Firstly, the system could be built in such a way that only smart objects are used to create passwords, such as speakers or voice assistants. These systems would need to be extended with a solution to recognize users' pointing interactions (e.g. camera) and the user would need to be equipped with a pointing device (e.g. mobile phone and RFID or beacon technology) to activate these objects as passwords. Based on our findings and prior work [10], we recommend pointing rather than a touch solution, as the latter was found to be less time efficient. A second approach may be to track the whole room and all user interactions via a real-time video recognition system, similar to the virtual reality solution for *RoomLock*, which would allow for all objects within that room to be used for passwords. Although this may provide a broader password space, it is also less favourable from a usable security perspective due to the additional amount of data that is being tracked and the continuity of such a tracking system.

5.5.3 Methodology: Comparing Interactions in RW and VR

These findings show that measuring workload is an important part of the methodology of investigating usable security interaction in 3D environments. Differences in workload arguably lead to an intensified focus on the main task (e.g. authentication), whilst neglecting necessary secondary tasks (e.g. awareness of surroundings) [33] and therefore ignoring possible ongoing shoulder surfing attacks. This needs further exploration as optimizing the design for cognitive load is an important feature in 3D authentication schemes.

6 Conclusion and Future Work

In this paper, we presented *RoomLock*, an authentication scheme specifically designed for authentication in the third dimension. It exploits the virtual 3D space, allowing users to select objects in a virtual room to create a 3D graphical password. *RoomLock* is based on analysis of prior work and features that are relevant to authentication in 3D.

Results from a usability and security study indicate that *RoomLock* is comparably usable and memorable, however the key strength lies in its high level of security. In our post-hoc threat model, we found no risk to shoulder surfing attacks and the immediate threat model revealed an attack rate of 12.5%. Regarding transferability between the two environments, we found that although, interactions took significantly longer in the virtual world compared to the real world, there were no differences in workload, which implies that users' perception of difficulty is not influenced by the environment, whereas performance is.

An obvious next step would be to investigate additional parameters for *RoomLock*, such as object size, length of password, and effect of personalized objects on memorability and entry time.

Future work may also review whether the findings are applicable to a MR (Mixed Reality) and AR (Augmented Reality) device, such as the Holo Lens [9]. From a technical and user experience point of view, *RoomLock* can be adapted without changes; the only difference being the interaction with the hands as pointers rather than with the controllers. However based on our findings there may be differences in workload due to the parallel view of the real and virtual world. Arguably, shoulder surfing risk should decrease, as the victim has a view of the real world compared to the VR experience.

Acknowledgments

The authors wish to thank An Ngo Tien and the ATH Usable Security group, whose projects contributed towards this paper.

References

[1] E. Arthur, P. Hancock, and S. Chrysler. The perception of spatial layout in real and virtual worlds. *Ergonomics*, 40(1):69–77, 1997.
[2] W. Barfield and T. A. Furness. *Virtual environments and advanced interface design*, vol. 55. Oxford University Press on Demand, 1995.
[3] R. Biddle, S. Chiasson, and P. C. Van Oorschot. Graphical passwords: Learning from the first twelve years. *ACM Computing Surveys (CSUR)*, 44(4):19, 2012.

[4] N. Burgess, E. A. Maguire, and J. O'Keefe. The human hippocampus and spatial and episodic memory. *Neuron*, 35(4):625–641, 2002.

[5] P. Corporation. The science behind passfacestm for windows, 2005.

[6] S. Das, G. Laput, C. Harrison, and J. I. Hong. Thumprint: Socially-inclusive local group authentication through shared secret knocks. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pp. 3764–3774. ACM, New York, NY, USA, 2017. doi: 10.1145/3025453.3025991

[7] S. Diefenbach and D. Ullrich. An experience perspective on intuitive interaction: Central components and the special effect of domain transfer distance. *Interacting with Computers*, 27(3):210–234, 2015.

[8] M. Eiband, M. Khamis, E. von Zezschwitz, H. Hussmann, and F. Alt. Understanding shoulder surfing in the wild: Stories from users and observers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 4254–4265. ACM, 2017.

[9] M. Fitzsimmons. Hands on: Microsoft hololens review, March 2017, Lastchecked: 2017-05-18. `http://www.techradar.com/reviews/wearables/microsoft-hololens-1281834/review`.

[10] C. George, M. Khamis, E. von Zezschwitz, M. Burger, H. Schmidt, F. Alt, and H. Hussmann. Seamless and secure vr: Adapting and evaluating established authentication systems for virtual reality. In *Proceedings of the Network and Distributed System Security Symposium (NDSS 2017)*, USEC '17. Internet Society, 2017. doi: 10.14722/usec.2017.23028

[11] A. Goode. Bring your own finger–how mobile is bringing biometrics to consumers. *Biometric Technology Today*, 2014(5):5–9, 2014.

[12] S. M. T. Haque, M. N. Al-Ameen, M. Wright, and S. Scielzo. Learning system-assigned passwords (up to 56 bits) in a single registration session with the methods of cognitive psychology. In *Proceedings of the Network and Distributed System Security Symposium (NDSS 2017)*, USEC '17. Internet Society, 2017. doi: 10.14722/usec.2017.23034

[13] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909

[14] E. Hayashi, O. Riva, K. Strauss, A. Brush, and S. Schechter. Goldilocks and the two mobile devices: going beyond all-or-nothing access to a device's applications. In *Proceedings of the Eighth Symposium on Usable Privacy and Security*, p. 2. ACM, 2012.

[15] L. Kagal, T. Finin, and A. Joshi. Trust-based security in pervasive computing environments. *Computer*, 34(12):154–157, Dec 2001. doi: 10.1109/2.970591

[16] A. Kendon. Do gestures communicate? a review. *Research on language and social interaction*, 27(3):175–200, 1994.

[17] L. Ledbetter. Oracle buys passlogix, 2017.

[18] E. L. Legge, C. R. Madan, E. T. Ng, and J. B. Caplan. Building a memory palace in minutes: Equivalent memory performance using virtual versus conventional environments with the method of loci. *Acta psychologica*, 141(3):380–390, 2012.

[19] E. A. Maguire, E. R. Valentine, J. M. Wilding, and N. Kapur. Routes to remembering: the brains behind superior memory. *Nature neuroscience*, 6(1):90–95, 2003.

[20] M. Mine et al. Virtual environment interaction techniques. *UNC Chapel Hill computer science technical report TR95-018*, pp. 507248–2, 1995.

[21] R. Parasuraman and M. Rizzo. *Neuroergonomics: The brain at work*. Oxford University Press, 2008.

[22] N. Pino. Htc vive review, November 2016, Lastchecked: 2017-01-22. `http://www.techradar.com/reviews/wearables/htc-vive-1286775/review`.

[23] A. P. Pons and P. Polak. Understanding user perspectives on biometric technology. *Commun. ACM*, 51(9):115–118, Sept. 2008. doi: 10.1145/1378727.1389971

[24] P. Rubin. Review: Oculus go, January 2018, Lastchecked: 2019-02-07. `https://www.wired.com/review/oculus-go/`.

[25] M. V. Sanchez-Vives and M. Slater. From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, 6(4):332–339, 2005.

[26] F. Schaub, M. Walch, B. Könings, and M. Weber. Exploring the design space of graphical passwords on smartphones. In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, SOUPS '13, pp. 11:1–11:14. ACM, New York, NY, USA, 2013. doi: 10.1145/2501604.2501615

[27] T. W. Schubert. The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. *Zeitschrift für Medienpsychologie*, 15(2):69–71, 2003. `http://www.igroup.org/pq/ipq/index.php`.

[28] M. Slater and S. Wilbur. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and virtual environments*, 6(6):603–616, 1997.

[29] M. Stokkenes, R. Ramachandra, and C. Busch. Biometric authentication protocols on smartphones: An overview. In *Proceedings of the 9th International Conference on Security of Information and Networks*, SIN '16, pp. 136–140. ACM, New York, NY, USA, 2016. doi: 10.1145/2947626.2951962

[30] E. von Zezschwitz, P. Dunphy, and A. De Luca. Patterns in the wild: A field study of the usability of pattern and pin-based authentication on mobile devices. In *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '13, pp. 261–270. ACM, New York, NY, USA, 2013. doi: 10.1145/2493190.2493231

[31] E. von Zezschwitz, M. Eiband, D. Buschek, S. Oberhuber, A. De Luca, F. Alt, and H. Hussmann. On quantifying the effective password space of grid-based unlock gestures. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, MUM '16, pp. 201–212. ACM, New York, NY, USA, 2016. doi: 10.1145/3012709.3012729

[32] E. von Zezschwitz, A. Koslow, A. De Luca, and H. Hussmann. Making graphic-based authentication secure against smudge attacks. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, IUI '13, pp. 277–286. ACM, New York, NY, USA, 2013. doi: 10.1145/2449396.2449432

[33] C. Wickens, A. Kramer, L. Vanasse, and E. Donchin. Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science*, 221(4615):1080–1082, 1983.

[34] D. K. Yadav, B. Ionascu, S. V. K. Ongole, A. Roy, and N. Memon. Design and analysis of shoulder surfing resistant pin based authentication mechanisms on google glass. In *International Conference on Financial Cryptography and Data Security*, pp. 281–297. Springer, 2015.